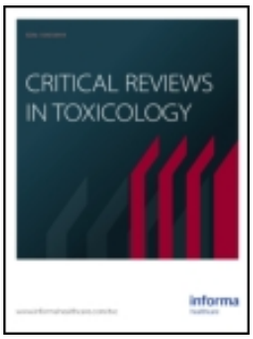

Taylor & Francis
Taylor & Francis Group

Critical Reviews in Toxicology

ISSN: 1040-8444 (Print) 1547-6898 (Online) Journal homepage: informahealthcare.com/journals/itxc20

Hypothesis-based weight-of-evidence evaluation of the neurodevelopmental effects of chlorpyrifos

Robyn L. Prueitt, Julie E. Goodman, Lisa A. Bailey & Lorenz R. Rhomberg

To cite this article: Robyn L. Prueitt, Julie E. Goodman, Lisa A. Bailey & Lorenz R. Rhomberg (2011) Hypothesis-based weight-of-evidence evaluation of the neurodevelopmental effects of chlorpyrifos, Critical Reviews in Toxicology, 41:10, 822-903, DOI: 10.3109/10408444.2011.616877

To link to this article: https://doi.org/10.3109/10408444.2011.616877

Critical Reviews in Toxicology, 2011; 41(10): 822–903

ISSN 1040-8444 print/ISSN 1547-6898 online
DOI: 10.3109/10408444.2011.616877

informa
healthcare

REVIEW ARTICLE

Hypothesis-based weight-of-evidence evaluation of the neurodevelopmental effects of chlorpyrifos

Robyn L. Prueitt, Julie E. Goodman, Lisa A. Bailey, and Lorenz R. Rhomberg

Gradient, Cambridge, Massachusetts, USA

Abstract

We used a hypothesis-based weight-of-evidence (HBWoE) approach to analyze the evidence regarding the hypothesis that chlorpyrifos can cause neurodevelopmental effects below the threshold for inhibition of acetylcholinesterase activity in the nervous system, which is an established mode of action for chlorpyrifos neurotoxicity. The epidemiology data do not consistently demonstrate associations between chlorpyrifos exposure and neurodevelopmental toxicity, and the animal toxicity data do not provide clear evidence that neurodevelopmental effects occur at doses below the threshold for acetylcholinesterase inhibition. The alternative mechanisms proposed to underlie potential neurodevelopmental effects in humans have been observed in the absence of acetylcholinesterase inhibition in a few *in vitro* studies but not in the developing brain *in vivo*. We provide perspective on the HBWoE approach compared with frameworks developed by the United States Environmental Protection Agency and the European Center for Ecotoxicology and Toxicology of Chemicals. We suggest that our HBWoE approach offers advantages over these frameworks in providing a better perspective on how to integrate all of the relevant data and how to use each line of evidence to inform the integration of other kinds of data or compare alternative hypotheses. Based on an HBWoE analysis, we conclude that a causal association between chlorpyrifos exposure and neurodevelopmental effects in the absence of acetylcholinesterase inhibition in the brain is not plausible in humans, and the few positive associations observed in epidemiology studies are most likely attributable to alternative explanations.

Keywords: Risk assessment, epidemiology, neurotoxicity, developmental toxicity, mechanism of action, acetylcholinesterase inhibition, pesticides, neurobehavior, cognitive and motor development, child behavior, ECETOC framework, US EPA framework

Contents

Address for Correspondence: Julie E. Goodman, Gradient, 20 University Rd, Cambridge, MA 02138, USA. Tel: 617-395-5000. E-mail: jgoodman@gradientcorp.com

(Received 13 April 2011; revised 18 July 2011; accepted 18 August 2011)

1. Introduction

Regulatory agencies are moving toward making greater use of human data in risk assessments, especially assessments of pesticides. The United States Environmental Protection Agency (US EPA) and the European Center for Ecotoxicology and Toxicology of Chemicals (ECETOC) have proposed frameworks for incorporating human data into chemical risk assessments (US EPA, 2010; ECETOC, 2009). We have recently developed a hypothesis-based weight-of-evidence (HBWoE) approach that not only parallels several key aspects of the US EPA and ECETOC frameworks but also provides guidance regarding how to integrate all of the data (positive, null, or of varying quality) that are relevant to determining human disease causation (Rhomberg *et al.*, 2010, 2011).

Human studies will likely be a large focus of US EPA's review of the pesticide chlorpyrifos, as there are now several epidemiology studies examining chlorpyrifos exposure and neurodevelopmental effects. The results of studies indicating neurodevelopmental effects in humans at low exposure levels are not consistent with the well-established animal and mode-of-action (MoA) data, which indicate neurological effects only at high chlorpyrifos exposures, and it will be a challenge to determine how to assess all of these data together.

In the present paper, we have two aims: first, to apply the HBWoE framework to a case study of chlorpyrifos, and second, to provide perspective on our approach compared with those put forth by others, by describing and evaluating the US EPA and ECETOC frameworks and contrasting their rationales with that of our own approach. The aim of our HBWoE analysis is to provide a critical review and synthesis of all the evidence regarding the hypothesized ability of chlorpyrifos to cause neurodevelopmental effects in humans in a transparent manner to determine whether they support a causal association at low chlorpyrifos exposures. This analysis not only provides insights about chlorpyrifos toxicity, it also addresses the larger issue of combining human, animal, and mechanistic data in risk assessments for pesticides and identifies ways to improve the US EPA and ECETOC frameworks.

2. Hypothesis-based weight of evidence

Incorporating human data into risk assessment is a critical aspect of evaluating the causes of human disease. It is important, however, to evaluate the question of human disease causation in the context of all relevant data, including epidemiology, animal toxicology, MoA (*e.g.*, *in vitro* and *in silico* approaches), and pharmacokinetics. Organizing data, evaluating data quality, and summarizing results of all the relevant studies are critical steps in evaluating all of the data relevant to the causal question.

An important further step in evaluating the relevant data is weighing all of the evidence in a clear, logical, and non-biased way so that judgments can be made based solely on the data at hand, rather than simply noting selected instances of consistency with (or contradiction of) pre-conceived ideas. Although "weight of evidence" (WoE) is often discussed as a necessary part of evaluating a causal association between a given disease and chemical exposures, there is little explicit guidance on how to weigh all of the evidence in a manner that can be documented and that for which the outcome can be used in risk management decisions. Weed (2005) points out that the term "weight of evidence" is often used loosely; he calls on practitioners to articulate what they mean by the phrase and to specify their approach. Clearly, professional judgment is involved, but it is not enough simply to name the evidence at hand and then announce one's conclusion. Flexibility in how evidence should be weighed is also necessary. Weighing all of the pertinent data, in all its diversity of study designs and complexity of bearing on the questions at hand, can be an overwhelming task when faced with a question of human disease causation. It is not a straightforward task to strike the proper balance between rigidly prescriptive guidelines (which tend to dictate scientific interpretations) and flexible, less structured guidelines that nonetheless provide some useful perspective on how, in practice, one should actually proceed and that provide adequate documentation of the basis for scientific professional judgments. There will likely be many proposed approaches to WoE; we believe, however, that there are several key aspects that should be central to a scientifically based WoE evaluation, and we have based our approach on them. These key aspects are:

1. Systematically review individual studies potentially relevant to causal question at hand (*e.g.*, epidemiology, MoA, pharmacokinetic, toxicology), with focus on evaluation of the quality of all individual studies (both negative and positive, of varying qualities).
2. Within a realm of investigation (*e.g.*, epidemiology, animal toxicology, or MoA studies), systematically examine the data for particular endpoints across studies, evaluating consistency, specificity, and reproducibility of outcomes.
3. Identify and articulate lines of argument (or "hypotheses"), newly proposed or those already put forth (if available), that bear on the available data. Discuss how available studies are used for each hypothesis to infer the existence, nature, or magnitude of human risk.
4. Evaluate the logic of the proposed hypotheses with respect to each line of evidence to determine how well the hypotheses are supported by the available data.
5. Evaluate the logic of the proposed hypotheses with respect to all lines of evidence holistically so that all of the data are considered and integrated and allowed to inform interpretation of one another.

6. Describe and compare (if more than one hypothesis has been put forth) the various alternate accounts of the observations at hand. That is, describe how well each overarching hypothesis is supported by all of the available data, discussing the uncertainties and inconsistencies in the data set and *ad hoc* assumptions required to support each hypothesis. This step involves presenting the lines of reasoning, based on the science and integration of the lines of evidence, so that the data will speak for themselves in supporting (or not supporting) the overarching hypotheses that have been put forth.
7. Formulate discussion and conclusion regarding the WoE, and proposed next steps.

These steps are intended to provide general guidance on how to weigh all of the evidence in a systematic way, but are also intended to be flexible. That is, every causal question has a different data set that will require a somewhat different specific approach for presentation and systematic review of the data at hand, but should generally follow these seven steps.

Analyses of various technical approaches to WoE have been offered by Krimsky (2005) and Linkov *et al.* (2009). Several additional frameworks have been put forth specifically as guidance for weighing evidence in the context of evaluating potential human disease causation. For example, US EPA's Office of Pesticide Programs (OPP) and the European Center for Ecotoxicology and Toxicology of Chemicals (ECETOC) have proposed frameworks for incorporating human data into chemical risk assessments (US EPA, 2010; ECETOC, 2009). Other human relevance frameworks have been put forth that provide guidance on incorporating MoA data into human risk assessment (Sonich-Mullin *et al.*, 2001; Meek *et al.*, 2003; Seed *et al.*, 2005; Boobis *et al.*, 2006; Boobis *et al.*, 2008), and US EPA (2010) and ECETOC (2009) have incorporated aspects of these guidelines into their frameworks.

We have developed the HBWoE framework, which has been described recently and applied to other chemical causation questions (Rhomberg *et al.*, 2010, 2011). It is hypothesis-based in the sense that it emphasizes articulation of the proposed bases for the relevance of the data to the causal question at hand, specifying the logic and reasoning. The approach weighs all of the data (*e.g.*, epidemiology, animal toxicology, MoA), both positive and negative, in terms of quality and relevance to humans in a way that allows each data set to inform the other, and further synthesizes all of the data to determine overall plausibility for causality in humans, considering uncertainties and inconsistencies in the data sets and *ad hoc* assumptions that may be required for some of the hypotheses put forth.

The HBWoE framework emphasizes articulation of the logic and reasoning that form the bases of various lines of argument (or "overarching hypotheses") that are either newly proposed or have already been put forth for a given question regarding human disease causation. A key aspect of the HBWoE framework is the importance of analysis of these lines of argument, or consideration of alternate "accounts" (or interpretations) of the available data and how each is supported by the available data. Hill (1965) makes explicit the importance of considering alternative "accounts" of the observations at hand in stating:

> None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non. What they can do with greater or less strength, is to help us to make up our minds on **the fundamental question – is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?** (Hill, 1965) [emphasis added]

The key outcome of the HBWoE framework is the evaluation and comparison of alternative and contrasting accounts. In the end, each account (that is, each tentative "story" as to why the facts are as they are) can be compared to other accounts. In this way, various competing overarching hypotheses can be weighed by comparing their relative success at explaining phenomena seen in the data, the relative reasonableness of *ad hoc* assumptions needed for each, and the relative naturalness and plausibility of the means whereby potentially refuting observations are reconciled with the account's central hypothesis. Although it is hard to reduce this evaluative process into checklists, scores, or enumerations, the hope is that, by not simply conducting such evaluations of alternative accounts but also by writing them down to be scrutinized and debated, the relative explanatory success of each account, and the relative "epistemological baggage" associated with defending each alternative interpretation, will be evident. This can then serve as the basis for assigning the relative degree of credence that should be given to an account that asserts the existence of a causal role of the exposures of interest in the disease *versus* accounts that ascribe any apparent patterns of association of exposure and disease that appear among the data to other, noncausal factors. In addition, from this assessment, one can more clearly define hypotheses and propose areas of research needed to fill data gaps for each account or to put their hypotheses to the test.

As part of the comparison of accounts, the HBWoE approach considers all data relevant to the causal question at hand, even negative data and (particularly when they are the bases for a particular line of argument) data of questionable quality or from studies with significant design shortcomings. In this last case, it is important to demonstrate the analysis and logic of how poor quality data have been interpreted within an account, how critical they are to the account's assertions, and the *ad hoc*

assumptions required to fit these data to the proposed hypothesis. In the HBWoE framework, such questionable data are automatically downweighted by their poor ability to discriminate between accounts, as their face-value interpretation is not markedly more compelling than alternative explanations that ascribe the outcomes to those extraneous factors or alternative possible causes that better-designed studies would have eliminated. That is, the results are relatively easily and credibly explained away as artifacts.

As described by Rhomberg *et al*. (2010, 2011), the goal of the HBWoE approach is broad in that the relative degrees of credence that should be placed in alternative possible interpretations of hypotheses are expressed in a way that shows how such credence is tied to specific scientific interpretations, considering consistencies, inconsistencies, and contradictions within and across the various data sets. The explanations in each account need not be proven—what is important is that one set out the following questions to be considered throughout the evaluation:

- What is being proposed as causal and generalizable phenomena (*i.e.*, what constitutes the basis for applying observations of biological perturbations or realized risks in other contexts to project potential risks to humans as they are exposed)?
- What is being proposed as the basis for deviations that lead to observations that do not fit the hypothesized causal model (*i.e.*, that would otherwise be counterexamples or refutations)?
- What assumptions are made that are *ad hoc* (*i.e.*, to explain particulars, but for which the evidence consists of their plausibility and the observations they are adduced to explain)?
- What further auxiliary assumptions have to be made, and how reasonable are they in view of our wider knowledge and understanding?
- What is relegated to error, happenstance, or other causes not relevant to the question at hand?
- For those events or processes proposed as critical for a given account, what other observable manifestations should they have? Are these other manifestations indeed found?
- If either the operation or necessity of the proposed critical events for a given account were disproven, how else would one explain the array of outcomes?

The HBWoE framework generally consists of the seven key aspects of WoE evaluations outlined above and in Table 1. First, the framework evaluates the intrinsic quality of the individual studies, and evaluates the data for consistency, specificity and reproducibility across various lines of evidence (*e.g.*, epidemiology, animal toxicology, and MoA studies), including both positive and negative studies and studies of varying quality. The next step involves articulation of various lines of argument that have been put forth within the scientific community to explain the observations at hand. The proffered explanations are based on the notion that true causal effects should be repeatable with some specificity and should be generally operating in all relevant test systems—or at least there should be reasoning as to why exceptions to this exist. In weighing the evidence, the framework focuses on critical evaluation of these various lines of argument, specifying the data on which each are based, and the reasoning for why these data are (or are not) informative about the human risk question at hand. These lines of argument are the "hypotheses" of the HBWoE framework, and they are articulated so that one can evaluate, throughout the process of weighing all of the evidence, how well they are in agreement with all of the data, how well they would explain patterns in the data if they were true, and what other consequences should have been observed if they were true and whether in fact these consequences are observed.

The HBWoE framework then traces the logic and reasoning within each line of evidence, in the context of the various hypotheses. The aim is to establish how well the hypotheses being examined comport with and help explain common patterns in the data, what data seem to constitute exceptions or contrary outcomes to the hypothesized causal principles, and what reasons for such exceptions might be proposed. The framework then traces through the logic regarding each line of evidence and how the animal tests, human studies, and MoA data inform interpretation of one another within the context of the various proposed hypotheses. The question is whether explanations or hypothesized causal factors proposed in one realm (*e.g.*, epidemiology) have aspects that should be observable in others (*e.g.*, MoA studies), enabling evaluation of whether signs of those causal processes do or do not appear where expected.

The final, and key, step to the HBWoE framework, as discussed above and in more detail by Rhomberg *et al*. (2010, 2011), is formulating alternate accounts of the observations at hand, and comparing these accounts. Clearly, there may be many accounts, but the major contending accounts will be those that require the fewest *ad hoc* explanations for why certain observations do not fit with the data at hand. As an explicit process to the HBWoE framework, the scientific judgment (or logical rationale) required for each account needs to be illustrated and discussed in narrative text to describe how the data are being weighed, and what *ad hoc* assumptions are required to account for some of the problematic facts within the observations at hand. Different methods can be applied (*e.g.*, organizational tables or figures), depending on the nature of the data, to organize and illustrate the consistencies and inconsistencies of the data as applied to various lines of evidence and various accounts. The point is to illustrate how one is tracing the logic through various competing accounts, and this will vary depending on the

Table 1. Comparison of frameworks for integrating human data into risk assessment.

Key aspects of a weight-of-evidence evaluation	US EPA framework	ECETOC framework	HBWoE framework
1. Systematically review individual studies potentially relevant to causal question at hand (*e.g.*, epidemiology, mode of action, pharmacokinetic, toxicology), both negative and positive, and of varying quality.	Yes (focus on epidemiology)	Yes (focus on epidemiology and animal toxicity studies)	Yes
2. Within a realm of investigation (*e.g.*, epidemiology, animal toxicology, or mode of action studies), systematically examine the data for particular endpoints across studies, evaluating consistency, specificity, and reproducibility of outcomes.	Yes (focus on epidemiology)	Yes (focus on epidemiology and animal toxicity studies)	Yes
3. Identify and articulate lines of argument (or "overarching hypotheses"), newly proposed or those already put forth (if available), that bear on the available data.	No explicit guidance	No explicit guidance	Yes
4. Evaluate the logic of the proposed hypotheses with respect to each line of evidence or realm of investigation (*e.g.*, separate evaluation of epidemiology, animal toxicology, and mode of action data).	No explicit guidance	No explicit guidance	Yes
5. Evaluate the logic of the proposed hypotheses with respect to all lines of evidence holistically so that all of the data are considered and integrated and allowed to inform interpretation of one another.	No explicit guidance (although importance of integrating is discussed)	No explicit guidance (although importance of integrating is discussed)	Yes
6. Describe and compare (if more than one hypothesis has been put forth) the various alternate accounts of the observations at hand. That is, describe how well each overarching hypothesis is supported by all of the available data, discussing the uncertainties and inconsistencies in the data set and *ad hoc* assumptions required to support each hypothesis.	No explicit guidance	No explicit guidance	Yes
7. Propose next steps (*e.g.*, sharpening of proposed hypothesis already put forth, propose additional testing to clarify data gaps).	Yes	Yes	Yes

data set, likely requiring illustration as well as narrative text. Therefore, the HBWoE framework is intended to be flexible so that each analysis can be constructed in a way that optimizes transparency and logic for the particular set of relevant data.

Below, we apply the HBWoE framework to evaluate the WoE regarding a causal association between exposure to chlorpyrifos and adverse effects on neurodevelopment in humans.

3. Chlorpyrifos case study

In this section, we apply the HBWoE framework in a case study of chlorpyrifos, to evaluate the evidence regarding the hypothesized ability of chlorpyrifos to cause neurodevelopmental effects in humans. The results of several epidemiology studies are not consistent with the well-established animal and MoA data, which indicate neurological effects only at high chlorpyrifos exposures. This case study will provide a critical review and synthesis of all of the relevant evidence in a transparent manner to determine whether they support a causal association between low chlorpyrifos exposures and neurodevelopmental effects. We begin with a discussion on the general background of chlorpyrifos, then we present the results of our evaluation of the epidemiology, animal toxicity, and mechanistic data.

3.1. Chlorpyrifos background

Chlorpyrifos is an organophosphorus (OP) insecticide that is the active component in pesticide formulations such as Dursban and Lorsban. Chlorpyrifos was widely used for agricultural and residential pest control until 2001, when restriction of its non-agricultural use began in the United States (US EPA, 2002). Current use of chlorpyrifos is mainly limited to controlling insect damage in agricultural settings worldwide. Human exposure to chlorpyrifos can occur through oral, dermal, and inhalation pathways. Inhalation and dermal exposures have likely been the predominant pathways for occupational exposure, and ingestion from residues in the diet is likely the predominant pathway for non-occupational exposures today (Eaton *et al.*, 2008). Human exposures to chlorpyrifos are estimated based on several biomarkers,

including various metabolites of chlorpyrifos that are described below.

Chlorpyrifos is well-absorbed after oral and inhalation exposures (Nolan *et al.*, 1984; Bakke *et al.*, 1976; Smith *et al.*, 1967; Ahdaya *et al.*, 1981), but dermal absorption is relatively low unless skin integrity is compromised (Aprea *et al.*, 1994; Shah *et al.*, 1987). Once absorbed into the body, chlorpyrifos is readily distributed to all organs and undergoes rapid metabolism. Oxidative desulfuration of chlorpyrifos *via* cytochrome P-450 (CYP450) enzymes to chlorpyrifos-oxon, the principal toxic metabolite, occurs predominantly in the liver, but extrahepatic metabolism has been reported, including in the brain (Chambers and Chambers, 1989). Chlorpyrifos-oxon is rapidly hydrolyzed by A-esterases, including paraoxonases such as PON1, to form diethylphosphate (DEP) and 3,5,6-trichloro-2-pyridinol (TCPy) (Sultatos and Murphy, 1983a,b). Because of this rapid metabolism, chlorpyrifos-oxon does not escape the liver once steady-state conditions are reached (Sultatos and Murphy, 1983a,b) and has not been detected in human blood or urine after oral administration (Timchalk *et al.*, 2002). Chlorpyrifos-oxon is detectable in rat blood, at concentrations close to the analytical limits of quantitation, but only after exposure to high doses (Timchalk *et al.*, 2002).

Chlorpyrifos itself undergoes oxidative dearylation *via* CYP450 enzymes to an unstable intermediate that is hydrolyzed to diethylthiophosphate (DETP) and TCPy (Timchalk *et al.*, 2002). TCPy is the major chlorpyrifos metabolite identified in the urine of both humans and animals (Bakke *et al.*, 1976; Nolan *et al.*, 1984), and its glucuronide and sulfate conjugates, as well as DEP and DETP, are also excreted in the urine. Detoxification of chlorpyrifos to DETP and TCPy occurs predominantly in the liver and plasma, and is also rapid and extensive. In humans, Nolan *et al.* (1984) estimated an elimination half-life of 27 hours for chlorpyrifos following oral or dermal exposure, and more than 90% of chlorpyrifos was eliminated within 48 hours in rats after a single-dose oral exposure (Bakke *et al.*, 1976; Smith *et al.*, 1967). The distribution and elimination of chlorpyrifos follow a two-compartment model, however, with the portion of chlorpyrifos that is partitioned into body fat or tightly bound to plasma proteins having much slower elimination. For example, Smith *et al.* (1967) reported an elimination half-life of 10–16 hours for chlorpyrifos in various rat organs except body fat, which had an estimated half-life of 62 hours.

Chlorpyrifos-oxon binds to and irreversibly inhibits cholinesterases, such as acetylcholinesterase (AChE), and inhibition of AChE in the nervous system is the mechanism through which chlorpyrifos toxicity is hypothesized to occur (Richardson, 1995). Chlorpyrifos itself can also inhibit cholinesterases, but has been reported to be two to five orders of magnitude less potent at inhibition of AChE than chlorpyrifos-oxon (Huff *et al.*, 1994; Das and Barone, 1999). The chlorpyrifos metabolites TCPy, DEP, and DETP are not considered to make a significant contribution to AChE inhibition and, thus, are not considered to be toxic. AChE is also associated with erythrocytes (red blood cells), and other esterases occur in several tissues at much higher concentrations than AChE, including plasma butyrylcholinesterase (BuChE) and various carboxylesterases in the liver and other organs (ATSDR, 1997; Eaton *et al.*, 2008). Chlorpyrifos-oxon also binds to and inhibits carboxylesterases and these, as well as erythrocyte and plasma cholinesterases, act as a metabolite "sink" to reduce the amount of chlorpyrifos-oxon that can reach the nervous system (Misulis *et al.*, 1993; Chanda *et al.*, 1997). Overall, chlorpyrifos toxicity through AChE inhibition results from a balance of activation and detoxification of chlorpyrifos-oxon through multiple pathways, including various CYP450 enzymes, PON1, circulating cholinesterases, and carboxylesterases.

AChE is an enzyme that terminates the action of the neurotransmitter acetylcholine at cholinergic synapses in the central and peripheral nervous system and at neuromuscular junctions (Palmer, 1980). Inhibition of AChE leads to an accumulation of acetylcholine at cholinergic synapses and overstimulation of nicotinic and muscarinic receptors throughout the body (Richardson, 1995; ATSDR, 1997; Eaton *et al.*, 2008). Acute cholinergic toxicity occurs when cholinesterase inhibition exceeds 70% (Clegg and van Gemert, 1999), and includes effects such as increased salivation and sweating, changes in blood pressure and heart rate, confusion, headache, nausea, diarrhea, muscle tremors, and, with very high doses, convulsions, respiratory failure, and death (ATSDR, 1997; Eaton *et al.*, 2008). These effects usually appear within a few minutes to 24 hours after exposure and are transient for non-fatal exposures, with clinical signs lasting for weeks following exposure in some cases (Lotti *et al.*, 1986).

Treatment of acute cholinergic toxicity is through the cholinergic muscarinic antagonist, atropine, which blocks the accumulation of acetylcholine on muscarinic receptors to relieve receptor hyperstimulation (Aiuto *et al.*, 1993; Namba *et al.*, 1971). Oximes, such as pralidoxime, can also be used for treatment if given shortly after exposure. Pralidoxime can displace chlorpyrifos-oxon from AChE and restore its activity, but only if the covalent bond between them has not undergone the process of "aging," in which the stability of the bond is enhanced (Eyer, 2003). In the absence of oximes, recovery of enzyme activity depends on synthesis of new enzyme, a process that may take days. In 20–50% of cases, an intermediate syndrome develops during or just after recovery from acute cholinergic toxicity (ATSDR, 1997; Eaton *et al.*, 2008). This syndrome involves weakness of the neck, limb, and respiratory muscles, and the underlying mechanism is not known. A delayed peripheral neuropathy can develop several weeks after cholinergic toxicity and the intermediate syndrome (Richardson, 1995: Moretto and Lotti, 1998; Albers *et al.*, 1999). The

clinical manifestations of this include motor weakness, with some involvement of peripheral sensory and autonomic function. These symptoms eventually stabilize, and recovery of strength and sensory function occurs, although residual sensory and autonomic dysfunction may persist for years after exposure cessation (ATSDR, 1997; Eaton *et al.*, 2008).

Several investigators have shown that young animals are more susceptible than adults to the acute toxicity of chlorpyrifos, with neonate animals having over an order of magnitude lower LD_{50} values than adults (Pope and Chakraborti, 1992; Pope *et al.*, 1991; Whitney *et al.*, 1995). This age-dependent susceptibility is likely attributable to different detoxication abilities between young animals and adults. In rodents, carboxylesterase activity is much lower in weanling animals than in adults (Karanth and Pope, 2000), and in both rodents and humans, PON1 activity is very low at birth and increases over time, reaching a plateau around postnatal day 21 in rodents and between 6 and 15 months of age in humans (Mueller *et al.*, 1983; Cole *et al.*, 2003; Li *et al.*, 1997). Developing organisms recover more quickly from cholinesterase inhibition than comparably-exposed adults, however, largely because of rapid synthesis of new cholinesterase molecules (Pope and Chakraborti, 1992; Pope *et al.*, 1991).

Human exposures to chlorpyrifos have been measured using several different metrics. Chlorpyrifos can be measured directly in blood, although usually in trace concentrations because of its rapid metabolism. The chlorpyrifos metabolites TCPy, DEP, and DETP can be measured in urine, but they have limitations as biomarkers of exposure to chlorpyrifos. TCPy in urine originates from exposure to not only chlorpyrifos, but from exposure to the pesticide chlorpyrifos-methyl and the herbicide triclopyr, as well as to pre-formed TCPy in the environment (MacIntosh *et al.*, 1999; Needham, 2005; Morgan *et al.*, 2005). Urinary DEP and DETP can also originate from exposure to other pesticides, such as diazinon and disulfoton, and to pre-formed, environmental DEP and DETP (Needham, 2005; Wessels *et al.*, 2003). Thus, these metabolites are not specific to chlorpyrifos and their use as an exposure metric, especially from environmental exposures, can overestimate chlorpyrifos exposure. Activities of erythrocyte AChE or plasma BuChE have also been used as biomarkers of chlorpyrifos exposure, but these activities are also not specific to chlorpyrifos, as other chemicals, including other OPs and N-methyl carbamate pesticides, inhibit cholinesterases (ATSDR, 1997; Barr and Angerer, 2006).

Controlled human exposure studies (Coulston *et al.*, 1972; Kisicki *et al.*, 1999; Nolan *et al.*, 1984) and occupational epidemiology studies (*e.g.*, Albers *et al.*, 2004a, 2004b, 2004c; 2007) have not reported clinical signs of cholinergic toxicity associated with low chlorpyrifos exposures in adults (*e.g.*, controlled human exposures used 0.01–2 mg/kg exposures). Because chlorpyrifos readily passes through the placenta, however, it has been hypothesized that exposure to chlorpyrifos may be associated with neurodevelopmental effects at doses below the threshold for AChE inhibition, through a non-enzymatic role of AChE in brain development or by other non-cholinergic mechanisms in the developing nervous system. Next, we describe the epidemiology studies that have been conducted to address the potential association of chlorpyrifos exposure with neurodevelopmental effects.

3.2. Epidemiology studies of neurodevelopmental effects

We evaluated the available epidemiology data that are relevant to determining whether there is sufficient evidence to support an association between chlorpyrifos exposure and adverse neurodevelopmental effects. Below, we provide a brief overview of the epidemiology literature followed by an endpoint-by-endpoint analysis of each neurodevelopmental outcome that has been investigated. Then, we critically evaluated the epidemiology data as a whole, considering many factors such as the weight of the exposure metric used, outcome assessed, clinical significance of reported effects, control of confounding factors, exposure-response relationships, and statistical limitations. For our analysis, we first conducted a literature search, using PubMed and Toxline, for all human studies measuring or estimating chlorpyrifos exposure and neurodevelopmental outcomes. Search terms included: "chlorpyrifos," "neurological," "neurobehavioral," "neurotoxicity," "behavior*," "birth outcomes," "cognitive," "intelligence," and epidemiol*." We also relied on the reference lists of several review articles (*e.g.*, Eaton *et al.*, 2008; Needham, 2005).

3.2.1. Overview of epidemiology studies

3.2.1.1. Cohort studies of chlorpyrifos Several cohort studies examining the association between chlorpyrifos exposure and neurodevelopmental outcomes in newborns and young children have been conducted, with multiple studies stemming from each cohort. Participants in these cohort studies were likely exposed to many classes of pesticides and other environmental chemicals, but residential or agricultural uses of chlorpyrifos were large sources of their pesticide exposure (Needham, 2005). Among the cohort studies, several different exposure metrics were used to assess chlorpyrifos exposure. Some studies measured chlorpyrifos directly in maternal prenatal and postnatal blood and in cord blood (Perera *et al.*, 2003; Whyatt *et al.*, 2004; Rauh *et al.*, 2006, 2011; Barr *et al.*, 2010). Other studies measured maternal urinary concentrations of TCPy, or both maternal and child urinary concentrations of total diethyl phosphate metabolites (DEPs), which include DEP and DETP (Eskenazi *et al.*, 2004, 2007; Berkowitz *et al.*, 2004; Young *et al.*, 2005; Engel *et al.*, 2007, 2011; Wolff *et al.*, 2007). Enzymatic activities of cholinesterases in whole blood and BuChE

in plasma were used in two studies as a general marker of OP exposure (Eskenazi *et al.*, 2004; Wolff *et al.*, 2007). Finally, one study used measurements of chlorpyrifos in ambient air through personal monitoring (Whyatt *et al.*, 2004).

The cohort studies evaluated multiple neurodevelopmental endpoints, which are described in more detail in a later section. Briefly, newborn head circumference was reported in the infants' medical records following delivery. Infant neurobehavioral capacities were measured with the Brazelton Neonatal Behavioral Assessment Scale (BNBAS). The BNBAS scores infant behavior in seven domains: habituation (ability to respond to stimuli while asleep), orientation (attention to visual and auditory stimuli and quality of alertness), motor performance, range of state (arousal and state lability), regulation of state (in the face of increasing levels of stimulation), autonomic stability (signs of stress related to homeostatic adjustments of the CNS), and primitive reflexes (Lester *et al.*, 1982). Cognitive and motor development were assessed using the Bayley Scales of Infant Development II (BSID-II). The BSID-II is a widely used test for identifying young children at risk for developmental delay, and it yields scores for Mental Development Index (MDI) and Psychomotor Development Index (PDI). Cognitive development was also assessed using the Weschler Intelligence Scales for Children (WISC-IV), which yields scores for four areas of mental functioning that are associated with overall IQ. The four indices are verbal comprehension (verbal concept formation), working memory (ability to memorize new information, concentrate, and manipulate information), perceptual reasoning (non-verbal and fluid reasoning), and processing speed (ability to focus attention and quickly order visual information). Behavioral outcomes such as Attention Deficit Hyperactivity Disorder (ADHD), Pervasive Development Disorder (PDD), and attention problems were measured through reporting by the mothers on the Child Behavior Checklist (CBCL). Several studies also evaluated birth weight and birth length, but we did not include these outcomes in our analysis because they are general measures of fetal growth rather than specific neurodevelopmental endpoints.

Each cohort and the studies evaluating neurodevelopmental endpoints are described below and summarized in Table 2. Each study, grouped by cohort, is presented in the rows of Table 2, with separate columns for each of the various exposure metrics and outcomes examined. This provides an overview of the exposure metrics and outcomes that were analyzed across studies and cohorts and is useful for the evaluation of the data regarding the association between chlorpyrifos exposure and neurodevelopmental effects presented in the studies.

Columbia cohort: A cohort of Dominican and African-American mother-newborn pairs living in inner-city neighborhoods in New York City was studied by researchers at Columbia University. Residential pesticide use is widespread among minority populations in New York City, with chlorpyrifos being the most heavily applied pesticide prior to the restriction of its residential use (Landrigan *et al.*, 1999; Whyatt *et al.*, 2002). In the Columbia cohort, 85% of the mothers reported using some form of pest control measure during pregnancy, including sticky traps, gels, can sprays, and pest bombs, and 35% reported using an exterminator (Whyatt *et al.*, 2002). Many of these measures were performed repeatedly (Whyatt, *et al.*, 2002, 2004), increasing the likelihood of repeated inhalation exposure. Deposition of chlorpyrifos on surface areas within the residences most likely led to dermal exposure, and potentially to exposure *via* ingestion beyond that from residues in the diet. The mothers delivered at New York Presbyterian Medical Center, Harlem Hospital, or their satellite clinics between 1998 and 2002. Mothers were eligible for the cohort if they were non-smokers, aged 18–35, were free of diabetes, hypertension, or known HIV, and resided in the area for a minimum of 1 year.

Perera *et al.* (2003) examined the association between chlorpyrifos exposure and newborn head circumference measured at birth in 113 mother-newborn pairs of this cohort. The authors measured chlorpyrifos levels in cord plasma collected at delivery and in maternal plasma collected within one day postpartum. Maternal and cord plasma chlorpyrifos levels were highly correlated ($r = 0.76$), so the authors only used the cord plasma levels in their analysis. Whyatt *et al.* (2004) also considered the association between cord plasma levels of chlorpyrifos with newborn head circumference in an expanded number of mother-newborn pairs (287), although they did not state whether all subjects from the Perera *et al.* (2003) study were included in their analysis. Whyatt *et al.* (2004) also conducted personal air monitoring of chlorpyrifos for 48 hours during the third trimester for 271 women and evaluated the association between these measurements and head circumference.

In an effort to evaluate cognitive and behavioral outcomes, Rauh *et al.* (2006) extended the follow-up through the first 3 years of life for a subset of 254 infants. The authors examined associations between the chlorpyrifos levels in cord plasma, as measured by Perera *et al.* (2003) and Whyatt *et al.* (2004), and cognitive and motor development, *via* scores for MDI and PDI at 12, 24, and 36 months. They also examined the association between chlorpyrifos exposure and behavioral outcomes, including ADHD, PDD, and attention problems, using the CBCL at 36 months of age. In a recent study, Rauh *et al.* (2011) used the WISC-IV to examine associations between cord plasma chlorpyrifos levels and cognitive development at 7 years of age for 265 children in the cohort.

CHAMACOS cohort: Researchers in the Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) at the University of California at Berkeley studied neurodevelopmental endpoints in a cohort of mother-newborn pairs living in an agricultural community in the Salinas Valley of California. Exposures to chlorpyrifos were primarily from agricultural spraying of

Table 2. Overview of cohort studies of chlorpyrifos exposure and neurodevelopmental outcomes.

		Exposure Metric							Outcome assessed					
Study	Subjects	Chlorpyrifos in cord blood	Chlorpyrifos in maternal blood	Urinary TCPy	Urinary DEPs	Cholinesterases in cord blood	Cholinesterases in maternal blood	Personal air monitoring	Age at outcome assessment	Head circumference	Infant neurobehavior (BNBAS)	Cognitive and motor development (BSID-II)	Cognitive development (WISC-IV)	Behavioral outcomes (ADHD, PDD, and attention problems) (CBCL)
Columbia Cohort (Infants born between 1998–2002)														
Perera *et al.*, 2003	113 African-American and Dominican mother-newborn pairs living in inner-city (NYC)	Plasma levels, collected at delivery	Plasma levels, collected within 1 day postpartum						Birth	X				
Whyatt *et al.*, 2004	Expanded number of mother-newborn pairs ($n = 287$) from Perera *et al.* (2003)	Plasma levels, collected at delivery	Plasma levels, collected within 2 days postpartum					Collected for 48 hours during third trimester	Birth	X				
Rauh *et al.*, 2006	254 children from Whyatt *et al.* (2004)	Plasma levels, collected at delivery	Plasma levels, collected within 1 day postpartum						12 months			X		
									24 months			X		
									36 months			X		X
Rauh *et al.*, 2011	265 children from Whyatt *et al.* (2004)	Plasma levels, collected at delivery	Plasma levels, collected within 2 days postpartum						7 years				X	
CHAMACOS Cohort (Infants born between 2000–2001)														
Eskenazi *et al.*, 2004	488 mother-newborn pairs living in agricultural community (Salinas Valley, CA)			Maternal levels, collected at 13 and 26 weeks gestation	Maternal levels, collected at 13 and 26 weeks gestation	ChE activity in whole blood; BuChE activity in plasma; collected at delivery	ChE activity in whole blood; BuChE activity in plasma; collected at 26 weeks gestation and delivery		Birth	X				
Young *et al.*, 2005	381 mother-newborn pairs from Eskenazi *et al.* (2004)				Maternal levels, collected at 14 and 26 weeks gestation and within one week of delivery				0 to 62 days		X			

(Continued)

Table 2. (Continued).

		Exposure Metric							Outcome assessed					
Study	Subjects	Chlorpyrifos in cord blood	Chlorpyrifos in maternal blood	Urinary TCPy	Urinary DEPs	Cholinesterases in cord blood	Cholinesterases in maternal blood	Personal air monitoring	Age at outcome assessment	Head circumfer-ence	Infant neuro-behavior (BNBAS)	Cognitive and motor develop-ment (BSID-II)	Cognitive development (WISC-IV)	Behavioral outcomes (ADHD, PDD, and attention problems) (CBCL)
Eskenazi *et al.*, 2007	396 children from Eskenazi *et al.* (2004)			Maternal levels, collected at 14 and 26 weeks gestation	Maternal levels, collected at 14 and 26 weeks gestation; child levels, collected at 6, 12, and 24 months of age				6 months			X		
									12 months			X		
									24 months			X		X
Mount Sinai Cohort (Infants born between 1998–2002)														
Berkowitz *et al.*, 2004	404 mother-newborn pairs living in NYC			Maternal levels, collected during third trimester					Birth	X				
Wolff *et al.*, 2007	404 mother-newborn pairs from Berkowitz *et al.* (2004)				Maternal levels, collected during third trimester		BuChE activity in plasma; collected during third trimester		Birth	X				
Engel *et al.*, 2007	311 mother-newborn pairs from Berkowitz *et al.* (2004)				Maternal levels, collected during third trimester				0 to 5 days		X			
Engel *et al.*, 2011	360 children from Berkowitz *et al.* (2004)				Maternal levels, collected during third trimester				12 months			X		
									24 months			X		
									7–9 years				X	
UMDNJ Cohort (Infants born between 2003–2004)														
Barr *et al.*, 2010	150 mother-newborn pairs scheduled for elective cesarean delivery at Saint Peter's University Hospital, New Brunswick, New Jersey	Serum levels, collected at delivery	Serum levels, collected just prior to delivery						Birth	X				

DEPs = diethyl phosphates; TCPy = 3,5,6-trichloro-2-pyridinol; ChE = cholinesterase; BuChE = butyrylcholinesterase.

several pesticides, pesticide drift from the spraying, and pesticide residues brought into homes on workers' clothing, with potential exposures from residential pesticide use and dietary exposure (Needham, 2005; Castorina *et al.*, 2003). A total of 488 pregnant women participated in the study. During pregnancy, 28% of the women worked in the fields, 14% worked at other agricultural jobs, and 85% lived in homes with agricultural workers (Eskenazi *et al.*, 2004). The mothers delivered at Natividad Medical Center during 2000 and 2001 and were excluded from the study if they were less than 18 years of age or had gestational or pre-existing diabetes, hypertension, twin births, or stillbirths.

Eskenazi *et al.* (2004) studied all 488 mother-newborn pairs, including 11 infants diagnosed with congenital anomalies at birth because, according to the authors, their exclusion did not materially affect the results. Concentrations of TCPy and DEPs were measured in maternal urine collected at two time periods during pregnancy (first period, mean 13 weeks gestation, range 4–29 weeks; second period, mean 26 weeks gestation, range 18–39 weeks) and these measurements were averaged for each mother. Maternal blood specimens were collected during the second period and cord blood was collected from newborns for measurements of total cholinesterase activity in whole blood and BuChE activity in plasma. The authors evaluated the association between these exposure estimates and head circumference measured at birth.

Young *et al.* (2005) further examined a subset of 381 mother-newborn pairs and evaluated the association between concentrations of DEPs in maternal urine collected at 14 and 26 weeks gestation and at one time point after delivery (usually within one week) and the BNBAS scores of infant behavior assessed at 2 months of age or younger.

Eskenazi *et al.* (2007) evaluated the association between chlorpyrifos exposure and MDI and PDI scores in a subset of 396 infants at 6, 12, and 24 months of age. The authors also examined the association between chlorpyrifos exposure and behavioral outcomes (ADHD, PDD, and attention problems) at 24 months of age. Infants with a medical condition that could affect assessment (Down syndrome, deafness, and hydrocephalus) were excluded from the study. The authors used measurements of TCPy and DEPs in maternal urine collected at 14 and 26 weeks gestation and DEPs in the children's urine collected at 6, 12, and 24 months of age, as surrogates for chlorpyrifos exposure.

Mount Sinai cohort: Researchers at Mount Sinai Medical School studied a multi-ethnic cohort of mother-infant pairs living in New York City (predominantly in East Harlem) who delivered between 1998 and 2002 at Mount Sinai Hospital. Approximately 71% of the mothers reported potential exposure to indoor application of pesticides (Berkowitz *et al.*, 2004). Mothers were excluded if they had serious chronic diseases such as diabetes, hypertension, or thyroid disease or if they developed a serious pregnancy complication that could affect fetal development. Mothers were also excluded if they consumed more than two alcoholic beverages per day or used illegal drugs, or if their child was born with a congenital malformation or severe prematurity (< 1,500 grams birth weight or < 32 weeks gestation).

In an initial study of this cohort, Berkowitz *et al.* (2004) evaluated the association between TCPy concentration in maternal urine collected during the third trimester and newborn head circumference measured at birth in 404 mother-infant pairs. Wolff *et al.* (2007) also studied head circumference in the 404 infants, using maternal urinary concentrations of DEPs and BuChE activity in plasma to estimate chlorpyrifos exposure.

In order to evaluate potential neurobehavioral outcomes, Engel *et al.* (2007) examined the association between maternal urinary concentrations of DEPs and behavior in a subset of 311 newborns, as assessed through administration of the BNBAS within five days of delivery.

Engel *et al.* (2011) also examined associations between maternal urinary DEP levels and cognitive and motor development *via* MDI and PDI scores for 177 children at 12 months of age and 247 children at 24 months of age. Cognitive development was also assessed using the WISC-IV in 114 children between seven and 9 years of age.

UMDNJ cohort: A cohort of 150 mother-newborn pairs was examined in a study led by researchers at the University of Medicine and Dentistry of New Jersey (UMDNJ; Barr *et al.*, 2010). More than half of the mothers reported using some type of pesticide during pregnancy (Barr *et al.*, 2010). The mothers underwent elective cesarean delivery at term (≥ 37 weeks) at Saint Peter's University Hospital in New Jersey between July, 2003 and May, 2004. Mothers were excluded if their hemoglobin levels were ≥ 8 mg/dL, they were taking medications that could potentially interfere with metabolism of environmental chemicals, or their fetus had congenital anomalies. Maternal blood was collected pre-operation and cord blood was collected within 15 minutes of birth for measurement of serum chlorpyrifos levels, and the authors evaluated the association between these levels and newborn head circumference.

3.2.1.2. Chlorpyrifos exposure in the cohort studies
Air concentrations of chlorpyrifos were measured for the Columbia and CHAMACOS cohorts. Whyatt *et al.* (2004) reported a mean and standard deviation of $15.3 \pm 31.8\ ng/m^3$ chlorpyrifos in personal air samples collected for 271 of the women in the Columbia cohort. Bradman *et al.* (2007) measured chlorpyrifos in indoor air in the homes of 20 children from the CHAMACOS cohort, reporting a median concentration of $11\ ng/m^3$ (range: 4.0–36). Using these measurements, Eaton *et al.* (2008) estimated the daily intake of chlorpyrifos from inhalation at 0.003–0.07 μg/kg-day for mothers in the Columbia cohort and 0.002–0.007 μg/kg-day for mothers in the CHAMACOS cohort. Eaton *et al.* (2008) noted

that, although chlorpyrifos exposures in air were not measured for the Mount Sinai cohort, other reports of exposure in this cohort indicate that the estimated daily intake from inhalation should be similar to that of the Columbia cohort. Based on FDA market basket surveys, Eaton *et al*. (2008) also estimated the daily intake of chlorpyrifos from consuming a typical US diet as 0.005 μg/kg-day for an average adult. If this intake is added to the estimated intake for inhalation exposure in the cohorts, with the assumption that inhalation contributes approximately one-third of the total exposure, the average daily exposure for mothers in the Columbia cohort is estimated at 0.008 μg/kg-day and for the CHAMACOS cohort at 0.007 μg/kg-day. These estimates are only slightly higher than the estimates for current exposures in the general population, which Eaton *et al*. (2008) estimate as 0.004–0.006 μg/kg-day for adults and less than 0.01 μg/kg-day for children, but they do not account for dermal exposure or for exposure *via* ingestion beyond typical dietary exposure.

Chlorpyrifos was measured in blood samples from the Columbia and UMDNJ cohorts. For the Columbia cohort, mean chlorpyrifos concentrations in cord plasma were 7.5 ± 7.5 pg/g for 113 subjects in the initial cohort (Perera *et al*., 2003) and 4.0 ± 6.1 pg/g for 287 subjects in the expanded cohort (Whyatt *et al*., 2004). These concentrations are similar to those measured in the general population, based on a mean chlorpyrifos level of 9 pg/g measured in serum from blood bank donors in Cincinnati, Ohio (Barr *et al*., 2002). For the UMDNJ cohort, Yan *et al*. (2009) reported mean chlorpyrifos concentrations of 0.55 pg/g in cord serum ($n = 148$) and 0.09 pg/g in maternal serum ($n = 138$). Barr *et al*. (2010) reported the same values, but with units of ng/g. This is likely an error, given that the values would be two orders of magnitude greater than those for the Columbia cohort if these units were correct, yet they were measured after the 2001 ban of residential chlorpyrifos use.

Urinary metabolites of chlorpyrifos were measured in each cohort. Median urinary concentrations of DEPs in maternal urine were reported as 22 nmol/L for the CHAMACOS cohort (Eskenazi *et al*., 2004), 12.97 nmol/L for the UMDNJ cohort (Yan *et al*., 2009), and ranged from 18.8–24.7 nmol/L for the Mount Sinai cohort (Wolff *et al*., 2007; Engel *et al*., 2007, 2011). Eskenazi *et al*. (2007) determined urinary levels of DEPs in 6- to 24-month-old children of the CHAMACOS cohort and reported geometric mean levels ranging from 10.5–15.2 nmol/L. Median concentrations of TCPy in maternal urine ranged from less than the detection limit to 0.61 μg/L for samples collected at various time points during the third trimester in the Columbia cohort (Whyatt *et al*., 2009). The median concentration of TCPy in maternal urine in the CHAMACOS cohort was reported as 3.3 μg/L (Eskenazi *et al*., 2004) and in the Mount Sinai cohort as 7.6 μg/L (Berkowitz *et al*., 2004). Yan *et al*. (2009) determined the mean concentration of TCPy in maternal urine for the UMDNJ cohort as 1.515 μg/L.

For most of the cohorts, urinary levels of TCPy are similar to those among the general US population, as determined by levels of these metabolites in the 1999–2000 National Health and Nutrition Examination Survey (NHANES; CDC, 2009). The median TCPy level in urine in the total NHANES population was 1.70 μg/L, which is slightly greater than the median reported for the Columbia and UMDNJ cohorts. The median TCPy level in the CHAMACOS cohort is equivalent to the 75th percentile reported for Mexican-Americans (3.20 μg/L) in the 1999–2000 NHANES data, and the median level in the Mount Sinai cohort is similar to the 90th percentile reported for the total NHANES population (7.30 μg/L).

3.2.1.3. Other human studies of chlorpyrifos There are many other studies of the effects of chlorpyrifos in humans besides the cohort studies described above. Several epidemiology studies have been conducted to examine the effects of chronic exposure in workers involved in the manufacture or application of chlorpyrifos (Albers *et al*., 2004a, 2000b, 2000c; 2007). Controlled exposure studies of chlorpyrifos have been conducted in healthy adult volunteers to determine the safety and pharmacokinetics of this compound in humans (Coulston *et al*., 1972; Kisicki *et al*., 1999; Nolan *et al*., 1984). There have also been many case studies of acute chlorpyrifos exposure reported after accidental or intentional poisoning incidents (as reviewed by ATSDR, 1997; Eaton *et al*., 2008). Because these studies do not evaluate potential neurodevelopmental effects of chlorpyrifos, they will not be discussed further.

3.2.2. Endpoint-by-endpoint analysis of neurodevelopmental effects

In this section, we discuss the individual neurodevelopmental outcomes analyzed in the cohort studies described above. For each outcome, we assessed the consistency of findings across studies, including consideration of the type of exposure metric, and whether any exposure-response relationships were evident.

3.2.2.1. Newborn head circumference The association between chlorpyrifos exposure and newborn head circumference has been investigated in six of the cohort studies (Table 3). Three of these studies used blood levels of chlorpyrifos as the exposure metric. In studies of the Columbia cohort, Perera *et al*. (2003) and Whyatt *et al*. (2004) reported no association between cord plasma levels of chlorpyrifos and head circumference using regression models with chlorpyrifos as a continuous variable ($p = 0.28$ and $p = 0.86$, respectively). Similarly, Barr *et al*. (2010) reported no association between increased chlorpyrifos levels in maternal ($p = 0.229$) or cord serum ($p = 0.989$) and head circumference in their study of the UMDNJ cohort, using exposure indicators of chlorpyrifos

concentrations > 75th percentile and ≤ 75th percentile in the regression models.

Three studies examined associations between maternal urinary concentrations of the metabolites TCPy and/or DEPs and head circumference. In a study of the CHAMACOS cohort, Eskenazi *et al.* (2004) reported no association between TCPy levels categorized above ($p = 0.85$) or below the median value ($p = 0.78$), compared to the referent group with non-detectable levels, and head circumference. The authors also reported no association between urinary levels of DEPs, which were analyzed as continuous variables on a $\log_{10}$ scale, and head circumference ($p = 0.07$). Berkowitz *et al.* (2004) reported no association between urinary levels of TCPy, above and below the limit of detection (LOD), and head circumference in the Mount Sinai cohort. The authors also examined maternal PON1 activity and reported a trend of decreased head circumference ($p = 0.014$) in mothers whose TCPy levels were above the LOD and whose PON1 activity was lowest, but the test for interaction among these was not statistically significant ($p > 0.05$). Wolff *et al.* (2007) reported no association between urinary levels of DEPs, as a continuous variable on a $\log_{10}$ scale, and head circumference ($p = 0.67$) in their study of the Mount Sinai cohort.

Two studies examined head circumference using cholinesterase activities, analyzed as continuous variables on a $\log_{10}$ scale, as estimates of exposure to OPs, including chlorpyrifos. For the CHAMACOS cohort, Eskenazi *et al.* (2004) reported no association between cholinesterase activity in maternal blood during pregnancy ($p = 0.45$) or at delivery ($p = 0.27$) and head circumference. Cholinesterase activity in cord blood was also not associated with this outcome ($p = 0.65$). The authors also reported no associations between head circumference and any of several measures of exposure, including BuChE activity in maternal plasma during pregnancy ($p = 0.58$), maternal plasma at delivery ($p = 0.73$), or cord plasma ($p = 0.91$). Wolff *et al.* (2007) reported no association between BuChE activity in maternal plasma and head circumference ($p = 0.728$) in the Mount Sinai cohort.

Only one study used air concentrations of chlorpyrifos as the exposure metric. Whyatt *et al.* (2004) reported no association in the Columbia cohort between maternal air concentrations of chlorpyrifos as assessed with personal monitors during the third trimester (as a continuous variable) and head circumference ($p = 0.59$), which is consistent with their results for cord plasma levels of chlorpyrifos.

Overall, studies of associations between chlorpyrifos exposure and newborn head circumference have reported consistently null results across cohorts and exposure measures. One study reported an association between maternal PON1 activity and head circumference when maternal TCPy levels in urine were considered jointly, but the test of interaction among these factors was not statistically significant.

3.2.2.2. Infant neurobehavior Two studies examined the association between maternal urinary concentrations of DEPs and infant neurobehavior as assessed by BNBAS scores (Table 4). In a study of the CHAMACOS cohort, Young *et al.* (2005) reported no association between the mean of urinary levels of DEPs at 14 and 26 weeks gestation as a continuous variable and each of the seven BNBAS cluster scores assessed within two months of birth, with the exception of the score for primitive reflexes. Increasing urinary DEPs were associated with an increased number of abnormal reflexes ($p < 0.05$). The authors stratified the results by age at BNBAS assessment and reported an association between levels of DEPs and abnormal reflexes for infants assessed after the first three days of life ($p < 0.05$) but not for those assessed within the first three days of life ($p > 0.05$). Among infants older than three days at assessment, the proportion with more than three abnormally-rated reflexes was examined by quintiles of average $\log_{10}$ total DEPs during pregnancy. A marginally statistically significant ($p = 0.05$) trend of increasing proportion of more than three abnormal reflexes with increasing DEPs quintiles was reported. For all infants in the study, the odds of having more than three abnormal reflexes (compared to three or fewer) increased with each 10-fold increase in maternal levels of DEPs during pregnancy (Odds Ratio [OR] = 3.4, 95% Confidence Interval [CI]: 1.2–9.9). The authors noted that there were no associations between DEPs measured in the post-delivery sample of maternal urine and performance in any BNBAS cluster, but the results were not shown.

Engel *et al.* (2007) used the BNBAS to assess neurobehavior in infants of the Mount Sinai cohort within five days of delivery. Urinary levels of DEPs during the third trimester were analyzed as a continuous variable on a $\log_{10}$ scale for all BNBAS clusters except for primitive reflexes, which were analyzed with Poisson regression because of the count nature of the data for this cluster. As with the CHAMACOS cohort results reported by Young *et al.* (2005), there were no associations between levels of DEPs and any of the BNBAS cluster scores, except for scores of abnormal primitive reflexes (Relative Risk [RR] = 1.49, 95% CI: 1.12–1.98, per 10-fold increase in DEPs). The authors also examined the association between DEPs and the number of abnormally-rated reflexes, and reported that maternal concentrations of DEPs above the median were associated with two or more abnormal reflexes (RR = 2.3, 95% CI: 1.1–5.0).

Overall, the two studies examining infant neurobehavior reported associations between increasing maternal urinary levels of DEPs and BNBAS scores for abnormal reflexes, but not with less-than-optimal scores for any other BNBAS cluster. No studies are available that assessed infant neurobehavior using measurements of chlorpyrifos itself as the exposure metric.

3.2.2.3. Cognitive and motor development The association between chlorpyrifos exposure and cognitive and motor development, as measured by MDI and PDI

Table 3. Results of cohort studies assessing head circumference.

Cohort	Study	Exposure metric/group	Exposure	N	Estimate	95% CI	p-value	Statistical method
Columbia	Perera *et al.*, 2003	Chlorpyrifos in cord plasma	Mean: 7.5 ± 7.5 pg/g	113	β = –0.005	NR	0.28	Multiple regression with log-transformed chlorpyrifos as a continuous variable
	Whyatt *et al.*, 2004	Chlorpyrifos in cord plasma	Mean: 4.0 ± 6.1 pg/g	287	β = –0.01	(–0.13, 0.11)	0.86	Multiple regression with log-transformed chlorpyrifos as a continuous variable
		Chlorpyrifos in maternal air	Mean: 15.3 ± 31.8 ng/m^3	271	β = –0.04	(–0.18, 0.10)	0.59	
CHAMACOS	Eskenazi *et al.*, 2004	Maternal urinary TCPy	Median: 3.3 μg/L					Linear regression with TCPy analyzed as categorical variables
		Detectable levels < median		220	β = 0.06	(–0.37, 0.49)	0.78	
		Detectable levels > median		221	β = 0.04	(–0.39, 0.47)	0.85	
		Maternal urinary DEPs	Median: 22 nmol/L	486	β = 0.28	(–0.02, 0.59)	0.07	Linear regression with log-transformed DEP levels as a continuous variable
		ChE (μmol/min/mL)						Linear regression with log-transformed cholinesterase levels as a continuous variable
		Maternal blood, pregnancy	Median: 5.7 μmol/min/mL	340	β = 0.06	(–0.09, 0.21)	0.45	
		Maternal blood, delivery	Median: 5.1 μmol/min/mL	357	β = –0.07	(–0.19, 0.05)	0.27	
		Cord blood	Median: 3.8 μmol/min/mL	292	β = –0.04	(–0.23, 0.14)	0.65	
		BuChE (μmol/min/mL)						
		Maternal blood, pregnancy	Median: 1.4 μmol/min/mL	340	β = 0.12	(–0.31, 0.56)	0.58	
		Maternal blood, delivery	Median: 1.4 μmol/min/mL	357	β = –0.07	(–0.45, 0.31)	0.73	
		Cord blood	Median: 1.2 μmol/min/mL	292	β = –0.03	(–0.50, 0.45)	0.91	
Mount Sinai	Berkowitz *et al.*, 2004	Maternal urinary TCPy	Median: 7.6 μg/L					Generalized linear models with TCPy analyzed as categorical variables
		TCPy < LOD		216	NR	NR	NS	
		TCPy > LOD		171				
		TCPy > LOD, Low PON1		47	NR	NR	0.014	
		TCPy > LOD, Medium PON1		57	NR	NR	NS	
		TCPy > LOD, High PON1		55	NR	NR	NS	
	Wolff *et al.*, 2007	Maternal urinary DEPs	Median: 18.8 nmol/L	318	$\beta \pm$ SE = –0.052 ± 0.12		0.67	Generalized linear models with log-transformed DEP levels or BuChE activity as continuous variables
		BuChE enzymatic activity	Range: <1.9–13.8 mg/ml * min	NR	$\beta \pm$ SE = 0.44 ± 1.27		0.728	
UMDNJ	Barr *et al.*, 2010	Chlorpyrifos in maternal serum	Mean: 0.09 ng/g					Multivariable linear regression with exposure indicators > *versus* ≤ 75th percentile
		> 75th percentile versus ≤ 75th percentile		138	NR	NR	0.229	
		Chlorpyrifos in cord serum	Mean: 0.55 ng/g					
		> 75th percentile versus ≤ 75th percentile		148	NR	NR	0.989	

Table 4. Results of cohort studies assessing infant neurobehavior using the BNBAS.

Cohort	Study	Exposure metric/ group	Exposure concentration	N	Estimate	95% CI	p-value	Statistical method
CHAMACOS	Young *et al.*, 2005	Maternal urinary DEPs	Median: 21 nmol/L	175	Habituation; $\beta = 0.33$	(–0.06, 0.72)		Regression models with log-transformed DEP levels as a continuous variable for all clusters except abnormal reflexes; Poisson regression for abnormal reflexes
				379	Orientation; $\beta = -0.32$	(–0.66, 0.03)		
				381	Motor Performance; $\beta = 0.10$	(–0.06, 0.27)		
				381	Range of State; $\beta = -0.02$	(–0.27, 0.24)		
				381	Regulation of State; $\beta = -0.15$	(–0.47, 0.17)		
				381	Autonomic Stability; $\beta = 0.06$	(–0.15, 0.27)		
				381	Reflexes; $\beta = 0.22$	(0.04, 0.40)	< 0.05	
		Maternal urinary DEPs, ≤ 3 days at assessment		109	Habituation; $\beta = 0.47$	(–0.05, 0.99)		Regression estimates stratified by age at BNBAS assessment (≤ 3 days)
				197	Orientation; $\beta = -0.11$	(–0.65, 0.43)		
				197	Motor Performance; $\beta = 0.08$	(–0.17, 0.33)		
				197	Range of State; $\beta = -0.21$	(–0.54, 0.12)		
				197	Regulation of State; $\beta = -0.08$	(–0.52, 0.37)		
				197	Autonomic Stability; $\beta = 0.31$	(0.01, 0.61)	< 0.05	
				197	Reflexes; $\beta = 0.08$	(–0.16, 0.32)		
		Maternal urinary DEPs, > 3 days at assessment		66	Habituation; $\beta = 0.20$	(–0.43, 0.83)		Regression estimates stratified by age at BNBAS assessment (> 3 days)
				182	Orientation; $\beta = -0.33$	(–0.73, 0.08)		
				184	Motor Performance; $\beta = 0.17$	(–0.05, 0.38)		
				184	Range of State; $\beta = 0.20$	(–0.21, 0.62)		
				184	Regulation of State; $\beta = -0.24$	(–0.72, 0.24)		
				184	Autonomic Stability; $\beta = -0.16$	(–0.47, 0.14)		
				184	Reflexes; $\beta = 0.37$	(0.09, 0.64)	< 0.05	

(Continued)

Table 4. (Continued).

Cohort	Study	Exposure metric/ group	Exposure concentration	N	Estimate	95% CI	p-value	Statistical method
		Maternal urinary DEPs, > 3 days at assessment, Quintile 1	0.51 to 0.90 nmol/L	37	χ^2 trend = 4.0		0.05	Regression estimates for proportion of > 3 abnormal reflexes, stratified by age at BNBAS assessment (> 3 days) with log-transformed DEP levels categorized by quintiles
		Quintile 2	0.90 to 1.10 nmol/L	37				
		Quintile 3	1.11 to 1.27 nmol/L	37				
		Quintile 4	1.28 to 1.58 nmol/L	37				
		Quintile 5	1.58 to 2.35 nmol/L	36				
		Maternal urinary DEPs		N = 381	Total DEPs, OR = 3.4	(1.2, 9.9)		Logistic regression for > 3 abnormal reflexes *versus* ≤ 3 abnormal reflexes; OR per 10-fold increase in log-transformed DEP levels
Mount Sinai	Engel *et al.*, 2007	Maternal urinary DEPs	Median: 24.7 nmol/L	144	Habituation; $\beta = 0.168$	(–0.23, 0.57)		Generalized linear models with log-transformed DEP levels as a continuous variable for all clusters except abnormal reflexes; Poisson regression for abnormal reflexes
				233	Orientation; $\beta = -0.106$	(–0.41, 0.20)		
				249	Motor Performance; $\beta = 0.049$	(–0.08, 0.17)		
				253	Range of State; $\beta = 0.035$	(–0.12, 0.19)		
				253	Regulation of State; $\beta = -0.047$	(–0.30, 0.21)		
				253	Autonomic Stability; $\beta = -0.154$	(–0.38, 0.08)		
				239	Reflexes; RR = 1.49	(1.12, 1.98)		
		Maternal urinary DEPs > median versus ≤ median		239	RR = 2.3	(1.1, 5.0)		Multivariable logistic regression with number of abnormal reflexes dichotomized at two or more

scores or by the WISC-IV, was assessed in four studies (Table 5). In a study of the Columbia cohort, Rauh *et al.* (2006) dichotomized the cord plasma chlorpyrifos levels measured by Perera *et al.* (2003) and Whyatt *et al.* (2004) to classify subjects into high exposure (> 6.17 pg/g) or lower exposure (≤ 6.17 pg/g) groups, because previous analyses had indicated associations with birth weight in this cohort at exposure levels greater than 6.17 pg/g (Perera *et al.*, 2003; Whyatt *et al.*, 2004). The authors reported no association between increased chlorpyrifos exposure measured at birth and MDI scores at 12, 24, or 36 months of age using adjusted multivariate regression analyses. An association between increased chlorpyrifos exposure and lower PDI scores at 36 months was reported ($p = 0.003$), but no associations were reported at 12 or 24 months. The risks of mental or psychomotor delays, as determined by MDI and PDI scores, were assessed using adjusted logistic regression. The authors reported increased risks for mental (OR = 2.37, 95% CI: 1.08–5.19) and motor (OR = 4.52, 95% CI: 1.61–12.70) delays for children in the high exposure group at 36 months and no increase in risk for either type of delay in the high exposure group at 12 and 24 months of age compared to those in the low exposure group.

Eskenazi *et al.* (2007) evaluated the association between chlorpyrifos exposure and MDI and PDI scores in children of the CHAMACOS cohort at 6, 12, and 24 months of age. The authors measured TCPy and DEPs in maternal urine collected at 14 and 26 weeks gestation and DEPs in child urine at the time of assessment with the BSID-II. The metabolite concentrations were $\log_{10}$-transformed and the maternal concentrations were averaged to create prenatal exposure values. A large proportion of the maternal urine samples had non-detectable levels of TCPy, so levels of this metabolite were categorized into three groups: < LOD for both measurements, and subdivided below and above the median of the average level for those with at least one detectable level. Adjusted multiple regression models revealed no associations between maternal TCPy levels and MDI or PDI scores at any age. The authors also examined the change in MDI and PDI scores associated with a 10-fold increase in DEPs. An overall pattern of negative associations between prenatal DEPs and MDI scores and positive associations between concurrent child DEPs and MDI scores was observed, but these associations were not statistically significant, with the exception of the positive association with concurrent DEPs at the 12-month assessment ($\beta = 1.89$, 95% CI: 0.21–3.58, $p \leq 0.05$). Neither prenatal nor concurrent urinary levels of DEPs were associated with PDI scores at any age.

Engel *et al.* (2011) used generalized linear models to examine associations between DEPs in maternal urine collected during the third trimester and MDI and PDI scores assessed at 12 and 24 months of age in children of the Mount Sinai cohort. The authors reported no associations between levels of DEPs and both MDI and PDI scores at either age.

Cognitive development was also assessed *via* the WISC-IV during the early school-age years in two studies. Rauh *et al.* (2011) examined the association between cord plasma chlorpyrifos, analyzed as a continuous variable, and log (ln)-transformed WISC-IV scores in the Columbia cohort at age seven using linear regression models adjusted for multiple covariates. The authors determined that a 1 pg/g increase in cord blood chlorpyrifos was associated with an estimated decrement in working memory scores ranging from 0.35 to 0.81 points ($p = 0.003$) and in full-scale IQ ranging from 0.20 and 0.40 points. Although the latter deficit was statistically significant ($p = 0.048$), the 95% confidence interval ranged from −0.006 to 0.001. Chlorpyrifos exposure was not associated with changes in scores for the WISC-IV indices of verbal comprehension, perceptual reasoning, or processing speed.

Engel *et al.* (2011) examined the association between maternal urinary levels of DEPs during the third trimester and performance on the WISC-IV in 7- to 9-year-old children from the Mount Sinai cohort. The authors used generalized linear models and reported no associations between DEPs and any of the WISC-IV indices.

Overall, the four studies examining chlorpyrifos exposure and cognitive and motor development are not consistent with respect to the exposure metric and timing of outcome assessment. Rauh *et al.* (2006) reported an association between increasing chlorpyrifos exposure and lower PDI scores at 36 months, as well as increased risks of mental and motor delays in highly-exposed children at this age, but Eskenazi *et al.* (2007) and Engel *et al.* (2011) assessed children only to 24 months of age. No associations between increasing chlorpyrifos exposure and lower MDI scores were reported at any age in each of the studies examining this endpoint, and both Eskenazi *et al.* (2007) and Engel *et al.* (2011) reported no associations with PDI scores. Rauh *et al.* (2011) reported a decrement in working memory scores on the WISC-IV associated with chlorpyrifos exposure in 7-year-old children, whereas Engel *et al.* (2011) reported no association with changes in scores on the WISC-IV in children between seven and 9 years of age.

3.2.2.4. Child behavioral outcomes Associations between chlorpyrifos exposure and behavioral outcomes in children were assessed in two of the studies that examined cognitive and motor development (Table 6). In an analysis of the Columbia cohort, Rauh *et al.* (2006) again used cord plasma chlorpyrifos levels dichotomized into high and low exposure categories and logistic regression to determine the risks of behavior problems (as assessed by the mothers' reporting on the CBCL) in children at 36 months of age. The authors reported associations between high chlorpyrifos exposure and attention problems (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). These risk estimates were based on 3.9–4.9% of the subjects, or up to 11 children, who scored in the

clinical (abnormal) range for these problems. Because of the small study size and because the fraction affected was so small, these risk estimates have large confidence intervals and are highly unstable.

For the CHAMACOS cohort, Eskenazi *et al*. (2007) examined associations between levels of TCPy and DEPs in maternal urine and CBCL outcomes in the clinical or borderline clinical range in children at 24 months of age. Urinary DEPs were also measured for the children at the time of assessment with the CBCL. The authors noted that there were no associations between maternal TCPy levels and attention problems in the borderline clinical range, ADHD in the borderline clinical range, or PDD in the clinical range (results were not shown). Maternal and concurrent child levels of DEPs were also not associated with attention problems or ADHD, but each 10-fold increase in concurrent child DEPs was associated with an increased risk of PDD scores in the clinical range (OR = 1.72, 95% CI: 1.12–2.64).

Overall, the two studies examining behavioral outcomes in children with chlorpyrifos exposure reported associations with PDD scores based on maternal reporting, although few subjects scored in the clinical range for this outcome in the study by Rauh *et al*. (2006). A low percentage of scores in the clinical range was also noted for the associations with attention problems and ADHD reported by Rauh *et al*. (2006).

3.2.3. Analysis of human data

To evaluate the relationship between chlorpyrifos exposure and neurodevelopmental effects, we first considered the overall consistency of the reported results across studies. Null results were reported for the outcome of head circumference in all of the cohorts, regardless of the exposure metric used. The two studies assessing infant neurobehavior in the CHAMACOS and Mount Sinai cohorts reported associations between increasing levels of DEPs and abnormal reflexes. The three studies of cognitive and motor development assessed *via* the BSID-II in the Columbia, CHAMACOS, and Mount Sinai cohorts were not consistent with respect to the exposure metric and timing of assessment, with associations being reported only in children at 36 months of age in the Columbia cohort, whereas children were only examined up to 24 months of age, with null results, in the CHAMACOS and Mount Sinai cohorts. Two of these same studies reported inconsistent results for child behavior outcomes, with the Columbia cohort study reporting associations with all three behavioral outcomes examined at 36 months, whereas the CHAMACOS cohort study only reported an association with PDD at 24 months of age with DEPs, but not TCPy, as the exposure metric. The two studies of cognitive development assessed *via* the WISC-IV were also inconsistent, with associations reported for decrements in working memory scores in the Columbia cohort but no associations with WISC-IV scores in the Mount Sinai cohort. Although it is presumed that studies of the same cohort represent different analyses of the same subjects, loss to follow-up and other factors led to different sample sizes in these studies. Because of this, results across studies of the same cohort are difficult to compare on an individual-by-individual basis. One cannot assess correlations between outcomes across these studies but can only look for trends within cohorts.

Because of the inconsistencies within and among studies, we critically examined the epidemiology data as a whole to assess whether the weight of the epidemiology evidence supports the hypothesis that chlorpyrifos is associated with neurobehavioral effects. This evaluation considered the reliability of the different types of exposure metrics used and whether results were dependent on the robustness of the exposure measurements. It also considered the validity of each neurodevelopmental outcome assessment and whether the results of these evaluations have clinical significance. We also determined whether potential confounding factors in these cohorts were addressed and whether any observed risks may have been attributable to them. Finally, we assessed whether there were consistent exposure-response associations within and across studies and potential statistical limitations among studies. Together, these analyses allowed for an assessment of which study results are likely to be the most valid, based on the weight of each of the factors examined, and whether they provide sufficient evidence to support the hypothesis that chlorpyrifos causes neurodevelopmental effects.

3.2.3.1. Exposure assessment The epidemiology studies examining neurodevelopmental effects potentially related to chlorpyrifos exposure used the following exposure metrics, most o f which were measured at only one point in time: chlorpyrifos levels in maternal and/or cord blood; maternal levels of TCPy in urine; maternal or child levels of DEPs in urine; cholinesterase activity in whole blood or plasma; and chlorpyrifos concentrations in ambient air.

Five of the cohort studies used chlorpyrifos levels in cord plasma or serum as an exposure metric (Perera *et al*., 2003; Whyatt *et al*., 2004; Rauh *et al*., 2006, 2011; Barr *et al*., 2010). Cord blood levels reflect the amount of chlorpyrifos absorbed by the mother and also the amount of absorbed dose transferred to the developing fetus (Needham, 2005), although, in being measured at birth, they most strongly reflect recent exposure rather than exposure earlier in development, which may differ if exposure levels are not consistent throughout pregnancy. As discussed in Section 3.1, chlorpyrifos is oxidized to chlorpyrifos-oxon, through which toxicity from cholinesterase inhibition is hypothesized to occur in the brain. Chlorpyrifos-oxon is rapidly hydrolyzed to TCPy or DEP, and has not been detected in human blood or urine after oral administration (Timchalk *et al*., 2002). For exposure assessment of chlorpyrifos, direct measurement of the parent compound more accurately reflects the chlorpyrifos dose in the brain than do measurements of its metabolites in urine, which are not on the hypothesized

Table 5. Results of cohort studies assessing cognitive and motor development using the BSID-II or WISC-IV.

Outcome assessed	Cohort	Study	Exposure metric/group	Exposure concentration	Age at assessment	N	Estimate	95% CI	p-value	Statistical method
MDI scores	Columbia	Rauh *et al.*, 2006	Chlorpyrifos in cord blood	Range: ND to 63 pg/g						Multivariate linear regression with log-transformed, dichotomized exposure variable
			High exposure (>6.17 pg/g) compared to low exposure (≤ 6.17 pg/g)		12 months	229	$\beta = -0.344 \pm 1.66$		0.836	
					24 months	225	$\beta = -1.480 \pm 2.03$		0.466	
					36 months	228	$\beta = -3.327 \pm 1.76$		0.6	
			High exposure (>6.17 pg/g) compared to low exposure (≤ 6.17 pg/g)		12 months	228	OR = 1.22	(0.48, 3.06)		Logistic regression with log-transformed, dichotomized exposure variable; odds of mental delay
					24 months	228	OR = 1.75	(0.86, 3.60)		
					36 months	228	OR = 2.37	(1.08, 5.19)		
	CHAMACOS	Eskenazi *et al.*, 2007	Maternal urinary TCPy	Median: 3.54 µg/L						Multiple linear regression with log-transformed metabolite levels as continuous variables
			< *Median detected*		6 months	395	$\beta = 0.24$	(–2.12, 2.60)		
			≥ *Median detected*				$\beta = 0.08$	(–2.29, 2.44)		
			< *Median detected*		12 months	393	$\beta = -0.45$	(–3.67, 2.76)		
			≥ *Median detected*				$\beta = -0.65$	(–3.88, 2.58)		
			< *Median detected*		24 months	369	$\beta = -1.02$	(–5.34, 3.31)		
			≥ *Median detected*				$\beta = -1.94$	(–6.26, 2.37)		
			Maternal urinary DEPs	Geometric mean: 18.1 nmol/L						
					6 months	395	$\beta = -0.16$	(–1.96, 1.65)		
					12 months	393	$\beta = -1.14$	(–3.51, 1.22)		
					24 months	369	$\beta = -0.85$	(–3.98, 2.27)		
			Concurrent child urinary DEPs	Geometric mean: 10.6 nmol/L	6 months	395	$\beta = 0.24$	(–0.78, 1.25)		
				Geometric mean: 15.2 nmol/L	12 months	393	$\beta = 1.89$	(0.21, 3.58)	≤ 0.05	
				Geometric mean: 10.5 nmol/L	24 months	369	$\beta = 1.02$	(–0.52, 2.57)		
	Mount Sinai	Engel *et al.*, 2011	Maternal urinary DEPs	Median: 20.2 nmol/L	12 months	149	$\beta = 0.03$	(–2.23, 2.29)		Generalized linear models with log-transformed DEP levels as a continuous variable
					24 months	208	$\beta = -1.47$	(–3.99, 1.04)		

(Continued)

Table 5. (Continued).

Outcome assessed	Cohort	Study	Exposure metric/group	Exposure concentration	Age at assessment	N	Estimate	95% CI	p-value	Statistical method
PDI scores	Columbia	Rauh *et al.*, 2006	Chlorpyrifos in cord blood	Range: ND to 63 pg/g						Multivariate linear regression with log-transformed, dichotomized exposure variable
			High exposure (>6.17 pg/g) compared to low exposure (≤ 6.17 pg/g)		12 months	228	$\beta = -3.30 \pm 2.11$		0.12	
					24 months	227	$\beta = 1.17 \pm 1.98$		0.56	
					36 months	228	$\beta = -6.46 \pm 2.18$		0.003	
			High exposure (>6.17 pg/g) compared to low exposure (≤ 6.17 pg/g)		12 months	228	OR = 1.88	(0.78, 4.53)		Logistic regression with log-transformed, dichotomized exposure variable; odds of psychomotor delay
					24 months	228	OR = 1.01	(0.37, 2.76)		
					36 months	228	OR = 4.52	(1.61, 12.70)		
	CHAMACOS	Eskenazi *et al.*, 2007	Maternal urinary TCPy	Median: 3.54 μg/L						Multiple linear regression with log-transformed metabolite levels as continuous variables
			< *Median detected*		6 months	395	$\beta = -0.56$	(−4.03, 2.91)		
			≥ *Median detected*				$\beta = -0.21$	(−3.69, 3.27)		
			< *Median detected*		12 months	393	$\beta = -0.70$	(−5.26, 3.86)		
			≥ *Median detected*				$\beta = -1.62$	(−6.20, 2.96)		
			< *Median detected*		24 months	369	$\beta = -2.65$	(−6.50, 1.21)		
			≥ *Median detected*				$\beta = -2.72$	(−6.57, 1.12)		
			Maternal urinary DEPs	Geometric mean: 18.1 nmol/L						
					6 months	395	$\beta = 0.02$	(−2.63, 2.67)		
					12 months	393	$\beta = 0.30$	(−3.03, 3.63)		
					24 months	369	$\beta = -0.86$	(−3.64, 1.92)		
			Concurrent child urinary DEPs	Geometric mean: 10.6 nmol/L	6 months	395	$\beta = 0.60$	(−0.89, 2.09)		
				Geometric mean: 15.2 nmol/L	12 months	393	$\beta = 1.91$	(−0.46, 4.27)		
				Geometric mean: 10.5 nmol/L	24 months	369	$\beta = 0.30$	(−1.07, 1.67)		

(Continued)

Table 5. (Continued).

Outcome assessed	Cohort	Study	Exposure metric/group	Exposure concentration	Age at assessment	N	Estimate	95% CI	p-value	Statistical method
	Mount Sinai	Engel *et al.*, 2011	Maternal urinary DEPs	Median: 20.2 nmol/L	12 months	149	β = –0.20	(–3.28, 2.87)		Generalized linear models with log-transformed DEP levels as a continuous variable
					24 months	210	β = 0.67	(–1.72, 3.06)		
WISC-IV scores	Columbia	Rauh *et al.*, 2011	Chlorpyrifos in cord blood	Range: ND to 63 pg/g	7 years	265	Full-Scale IQ; β = –0.003	(–0.006, 0.001)	0.048	Multivariate linear regression with exposure as a continuous variable and ln-transformed WISC-IV scores
							Working Memory; β = –0.006	(–0.010, –0.002)	0.003	
							Verbal Comprehension; β = –0.002	(–0.005, 0.001)	0.208	
							Perceptual Reasoning; β = –0.002	(–0.006, 0.002)	0.29	
							Processing Speed; β = 0.001	(–0.004, 0.005)	0.728	
	Mount Sinai	Engel *et al.*, 2011	Maternal urinary DEPs	Median: 20.2 nmol/L	7–9 years	114	Full-Scale IQ; β = –3.15	(–7.19, 0.89)		Generalized linear models with log-transformed DEP levels as a continuous variable
							Working Memory; β = –3.48	(–7.29, 0.34)		
							Verbal Comprehension; β = –0.08	(–3.91, 3.76)		
							Perceptual Reasoning; β = –4.37	(–9.10, 0.36)		
							Processing Speed; β = –2.11	(–6.81, 2.59)		

Table 6. Results of cohort studies assessing child behavioral outcomes using the CBCL.

Outcome assessed	Cohort	Study	Exposure metric/group	Exposure concentration	Age at assessment	N	Estimate	95% CI	p-value	Statistical method
ADHD	Columbia	Rauh *et al.*, 2006	Chlorpyrifos in cord blood	Range: ND to 63 pg/g						Logistic regression with log-transformed, dichotomized exposure variable
			High exposure (> 6.17 pg/g) compared to low exposure (≤ 6.17 pg/g)		36 months	228	OR = 6.50	(1.09, 38.69)		
	CHAMACOS	Eskenazi *et al.*, 2007	Maternal urinary DEPs	Geometric mean: 18.1 nmol/L	24 months	356	OR = 0.59	(0.21, 1.68)		Multiple regression; OR per 10-fold increase in log-transformed DEP levels
			Concurrent child urinary DEPs	Geometric mean: 10.5 nmol/L	24 months	356	OR = 1.18	(0.72, 1.94)		
Attention problems	Columbia	Rauh *et al.*, 2006	Chlorpyrifos in cord blood	Range: ND to 63 pg/g						Logistic regression with log-transformed, dichotomized exposure variable
			High exposure (> 6.17 pg/g) compared to low exposure (≤ 6.17 pg/g)		36 months	228	OR = 11.26	(1.79, 70.99)		
	CHAMACOS	Eskenazi *et al.*, 2007	Maternal urinary DEPs	Geometric mean: 18.1 nmol/L	24 months	356	OR = 0.78	(0.26, 2.31)		Multiple regression; OR per 10-fold increase in log-transformed DEP levels
			Concurrent child urinary DEPs	Geometric mean: 10.5 nmol/L	24 months	356	OR = 1.02	(0.61, 1.71)		
PDD	Columbia	Rauh *et al.*, 2006	Chlorpyrifos in cord blood	Range: ND to 63 pg/g						Logistic regression with log-transformed, dichotomized exposure variable
			High exposure (> 6.17 pg/g) compared to low exposure (≤ 6.17 pg/g)		36 months	228	OR = 5.39	(1.21, 24.11)		
	CHAMACOS	Eskenazi *et al.*, 2007	Maternal urinary DEPs	Geometric mean: 18.1 nmol/L	24 months	355	OR = 0.88	(0.37, 2.07)		Multiple regression; OR per 10-fold increase in log-transformed DEP levels
			Concurrent child urinary DEPs	Geometric mean: 10.5 nmol/L	24 months	355	OR = 1.72	(1.12, 2.64)	≤ 0.05	

causal pathway, or measurement of cholinesterase activity, which is affected by other compounds in addition to chlorpyrifos.

The study by Barr *et al.* (2010) also used maternal levels of chlorpyrifos in serum as an exposure metric. Maternal blood levels of chlorpyrifos may be a surrogate for fetal levels of chlorpyrifos, because chlorpyrifos is distributed through the body in the lipids of the blood, brain, and other tissues (Needham, 2005). They are not as good a metric as cord blood levels, however, because they do not reflect the absorbed dose transferred to the fetus. If the exposure route is ingestion, much of the absorbed chlorpyrifos may undergo first-pass metabolism to TCPy before it reaches the maternal systemic blood supply and, hence, the placenta (Needham, 2005). In addition, blood chlorpyrifos concentrations depend somewhat on the equilibrium between concentrations in adipose tissue and blood (Needham, 2005), and blood lipid levels can increase up to four-fold during pregnancy, resulting in the distribution of chlorpyrifos in blood being affected by these lipid changes (McMullin *et al.*, 2008). If not accounted for, this can result in overestimation of chlorpyrifos concentrations in maternal and fetal serum or plasma. Concentrations of chlorpyrifos in blood can be evaluated on both a concentration basis and a lipid basis to adjust for this, but this was not done in any of the cohort studies.

Whyatt *et al.* (2004) used personal monitoring of prenatal chlorpyrifos concentrations in ambient air for 48 hours during the third trimester in their study of the Columbia cohort. Exposure to chlorpyrifos can come from multiple sources (diet, residential and workplace use) and *via* multiple routes (ingestion, inhalation, dermal absorption), however. Although inhalation is a likely route of exposure in the Columbia cohort, air concentrations do not reflect the amount of chlorpyrifos absorbed from all of the potential sources and routes. Because of this, they are not as good of an exposure measure as biomarkers that are internal dosimeters.

Three studies used measurements of TCPy in urine as an exposure metric (Eskenazi *et al.*, 2004, 2007; Berkowitz *et al.*, 2004). Urinary TCPy is not a reliable biomarker of exposure to chlorpyrifos because TCPy in urine originates from exposure to not only chlorpyrifos, but to chlorpyrifos-methyl, triclopyr, and TCPy itself (MacIntosh *et al.*, 1999; Needham, 2005; Morgan *et al.*, 2005). Thus, measures of TCPy in urine can overestimate exposure to chlorpyrifos in both the mother and fetus. Exposure to pre-formed TCPy can occur through multiple environmental sources, including the diet (Morgan *et al.*, 2005). TCPy is not cholinesterase-inhibiting (Morgan *et al.*, 2005) and has not been shown to be associated with adverse effects (Eaton *et al.*, 2008).

Six of the studies used urinary levels of DEPs as the exposure metric (Eskenazi *et al.*, 2004, 2007; Engel *et al.*, 2007, 2011; Young *et al.*, 2005; Wolff *et al.*, 2007). Urinary DEPs originate from exposure to diazinon, disulfoton, and DEPs themselves in addition to chlorpyrifos, so their measurement is also not a specific biomarker of chlorpyrifos exposure and can overestimate chlorpyrifos levels in the mother and fetus.

Two studies used cholinesterase activities in whole blood or plasma as an exposure metric (Eskenazi *et al.*, 2004; Wolff *et al.*, 2007). Although the interrelationship of the different exposure metrics is not well studied, Eskenazi *et al.* (2004) found no correlation between concurrent measurements of DEPs in urine and either total cholinesterase activity in blood or BuChE activity in plasma in the CHAMACOS cohort. Measurement of cholinesterase activity is not a chlorpyrifos-specific exposure metric because other chemicals, including other OPs and N-methyl carbamate pesticides, inhibit cholinesterases. Another drawback of monitoring cholinesterase activity is that large doses of chlorpyrifos are required for significant inhibition; therefore, such measures are more appropriately used as indicators of toxicity at high exposure levels and are relatively insensitive at the low exposure levels observed in the cohort studies (Wessels *et al.*, 2003).

Most of the exposure metrics used in the cohort studies were measured at only one point in time. Exposures to chlorpyrifos and other pesticides can be transient and highly variable from day to day, so a single measurement may not represent average exposure over time or exposure at some earlier or later time. For example, large intraindividual variability in maternal TCPy and DEP metabolite levels was reported in the Columbia and CHAMACOS cohorts (Whyatt *et al.*, 2009; Eskenazi *et al.*, 2004, 2007; Young *et al.*, 2005), and Eskenazi *et al.* (2007) reported that urinary DEPs in children measured more than a few days apart are uncorrelated, suggesting considerable intraindividual variability in each study.

Another issue contributing to exposure measurement error with the use of pesticide biomarkers is that OPs and their metabolites have a short residence time in the body. Once absorbed, they are rapidly eliminated with biologic half-lives on the order of hours to days in adults (Barr *et al.*, 2002). Thus, any measure of chlorpyrifos or its metabolites in blood or urine at a single time point reflects exposure during the brief period of time prior to measurement and may not accurately reflect exposure throughout the entire critical period of neurodevelopment unless exposure was continuous and yielded a steady-state concentration. For example, chlorpyrifos measurements at or near the time of delivery, such as in cord blood, would reflect only exposure during late pregnancy. The period of vulnerability to chlorpyrifos may begin earlier in pregnancy and extend through the period of synaptic modeling, which continues well into childhood and adolescence (Slotkin, 2004), but the data do not allow conclusions regarding earlier or later exposures. Even a small difference between the measured levels and the actual levels prior to conception, during early pregnancy, or during early childhood could lead to a relatively high degree of exposure measurement error, biasing the results, especially if the day-to-day variation is

substantial compared to the amount of variation among subjects in their single measurements. As is the case for biomarkers, the single prenatal air sample measured in the study by Whyatt *et al*. (2004) may not be representative of average exposure during a particular trimester or the entire pregnancy.

In some studies, many of the measured biomarker concentrations were at or near the LOD (*e.g.*, Barr *et al*., 2010; Berkowitz *et al*., 2004; Eskenazi *et al*., 2004, 2007). The uncertainty associated with such measurements can lead to exposure measurement error, biasing results toward or away from the null.

Direct measurement of chlorpyrifos in cord blood is the most reliable exposure metric, and studies using this method should carry more weight in assessing the hypothesis that chlorpyrifos exposure causes neurodevelopmental effects.

3.2.3.2. Outcome assessment The epidemiology studies considered the following neurodevelopmental outcomes: head circumference; infant neurobehavior as assessed by the BNBAS; cognitive development as assessed by MDI scores on the BSID-II and by the WISC-IV; motor development as assessed by PDI scores on the BSID-II; and behavioral outcomes (attention problems, ADHD, and PDD) as assessed by the CBCL. Measurements of these outcomes can be informative for assessing neurodevelopmental effects, but the cohort study results are limited by the sensitivity and predictive ability of these standardized developmental tests in the first few years of life.

Head circumference: Head circumference was examined in six of the cohort studies (Perera *et al*., 2003; Whyatt *et al*., 2004; Eskenazi *et al*., 2004; Berkowitz *et al*., 2004; Wolff *et al*., 2007; Barr *et al*., 2010). Head circumference correlates with brain weight (Lemons *et al*., 1981) and some studies have reported that reduction in head circumference correlates with lower IQ and poorer cognitive functioning in childhood (Hack *et al*., 1991; Lasky *et al*., 1981; Ounsted *et al*., 1988; Rushton and Ankney, 2009), whereas others have reported no strong influence of brain volume on overall cognitive performance (Schoenemann *et al*., 2000). The studies that examined head circumference compared cord blood levels to head circumference only at birth, so the exposures measured do not necessarily precede the outcome measured, and this single measure can lead to misclassification of outcome.

Infant neurobehavior: Infant neurobehavior was examined in two of the cohort studies (Young *et al*., 2005; Engel *et al*., 2007). Both cohort studies measured infant behavior using the BNBAS, which consists of 28 behavioral items scored on a nine-point scale and 18 reflex items scored on a four-point scale. Each of these items were reduced to seven clusters based on the scoring method developed by Lester *et al*. (1982): habituation, orientation, motor performance, range of state, regulation of state, autonomic stability, and primitive reflex. The six behavioral cluster scores are calculated such that higher scores represent more optimal functioning, whereas the reflex cluster score is the total number of reflexes coded as abnormal, so that higher scores indicate less optimal functioning. Young *et al*. (2005) stated that the 18 reflex items of the BNBAS are not designed to provide a neurological diagnosis, but can potentially identify gross neurologic abnormalities, as more than three abnormally rated reflexes may be clinically relevant (Brazelton and Nugent, 1995).

Young *et al*. (2005) stated that the BNBAS can be appropriately administered through the first two months of life, and the authors assessed infants between 0 and 62 days of age. Engel *et al*. (2007) conducted assessments within five days of birth and noted that 23% of the initial cohort was not assessed for various reasons, including weekend delivery. This is a potential source of selection bias, as any factors related to weekend delivery, such as fewer induced deliveries, may be underrepresented among the subjects who were tested. Engel *et al*. (2007) stated that even if those factors are related to exposure or disease status, this alone does not impose a bias on their findings, as fewer induced deliveries with pesticide exposure alone would affect precision but would not affect the validity of the estimates.

In both studies, each infant was assessed only once. The prognostic utility of a single assessment of infant reflexes shortly after birth is unclear and the study results may be subject to misclassification of outcome from the use of a single measurement of neurobehavior.

Cognitive and motor development: Three of the cohort studies examined associations between chlorpyrifos exposure and cognitive and motor development, as measured by the BSID-II within the first 3 years of life (Eskenazi *et al*., 2007; Rauh *et al*., 2006; Engel *et al*., 2011). The BSID-II is used to identify young children at risk for mental and motor delay. Rauh *et al*. (2006) stated that when administered at 3 years of age, the BSID-II demonstrates moderate predictive power for subsequent intelligence and school performance but is only clinically useful for children performing in the subnormal range.

Two studies assessed cognitive development between the ages of seven and 9 years through scores on the WISC-IV (Rauh *et al*., 2011; Engel *et al*., 2011). The WISC-IV measures four indices of cognitive functioning (verbal comprehension, working memory, perceptual reasoning, and processing speed) that are combined to yield a full-scale IQ score. Rauh *et al*. (2011) noted that WISC-IV scores can be influenced by factors such as socioeconomic status (SES) and child behavior problems.

Behavioral outcomes: Two studies examined behavioral outcomes as measured by scores on the 99-item CBCL for ages 18 months to 5 years (Rauh *et al*., 2006; Eskenazi *et al*., 2007). The CBCL is a widely used measure to assess children's emotional and behavioral problems and competencies during the previous two months, and its validity and reliability have been documented

(Achenbach *et al.*, 2003). It was administered to mothers in the Columbia cohort when the children were 36 months of age (Rauh *et al.*, 2006) and to mothers in the CHAMACOS cohort when their children were 24 months of age (Eskenazi *et al.*, 2007). Both studies examined the results of three scales of the CBCL: attention problems, ADHD, and PDD. The attention problem scale rates behaviors related to concentration and sitting still. The ADHD scale includes additional attention items such as "gets into everything," and criteria for this scale are derived from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV), which is used as a diagnostic tool (APA, 2000, as cited by Eskenazi *et al.*, 2007). The PDD scale criteria are also derived from the DSM-IV, and they rate behaviors that are consistent with Asperger's Disorder and Autistic Disorder, and include items such as avoiding eye contact, rocking of the head and body, and unresponsiveness to affection. Scores on each scale are considered of borderline clinical significance if they are > 93rd percentile of the national norms and of clinical significance if > 97th percentile (Achenbach and Rescorla, 2000).

The single assessment of behavioral outcomes at one age per study may lead to outcome misclassification, particularly when the measure is based on maternal report, which is subject to reporting bias. Rauh *et al.* (2006) noted that the DSM-IV, from which the CBCL criteria for ADHD are derived, has low sensitivity for assessing the inattentiveness of preschool-aged children, limiting the usefulness of this outcome measure.

3.2.3.3. Clinical significance To determine whether any reported outcomes associated with chlorpyrifos exposure are adverse effects, one should assess whether they are clinically significant (Goodman *et al.*, 2010). That is, whether or not a statistically significant association has been observed, one should determine whether an outcome actually constitutes an adverse effect. If the outcomes observed in the cohort studies are clinically significant, they are considered to constitute an adverse effect.

Two studies used the BNBAS to assess seven clusters of infant neurobehavior, including abnormal reflexes (Young *et al.*, 2005; Engel *et al.*, 2007). More than three abnormally rated reflexes may be clinically relevant, often resulting in a more intensive neurologic examination and possible intervention (Brazelton and Nugent, 1995).

Three studies examined associations between chlorpyrifos exposure and cognitive and motor development as assessed by MDI and PDI scores on the BSID-II (Rauh *et al.*, 2006; Eskenazi *et al.*, 2007; Engel *et al.*, 2011). Scores for the BSID-II have a mean (standard deviation [SD]) of 100 (15) for "normal," so, assuming a normal distribution, in 68% of a standard population, the scores range from 85 to 115. As noted above, the BSID-II is only clinically useful for predicting subsequent intelligence and school performance in children scoring in the subnormal range (< 85).

Two studies used the WISC-IV to examine associations between chlorpyrifos exposure and cognitive development (Rauh *et al.*, 2011; Engel *et al.*, 2011). Like the BSID-II, scores for the WISC-IV have a mean and standard deviation of 100 and 15, respectively, so "normal" scores range from 85 to 115.

Two studies assessed the association between chlorpyrifos exposure and the risks of scoring in the clinical range of the CBCL for attention problems, ADHD, and PDD (Rauh *et al.*, 2006; Eskenazi *et al.*, 2007). Scores on each of these scales are considered of borderline clinical significance if they are > 93rd percentile of the national norms and of clinical significance if > 97th percentile. Eskenazi *et al.* (2007) reported a small percentage of children scoring in the clinical range for attention problems and ADHD (2.0–2.8%), so they used the less-conservative borderline cut-points for these outcomes. The true clinical significance of scores for DSM-IV-oriented scales, such as ADHD and PDD, is still unknown because scores in the clinical range of these scales based on maternal report are not directly equivalent to a DSM diagnosis (Achenbach and Rescorla, 2000).

3.2.3.4. Confounding and bias Many genetic and environmental factors are hypothesized to affect neurodevelopmental outcomes, and these factors could be correlated with exposure to chlorpyrifos or other pesticides, generating confounding. Only a small number of these factors were considered as confounders in the cohort studies. Because several studies examined the same cohorts, confounding factors that affect outcomes in a given cohort will do so across all studies of that cohort. Lack of adjustment for these factors in statistical models decreases the likelihood that any observed effects are attributable to chlorpyrifos exposure.

Each cohort was exposed to multiple types of pesticides besides chlorpyrifos. While studies of the Columbia and Mount Sinai cohorts used chlorpyrifos levels as the exposure metric, confounding by other pesticides could have occurred. Other studies relied on non-specific metabolite levels that reflect exposure to other pesticides in addition to chlorpyrifos, so it cannot be known whether any observed associations are attributable to chlorpyrifos, these other pesticides, or some other confounder. Chlorpyrifos exposures were often correlated with other pesticides, and only one of the studies controlled for other pesticide exposures to ensure that the main effects of chlorpyrifos were not attributable to other pesticides. Whyatt *et al.* (2004) controlled for diazinon and the carbamate insecticide propoxur (*via* its metabolite, 2-isopropoxyphenol) in models predicting the associations between chlorpyrifos exposure and birth weight and birth length, but did not control for these pesticides in models for head circumference. For these reasons, any observed associations in the cohort studies may be attributable to pesticides other than chlorpyrifos.

Exposure to lead has been associated with adverse cognitive, motor, and behavioral outcomes in children, and cohort studies of these effects in early childhood considered potential confounding by lead exposure. Rauh *et al*. (2006, 2011) measured lead levels as well as chlorpyrifos levels in cord blood. They reported that lead and chlorpyrifos levels were not correlated, so they did not include lead as a covariate in their regression models for each outcome. Eskenazi *et al*. (2007) also measured lead biomarkers and modeled them simultaneously with pesticide exposures for MDI scores, for which no associations were reported with TCPy and DEPs, but did not include lead exposure in the models for PDD, for which an association was reported with levels of DEPs.

Exposure to environmental tobacco smoke (ETS) has also been associated with adverse cognitive, motor, and behavioral outcomes. In studies of the Columbia cohort, the authors attempted to eliminate active smokers by excluding women with plasma cotinine levels > 25 ng/ml from the analyses (Perera *et al*., 2003). Approximately 43% of mothers reported a smoker in the home and also had cotinine values that reflected tobacco exposure. Cotinine levels have a half-life during pregnancy of 8.8 hours, and maternal plasma samples were obtained within one (Perera *et al*., 2003) or two days (Whyatt *et al*., 2004) after delivery, increasing the likelihood that cotinine status from this measure is considerably underestimated. All mothers also had detectable inhalation levels of one or more PAH, but Perera *et al*. (2003) reported no significant interactions between PAHs and chlorpyrifos.

Cotinine levels were not measured in the CHAMACOS and Mount Sinai cohorts, and information regarding smoking use in the mothers was obtained solely by questionnaires, a practice that is subject to reporting bias in that mothers may have underreported this factor because it is not socially desirable. There was no information collected on smoking in the study of the UMDNJ cohort (Barr *et al*., 2010).

Maternal alcohol has been shown to be a significant predictor of adverse behaviors in children, and attention problems in particular (Sood *et al*., 2001). Information regarding alcohol use was obtained solely by questionnaires in each cohort, which, as noted above, is subject to reporting bias. Perera *et al*. (2003) noted that 24% of the mothers in the Columbia cohort reported alcohol use during pregnancy, and in analyses of behavioral outcomes in this cohort, Rauh *et al*. (2006) did not control for alcohol exposure. Mothers in the Mount Sinai cohort were excluded if they drank more than two alcoholic drinks per day, which eliminates only those with heavy consumption and introduces uncertainty regarding alcohol exposure levels across the remainder of the cohort. There was no information collected on alcohol use in the study of the UMDNJ cohort (Barr *et al*., 2010).

The cohorts in which associations were reported all come from populations with low SES. The effects of low SES on fetal and child neurodevelopment have been demonstrated in a range of populations (Rauh *et al*., 2004); thus, it is expected that children in these cohorts should have lower scores in the outcome measures examined. Both Rauh *et al*. (2006) and Eskenazi *et al*. (2007) observed a large increase in the percentage of children with deficits in MDI scores at the final time point in their study in which the children were examined (36 months and 24 months, respectively), which could be attributed to a lack of stimulating environments, which is common in low SES populations, or to this measure not being clinically valid in Spanish or Latino immigrant communities, as the BSID-II is confounded by significant language demands (Youngstrom *et al*., 2010). Engel *et al*. (2007) noted that the exclusions of mothers after study entry because they moved out of the area, were lost to follow-up, or lacked prenatal biological specimens may be a potential source of selection bias because these exclusion factors largely reflect socioeconomic condition. These reasons for exclusion may also be associated with other lifestyle factors that are correlated with chlorpyrifos exposure, providing other possibilities for confounding. Of the studies reporting associations with neurodevelopmental outcomes, all examined at least some confounding factors related to SES (*e.g.*, mother's education; household income; quality of home environment) and those that were associated with outcomes were included as covariates in the final models.

Maternal IQ is also a potential confounder for cognitive outcomes, and it was controlled for only in the studies by Rauh *et al*. (2006, 2011). Rauh *et al*. (2006) used the sample mean to substitute for IQ scores that were missing for 29 of the 254 mothers, however, and such imputation of missing data adds considerable uncertainty to this factor.

All of the studies, with the exception of the UMDNJ cohort study, were conducted during a time period spanning the phase-out of residential uses of chlorpyrifos that began on January 1, 2001. After the ban, it was possible that some families continued to use chlorpyrifos products purchased before the ban, so it is difficult to determine the date when exposure stopped. It is possible that other factors distinguished the pre- and post-ban periods, and that any of the reported associations with neurodevelopmental outcomes may have been attributable to these factors. This possibility was addressed only in studies of the Columbia cohort. Whyatt *et al*. (2004) reported that levels of chlorpyrifos in ambient air and blood samples from this cohort decreased substantially between 1998 and 2002. When their analysis with chlorpyrifos air data was stratified among newborns born before *versus* on or after January 1, 2001, there was still no association with head circumference. The authors did not do this analysis using chlorpyrifos levels in blood as the exposure metric. Rauh *et al*. (2006) examined associations with MDI and PDI scores with chlorpyrifos exposures pre-ban, mid-ban, and post-ban, and observed statistically significant increases in MDI and PDI scores

from pre-ban to mid-ban, but slight decreases in these scores post-ban that were not statistically significant, indicating no improvements in mental or motor function after the ban. These results indicate that there were no confounding factors that distinguished the two periods for this cohort.

Although each cohort study adjusted for several different confounders, other factors that could affect the results may not have been accounted for, which increases the likelihood that there are alternative explanations for the observed outcomes other than exposure to chlorpyrifos. Thus, studies that do not adequately account for potential confounders carry less weight in addressing the hypothesis that chlorpyrifos causes neurodevelopmental effects.

3.2.3.5. Exposure-response If chlorpyrifos is a causal factor for neurodevelopmental effects, one would expect the risks for these effects to increase with exposure both within and among studies. Most of the chlorpyrifos cohort studies did not fully assess the exposure-response relationship. While many studies analyzed the exposure metric as a continuous variable (Perera *et al.*, 2003; Whyatt *et al.*, 2004; Wolff *et al.*, 2007; Young *et al.*, 2005; Engel *et al.*, 2007, 2011; Eskenazi *et al.*, 2004, 2007; Rauh *et al.*, 2011), some studies dichotomized exposure levels (Barr *et al.*, 2010; Eskenazi *et al.*, 2004, 2007; Berkowitz *et al.*, 2004; Rauh *et al.*, 2006). This sacrificed exposure information and does not allow full exploration of subtle relationships between exposure and outcome.

Only two studies performed analyses assessing exposure-response relationships. Rauh *et al.* (2006) originally categorized exposure into tertiles, with an additional referent group having undetectable chlorpyrifos levels, and they observed lower mean MDI and PDI scores in the highest tertile and the referents compared with the low and middle tertiles, indicating a lack of an exposure-response relationship. The authors reported these results as "preliminary" and used the results from a *post hoc* dichotomization of exposure levels above and below the cut-off for the highest tertile in the main analyses.

Young *et al.* (2005) examined the exposure-response relationship for the proportion of infants with more than three abnormally rated reflexes across quintiles of maternal levels of DEPs. Although there was an increasing trend of this outcome with increasing quintiles of exposure, the result was only marginally statistically significant ($p = 0.05$).

It is not possible to assess clearly whether an exposure-response relationship exists among the studies because most studies reporting an association with a specific outcome did not use the same exposure metric. For example, studies of the Columbia cohort used blood levels of chlorpyrifos, whereas those of the CHAMACOS cohort used urinary metabolites. Overall, there is a lack of clear exposure-response information in most of the cohort studies, which limits their ability to aid in assessing the hypothesis that chlorpyrifos causes neurodevelopmental effects.

3.2.3.6. Statistical analyses The statistical strength of epidemiology studies is related, in part, to the number of study subjects, or sample size. The sample sizes in the chlorpyrifos cohort studies ranged from 113 to 486, depending on the exposure metric and outcome, as some studies reported missing measurements for particular metrics or the inability to assess outcomes in some of the children for various reasons. Studies with sample sizes near the high end of this range have a higher likelihood of observing an association, if indeed there truly is one, than those with sample sizes near the low end of the range. If studies with smaller sample sizes report associations when studies with larger sample sizes do not, it may be that the reported associations are statistical anomalies. Of the five studies that reported associations between chlorpyrifos exposure and neurodevelopmental outcomes, three had sample sizes in the lower half of this range (Engel *et al.*, 2007; Rauh *et al.*, 2006, 2011), and two were in the upper half (Young *et al.*, 2005; Eskenazi *et al.*, 2007).

When data are analyzed in multiple ways or for multiple outcomes, the meaning of the statistical tests that are used becomes distorted by the multiple comparisons problem, such that if enough tests are run, it becomes more likely that several results will be statistically significant by chance. For example, Young *et al.* (2005) performed several different analyses of data relating to BNBAS scores and reported statistically significant associations for abnormal reflexes in infants greater than three days of age, and also for autonomic stability in infants ≤ 3 days of age. The latter result is contrary to the *a priori* hypotheses of a detrimental impact of chlorpyrifos, however, and the authors noted that this may be the result of multiple testing, as there is no explanation for a biologically protective effect of chlorpyrifos on infant neurobehavior. It is also possible that their reported association for abnormal reflexes could be the result of multiple testing as well, particularly because the magnitudes of both associations were almost the same. Eskenazi *et al.* (2007) and Rauh *et al.* (2006) also conducted many different analyses and did not adjust for multiple comparisons.

The cohort studies that dichotomized exposure values into low and high groups used cut-off values with no biological basis. For example, Berkowitz *et al.* (2004) divided urinary TCPy concentrations by the LOD of their analytical method because of the large number of concentrations below this limit. Thus, this cut-off value was chosen on an analytical, rather than biological, basis and may have no biologic relevance. The authors did not report the use of other cut-offs with which to compare their results. Other studies used cut-off points such as the median (Eskenazi *et al.*, 2004, 2007; Engel *et al.*, 2007), the 75th percentile (Barr *et al.*, 2010), or the cut-off

value for the highest tertile (Rauh *et al.*, 2006). None of these cut-offs have a biological basis. Sensitivity analyses can be used to assess whether results are dependent on where the cut-off value is chosen, but this type of analysis was not routinely performed in the chlorpyrifos cohort studies.

Overall, studies with small sample sizes, potential issues with multiple comparisons, and arbitrary cut-offs for exposure with no biological relevance should carry less weight in assessing the hypothesis that chlorpyrifos causes neurodevelopmental effects because they may support the alternative hypothesis that the observed associations are statistical anomalies.

3.2.4. Conclusions for human data

There are many inconsistencies in the results both within and among the cohort studies examining the association between chlorpyrifos exposure and neurodevelopmental outcomes in newborns and young children. We assessed which study results are likely to be the most valid, considering the exposure metric used, outcome assessed, clinical significance of reported effects, control of confounding factors, exposure-response relationships, and statistical limitations. Based on these factors, we determined whether the studies provide sufficient evidence to support the hypothesis that chlorpyrifos causes neurodevelopmental effects, or whether there are alternative hypotheses that are more likely to explain the results.

Newborn head circumference was examined as an outcome in at least one study of each of the four cohorts. All studies examining this outcome reported no association, regardless of the exposure metric used. Each study measured exposure at only one point in time, which could lead to a high degree of exposure measurement error. Three of the studies used the most reliable exposure metric, which is direct measurement of chlorpyrifos in blood (Perera *et al.*, 2003; Whyatt *et al.*, 2004; Barr *et al.*, 2010). None of the studies adequately controlled for exposure to other pesticides or ETS, maternal alcohol use, low SES, or other potential confounding factors. While a few of the studies had very small sample sizes (*e.g.*, < 150; Perera *et al.*, 2003; Barr *et al.*, 2010), others had relatively larger sample sizes (*e.g.*, 486; Eskenazi *et al.*, 2004). Regardless of the weight of each of these factors, null results were reported across all studies of newborn head circumference, increasing the likelihood that the overall findings are robust and that there is no association between chlorpyrifos exposure and decreased newborn head circumference.

The two studies of effects on infant neurobehavior, as assessed by the BNBAS, reported associations between chlorpyrifos exposure and abnormal reflexes in the CHAMACOS (Young *et al.*, 2005) and Mount Sinai (Engel *et al.*, 2007) cohorts. These studies used only urinary metabolite levels, measured at one point in time, as the exposure metric. Outcome was only assessed at one time point as well, and Young *et al.* (2005) reported potential impacts of the wide variation in the age of infants in their assessment on their results. Their report of a biologically protective effect of exposure for one of the seven cluster scores increases the plausibility of the interpretation that their results are attributable to multiple testing, rather than a true effect. The prognostic utility of a single measurement of infant reflexes shortly after delivery is unclear. Although both studies showed effects related to the same cluster of the BNBAS, these studies suffer from potential exposure measurement error from the use of urinary DEPs as the exposure metric, and misclassification of outcome from a single assessment of neurobehavior. Because of these uncertainties, their results may also support the hypothesis that other factors are causal for this outcome or that the observed associations are statistical anomalies; thus, they should carry less weight in the assessment of neurodevelopmental effects from chlorpyrifos exposure.

The three studies that examined cognitive and motor development *via* the BSID-II used different exposure metrics, and reported no associations between chlorpyrifos exposure and MDI or PDI scores up to 24 months of age in the Columbia (Rauh *et al.*, 2006), CHAMACOS (Eskenazi *et al.*, 2007), and Mount Sinai (Engel *et al.*, 2011) cohorts. Only Rauh *et al.* (2006) examined these outcomes in children at 36 months of age, using a single measure of chlorpyrifos in cord blood as the exposure metric. At this age, associations were reported for lower PDI scores and for mental and motor delays. It is unclear whether these reported effects are clinically significant, however. The difference in mean PDI scores was a modest –6.5 points between the low and high exposure groups, and the mean score of the high-exposure group was 95.69, which is well within the normal range (> 85). Similarly, the difference in mean MDI scores between low and high exposure groups was –3.3 points, with a mean score of 87.39 for the high exposure group. Both studies conducted many different analyses without adjustment for multiple comparisons, did not provide evidence of an exposure-response relationship, and did not adequately control for confounding by other exposures and SES. Although Rauh *et al.* (2006) used a more reliable exposure metric, Eskenazi *et al.* (2007) had a sample size that was almost twice as large. Two studies also examined cognitive development through scores on the WISC-IV administered during the early school-age years. Rauh *et al.* (2011) reported an association between chlorpyrifos concentrations in cord blood and decrements in working memory scores at age seven in the Columbia cohort. By contrast, Engel *et al.* (2011) reported no associations between maternal urinary DEPs and scores on any WISC-IV index in 7- to 9-year-olds in the Mount Sinai cohort. Because of the lack of confirmation of the associations at the same age and the methodological issues described, the studies of cognitive and motor development do not carry enough weight to decrease the likelihood that there are alternative explanations for the observed outcomes other than

exposure to chlorpyrifos or that the observed associations are statistical anomalies.

Two of the same studies that measured cognitive and motor development using the BSID-II also examined behavior outcomes reported by mothers on the CBCL, although at different ages (36 months in the study by Rauh *et al.*, 2006; 24 months in the study by Eskenazi *et al.*, 2007). The only consistent association between these two studies was with increased risk of PDD. Rauh *et al.* (2006) reported associations between chlorpyrifos levels in cord blood and all three scales examined (attention problems, ADHD, and PDD) at 36 months in the Columbia cohort. Eskenazi *et al.* (2007) reported no associations between maternal TCPy or DEPs and any of these outcomes, but did report an association between concurrent child DEPs and increased risk of PDD scores in the clinical range in the CHAMACOS cohort. In both studies, very few children scored in the clinical ranges of the CBCL, limiting the clinical significance of the results. Although associations between exposure and risk of PDD were reported in both cohorts, different exposure metrics were used, and children were assessed at different ages. The single assessment of behavioral outcomes at one age per study may lead to outcome misclassification, particularly when the measure is based on reporting by mothers, which is subject to reporting bias. The study using the more robust exposure metric at only one time point (Rauh *et al.*, 2006) reported associations with all three scales examined, whereas the study with the larger sample size but using a less-reliable exposure metric at two time points (Eskenazi *et al.*, 2007) only reported an association with PDD. As with their assessment of effects on cognitive and motor development, these studies do not report consistent results and do not carry enough weight to decrease the likelihood that there are alternative explanations for the observed outcomes other than exposure to chlorpyrifos or that the observed associations are statistical anomalies.

In conclusion, the chlorpyrifos cohort studies do not report consistent results. There are very few studies of each specific neurodevelopmental outcome, limiting the ability to look for consistency across studies or cohorts. The studies with more robust factors, such as reliable exposure metrics or larger sample sizes, do not appear to be more likely to report associations with adverse neurodevelopmental effects. Overall, the epidemiology data are not sufficiently robust to support the hypothesis that chlorpyrifos is a causal factor in neurodevelopmental effects. In the following sections, we will describe the toxicology and mechanistic data regarding chlorpyrifos. We then use the HBWoE approach to evaluate the weight of the evidence regarding a causal association between chlorpyrifos exposure and adverse neurodevelopmental effects.

3.3. Neurodevelopmental toxicity studies in animals

We evaluated the available neurodevelopmental toxicity animal data that are relevant to determining whether sufficient evidence is available to support the hypothesis that chlorpyrifos causes adverse neurodevelopmental effects. Below, we provide a brief overview of these data followed by an endpoint-by-endpoint analysis of the different neurodevelopmental outcomes that have been investigated. Then, we critically assess the data as a whole for consistency of the observed outcomes, evaluate the evidence for exposure-response relationships, and discuss the weight of studies based on factors related to study design. To conduct our analysis, we searched PubMed and Toxline for animal studies that investigated neurodevelopmental effects of chlorpyrifos exposure. Search terms included: "chlorpyrifos," "animal" "cognitive," "neurodevelopmental," "behavior*," "motor," "reflex," "learning," and "memory." We also relied on the reference lists within the studies that we found in the literature search. We limited our evaluation to studies that examined neurodevelopmental effects that are similar to those examined in the cohort studies (*i.e.*, infant neurobehavior, motor and cognitive effects, and general behavioral effects). We identified 22 studies that investigated these potential neurodevelopmental effects in rats and mice.

3.3.1. Overview of animal studies

Many studies have examined the potential association between chlorpyrifos and neurodevelopmental responses in rodents. In the studies that we reviewed, the exposure periods generally ranged from the first gestational day (GD) to the post-weanling period, up to postnatal day (PND) 25 (*i.e.*, the 25th day after birth or extraction from the womb) and encompassed either prenatal time points alone, prenatal and postnatal time points, or postnatal time points alone. Across studies, exposure concentrations ranged from 0.03 to 40 mg/kg chlorpyrifos and were delivered mostly *via* subcutaneous injection or oral gavage, although one study exposed rats dermally and another used intraperitoneal injection as the exposure route. Inhalation exposures to chlorpyrifos were not examined in the animal studies. The neurodevelopmental tests of these exposures were conducted at various time points, ranging from immediately following exposure to adulthood, or even after mating and subsequent reproduction. These tests examined potential effects of chlorpyrifos on social behavior (including maternal behavior), emotion and anxiety, motor function (including locomotor activity, neuromuscular and neuromotor function, and sensorimotor reflexes), and cognitive function (*i.e.*, learning and memory). Table 7 summarizes these tests and their psychobiological significance. In addition, many of the studies described below that investigated the aforementioned effects also assessed brain AChE activity.

3.3.2. Endpoint-by-endpoint analysis of neurodevelopmental effects in animals

In this section, we discuss the animal studies that evaluated the individual neurodevelopmental outcomes

Table 7. Neurodevelopmental tests used in animal studies of chlorpyrifos.

Test	Description	Psychobiological significance	References
Social behavior			
Sociability	An unfamiliar animal of the same strain, sex, and age is placed in a side chamber to serve as a social stimulus; this allows nose contact through the bars, but prevents fighting. The other side chamber contains an empty wire cage identical to that enclosing the social stimulus, serving as a control for exploring a novel object/environment.	Sociability and social preference	Venerosi *et al.* (2008)
Socioagonistic behavior	The animal undergoes an agonistic encounter with an unfamiliar partner of the same sex and age and the behavior is video-recorded. Nonsocial and social behaviors are scored; these usually include general activity, such as exploring, rearing, and digging, as well as social responses, including attack, often accompanied by biting attempts; aggressive grooming; tail rattling; offensive upright posture; defensive upright posture; and submissive upright posture.	Social aggression	Ricceri *et al.* (2006)
Social novelty preference	The test animal is restricted in a chamber and an unfamiliar partner is introduced. The test mouse is allowed to choose between a familiar partner and the novel, unfamiliar partner. The mouse is observed for choice of partner and time spent with partner.	Sociability and social preference	Venerosi *et al.* (2008)
Maternal behavior			
Maternal aggressive behavior	After delivery, pups are removed and dam is exposed to an intruder male in its home cage. Maternal behaviors are recorded.	Maternal aggression against an intruder	Venerosi *et al.* (2008)
Pup-directed behavior	Maternal behavior with pups, including pup retrieval, sniffing, crouching over the pups, and nest-building is observed and recorded.	Maternal behavior	Ricceri *et al.* (2006); Venerosi *et al.* (2008)
Emotion/Anxiety			
Light/dark box test	The light/dark box test consists of a box with two compartments: a small, opaque, dark compartment closed by a red glass and a large, transluscent and illuminated white compartment. The floor of the light compartment is divided into identical squares to allow determination of exploratory behavior. The compartments are connected by an opening to allow mice to freely move between them. A spotlight illuminates the apparatus to provide an aversive stimulus. Animals are individually placed in the center of the light compartment and allowed to explore the compartments. The time spent in each compartment, the time spent in the central square of the light compartment, and the number of crossings between compartments is recorded. This apparatus provides an unfamiliar environment for rodents where a natural conflict occurs between the tendency to explore a novel environment (spontaneous exploratory behavior) and the innate aversion to an open field and brightly illuminated area. As the light, but not the dark, compartment provides stress-like conditions for mice, the percentage of time spent by the mice in the dark compartment is used as a measure of anxiety in mice. In addition, exploratory behavior in the light compartment (time spent in the central square relative to the entire time spent in the light compartment) is also indicative of anxiety. The number of crossings between compartments provides an estimation of the anxiety level if measured without changes of spontaneous locomotor activity.	Anxiety	Venerosi *et al.* (2008, 2010); Braquenier *et al.* (2010)
Ultrasonic vocalization	During the postnatal period, pups emit ultrasonic calls to solicit parental care from the dam. These calls peak at PND 7–9. The most frequently used paradigm to elicit ultrasonic calls is to isolate a pup from the dam for a few minutes. Parameters that are analyzed include number, duration, and frequency of calls.	Homing/Anxiety	Venerosi *et al.* (2009); Laviola *et al.* (2006)

(Continued)

Table 7. (Continued).

Test	Description	Psychobiological significance	References
Elevated plus-maze	The elevated plus-maze is a conflict-avoidance test. Rodents prefer a dark, enclosed, small space over a brightly lit, open, large space, but they are also highly exploratory. This maze represents the natural conflict between the tendency of mice to explore a novel environment and the tendency to avoid a brightly lit open field. The behavior is also influenced by the fear of heights. The elevated plus-maze is in the shape of a "+". The four arms extend from a central platform. Two alternate arms are dark and enclosed, while two alternate arms are open, lit, and available with or without ledges. Additionally, the surface of all arms is raised 1 m above the floor. The intensity of the fear drive (anxiety) is measured by the ratio of entries in open and enclosed arms. The open and enclosed arms of the maze generate exploratory behavior and the avoidance of elevated open arms is an indication of the intensity of anxiety. Furthermore, head dipping over the edges of the open arms, risk-assessment, self grooming, and defecation can be recorded.	Anxiety	Braquenier *et al.* (2010); Ricceri *et al.* (2006); Icenogle *et al.* (2004); Aldridge *et al.* (2004)
Forced swimming test	The animal is observed during a swimming session and behavioral parameters such as floating (the animal floats in the water making only those movements necessary to keep its head above water), struggling (vigorous movements of all four limbs with the forelimbs breaking the surface of the water) and swimming (all four limbs move under the surface of the water with the head kept above water) are observed. The frequency and duration of each pattern are scored.	Mood or adaptation/learning	Venerosi *et al.* (2010)
Reflexes and motor activity			
Open field locomotor test	Assessment takes place in either a circular platform arena or square arena. The animal is placed in the center of the open field arena and allowed to freely move about while being measured by an automated tracking system. Measurements taken include, but are not limited to, general path length, rearing (standing on hind limbs) behaviors, and exploratory behaviors.	Locomotor activity (general activity levels, gross locomotor activity, and exploration habits. Assessment over several days allows habituation to the environment to be evaluated.)	Carr *et al.* (2001); Ricceri *et al.* (2003, 2006)
Geotaxic response	The animal is placed head-downward and given a period of time to turn itself 180 degrees. Animals are scored for completing the 180 degree turn.	Sensorimotor reflexes	Dam *et al.* (2000)
Cliff avoidance	This test assesses pup withdrawal from the edge of a flat surface when its snout and forepaws are placed over a real or virtual cliff; time of aversion can be recorded.	Sensorimotor reflexes	Venerosi *et al.* (2009)
Acoustic startle test	The animal is subjected to a loud sound or click. A subsequent freezing reaction and blink indicate functional hearing abilities.	Auditory reflexes (hearing ability)	Icenogle *et al.* (2004); Johnson *et al.* (2009)
Free-fall and surface righting	Free-fall: The ability of a pup to correct its posture such that it lands on all four feet when released from a supine position above a soft surface.	Posture reflex	Dam *et al.* (2000); Johnson *et al.* (2009)
	Surface: The ability of a pup to right itself to an upright position from a supine position.		
Inclined plane	The animal is placed on a flat plane in the horizontal position with its head facing the side of the board to be raised. The angle at which the animal begins to slip downward is recorded.	Neuromotor performance	Abou-Donia *et al.* (2006)
Grip strength and forepaw grip time	Forelimb strength is measured as tension force using a computerized grip strength meter (GSM). The animal is lifted over the baseplate of the GSM by the tail so that its forepaws are allowed to grasp onto the steel grip. The mouse is then gently pulled backward by the tail until the grip is released. The GSM measures the maximal force before the mouse releases the bar. Forepaw strength is assessed by having animals grip a wooden dowel that is held horizontally and raised so that the animal supports its weight. The time to release grip is recorded in seconds.	Neuromuscular function and muscular strength	Abou-Donia *et al.* (2006)

(Continued)

Table 7. (Continued).

Test	Description	Psychobiological significance	References
Beam-walking	This test assesses the animal's performance in crossing a narrow wooden beam.	Motor coordination and balance	Abou-Donia *et al.* (2006)
Rotorod test	This test assesses the number of times pups fall off a rotating rod during a 2-minute period.	Neuromotor performance	Muto *et al.* (1992)
Memory and Learning			
T-maze	The animal is placed in the start arm of the T-maze and is allowed to explore one of the two goal arms. The mouse is then returned to the start for the next trial and will typically choose to explore the alternate arm. Because this behavior is not reinforced it is considered "spontaneous."	Exploratory behavior and/or memory (measures exploratory behavior when the intervals between the trials are short, but also serves a spatial memory test when the interval between trials is lengthened)	Maurissen *et al.* (2000); Icenogle *et al.* (2004); Levin *et al.* (2001, 2002)
Multi-arm radial maze	Animals are trained to visit a pattern of arms in a multi-arm arm maze to receive a reward (*e.g.*, food). To perform well, the animal must keep the egocentric or allocentric/spatial information regarding which arms of the maze it has already visited during the course of the task. Assessment of choice accuracy for information presented within the day's test session and learning that has accumulated across days of testing are often referred to as working memory and reference memory tasks, respectively.	Working and reference memory errors	Johnson *et al.* (2009); Icenogle *et al.* (2004); Levin *et al.* (2001, 2002); Haviland *et al.* (2010)
Morris swim test	Animals are trained to locate a hidden platform within a pool of opaque water using spatial cues around the pool as a guide. The tank is surrounded by curtains that contain contrasting shapes to serve as spatial cues. Free swims (trials in which the platform is removed) are given throughout training to assess learning. Cued or visible platform trials are typically given at the end of training and serve as a sensorimotor index for each animal.	Spatial learning and memory	Jett *et al.* (2001)
Figure-8 apparatus	The apparatus consists of a continuous, enclosed alley in the shape of a Figure-8, and two blind alleys extending from either side. A single photobeam is located in each blind alley and three photobeams are located in the loops of the Figure-8 for a total of eight photobeams. Animals are allowed to roam in the maze for a set time and photobeam breaks are recorded to measure locomotor activity. To assess the speed of habituation, the blocks can be analyzed for linear and quadratic trends.	Motor activity and habituation	Icenogle *et al.* (2004)
Foraging maze	The foraging maze consists of a central platform elevated from the floor on a small movable table, containing eight arms projecting radially. At the end of each radial arm, two choice arms project in a 'Tee' format. At the end of each Tee, a low barrier is placed on both sides that serves as a food cup, and conceals a large bait under a screen to mask olfactory clues. The maze has opaque sides and clear overhead plastic and is housed in a lighted room with multiple extra-maze visual cues overhead. A fasted animal is placed in a cylinder in the central starting area and then allowed to roam freely about the maze for a set time or until all baits have been discovered. Radial arm choice and pre-reward and post-reward arm entries are recorded.	Learning, memory, and foraging behavior	Haviland *et al.* (2010)
Passive avoidance	This task exploits the rodent's natural preference for dark environments. In this task the animal has to choose between responding to obtain a positive reinforcement (entering into a dark chamber) and not responding to avoid an electric footshock (not entering the dark chamber). The latency to refrain from entering into the dark compartment serves as an index of conditioned suppression and the ability to avoid, and allows memory to be assessed.	Conditioned suppression of instrumental behavioral responses (memory)	Ricceri *et al.* (2003)
Acclimatization	The test animal is first placed in a middle chamber and the doorways into side chambers are opened to allow exploration of the whole cage. Animals are observed for time of exploration of the chamber.	Exploration and habituation	Venerosi *et al.* (2008)

described above. We distinguish studies by exposure route, as studies with oral exposures are considered more relevant for determining risks to human health and are given more weight than studies with exposures *via* injection. We also distinguish studies by the developmental period of exposure to allow for an assessment of the consistency of findings across studies in which exposures occurred during the same neurodevelopmental window. If chlorpyrifos acts as a neurodevelopmental toxicant in the absence of inhibition of AChE activity in the nervous system, results should be consistent in terms of dose-response relationships and patterns of response over neurodevelopmental windows of exposure, both within and across studies, and indicative of effects at doses below those which induce brain AChE inhibition. Other considerations to bear in mind are the adequacy of the study design (*i.e.*, sufficient number of animals, relevant dose levels, and appropriate methodologies), whether the meaning of the statistical tests become distorted by the multiple comparisons problem when data are analyzed for multiple outcomes, and the biological significance of responses.

3.3.2.1. Social and maternal behavior Three studies conducted by the same research group assessed potential chlorpyrifos-associated effects on social and maternal behavior in mice. One study investigated both prenatal exposure *via* oral dosing of dams during gestation and direct postnatal exposure of pups *via* subcutaneous injection (Riccerri *et al.*, 2006), and two studies examined only postnatal, subcutaneous exposures of pups (Riccerri *et al.*, 2003; Venerosi *et al.*, 2008). The results of these studies are described below and summarized in Table 8.

Socioagonistic behavior: Two studies assessed the effects of chlorpyrifos on socioagonistic behavior in mice. Ricceri *et al.* (2006) exposed mouse dams to 0, 3, or 6 mg/kg-day chlorpyrifos by oral gavage during GD 15–18, and offspring were treated with subcutaneous injection of 0, 1, or 3 mg/kg-day chlorpyrifos during PND 11–14. The authors reported that prenatal exposure with 6 mg/kg-day chlorpyrifos increased offensive upright posture in male offspring during PND 75–80, but showed no effect on the other socioagonistic behaviors examined (defensive or submissive upright posture, attack, aggressive grooming of partner, or tail rattling). There was no postnatal treatment effect nor was there an interaction between prenatal and postnatal treatments for this endpoint. The only socioagonistic effects associated with postnatal exposure were increases in the frequency and duration of attacks by male mice exposed to 3 mg/kg-day chlorpyrifos. In addition, there were no effects on any of the general social interaction behaviors examined, such as rearing, digging, moving around the cage, or self-grooming, with any treatment. The authors also reported that there were no effects of chlorpyrifos exposure during the prenatal, postnatal, or both periods on brain AChE activity on GD 19 or PND 15.

Ricceri *et al.* (2006) reported no effect of chlorpyrifos exposure on rearing, wall rearing, digging, moving around cage, or self-grooming on PND 70 in the offspring of mouse dams exposed prenatally *via* oral gavage to 0, 3, or 6 mg/kg-day chlorpyrifos during GD 15–18 followed by postnatal exposure *via* subcutaneous injection of pups to 0, 1, or 3 mg/kg-day chlorpyrifos during PND 11–14. The authors also reported no effects on brain AChE activity on either GD 19 or PND 15 with any chlorpyrifos exposure.

Ricceri *et al.* (2003) exposed mouse pups to 0, 1, or 3 mg/kg-day chlorpyrifos by subcutaneous injection from PND 1–4 or PND 11–14 and reported an increase in aggressive grooming of social partner in males on PND 45 with both doses during the earlier treatment period. Other markers of agonistic behavior– such as attack, tail rattling, offensive upright posture, and defensive postures– were not individually reported. When all aggressive responses were pooled together for the PND 1–4 exposure period, animals in the 1 mg/kg-day group had an increased aggressive response frequency in the first two (out of four) five-minute observation blocks, while animals in the 3 mg/kg-day group did so only in the last two five-minute blocks of observation. Chlorpyrifos exposures during PND 11–14 resulted in an enhancement of the agonistic behavior observed in animals with exposure during PND 1–4, as greater increases in aggressive response frequency were observed with both doses, but were higher with 1 mg/kg-day *vs*. 3 mg/kg-day chlorpyrifos exposure. Despite these indications of chlorpyrifos-associated agonistic behavior, the authors reported that treatment during PND 1–4 or PND 11–14 had no effect on investigative and affiliative behaviors at 45 days of age. In addition, chlorpyrifos treatment during the earlier exposure period had no effect on soliciting behavior, whereas treatment during the later exposure period produced mixed results, as the authors reported an increased frequency of the "push-under" but not the "crawl" response after treatment with 3 mg/kg-day chlorpyrifos. The authors reported a transient inhibitory effect of chlorpyrifos on brain AChE activity (approximately 20% inhibition relative to controls) on PND 4 in mice exposed to 1 or 3 mg/kg-day chlorpyrifos during PND 1–4 but not in mice exposed to 3 mg/kg-day with treatment during PND 11–14 or PND 32–35.

Maternal behavior: Two studies examined the potential effects of chlorpyrifos on maternal aggression and pup-induced maternal behavior. Ricceri *et al.* (2006) reported neither a main effect of prenatal chlorpyrifos nor an interaction between prenatal and postnatal treatments for foster pup-directed behaviors on PND 90 in virgin female offspring of mouse dams exposed to 0, 3, or 6 mg/kg-day chlorpyrifos *via* oral gavage during GD 15–18, followed by subcutaneous exposure to 0, 1, or 3 mg/kg-day chlorpyrifos during PND 11–14. In addition, there were no effects of chlorpyrifos on latency, frequency, and duration of nest building and pup retrieval to the nest. The authors reported an increase in

Table 8. Rodent studies of social and maternal behavior.

Study	Species	Animals per dose	Exposure			Outcome						Effects on AChE activity in brain
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	
Ricceri *et al.* (2006)	Mouse, CD-1	10	Oral gavage	0, 3, 6	GD 15–18	Male	PND 75–80	Socioagonistic behavior	Increased offensive upright posture	Defensive or submissive upright posture; frequency and duration of attack response; aggressive grooming of partner, tail rattling	Increase in one marker of socioagonistic behavior with 6 mg/kg-day	Null
			s.c.	0, 1, 3	PND 11–14	Male	PND 75–80	Socioagonistic behavior	Increased frequency and duration of attack response	Offensive upright posture; defensive upright or submissive posture; aggressive grooming of partner, tail rattling	Increase in two markers of socioagonistic behavior with 3 mg/kg-day	
			Oral gavage	0, 3, 6	GD 15–18	Male	PND 75–80	Social interactions	None	Rearing; digging; moving around cage; self-grooming	Null	
			s.c.	0, 1, 3	PND 11–14							
			Oral gavage	0, 3, 6	GD 15–18	Female	PND 90	Maternal behavior	None	Nest building latency, frequency, and duration; pup retrieval to nest; crouch response frequency and duration; pup licking frequency and duration; pup sniffing	Null	
			s.c.	0, 1, 3	PND 11–14	Female	PND 90	Maternal behavior	Increased frequency and duration of crouch response; decreased licking frequency; increased licking duration; decreased sniffing	Nest building latency, frequency, and duration; pup retrieval to nest	Alterations of 5 markers of maternal behavior with 1 and 3 mg/kg-day	

(Continued)

Table 8. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						Effects on AChE activity in brain
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	
Ricceri *et al.* (2003)	Mouse, CD-1	7 to 10	s.c.	0, 1, 3	PND 1–4	Male, Female	PND 45	Social interactions	None	Social interactions with an unfamiliar untreated mouse for 20 minutes	Null	~20% inhibition with 1 and 3 mg/kg-day
					PND 11–14							Null
					PND 1–4	Male, Female	PND 45	Social investigative behaviors	None	Anogenital sniffing; nose sniffing; body sniffing; following partner around cage; mutual sniffing of anogenital area)	Null	~20% inhibition with 1 and 3 mg/kg-
					PND 11–14							Null
					PND 1–4	Male, Female	PND 45	Affiliative behaviors	None	Social rest, usually next to partner; grooming partner's body)	Null	~20% inhibition with 1 and 3 mg/kg-day
					PND 11–14							Null
					PND 1–4	Male, Female	PND 45	Soliciting behaviors	None	Push-under response (pushing snout or whole anterior body under partners'); crawl behavior (crawling over partner)	Null	~20% inhibition with 1 and 3 mg/kg-day
					PND 11–14				Increased frequency of the push-under response	Crawl behavior	Alteration of one marker of soliciting behavior with 3 mg/kg-day	Null
					PND 1–4	Male	PND 45	Socioagonistic behavior	Increased frequency of aggressive grooming of partner; increased frequency of pooled aggressive responses in two of four observation blocks	Individual aggressive behaviors: attack response; tail rattling; offensive upright posture; defensive upright or submissive posture	Increase in one marker of socioagonistic behavior and transient increase in overall socioagonistic behavior with 1 and 3 mg/kg-day	~20% inhibition with 1 and 3 mg/kg-day

(Continued)

Table 8. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						Effects on AChE activity in brain
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	
					PND 11–14				Increased frequency of pooled aggressive responses in all observation blocks	Individual aggressive behaviors: attack response; aggressive grooming of partner; tail rattling; offensive upright posture; defensive upright or submissive posture	Increased socioagonistic behavior with 1 and 3 mg/kg-day	Null
Venerosi *et al.* (2008)	Mouse, CD-1	12 to 15	s.c.	0, 3	PND 11–14	Male, Female	PND 40–45	Acclimatization	Decreased distance moved when allowed to freely explore cage in first of three sessions	None	Altered acclimatization with 3 mg/kg-day	Not examined
							PND 40–45	Social novelty preference and sociability	None	Time spent with partner 1 *vs.* partner 2; time spent with a partner *vs.* empty chamber	Null	
						Female	Birth to PPD 7	Nest-building activity	Longer latency to nest building	Amount of materials used for nest; nest qualitative features	Alteration of one marker of maternal behavior with 3 mg/kg-day	
							PPD 1	Maternal behavior	Shorter latency to licking	Licking duration; licking frequency; time spent with pups; nursing; pup retrieval; pup sniffing	Alteration of one marker of maternal behavior with 3 mg/kg-day	
							PPD 7	Maternal aggression against male intruder	Decreased defensive posture duration; increased digging duration; increase in pooled investigative behaviors	Attack frequency; individual investigative behaviors: circling, social sniffing, and following)	Decrease in 3 markers of maternal aggression with 3 mg/kg-day	

Notes: CPF = chlorpyrifos; GD = gestational day; PND = postnatal day; PPD = postpartum day; s.c. = subcutaneous injection.

frequency and duration of crouch response (*i.e.*, crouching over the pups), decreased frequency and increased duration of licking, and decreased sniffing with both 1 and 3 mg/kg-day postnatal exposures. These observed behaviors were not associated with altered brain AChE activity. In a subsequent study, Venerosi *et al.* (2008) exposed mice to 0 or 3 mg/kg-day chlorpyrifos by subcutaneous injection during PND 11–14 and reported some changes in maternal behavior on post-partum day (PPD) 1 in chlorpyrifos-treated females that were mated on PND 60. For example, treated females had a shorter latency to licking of pups, but licking duration and frequency were not different from controls. Moreover, time spent with pups, nursing, pup retrieval, and pup sniffing were not affected by chlorpyrifos treatment. In assessing maternal aggressive behavior on PPD 7, the authors reported that treated dams displayed fewer defensive postures, had a longer digging duration, and showed increased investigative behavior, but there were no differences in frequency of attacks compared to untreated controls. In addition, treated dams had a longer latency to build nests, but there was no effect of treatment on either the amount of nest material used or nest quality features on PPD 7.

Summary: The two studies examining social behavior each reported effects on a few isolated markers of socioagonistic behavior in mice, although the same specific markers were not affected in both studies. These effects were observed with prenatal (GD 15–18), oral chlorpyrifos exposure of 6 mg/kg-day or with subcutaneous postnatal (PND 1–4 or PND 11–14) exposures of 1 or 3 mg/kg-day, and AChE inhibition was observed with the postnatal exposures in only one of the studies.

The two studies that assessed maternal behavior reported effects on a few markers with postnatal (PND 11–14) chlorpyrifos exposure *via* subcutaneous injection at doses of 1 or 3 mg/kg-day, with no associated inhibition of AChE activity, but reported no effects on any markers with prenatal (GD 15–18) oral exposures up to 6 mg/kg-day. Similar to the studies of socioagonistic behavior, the same specific markers of maternal behavior that were examined in both studies were not affected in the same way.

3.3.2.2. Emotion and anxiety Nine studies assessed the effects of chlorpyrifos exposure on emotional changes or anxiety in rodents using ultrasonic vocalization monitoring, the elevated plus-maze test, the light-dark box test, and the forced swim test. Three studies examined oral exposures, either prenatally *via* dosing of dams during gestation (Venerosi *et al.*, 2009, 2010) or perinatally *via* gestational exposure that continued after birth through nursing (Braquenier *et al.*, 2010). Six studies investigated subcutaneous exposures either prenatally (Icenogle *et al.*, 2004; Laviola *et al.*, 2006) or postnatally (Aldridge *et al.*, 2005a; Ricceri *et al.*, 2003; Venerosi *et al.*, 2008), with one study including prenatal exposure *via* oral dosing of dams during gestation for a subset of animals (Ricceri *et al.*, 2006). These studies are described below and summarized in Table 9.

Anxiety and distress calling in pups: One study with oral exposure and two studies with subcutaneous exposure examined anxiety and distress calling in rodent pups as measured by ultrasonic vocalization. Venerosi *et al.* (2009) exposed mouse dams to 0 or 6 mg/kg-day chlorpyrifos *via* oral gavage during GD 15–18 and monitored ultrasonic vocalization within 10 days of birth. The authors reported that chlorpyrifos-treated offspring emitted fewer and shorter ultrasonic calls, with longer latency to call and higher frequency of calls, compared to controls.

By contrast, the studies with subcutaneous exposures reported no effects on anxiety or distress calling in mouse pups. Laviola *et al.* (2006) reported that offspring of mouse dams exposed to 5 mg/kg-day chlorpyrifos-oxon during GD 14–16 did not exhibit altered ultrasonic vocalization relative to controls when tested within several days of birth. Consistent with this study, Ricceri *et al.* (2003) reported no effects on ultrasonic vocalization and homing in mouse pups with postnatal exposures to 1 or 3 mg/kg-day chlorpyrifos during PND 1–4 when tested a few days after exposure. Brain AChE activity was transiently inhibited approximately 20% compared to controls on PND 4 with exposures to both doses in these animals.

Anxiety assessed by the elevated plus-maze test: One study with oral exposure and three studies with subcutaneous exposure examined anxiety in adolescent or adult rodents using the elevated plus-maze test. Braquenier *et al.* (2010) assessed effects on anxiety using the elevated plus maze test in 80-day-old female offspring of mouse dams exposed by oral gavage to 0.2, 1, or 5 mg/kg-day chlorpyrifos during GD 15 - PND 14. The authors reported that treated mice were more anxious than controls, as indicated by a decrease in the percent time spent in the open arms of the test and a lower proportion of entries in the open arms. This anxiety was not chlorpyrifos-exposure-dependent, however, as the effects were observed with the 1 mg/kg-day dose but not the 0.2 or 5 mg/kg-day doses, and there is no evidence to support a non-monotonic dose-response curve for this endpoint. In addition, the authors reported a 14% inhibition of brain AChE activity on PND 1 in the offspring of dams exposed orally to 5 mg/kg-day chlorpyrifos from GD 15 until birth compared to controls.

Studies that assessed anxiety with the elevated plus-maze test after subcutaneous exposures of chlorpyrifos reported no increases in anxiety across prenatal and postnatal exposure periods. Icenogle *et al.* (2004) exposed rat dams to 0, 1, or 5 mg/kg-day chlorpyrifos by subcutaneous injection during GD 9–12 and reported no effects on anxiety in 4- to 8-week-old offspring in the elevated plus-maze test, but did report an indication of hyperactivity with the 5 mg/kg-day dose. Similarly, Ricceri *et al.* (2006) reported that after prenatal oral exposure to chlorpyrifos (GD 15–18; 0, 3, or 6 mg/kg-day dosing

of dams) followed by postnatal subcutaneous exposure to chlorpyrifos (PND 11–14; 0, 1, or 3 mg/kg-day), there were no prenatal by postnatal exposure interactions in the elevated plus-maze test when mice were four months of age. With prenatal exposure, a decrease in head dipping frequency (indicating less anxiety) was reported for males exposed to 3 mg/kg-day chlorpyrifos. With postnatal exposure, females spent more time in the maze's open arms (also indicating less anxiety) when exposed to 3 mg/kg-day chlorpyrifos. These effects were not associated with altered brain AChE activity. Aldridge *et al.* (2005a) exposed rat pups to 0 or 1 mg/kg-day chlorpyrifos *via* subcutaneous injection during PND 1–4 and reported no treatment-related effects on anxiety in the elevated plus-maze test on PND 52–53 in females and *less* anxious behavior (by spending more time in the open arms) in treated males compared to controls.

Anxiety assessed by the light/dark box test: Two studies with oral exposures and one study with subcutaneous exposure examined anxiety in adult rodents using the light/dark box test. Venerosi *et al.* (2010) exposed mouse dams *via* oral gavage to 0 or 6 mg/kg-day chlorpyrifos during GD 15–18 and offspring were tested on PND 90. The authors reported that chlorpyrifos-treated females, but not males, showed some indication of elevated anxiety by spending more time in the tunnel connecting the light and dark compartments. Treated mice of both sexes showed no effects on several other parameters of this test, however, such as time spent in either the light or dark compartment, latency to enter the dark compartment, and risk assessment and exploratory behavior in the light compartment.

Braquenier *et al.* (2010) assessed effects on anxiety using the light/dark box test in 72-day-old female offspring of mouse dams exposed by oral gavage to 0.2, 1, or 5 mg/kg-day chlorpyrifos during GD 15–PND 14. The authors reported that treated mice were more anxious than controls, indicated by the preference to spend less time in the center of the light compartment and a decreased number of compartment switches. This anxiety was not chlorpyrifos-exposure-dependent, however, as the effects were observed with the 1 mg/kg-day dose but not the 0.2 or 5 mg/kg-day doses. In addition, the authors reported a 14% inhibition of brain AChE activity on PND 1 in the offspring of dams exposed orally to 5 mg/kg-day chlorpyrifos from GD 15 until birth.

In contrast to the oral studies, Venerosi *et al.* (2008) reported that mouse dams that were subcutaneously exposed to 3 mg/kg-day chlorpyrifos during PND 11–14 were *less* anxious than control dams, as they were more likely to enter the light compartment and do it faster than untreated controls on PPD 2 after mating on PND 60.

Mood assessed by the forced swim test: Venerosi *et al.* (2010) exposed mouse dams *via* oral gavage to 0 or 6 mg/kg-day chlorpyrifos during GD 15–18 and reported no treatment-related effects on the performance of offspring in the forced swim test, which assesses mood by monitoring swimming, struggling, and floating in water, on PND 90.

Summary: One of the three studies examining anxiety and distress calling in pups reported alterations in calling, with oral prenatal exposure of 6 mg/kg-day chlorpyrifos. By contrast, the two studies of this endpoint with subcutaneous exposure did not report effects with chlorpyrifos doses up to 3 mg/kg-day, or with chlorpyrifos-oxon at 5 mg/kg-day.

Of the four studies that assessed anxiety using the elevated plus-maze test, only the study with oral exposure spanning the prenatal and postnatal periods reported increased anxiety (Braquenier *et al.*, 2010). This effect was only studied in females and was not dose-dependent, as it was observed with exposure to 1 mg/kg-day, but not with 0.2 or 5 mg/kg-day, and inhibition of AChE activity in the brain was observed only with exposure to the highest dose. By contrast, two studies with subcutaneous exposure during prenatal periods reported no effects on anxiety in the elevated plus-maze test at doses up to 6 mg/kg-day, one study with prenatal oral exposure reported decreased anxiety in males at a dose of 3 mg/kg-day, and two studies with subcutaneous exposure during postnatal periods reported decreased anxiety in females dosed with 3 mg/kg-day or in males dosed with 1 mg/kg-day.

Consistent with the studies using the elevated plus-maze test, the two studies that assessed anxiety with the light/dark box test after oral exposures during prenatal or both prenatal and postnatal periods reported increased anxiety in females at doses of 1 or 6 mg/kg-day, but not at doses of 0.2 or 5 mg/kg-day. The specific measures of increased anxiety in this test were not the same in both studies, however. By contrast, the one study that assessed the effects of postnatal subcutaneous chlorpyrifos exposure in the light/box test reported *decreased* anxiety in females at a dose of 3 mg/kg-day. Only one study assessed mood in the forced swim test, and no effects were reported in this study after prenatal oral exposure to 6 mg/kg-day chlorpyrifos.

Overall, increased anxiety was only observed in studies with oral chlorpyrifos exposures of 6 mg/kg-day during GD 15–18 or of 1 mg/kg-day during GD 15–PND 14, although effects on anxiety with the latter exposure were not dose-dependent, as no effects were observed with a dose of 5 mg/kg-day that inhibited AChE activity in the brain.

3.3.2.3. Motor function Seventeen studies assessed the effects of chlorpyrifos exposure on motor function in rodents, using tests of locomotor activity, neuromuscular or neuromotor function, and sensorimotor reflexes. Seven studies examined oral exposures, either prenatally *via* dosing of dams during gestation (Venerosi *et al.*, 2009), with one study including postnatal exposure *via* subcutaneous injection for a subset of animals (Ricceri *et al.*, 2006), perinatally *via* gestational exposure that continued after birth through nursing (Maurissen *et al.*,

Table 9. Rodent studies of effects on emotion and anxiety.

Study	Species	Animals per dose	Exposure			Outcome						Effects on AChE activity in brain
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	
Venerosi *et al.* (2009)	Mouse, CD-1	15 per sex	Oral gavage	0, 6	GD 15-18	Male, Female	PND 4, 7, 10	Anxiety *via* ultrasonic vocalization	Fewer calls, shorter call duration, longer latency to calls, and higher peak frequency of calls (PND 10 only)	Peak amplitude of calls	Increased anxiety on PND 10 with 6 mg/kg-day	NR
Venerosi *et al.* (2010)	Mouse, CD-1	10 to 12	Oral gavage	0, 6	GD 15-18	Male, Female	PND 90	Anxiety *via* the light/dark box test	Increased percentage of time spent in the tunnel connecting the two compartments of the apparatus (females only)	Latency to enter dark compartment; time spent in dark or light compartments; risk assessment and exploratory behavior in light compartment	Increase in one marker of anxiety for females only, with 6 mg/kg-day	NR
		5 to 7					PND 90	Mood *via* the forced swim test	None	Swimming, struggling, and floating in water	Null	
Braquenier *et al.* (2010)	Mouse, CD-1	8 to 10	Oral gavage	0, 0.2, 1, 5	GD 15 - PND 14	Female	PND 80	Anxiety *via* the elevated plus-maze	Decreased percentage of time spent in open arms; decreased proportion of entries in open arms	None	Increased anxiety with 1 mg/kg-day	14% inhibition with 5 mg/kg-day
							PND 72	Anxiety *via* the light/dark box test	Decreased percentage of time spent in center of light compartment; decreased number of compartment switches	Percentage of time in dark compartment	Increased anxiety with 1 mg/kg-day	
Ricceri *et al.* 2006	Mouse, CD-1	10	Oral gavage	0, 3, 6	GD 15-18	Male, Female	4 months	Anxiety *via* the elevated plus-maze	Decreased head dipping frequency (males only)	Frequency of arm entries; duration of time spent in open, central, and closed arms; frequency and duration of risk assessment posture	Decreased anxiety in males with 3 mg/kg-day	Null

(Continued)

Table 9. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
			s.c.	0, 1, 3	PND 11-14	Male, Female	4 months	Anxiety *via* the elevated plus-maze	Increased percentage of time spent in open arms (females only)	Frequency of arm entries; duration of time spent in central and closed arms; head dipping frequency; frequency and duration of risk assessment posture	Decreased anxiety in females with 3 mg /kg-day	
Icenogle *et al.* (2004)	Rat, Sprague-Dawley	10	s.c.	0, 1, 5	GD 9-12	Male, Female	4-8 weeks	Anxiety and hyperactivity *via* the elevated plus-maze	Increased number of center crosses	Duration of time spent in open and closed arms	Increased hyperactivity with 5 mg/kg-day	NR
Laviola *et al.* (2006)	Mouse, C57/B16	4 to 6	s.c.	0, 5 (chlorpyrifos-oxon)	GD 14-16	Male, Female	PND 3, 7, 11	Anxiety *via* ultrasonic vocalization	None	Number of calls	Null	NR
Aldridge *et al.* (2005)	Rat, Sprague-Dawley	9 per sex	s.c.	0, 1	PND 1-4	Male, Female	PND 52-53	Anxiety *via* the elevated plus-maze	Increased percentage of time spent in open arms and increased number of center crosses (males only)	Duration of time spent in closed arms	Decreased anxiety and increased hyperactivity in males with 1 mg/kg-day	NR
Ricceri *et al.* (2003)	Mouse, CD-1	10	s.c.	0, 1, 3	PND 1-4	Male, Female	PND 5, 8, 11 PND 10	Anxiety *via* ultrasonic vocalization Homing	None None	Number of calls Latency to reach nest area; time spent in nest area; locomotor activity during test	Null Null	~20% inhibition with 1 and 3 mg/kg-day

(Continued)

Table 9. (Continued).

Study	Species	Animals per dose	Exposure			Outcome					
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
Venerosi *et al.* (2008)	Mouse, CD-1	9 to 11	s.c.	0, 3	PND 11-14	Female	PPD 2 (after mating on PND 60)	Anxiety *via* the light/ dark box test	Increased proportion of dams entered light compartment; decreased latency to enter light compartment	Percentage of time in dark or light compartment; number of transitions across tunnel connecting the two compartments	Decreased anxiety in dams with 3 mg/kg-day	NR

Notes: CPF = chlorpyrifos; GD = gestational day; NR = not reported; PND = postnatal day; PPD = postpartum day; s.c. = subcutaneous injection.

2000; Braquenier *et al.*, 2010), or postnatally by direct oral exposure to pups (Moser *et al.*, 1998; Carr *et al.*, 2001; Johnson *et al.*, 2009). Eight studies investigated subcutaneous exposures administered either prenatally (Chanda and Pope, 1996; Icenogle *et al.*, 2004; Levin *et al.*, 2002; Laviola *et al.*, 2006) or postnatally (Ricceri *et al.*, 2003; Dam *et al.*, 2000; Levin *et al.*, 2001; Chakraborti *et al.*, 1993). Finally, one study examined intraperitoneal exposure (Muto *et al.*, 1992) and another study examined dermal exposure (Abou-Donia *et al.*, 2006) during gestation. These studies are described below and summarized in Table 10.

Locomotor activity: Six studies assessed locomotor activity in rodents after oral exposure to chlorpyrifos. Only one of these studies examined exposure during the prenatal period alone. Venerosi *et al.* (2009) reported no effects on locomotor activity in the 12-day-old offspring of mouse dams exposed to 6 mg/kg-day chlorpyrifos *via* oral gavage during GD 15–18.

Three studies monitored locomotor activities after oral chlorpyrifos exposures that spanned the prenatal and postnatal periods. In a study that complied with US EPA Pesticide Assessment Guidelines and Good Laboratory Practice (GLP) regulations, Maurissen *et al.* (2000) exposed rat dams *via* oral gavage to 0, 0.3, 1, or 5 mg/kg-day chlorpyrifos during GD 6– PND 10 and reported no effects on locomotor activity in offspring assessed at various time points between PND 13 and PND 60. Consistent with this study, Braquenier *et al.* (2010) reported no effects on locomotor activity in the offspring of mouse dams orally exposed to 0, 0.2, 1, or 5 mg/kg-day chlorpyrifos during GD 15– PND 14. The authors also assessed brain AChE activity on PND 1 and reported statistically significant inhibition of brain AChE activity only in the offspring of dams exposed to the highest chlorpyrifos dose (5 mg/kg-day).

Ricceri *et al.* (2006) examined motor activity in 70-day-old offspring of mouse dams treated with 0, 3, or 6 mg/kg-day chlorpyrifos by oral gavage during GD 15–18 and then subsequently treated with subcutaneous injection of 0, 1, or 3 mg/kg-day chlorpyrifos during PND 11–14. The authors reported increased motor activity in animals prenatally treated with 6, but not 3 mg/kg-day chlorpyrifos and postnatally treated with either vehicle or 1, but not 3 mg/kg-day chlorpyrifos compared to control offspring of dams that received only vehicle. These results demonstrate a lack of an exposure-response relationship. The authors reported a three-way interaction between prenatal treatment, postnatal treatment, and five-minute activity-monitoring blocks, however. In addition, the authors reported a hyperactivating effect of chlorpyrifos exposure in the first of four five-minute sessions of the motor activity test in mice with postnatal exposure to 3 mg/kg-day chlorpyrifos, but this was limited to the offspring of dams that received either vehicle or 3 mg/kg-day chlorpyrifos, but not 6 mg/kg-day chlorpyrifos, demonstrating transient effects with no exposure-response relationship. In contrast to Braquenier *et al.* (2010),

Table 10. Rodent studies of effects on sensorimotor reflexes and motor activity.

			Exposure			Outcome						
Study	Species	Animals per dose	Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
Venerosi *et al.* (2009)	Mouse, CD-1	13 to 18 per sex	Oral gavage	0, 6	GD 15-18	Male, Female	PND 12	Spontaneous motor behavior	Decreased frequency and duration of pivoting; increased duration of immobility	Frequency and duration of crossing (locomotor activity), head movement, wall climbing, and grooming	Decrease in three markers of spontaneous motor behavior with 6 mg/kg-day	NR
							PND 3, 6, 9, 12, 15	Sensorimotor maturation	None	Righting reflex; grasping reflex; cliff aversion; forelimb and hind limb reflexes; pole grasping	Null	
Ricceri *et al.* (2006)	Mouse, CD-1	10	Oral gavage	0, 3, 6	GD 15-18	Male	PND 70	Locomotor activity *via* open field test	Increased crossing frequency	Crossing duration; rearing; wall rearing; self-grooming	Increase in one marker of locomotor activity with prenatal exposure to 6 mg/kg-day followed by postnatal exposure to 0 or 1 mg/kg-day, and with prenatal exposure to 0 or 3 mg/kg-day followed by postnatal exposure to 3 mg/kg-day	Null
			s.c.	0, 1, 3	PND 11-14							
Maurissen *et al.* (2000)	Rat, Sprague-Dawley	20 per sex	Oral gavage	0, 0.3, 1, 5	GD 6 - PND 10	Male, Female	PND 13, 17, 21, 60	Locomotor activity	None	Locomotor activity	Null	NR
							PND 22, 61	Auditory reflexes *via* acoustic startle test	None	Auditory reflexes	Null	
Braquenier *et al.* (2010)	Mouse, CD-1	8 to 10	Oral gavage	0, 0.2, 1, 5	GD 15 - PND 14	Female	> PND 21 for 8 days	Locomotor activity	None	Locomotor activity	Null	14% inhibition with 5 mg/kg-day
Moser *et al.* (1998)	Rat, Long Evans hooded	5 per sex	Oral gavage	0, 5, 20	PND 17	Male, Female	3.5 or 6.5 hours post-exposure	Locomotor activity *via* Figure-8 apparatus	Decreased locomotor activity	None	Decreased locomotor activity with 20 mg/kg-day	50–90% inhibition with 5 and 20 mg/kg-day
								Sensorimotor reflexes	Righting deficits; gait changes	None	Deficits in sensorimotor reflexes with 20 mg/kg-day	

(Continued)

Table 10. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
Carr *et al.* (2001)	Rat, Sprague-Dawley	10	Oral gavage	Low: 3 (PND 1–21)	PND 1–21 (every other day)	Male, Female	PND 10, 12, 14, 16, 18, 20, 25, 30	Locomotor activity *via* open field test	Decreased locomotor activity (PND 25 and 30 only; medium and high dose groups only)	None	Decreased locomotor activity with 6 and 12 mg/kg-day	17–70% inhibition with 3, 6, and 12 mg/kg-day
				Medium: 3 (PND 1–5), 6 (PND 7–21)								
				High: 3 (PND 1–5), 6 (PND 7–13), 12 (PND 15–21)								
Johnson *et al.* (2009)	Rat, Sprague-Dawley	9 to 14 per sex	Oral gavage	Low: 1 (PND 1–5), 1 (PND 6–13), 1 (PND 14–20)	PND 1–20	Male, Female	NR	Sensorimotor reflexes	None	Surface righting; free-fall righting; negative geotaxis; cliff avoidance; auditory reflexes	Null	14–53% inhibition with 1 to 6 mg/kg-day
				Medium: 1 (PND 1–5), 2 (PND 6–13), 4 (PND 14–20)								
				High: 1.5 (PND 1–5), 3 (PND 6–13), 6 (PND 14–20)								
Chanda and Pope (1996)	Rat, Sprague-Dawley	7 to 8	s.c.	25	GD 12–19	Male, Female	PND 1, 3	Sensorimotor reflexes	Decreased cliff avoidance; surface righting deficits	None	Deficits in sensorimotor reflexes with 25 mg/kg-day	60% inhibition with 25 mg/kg-day
Icenogle *et al.* (2004)	Rat, Sprague-Dawley	10 per sex	s.c.	0, 1, 5	GD 9–12	Male, Female	4–6 weeks	Locomotor activity *via* Figure-8 apparatus	Decreased locomotor activity (2 of 12 test sessions only)	None	Transient decrease in locomotor activity with 5 mg/kg-day	NR
							NR	Auditory reflexes *via* acoustic startle test	None	Auditory reflexes	Null	
Levin *et al.* (2002)	Rat, Sprague-Dawley	10 per sex	s.c.	0, 1,	GD 17–20	Male, Female	4–6 weeks	Locomotor activity *via* Figure-8 apparatus	None	Locomotor activity	Null	NR
Laviola *et al.* (2006)	Mouse, C57/B16	4 to 6	s.c.	0, 5 (chlorpyrifos-oxon)	GD 14–16	Male, Female	PND 3,7,11	Neuromuscular function *via* grasping reflex test	Increased fall angle (PND 3 only)	None	Increased maturation of grasping reflex at 5 mg/kg-day chlorpyrifos-oxon	NR
								Sensorimotor reflexes	None	Righting reflex	Null	

(Continued)

Table 10. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
Ricceri *et al.* (2003)	Mouse, CD-1	10	s.c.	0, 1, 3	PND 1-4	Male, Female	PND 25	Locomotor activity *via* open field test	None	Locomotor activity	Null	~20% inhibition with 1 and 3 mg/kg-day
					PND 11-14				Increased locomotor activity	None	Increased locomotor activity with 1 and 3 mg/kg-day	Null
					PND 1-4	Male, Female	PND 45	Motor behavior	Decreased self-grooming	Exploring; rearing; digging	Decrease in one marker of motor behavior at 1 and 3 mg/kg-day	~20% inhibition with 1 and 3 mg/kg-day
					PND 11-14				None	Exploring; rearing; digging; self-grooming	Null	Null
Dam *et al.* (2000)	Rat, Sprague-Dawley	23 to 24 per sex	s.c.	0, 1	PND 1-4	Male, Female	PND 3, 4	Sensorimotor reflexes	Deficits in reflex righting (females only)	None	Deficit in one marker of sensorimotor reflexes in females with 1 mg/kg-day	20-60% inhibition with 1 mg/kg-day
		23 to 24 per sex	s.c.	0, 1	PND 1-4	Male, Female	PND 5, 8	Sensorimotor reflexes	Deficits in geotaxic response (females only)	None	Deficit in one marker of sensorimotor reflexes in females with 1 mg/kg-day	20-60% inhibition with 1 mg/kg-day
		5 to 7 per sex	s.c.	0, 1	PND 1-4	Male, Female	PND 21, 30	Locomotor activity *via* open field test	Decreased locomotor activity and rearing (males only)	Grooming	Decrease in two markers of locomotor activity in males with 1 mg/kg-day	20-60% inhibition with 1 mg/kg-day
				0, 5	PND 11-14				Increased rearing (males on PND 30 only)	Locomotor activity; grooming	Increase in one marker of locomotor activity in males with 5 mg/kg-day	20-60% inhibition with 5 mg/kg-day

(Continued)

Table 10. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
Levin *et al.* (2001)	Rat, Sprague-Dawley	10	s.c.	0, 1	PND 1-4	Male, Female	4-6 weeks	Locomotor activity *via* Figure-8	None	Locomotor activity	Null	NR
				0, 5	PND 11-14							
Chakraborti *et al.* (1993)	Rat, Sprague-Dawley	10	s.c.	0, 40	PND 7-10	Male, Female	2, 4, 6, 8 weeks post-exposure	Locomotor activity *via* open field test	None	Locomotor activity	Null	55-60% inhibition with 40 mg/kg-day
Muto *et al.* (1992)	Rat, species unspecified	12 to 20	i.p.	0, 0.03, 0.1, 0.3 (as Dursban)	GD 0-7	Not specified	PND 16	Neuromotor function *via* the rotorod test	Increased number of falls	None	Decreased neuromotor function with 0.03 to 0.3 mg/kg-day	
					GD 7-21			Neuromotor function *via* the rotorod test	Increased number of falls	None	Decreased neuromotor function with 0.3 mg/kg-day	
					PND 3, 10, or 12			Neuromotor function *via* the rotorod test	Increased number of falls	None	Decreased neuromotor function with 0.1 and 0.3 mg/kg-day	NR
								General motor behavior	None	General motor behavior	Null	
								Neuro muscular function *via* inclined plane test	None	Neuro muscular function	Null	
Abou-Donia *et al.* (2006)	Rat, Sprague-Dawley	10	Dermal	0, 1	GD 4-20	Male, Female	PND 90	Motor coordination/balance *via* beam-walking test	None	Beam-walking score	Null	
								Neuro muscular function *via* inclined plane test	Decreased incline angle (females only)	None	Deficits in one marker of neuromuscular function in females with 1 mg/kg-day	25% increase in females with 1 mg/kg-day
								Neuro muscular function	Decreased forepaw grip time	None	Deficits in one marker of neuromuscular function with 1 mg/kg-day	

Notes: CPF = chlorpyrifos; GD = gestational day; NR = not reported; PND = postnatal day; s.c. = subcutaneous injection.

Ricceri *et al*. (2006) reported no chlorpyrifos-associated effects on brain AChE activity even at a higher prenatal exposure concentration (6 mg/kg-day).

Two studies examined locomotor activity after oral exposure to chlorpyrifos during postnatal periods. Moser *et al*. (1998) reported no effects on motor activity in rats exposed to 5 mg/kg chlorpyrifos *via* oral gavage on PND 17 and tested a few hours post-exposure, although 50–60% inhibition of brain AChE was observed with this exposure. A higher concentration of chlorpyrifos (20 mg/kg) in this study produced both brain AChE inhibition (70–90%) and decreased motor activity. Similarly, Carr *et al*. (2001) reported decreased locomotor activity in 25- and 30-day-old rats exposed to 6 or 12 mg/kg-day chlorpyrifos *via* oral gavage during PND 1–21, but not with exposure to 3 mg/kg-day. Dose-dependent inhibition of brain AChE activity, ranging from 17–70%, was observed at all exposure concentrations used in this study.

Six studies examined effects of subcutaneous exposure to chlorpyrifos on locomotor activity, two of which involved exposures during prenatal periods. Icenogle *et al*. (2004) reported a decrease in locomotor activity in offspring of rat dams exposed during GD 9–12 *via* subcutaneous injection to chlorpyrifos at a dose of 5 mg/kg-day, but not 1 mg/kg-day, when tested during 4–6 weeks of age. By contrast, Levin *et al*. (2002) reported no change in locomotor activity in the offspring of rat dams exposed on GD 17–20 to the same doses of chlorpyrifos by the same route and tested at the same age.

Four studies assessed the effects of subcutaneous exposure to chlorpyrifos during postnatal periods. Ricceri *et al*. (2003) reported an increase in locomotor activity in 25-day-old mice after subcutaneous exposure to 1 or 3 mg/kg-day chlorpyrifos during PND 11–14 but not during PND 1–4. Reduced self-grooming was observed after exposure to both doses during PND 1–4, but not during PND 11–14. The authors also assessed brain AChE activity after exposure but found contrasting results to the assessment of locomotor activity, in that an inhibitory effect on AChE of approximately 20% was observed after exposure to either 1 or 3 mg/kg-day chlorpyrifos during PND 1–4, but not after exposure during PND 11–14 or PND 32–35.

Dam *et al*. (2000) exposed rat pups *via* subcutaneous injection to 0 or 1 mg/kg-day chlorpyrifos during PND 1–4 or to 0 or 5 mg/kg-day chlorpyrifos during PND 11–14 and assessed locomotor activity on PND21 and PND30. The authors reported decreased locomotor activity and decreased rearing in males exposed to 1 mg/kg-day chlorpyrifos during PND 1–4 and no effects on locomotor activity in rats of either sex with exposure to 5 mg/kg-day chlorpyrifos on PND 11–14, although increased rearing was observed in males on PND 30. There were no alterations in grooming for either dose group. Dam *et al*. (2000) also exposed rats to 1 mg/kg chlorpyrifos on PND 1 or 5 mg/kg on PND 11 and assessed brain AChE activity two and four hours post-exposure. AChE inhibition of 20–60% was observed two hours after both exposures, with greater effects (60% inhibition) for males exposed to 1 mg/kg chlorpyrifos on PND 1, but the effects either diminished or disappeared by four hours post-exposure.

In contrast to the study by Dam *et al*. (2000), a subsequent study from the same research group reported no effects on motor activity in 4- to 6-week-old rats exposed to either 1 mg/kg-day chlorpyrifos during PND 1–4 or to 5 mg/kg-day during PND 11–14 *via* subcutaneous injection (Levin *et al*., 2001). Consistent with this study, rat pups exposed to a much higher concentration of chlorpyrifos (40 mg/kg-day) *via* subcutaneous injection during PND 7–10 showed no changes in motor activity during an 8-week follow-up period, although brain AChE activity was inhibited by 55–60% four days after exposure cessation and by 20–32% two weeks after exposure cessation (Chakraborti *et al*., 1993).

Neuromuscular and neuromotor function: Three studies examined effects on neuromuscular and/or neuromotor function, using either intraperitoneal injection or dermal application of chlorpyrifos, or subcutaneous injection of chlorpyrifos-oxon. Muto *et al.* (1992) assessed neuromotor function in the rotorod test with the 16-day-old offspring of rat dams exposed to 0 (saline), 0.03, 0.1, or 0.3 mg/kg-day chlorpyrifos as Dursban pesticide (1% chlorpyrifos, 6% xylene, 93% water, according to the authors) *via* intraperitoneal injection during GD 0–7 or GD 7–21. The authors reported deficits in neuromotor function, as evidenced by an increased number of falls in the rotorod test, in offspring exposed to all doses during GD 0–7 and only the highest dose during GD 7–21. This study also examined the effects of direct postnatal exposure of rat pups to 0, 0.1, or 0.3 mg/kg-day chlorpyrifos (as Dursban) *via* intraperitoneal injection on PND 3, 10, or 12. Rats exposed to the higher dose at each of the three time points had an increased number of falls in the rotorod test on PND 16, as did rats in the lower dose group, but only when exposed on PND 12, but not on PND 3 or 10. Postnatal exposure to either dose produced no effects on general motor behavior or in the incline plane test of neuromuscular function.

Abou-Donia *et al.* (2006) exposed rat dams to 1 mg/kg-day chlorpyrifos during GD 4–20 by dermal application and reported that motor coordination and balance were not affected in the beam walking test in offspring of treated dams compared to controls on PND 90. The authors also reported that females, but not males, showed deficits in neuromuscular function in the incline plane test, however, and both sexes showed deficits in forepaw grip time, another test of neuromuscular function. In addition, the authors reported a 25% *increase* in brain AChE activity on PND 90 in treated females relative to controls. Given the expected recovery of AChE activity after exposure to chlorpyrifos from the synthesis of new AChE molecules, it is expected that any inhibition of AChE activity would have been observed shortly after this exposure, so the observed increase in activity 90 days post-exposure is likely unrelated to chlorpyrifos exposure.

Laviola *et al.* (2006) exposed mouse dams to 0 or 5 mg/kg-day chlorpyrifos-oxon *via* subcutaneous injection during GD 14–16 and assessed neuromuscular function in the grasping reflex tests in offspring on PND 3, 7, and 11. The authors reported that treated mice showed an increase in fall angle (*i.e.*, they held on longer when the supporting surface was tilted) compared to controls, but only on PND 3.

Sensorimotor reflexes: Four studies examined the effects of oral chlorpyrifos exposure on measures of sensorimotor reflexes. Venerosi *et al.* (2009) reported no effects on sensorimotor maturation (grasping, righting, and cliff avoidance) in offspring of mouse dams exposed *via* oral gavage to 6 mg/kg-day chlorpyrifos during GD 15–18 when assessed within the first two weeks of life. The authors also reported that chlorpyrifos treatment resulted in shorter and less frequent pivoting behavior and increased immobility, but no changes in other spontaneous motor behavior such as crossing, head moving, wall climbing, or grooming.

Maurissen *et al.* (2000) reported no effect on auditory reflexes (as assessed using the acoustic startle test) of 22- and 61-day-old offspring of rat dams orally exposed to 0.3, 1, or 5 mg/kg-day chlorpyrifos beginning on GD 6 and continuing through PND 10.

Johnson *et al.* (2009) orally exposed rat pups to one of three treatments: 1 mg/kg-day chlorpyrifos during PND 1–20 (low exposure); 1 mg/kg-day chlorpyrifos during PND 1–5, 2 mg/kg-day chlorpyrifos during PND 6–13, and 4 mg/kg-day chlorpyrifos during PND 7–20 (medium exposure); or 1.5 mg/kg-day chlorpyrifos during PND 1–5, 3 mg/kg-day chlorpyrifos during PND 6–13, and 6 mg/kg-day chlorpyrifos during PND 7–20 (high exposure). The authors reported no effects on sensorimotor reflexes (surface and free-fall righting, negative geotaxis, cliff avoidance, and auditory reflexes) with any of the three treatments. They also reported a dose-dependent (14–53%) inhibition of AChE activity in the brain immediately after exposure in each treatment group. Approximately 20% inhibition of AChE activity persisted for up to 20 days post-exposure in the medium and high exposure groups.

Moser *et al.* (1998) treated 17-day-old rats with 0, 5, or 20 mg/kg chlorpyrifos *via* oral gavage and reported that exposure to 20 mg/kg was associated with altered sensorimotor reflexes compared to controls when animals were tested within a few hours of exposure. Further, the authors found that exposure to both doses of chlorpyrifos was associated with AChE inhibition in the brain (50–60% inhibition for the 5 mg/kg dose group and 70–90% inhibition for the 20 mg/kg dose group).

Four studies examined effects on sensorimotor reflexes with subcutaneous exposure to chlorpyrifos. Icenogle *et al.* (2004) reported no effects on auditory reflexes in the offspring of rats subcutaneously exposed to 1 or 5 mg/kg-day chlorpyrifos during GD 9–12. In a study with a much higher exposure, Chanda and Pope (1996) reported deficits in sensorimotor reflexes on PND 1 and PND 3 in the offspring of rat dams exposed subcutaneously to 25 mg/kg-day chlorpyrifos during GD 12–19. The effects appeared to decline rapidly with age, as they were decreased on PND 3 *vs.* PND 1. A 60% inhibition of brain AChE activity was observed in treated animals on GD 20. One study used prenatal subcutaneous exposure to chlorpyrifos-oxon instead of chlorpyrifos. Laviola *et al.* (2006) reported that offspring of mouse dams exposed to 5 mg/kg-day chlorpyrifos-oxon via subcutaneous injection during GD 14–16 were not affected in terms of righting reflex on PND 3, 7, or 11.

Dam *et al.* (2000) examined effects on sensorimotor reflexes in rat pups exposed *via* subcutaneous injection to 0 or 1 mg/kg-day chlorpyrifos during PND 1–4 or to 0 or 5 mg/kg-day chlorpyrifos during PND 11–14. The authors reported that female pups exposed to 1 mg/kg-day during PND 1–4 showed deficits in reflex righting, which was assessed on PND 3–4, and in geotaxic response, which was tested on PND 5–8. Brain AChE activity was transiently inhibited by 20–60% with exposure to 1 mg/kg-day chlorpyrifos on PND 1 or by 20–30% with exposure to 5 mg/kg-day chlorpyrifos on PND 11 in this study, with approximately three-fold greater inhibition in males compared to females.

Summary: Three of four studies that examined oral chlorpyrifos exposures that began during gestation reported no effects on locomotor activity at doses up to 6 mg/kg-day. One study reported increased locomotor activity without inhibition of brain AChE activity after prenatal oral exposure to 6, but not 3, mg/kg-day chlorpyrifos and postnatal subcutaneous exposure to either vehicle or 1, but not 3, mg/kg-day chlorpyrifos, demonstrating a lack of an exposure-response relationship. The two studies that examined postnatal oral exposures reported decreased locomotor activity at doses of 6 mg/kg-day and higher, and inhibition of brain AChE activity with these doses. One study with prenatal subcutaneous exposure reported decreased locomotor activity with 5 mg/kg-day, whereas the other study, later in gestation, reported no effects with the same dose. Results were largely null in the four studies that used subcutaneous postnatal exposures, although one study reported increased locomotor activity with exposure during PND 11–14 but not PND 1–4 with 1 and 3 mg/kg-day chlorpyrifos, and another reported decreased locomotor activity with exposure to 1 mg/kg-day during PND 1–4 but not with exposure to 5 mg/kg-day during PND 11–14. Brain AChE activity was inhibited with all the exposures in which these effects were observed.

Two studies reported deficits in neuromuscular and neuromotor function with chlorpyrifos exposure. One of these studies reported effects with intraperitoneal exposure to chlorpyrifos (as Dursban) at concentrations as low as 0.03 mg/kg-day prenatally and 0.1 mg/kg-day postnatally, but it is possible that the effects were attributable to xylene, which was contained in the Dursban mixture and was not controlled for in the study. The other study reported deficits in neuromuscular function with dermal exposure to 1 mg/kg-day chlorpyrifos. A third

study reported effects on one parameter of neuromuscular function with subcutaneous exposure to 5 mg/kg-day chlorpyrifos-oxon.

Two of the four studies of sensorimotor reflexes that used oral chlorpyrifos exposures reported alterations in measures of this endpoint with a prenatal dose of 6 mg/kg-day and postnatal doses ≥ 20 mg/kg-day. Inhibition of brain AChE activity was observed with postnatal doses of ≥ 5 mg/kg-day in one oral exposure study and with doses as low as 1 mg/kg-day in another oral study. Two of four studies examining sensorimotor reflexes with subcutaneous chlorpyrifos exposures reported deficits in measures of this endpoint with a prenatal dose of 25 mg/kg-day and a postnatal dose of 1 mg/kg-day, and inhibition of AChE activity in the brain was also observed with these effects.

Overall, studies with postnatal exposures often indicated that effects on motor function are larger at a younger age, require higher exposure concentrations to produce effects with advancing age, and tend to be transient, as they usually persisted from a few hours to several days. In almost every study that examined AChE activity and reported effects on motor function, inhibition of brain AChE was observed at the same chlorpyrifos doses as those associated with the effects on motor function.

3.3.2.4. Cognitive function Ten studies examined the effects of chlorpyrifos exposure on cognitive function in rodents *via* one or more tests that assessed exploratory behavior, learning, or memory. Two of these studies investigated oral exposures, either perinatally *via* gestational exposure that continued after birth through nursing (Maurissen *et al.*, 2000) or postnatally (Johnson *et al.*, 2009). Eight studies examined subcutaneous exposures administrated either prenatally (Icenogle *et al.*, 2004; Levin *et al.*, 2002; Haviland *et al.*, 2010) or postnatally (Aldridge *et al.*, 2005a; Levin *et al.*, 2001; Ricceri *et al.*, 2003; Venerosi *et al.*, 2008; Jett *et al.*, 2001). These studies are described below and summarized in Table 11.

Learning and memory assessed by the T-maze test: One study examined effects on learning and memory in the T-maze test with oral exposure to chlorpyrifos. Maurissen *et al.* (2000) exposed rat dams *via* oral gavage to 0, 0.3, 1, or 5 mg/kg-day chlorpyrifos during GD 6– PND 10, and offspring were assessed in the T-maze during PND 22–24 or PND 61–91. The authors reported no treatment-related effects on learning and memory.

Two studies assessed rodents using the T-maze test after prenatal subcutaneous exposure to chlorpyrifos. Icenogle *et al.* (2004) reported transient effects (only in the first of five observation sessions) on exploratory behavior, as reflected by shorter spontaneous alternation latency (*i.e.*, hyperactivity), in 4- to 8-week-old offspring of rat dams exposed to 1 or 5 mg/kg-day chlorpyrifos *via* subcutaneous injection during GD 9–12. Similarly, Levin *et al.* (2002) reported a transient decrease in alternation latency that resolved with repeated trials of the test in the offspring of rat dams exposed subcutaneously to 1 or 5 mg/kg-day chlorpyrifos during GD 17–20. Alternation between T-maze arms, as a measure of exploratory behavior, was not associated with chlorpyrifos exposure in this study.

One study examined the effects of subcutaneous exposure to chlorpyrifos during the postnatal period on performance in the T-maze test. Levin *et al.* (2001) exposed rat pups subcutaneously to 1 mg/kg-day chlorpyrifos during PND 1–4 or to 5 mg/kg-day chlorpyrifos during PND 11–14 and performance in the T-maze was assessed when the rats were four to six weeks of age. The authors reported no treatment-related effects on alternation frequency but found that males responded with a longer alternation latency compared to controls in the third of 12 test sessions when exposed to 5 mg/kg-day during PND 11–14.

Memory assessed by the radial arm maze: One study assessed memory using the radial arm maze after oral exposure to chlorpyrifos. Johnson *et al.* (2009) reported no effects on working memory errors in female rats orally exposed during PND 1–21 to chlorpyrifos concentrations as high as 6 mg/kg-day in any of the four weeks of testing (four days/week) that began on PND 36, but observed fewer working memory errors in female rats exposed to the "medium" concentration range of 1–4 mg/kg-day chlorpyrifos during PND 1–21 (but not to the "low" exposure of 1 mg/kg-day or to the "high" exposure range of 1.5–6 mg/kg-day) when all the days in the four weeks were averaged. By contrast, male rats in the high exposure group made more working memory errors during all weeks and those in the low and medium exposure groups made more working memory errors during the fourth week of testing. With the medium and high chlorpyrifos exposures, female rats made fewer reference memory errors relative to controls and males made more such errors during week two of testing. The authors also reported a dose-dependent (14–53%) inhibition of AChE activity in the brain immediately after exposure in each treatment group. Inhibition of AChE activity (approximately 20%) persisted for up to 20 days post-exposure in the medium and high exposure groups.

Three studies examined the effects of subcutaneous exposure to chlorpyrifos during the postnatal period on performance in the radial arm maze. Icenogle *et al.* (2004) assessed 8- to 13-week-old offspring of rat dams exposed to 1 or 5 mg/kg-day chlorpyrifos *via* subcutaneous injection during GD 9–12. The authors reported an indication of increased working and reference memory errors with the 5 mg/kg-day exposure, although in most sessions (12 of 18), there were no differences between treated rats and controls. Levin *et al.* (2002) exposed rat dams subcutaneously to 1 or 5 mg/kg-day chlorpyrifos during GD 17–20 and reported no treatment-related effects on error frequency in their 8- to 13-week-old offspring when each of the 18 sessions were considered, but the mean number of both working and reference memory errors for all 18 sessions taken together was higher in females of the 1 mg/kg-day dose group compared to controls. These effects were not observed in females of the higher dose group or in males at

Table 11. Rodent studies of effects on cognitive function.

Study	Species	Animals per dose	Exposure			Outcome						
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
Maurissen *et al.* (2000)	Rat, Sprague–Dawley	8 per sex	Oral gavage	0, 0.3, 1, 5	GD 6 – PND 10	Male, Female	PND 22–24, PND 61–91	Learning and memory *via* the T-maze test	None	Spatial delayed alternation (learning acquisition; short-term retention; non-mnemonic factors)	Null	NR
Johnson *et al.* (2009)	Rat, Sprague–Dawley	9 to 14 per sex	Oral gavage	Low: 1 (PND 1–5), 1 (PND 6–13), 1 (PND 14–20)	PND 1–20	Male, Female	PND 36–60	Memory *via* the 12-arm radial maze	Increased number of working memory errors (males only; one of four testing weeks)	Reference memory errors	Transient deficit in one marker of memory in males with 1 mg/kg–day	14–53% inhibition with 1 to 6 mg/kg– day
				Medium: 1 (PND 1–5), 2 (PND 6–13), 4 (PND 14–20)					Fewer working and reference memory errors (females only); increased number of working and reference memory errors (males only; one of four testing weeks)	None	Transient memory deficits in males and memory gains in females and with 1 to 4 mg/kg–day	
				High: 1.5 (PND 1–5), 3 (PND 6–13), 6 (PND 14–20)					Increased number of working memory errors (males only); fewer reference memory errors (females only); increased number of reference memory errors (males only; one of four testing weeks)	None	Memory deficits in males and memory gains in females with 1.5 to 6 mg/kg–day	
Icenogle *et al.* (2004)	Rat, Sprague–Dawley	10 per sex	s.c.	0, 1, 5	GD 9–12	Male, Female	4–8 weeks	Exploratory behavior *via* the T-maze test	Shorter spontaneous alternation latency (first of five sessions only)	None	Increased hyperactivity with 1.5 mg/kg–day	NR
							8–13 weeks	Memory *via* the 16–arm radial maze	Increased working memory errors (two of six sessions); increased reference memory errors (one of six sessions)	None	Transient memory deficits with 5 mg/kg–day	

(Continued)

Table 11. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
							4–8 weeks	Habituation *via* Figure–8 apparatus	Faster linear trend of habituation	None	Faster habituation with 5 mg/kg-day	
Levin *et al.* (2002)	Rat, Sprague–Dawley	10 per sex	s.c.	0, 1, 5	GD 17–20	Male, Female	4–6 weeks	Exploratory behavior *via* the T–maze test	Decrease in alternation latency (one or two of five test trials)	Alternation between maze arms	Transiently increased hyperactivity with 1 and 5 mg/kg–day	NR
							8–13 weeks	Memory *via* the 16–arm radial maze	Increased mean number of working and reference memory errors across 18 sessions (females only)	Memory error frequency	Memory deficits in females with 1 mg/kg–day	
							4–6 weeks	Habituation *via* Figure–8 apparatus	Slower linear trend of habituation (females only)	Habituation time	Slower habituation in females with 1 and 5 mg/kg–day	
Haviland *et al.* (2010)	Mouse, Swiss Webster	8	s.c.	0, 1, 5	GD 17–20	Male, Female	PND 60–81	Learning *via* the foraging maze	Decreased food recognition learning (two or three of nine sessions; females only); increased food recognition learning (four of nine sessions; males only); decreased food position learning (two of nine sessions; females only); increased food position learning (two of nine sessions; males only)	Foraging rate	Transiently decreased foraging ability in females with 1 and 5 mg/kg–day; transiently increased foraging ability in males with 1 and 5 mg/kg–day	NR
								Memory *via* the 8–arm radial maze	Increased reference memory errors (two of nine sessions); decreased reference memory errors (two of nine sessions; females only)	None	Transient memory deficits in males with 1 or 5 mg/kg–day; transient memory deficits in females with 1 mg/kg–day; transient memory gains in females with 1 or 5 mg/kg–day	

(Continued)

Table 11. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
Aldridge *et al.* (2005)	Rat, Sprague-Dawley	9 per sex	s.c.	0, 1	PND 1-4	Male, Female	PND 64-99	16-arm radial maze	Increased mean number of reference memory errors across 18 sessions (males only)	Working memory errors	Memory deficits in males with 1 mg/kg-day	NR
Levin *et al.* (2001)	Rat, Sprague-Dawley	10	s.c.	0, 1	PND 1-4	Male, Female	4-6 weeks	Exploratory behavior *via* the T-maze test	None	Alternation frequency; alternation latency	Null	NR
				0, 5	PND 11-14				Increase in alternation latency (one of 12 sessions; males only)	Alternation frequency	Transiently decreased hyperactivity in males with 5mg/kg-day	
				0, 1	PND 1-4		8-13 weeks	Memory *via* the 16-arm radial maze	Increased number of working and reference memory errors (over first three of 18 sessions; males only); increased mean number of working and reference memory errors (females)	None	Memory deficits with 1 mg/kg-day	
				0, 5	PND 11-14				None	Working and reference memory errors	Null	
				0, 1	PND 1-4		4-6 weeks	Habituation *via* Figure-8 apparatus	None	Habituation	Null	
				0, 5	PND 11-14				Slower linear trend of habituation	None	Slower habituation with 5 mg/kg-day	
Ricceri *et al.* (2003)	Mouse, CD-1	10	s.c.	0, 3	PND 1-4	Male, Female	PND 35-38	Novelty-seeking behavior	Increased activity rate in novel compartment (one of five sessions)	Novelty preference; latency to enter novel compartment; activity rate in familiar compartment	Transient alteration in one marker of novelty-seeking behavior with 3 mg/kg-day	~20% inhibition with 1 and 3 mg/kg-day
					PND 11-14				Increased activity rate in novel compartment (two of five sessions)	Novelty preference; latency to enter novel compartment; activity rate in familiar compartment		Null

(Continued)

Table 11. (Continued).

Study	Species	Animals per dose	Exposure			Outcome						
			Route	Concentration (mg CPF/kg b.w./day)	Developmental period	Sex	Test period	Endpoint examined	Statistically significant results	Null endpoints	Implication of effects	Effects on AChE activity in brain
				0, 1, 3	PND 1-4	Male, Female	PND 60-81	Passive avoidance learning	None	Memory	Null	~20% inhibition with 1 and 3 mg/kg-day
					PND 11-14							Null
Venerosi *et al.* (2008)	Mouse, CD-1	9 to 11	s.c.	0, 3	PND 11-14	Male, Female	PND 40-45	Acclimatization	Decreased exploration (one of three sessions)	None	Transient decrease in exploration with 3 mg/kg-day	NR
Jett *et al.* (2001)	Rat, Long Evans	8 to 10 per sex	s.c.	0, 0.3, 7	PND 7, 11, 15 (pre weaning)	Male, Female	PND 24-28	Cognitive function *via* the Morris swim test	Increased escape latency (two of four testing days); learning deficiency in probe test	Swimming speed	Transient learning deficits with 7 mg/kg-day	Null
		3 to 4 per sex			PND 22, 26 (post-weaning)				Increased escape latency (two of four testing days); learning deficiency in probe test	Swimming speed	Transient learning deficits with 0.3 and 7 mg/kg-day	Null

Notes: CPF = chlorpyrifos; GD = gestational day; NR = not reported; PND = postnatal day; s.c. = subcutaneous injection.

either dose. Haviland *et al.* (2010) exposed mouse dams to 0, 1, or 5 mg/kg-day chlorpyrifos *via* subcutaneous injection during GD 17–20 and assessed their offspring in the radial arm maze during PND 60–81. The authors reported an increase in reference memory errors compared to controls for treated males and females in only two of nine sessions over the 21-day testing period, but these errors were not exposure-dependent, and treated females made fewer errors compared to controls in two of the sessions.

Two studies assessed memory using the radial arm maze after subcutaneous exposure to chlorpyrifos during postnatal periods. Levin *et al.* (2001) exposed rat pups to 1 mg/kg-day chlorpyrifos *via* subcutaneous injection during PND 1–4 or to 5 mg/kg-day chlorpyrifos during PND 11–14. When assessed at eight to 13 weeks of age, males exposed to 1 mg/kg-day chlorpyrifos during PND 1–4 made more working memory errors than controls when the first three sessions were averaged, but not in any of the other 15 sessions. With this same exposure, female rats made more working memory errors only when all sessions were averaged together. A similar response was observed for reference memory errors with exposure during PND 1–4, and no effects were reported with exposure during PND 11–14. A subsequent study by the same research group (Aldridge *et al.*, 2005a) reported an association between exposure to 1 mg/kg-day chlorpyrifos *via* subcutaneous injection during PND 1–4 and increased reference memory errors in 9-week-old male rats when all 18 observation sessions of the radial arm maze were averaged together, with no effects in females.

Habituation assessed by the Figure-8 apparatus: Two studies investigated habituation with the Figure-8 apparatus after prenatal exposure to chlorpyrifos *via* subcutaneous injection. Icenogle *et al.* (2004) reported faster habituation in the Figure-8 apparatus for the 4- to 8-week-old offspring of rat dams exposed to 5 mg/kg-day chlorpyrifos *via* subcutaneous injection during GD 9–12, but not with exposure to 1 mg/kg-day. By contrast, Levin *et al.* (2002) reported no associations between chlorpyrifos exposure and motor activity level or habituation time in all 12 of the five-minute blocks of the Figure-8 apparatus test in the offspring of rat dams exposed subcutaneously to 1 or 5 mg/kg-day chlorpyrifos during GD 17–20. Exposure to both doses was associated with lower linear trends of habituation (*i.e.*, slower) *vs.* controls in females, but not males, however.

One study assessed habituation after postnatal subcutaneous exposure to chlorpyrifos. Levin *et al.* (2001) reported a decreased motor activity linear trend in the Figure-8 apparatus, an indication of slower habituation, in 4- to 6-week-old rats exposed to 5 mg/kg-day chlorpyrifos during PND 11–14, but not in rats exposed to 1 mg/kg-day chlorpyrifos during PND 1–4.

Other tests of cognitive function: One study with subcutaneous exposure during the prenatal period assessed learning in the nine-session foraging maze. Haviland *et al.* (2010) exposed mouse dams to 0, 1, or 5 mg/kg-day chlorpyrifos *via* subcutaneous injection during GD 17–20 and assessed their offspring during PND 60–81. Treated females showed decreased food recognition learning in sessions 5 and 6 with both chlorpyrifos doses and also in sessions 4, 8, and 9 with the high dose. In males, food recognition learning was *increased* in sessions 3, 4, 5, and 7 in the low, but not the high, dose group. Food position learning was decreased in females in sessions 5 and 8, only in the high dose group, and was *increased* in males in sessions 2 and 7, only in the low dose group. In addition, chlorpyrifos exposure at either dose was not associated with foraging activity.

Three studies with postnatal exposure *via* subcutaneous injection assessed various markers of cognitive function. Ricceri *et al.* (2003) reported that 60-day-old mice that received subcutaneous injections of 1 or 3 mg/kg-day chlorpyrifos during PND 1–4 or PND 11–14 did not show an effect in passive avoidance learning, a test that assesses memory *via* observing conditioned suppression of behavioral responses. In assessing the novelty-seeking behavior of mice in the high dose group on PND 35–38, the authors reported no treatment effect on novelty preference or latency to enter the novel compartment. The authors reported an increased activity rate in the novel compartment with treatment during both postnatal time periods, but this was limited to one of five sessions performed after the earlier treatment and two of five sessions performed after the later one. In addition, the authors reported approximately 20% inhibition of brain AChE in mice exposed to either 1 or 3 mg/kg-day during PND 1–4, but not during PND 11–14 or PND 32–35. In a subsequent experiment from the same laboratory, Venerosi *et al.* (2008) reported that mice exposed to 3 mg/kg-day chlorpyrifos *via* subcutaneous injection during PND 11–14 explored a new cage less than control mice on PND 40–45, although this difference in exploration was limited to the first of three sessions of a five-minute observation period (*i.e.*, the first ~ 1.7 minutes).

Jett *et al.* (2001) exposed rats to 0, 0.3, or 7 mg/kg chlorpyrifos *via* subcutaneous injection on PND 7, 11, and 15 (pre-weaning group) or on PND 22 and 26 (post-weaning group) and tested cognitive function with the Morris swim test on PND 24 through 28. The authors reported that rats exposed to 7 mg/kg chlorpyrifos in the pre-weaning group took longer to find the platform in the Morris swim test on PND 24 and PND 28 than did controls, but no effects were observed on the other testing days. Rats exposed to both chlorpyrifos doses in the post-weaning group also took longer to find the platform, but only on PND 26 and 28. The authors also administered the probe version of the Morris swim test on the last testing day (PND 28). In this test, the platform is removed and the degree of learning is determined by the amount of time spent in close proximity to the learned platform position. Pre-weaning treatment with 7 mg/kg and post-weaning treatment with both chlorpyrifos doses were associated with learning deficiencies in the probe test, but the effect magnitude did not change with exposure in the post-weaning group. Swimming speed

was not altered by chlorpyrifos treatment in either the pre- or post-weaning group. The authors also assessed brain AChE activity within a few days of exposure (PND 28) and observed no inhibition of this activity with any chlorpyrifos treatment.

Summary: One study assessing learning and memory in the T-maze test after oral exposure to chlorpyrifos during gestation and lactation reported no effects. Each of the three studies with subcutaneous exposure that used this measure of cognitive function reported transient effects on the latency of alternation. Alternation latency was transiently decreased with prenatal exposure to 1 or 5 mg/kg-day and was transiently increased with postnatal exposure to 5 mg/kg-day.

In studies that assessed memory in the radial arm maze, any treatment-related effects were usually transient, as they were observed in only a few sessions of this test and often only reached statistical significance when all test sessions were averaged together. One study with postnatal oral exposure reported fewer working and reference memory errors in female rats and more such errors in male rats in certain sessions of the test at doses ranging from 1 to 6 mg/kg-day chlorpyrifos, and AChE activity in the brain was also inhibited in each dose group. All three studies of prenatal subcutaneous exposure reported increases in working and reference memory errors in some test sessions with chlorpyrifos exposures of 1 or 5 mg/kg-day, but these did not always show an exposure-response relationship, and one of the studies reported fewer working memory errors in females in some test sessions. Both studies with subcutaneous chlorpyrifos exposure during postnatal time periods reported increases in reference and memory errors when some or all test sessions were averaged with a dose of 1 mg/kg-day during PND 1–4, but not during PND 11–14.

Habituation was tested in the Figure-8 apparatus only after subcutaneous exposures to chlorpyrifos. Prenatal exposure during GD 9–12 was associated with faster habituation at a dose of 5 mg/kg-day in one study. By contrast, exposure to 1 or 5 mg/kg-day chlorpyrifos during GD 17–20 was associated with a trend of slower habituation in another study. One study examined postnatal chlorpyrifos exposure and reported slower habituation after exposure to 5 mg/kg-day chlorpyrifos during PND 11–14, but no effects with exposure to 1 mg/kg-day during PND 1–4.

Four studies assessed cognitive function in rodents after subcutaneous exposure to chlorpyrifos using various other tests, and effects were observed in only a few test sessions in each study. One study assessed learning in the foraging maze after prenatal exposure and reported decreased food recognition and position learning compared to controls in females in some test sessions with exposure to 1 or 5 mg/kg-day chlorpyrifos, and increased food recognition and position learning in males in certain test sessions with exposure to 1 mg/kg-day, but not 5 mg/kg-day chlorpyrifos. In the novelty seeking test, mice had an increased activity rate in the novel compartment in one or two test sessions compared to controls when treated with 3 mg/kg-day chlorpyrifos during PND 1–4 or PND 11–14, and inhibition of AChE activity in the brain was also observed after treatment with this dose during PND 1–4, but not during PND 11–14. Postnatal exposure to 3 mg/kg-day chlorpyrifos was also associated with less exploratory behavior in a new cage in one of three observation sessions. Postnatal exposure to 7 mg/kg-day chlorpyrifos prior to weaning was associated with learning difficulties in the Morris swim test, as was exposure to 0.3 or 7 mg/kg-day administered after weaning. These effects were observed only on two of the four testing days and were not associated with inhibition of AChE activity.

Overall, when chlorpyrifos-associated cognitive effects were reported in the above studies, there was no clear trend associated with the exposure route or the developmental phase during which exposure occurred. Effects were almost always transient, as they were observed only during a few sessions of each test administered, and were often observed to be in the opposite direction in the same test (*e.g.*, more memory errors *vs*. fewer memory errors compared to controls in the radial arm maze). Inhibition of AChE activity in the brain was only assessed in a few of the studies, all of which examined effects of chlorpyrifos exposures during postnatal periods. While some of these studies reported cognitive effects in conjunction with AChE inhibition, others reported effects in different cognitive tests in the absence of AChE inhibition.

3.3.3. Analysis of animal data

In the following sections, we critically examine the animal data as a whole to assess whether the evidence supports the hypothesis that chlorpyrifos is associated with neurodevelopmental effects. This evaluation considers the exposure route, the adequacy of study design, the consistency of reported outcomes and exposure-response relationships within and across studies, and the biological significance of responses.

3.3.3.1. Adequacy of study design The most rigorous studies have a sufficient number of animals, use dose levels and routes of exposure that are relevant to human exposures, and use appropriate laboratory and statistical methodologies. In our evaluation, we considered the rigor of all the animal studies. The outcome of a study has no bearing on how rigorous the study is, and null results should not be dismissed if they are obtained from a robust and well-conducted study. Regardless of their results, the most rigorous studies carry the most weight.

Studies with a sufficient number of animals carry more weight, as they have more statistical power to detect small differences in outcomes among treatment groups, such that a null finding is more likely to represent a true lack of an effect rather than a failure to detect a true difference. While many of the neurodevelopmental studies of chlorpyrifos used at least 20 animals per dose group, several used very small numbers of animals, including the study by Ricceri *et al*. (2006), which used only 10 animals per

dose group for examining effects on social and maternal behavior and motor function; the study by Chanda and Pope (1996), which used only 7–8 animals per dose group to examine effects on motor function; and the study by Laviola *et al*. (2006), which used only 4 to 6 animals per dose group to study effects on anxiety, neuromuscular function, and sensorimotor reflexes. Using a sufficient number of animals of both sexes is also important, as sex differences in effects are often observed with studies in rodents.

The use of relevant dose levels is also an important consideration, particularly when examining neurodevelopmental effects of chlorpyrifos which are hypothesized to occur at very low doses that are not associated with systemic toxicity and inhibition of AChE activity in the brain. Several of the studies of neurodevelopmental effects used high exposures that have been shown to cause inhibition of brain AChE activity in many other studies, and these should carry less weight than those studies that examined chlorpyrifos doses below those known to cause AChE inhibition. It should be noted that the lowest observed adverse effect level (LOAEL) of 1 mg/kg-day for AChE inhibition in the brain in the animal studies reviewed here is five orders of magnitude higher than the estimated chlorpyrifos exposures of the mothers in the Columbia and CHAMACOS cohorts.

Oral exposures to chlorpyrifos *via* dietary dosing are the most relevant to current human exposures in non-occupational settings. Inhalation is also a relevant exposure route for humans, mainly for occupational exposures since the restriction of the residential use of chlorpyrifos in 2001. The majority of studies examining neurodevelopmental effects in rodents after chlorpyrifos exposure used oral gavage or subcutaneous injection as the exposure route, and none of the studies used dietary or inhalation exposures. Subcutaneous injection is not a relevant exposure route for humans, although it is similar to inhalation or dermal routes in that it avoids the extensive first-pass detoxifying metabolism that occurs in the liver after oral exposure and, thus, could produce higher systemic doses of chlorpyrifos compared to oral exposures. Because the oral exposure route is more relevant to humans, studies with oral exposures are more applicable to determining risks to human health and are given more weight than studies using exposure *via* injection methods.

Three studies that examined oral exposures to chlorpyrifos and used numbers of animals in the high range across studies (10–18 per sex, per dose) reported mostly null effects. Maurissen *et al*. (2000) reported no effects on locomotor activity or auditory reflexes in groups of 20 male and 20 female rats exposed orally to chlorpyrifos at concentrations ranging from 0.3–5 mg/kg-day during the entire perinatal period (GD 6 through PND 10). This is the only rodent study of neurodevelopmental effects in our evaluation that was conducted under US EPA Pesticide Assessment Guidelines and GLP regulations, which means there were specific control measures taken to help ensure the consistency and reliability of the results. Johnson *et al*. (2009) exposed rat pups (9–14 per sex, per dose) *via* oral gavage to chlorpyrifos doses ranging from 1–6 mg/kg-day during various postnatal periods and reported no effects on sensorimotor reflexes, *decreases* in working and reference memory errors in females in the radial arm maze, and increases in working and reference memory errors in males. The effects on memory were observed when doses spanned the range of 1–6 mg/kg-day, but not when the dose was consistently 1 mg/kg-day, and inhibition of AChE activity in the brain was observed after treatment with all doses. Venerosi *et al*. (2009) reported no effects on anxiety, sensorimotor maturation, or locomotor activity in groups of mice (13–18 per sex, per dose) orally exposed to the relatively high dose of 6 mg/kg-day chlorpyrifos during GD 15–18.

Two studies that also used a relatively high number of animals, but the less relevant exposure route of subcutaneous injection, reported some effects on certain neurodevelopmental markers, but most were also observed in conjunction with inhibition of AChE activity. Ricceri *et al*. (2003) subcutaneously exposed mouse pups (7–17 per sex, per dose) to 1 or 3 mg/kg-day chlorpyrifos during PND 1–4 or PND 11–14 and reported increases in some markers of socioagonistic behavior with both doses during both exposure periods, increased locomotor activity with both doses during the latter exposure period, but not the former, and an increase in one aspect of novelty-seeking behavior with the higher dose during both exposure periods. Inhibition of AChE activity was observed with both doses, but only in mice exposed during the earlier period. Dam *et al*. (2000) exposed 23–24 rat pups per sex *via* subcutaneous injection to 1 mg/kg-day chlorpyrifos during PND 1–4 and reported deficits in reflex righting and geotaxic response in females only, as well as inhibition of AChE activity in the brain.

Together, studies with the most weight report largely null effects across various neurodevelopmental tests. The studies that do report treatment-related effects often report inhibition of AChE activity in the brain at the same doses associated with the neurodevelopmental effects, suggesting that the effects occur *via* inhibition of AChE activity.

3.3.3.2. Consistency of outcomes within and across studies Many of the studies measured several different endpoints, with some of these being assessed at many timepoints, leading to many statistical comparisons. As noted above, when multiple endpoints are examined in the same study, the probability of finding apparent effects when there are none increases, and it becomes more likely that several results will be statistically significant by chance. This is also true across studies, as the probability of chance findings increases as the number of studies examining the same effects increases. The remedy to this problem is to look for consistency in the impacts of exposure on each specific endpoint and endpoints expected

to occur *via* a similar mode of action, both within and across studies. If the majority of studies report no statistically significant results for an effect, but a few studies do report effects, there is a high probability that the positive results are not treatment-related but are in fact due to chance or another factor (Goodman *et al.*, 2010).

The studies examining social and maternal behavior reported effects on a few isolated markers of these outcomes, although the same specific markers were not affected in the same way across studies. Effects on socioagonistic behavior were observed with both oral prenatal exposure of 6 mg/kg-day chlorpyrifos and subcutaneous postnatal exposures of 1 or 3 mg/kg-day, and those on maternal behavior were observed with subcutaneous postnatal exposure of 1 or 3 mg/kg-day chlorpyrifos but not with oral prenatal exposures up to 6 mg/kg-day.

The studies examining anxiety-related outcomes did not report consistent results across studies. Only one of three studies examining anxiety and distress calling in pups reported alterations in calling, with oral prenatal exposure to 6 mg/kg-day chlorpyrifos, whereas the other two reported no effects with subcutaneous exposure with chlorpyrifos doses up to 3 mg/kg-day, or with chlorpyrifos-oxon at 5 mg/kg-day. Of the four studies that assessed anxiety in the elevated plus-maze test, only one study with oral exposure spanning the prenatal and postnatal periods reported increased anxiety in females, and the effects were not dose-dependent. By contrast, one study with prenatal oral exposure reported decreased anxiety in males dosed with 3 mg/kg-day, two studies with prenatal subcutaneous exposure reported no effects at doses up to 6 mg/kg-day, and two studies with postnatal subcutaneous exposure reported decreased anxiety in females dosed with 3 mg/kg-day or in males dosed with 1 mg/kg-day. The two studies that assessed anxiety with the light/dark box test after oral exposures during prenatal or both prenatal and postnatal periods reported increased anxiety in females at doses of 1 or 6 mg/kg-day, but not at doses of 0.2 or 5 mg/kg-day, and the specific measures of increased anxiety were not the same in both studies. By contrast, the one study that examined postnatal subcutaneous exposure reported *decreased* anxiety in females at a dose of 3 mg/kg-day. The one study that assessed mood in the forced swim test reported no effects after prenatal oral exposure to 6 mg/kg-day chlorpyrifos.

The results of studies examining effects of chlorpyrifos on motor function were largely null, with certain effects observed in only a few studies and usually at very high doses. Three of four studies that examined oral chlorpyrifos exposures that began during gestation reported no effects on locomotor activity at doses up to 6 mg/kg-day, whereas the fourth study reported increased locomotor activity after prenatal oral exposure to 6, but not 3, mg/kg-day chlorpyrifos and postnatal subcutaneous exposure to either vehicle or 1, but not 3, mg/kg-day chlorpyrifos, demonstrating no exposure-response relationship. The two studies that examined the effects of postnatal oral exposures reported decreased locomotor activity at doses of 6 mg/kg-day and higher. One study with prenatal subcutaneous exposure reported decreased locomotor activity with 5 mg/kg-day, whereas the other study, later in gestation, reported no effects with the same dose. Of the four studies that used subcutaneous postnatal exposures, two studies reported no effects, one study reported increased locomotor activity with exposure to 1 and 3 mg/kg-day chlorpyrifos during PND 11–14, but not PND 1–4, and another reported decreased locomotor activity with exposure to 1 mg/kg-day during PND 1–4 but not with exposure to 5 mg/kg-day during PND 11–14. Of the two studies examining neuromuscular and neuromotor function associated with chlorpyrifos exposure, one reported deficits in neuromotor function in the rotorod test with intraperitoneal exposure to chlorpyrifos (as Dursban) at concentrations as low as 0.03 mg/kg-day prenatally and 0.1 mg/kg-day postnatally, but no effects on general motor behavior or in the incline plane test of neuromuscular function. It is possible that the effects observed in this study were attributable to the uncontrolled exposure to xylene in the Dursban mixture. The other study reported deficits in neuromuscular function with dermal exposure to 1 mg/kg-day chlorpyrifos, but motor coordination and balance were not affected. Two of the four studies of sensorimotor reflexes that used oral chlorpyrifos exposures reported no effects, whereas the other two reported alterations in measures of this endpoint with a prenatal dose of 6 mg/kg-day and postnatal doses ≥ 20 mg/kg-day. Two of five studies examining sensorimotor reflexes with subcutaneous chlorpyrifos exposures reported no effects, whereas three others reported deficits in measures of this endpoint with a prenatal dose of 25 mg/kg-day and postnatal doses ranging from 1 to 5 mg/kg-day.

In studies of chlorpyrifos-associated cognitive effects, there was no clear trend associated with the exposure route or the developmental phase during which exposure occurred. Effects were almost always transient, as they were observed only during a few sessions of each test administered, and they were often observed to be in the opposite direction in the same test. No effects were reported in one study of learning and memory in the T-maze test after oral exposure to chlorpyrifos during gestation and lactation, and the three studies with subcutaneous exposure reported a transient decrease in alternation latency with prenatal exposure to 1 or 5 mg/kg-day and a transient *increase* in this endpoint with postnatal exposure to 5 mg/kg-day during PND 11–14, but not during PND 1–4. In studies that assessed memory in the radial arm maze, treatment-related effects were usually observed in only a few sessions of this test and often only reached statistical significance when all test sessions were averaged together. One study with postnatal oral exposure reported *fewer* working and reference memory errors in female rats and more such errors in male rats in certain sessions of the test at doses ranging from 1 to 6 mg/kg-day chlorpyrifos. All three studies of

prenatal subcutaneous exposure reported increases in working and reference memory errors in some test sessions with chlorpyrifos exposures of 1 or 5 mg/kg-day, but these did not always show an exposure-response relationship, and one of the studies reported *fewer* working memory errors in females in some test sessions. Both studies with subcutaneous chlorpyrifos exposure during postnatal time periods reported increases in reference and memory errors when some or all test sessions were averaged with a dose of 1 mg/kg-day during PND 1–4, but not during PND 11–14. Three studies assessed habituation with the Figure-8 apparatus after subcutaneous exposures to chlorpyrifos. Prenatal exposure during GD 9–12 was associated with faster habituation at a dose of 5 mg/kg-day in one study, whereas exposure to 1 or 5 mg/kg-day chlorpyrifos during GD 17–20 was associated with a trend of slower habituation in another study. The third study reported slower habituation after exposure to 5 mg/kg-day chlorpyrifos during PND 11–14, but no effects with exposure to 1 mg/kg-day during PND 1–4.

Four studies examined other tests of cognitive function after subcutaneous exposures, and effects were observed in only a few test sessions in each study. One study assessed learning in the foraging maze and reported decreased food recognition and position learning in females compared to controls in some test sessions with exposure to 1 or 5 mg/kg-day chlorpyrifos, and *increased* food recognition and position learning in males in certain test sessions with exposure to 1 mg/kg-day, but not 5 mg/kg-day chlorpyrifos. In the novelty seeking test, mice had an increased activity rate in the novel compartment compared to controls in one or two test sessions when treated with 3 mg/kg-day chlorpyrifos during PND 1–4 or PND 11–14, and postnatal exposure to this same dose was also associated with less exploratory behavior in a new cage in one of three observation sessions in another study. Postnatal exposure to 7 mg/kg-day chlorpyrifos prior to weaning was associated with learning difficulties in the Morris swim test on two of four testing days, as was exposure to 0.3 or 7 mg/kg-day administered after weaning.

In addition to inconsistencies of the same effects across studies, there were also inconsistencies across related endpoints at similar doses. It is assumed that if chlorpyrifos causes adverse neurodevelopmental effects, one should see increases in adverse social and maternal behaviors, anxiety, hyperactivity (including increased locomotor activity), and adverse cognitive effects such as memory errors, but this was not always the case within or across studies. For example, although Ricceri *et al.* (2006) reported treatment-related effects on socioagonistic and maternal behavior, as well as increased locomotor activity, the authors reported either no effects or *decreased* anxiety depending on the period of exposure. Icenogle *et al.* (2004) reported cognitive deficits but also reported null effects on anxiety and auditory reflexes and *decreased* locomotor activity. Aldridge *et al.* (2005a) reported increased reference memory errors in male rats but *less* anxiety in these animals. Finally, two studies by Levin *et al.* (2001, 2002) reported slower habituation in the Figure-8 apparatus, which is indicative of increased locomotor activity, but no effects in a direct test of locomotor activity at the same doses.

Taken together, the studies assessing potential neurodevelopmental effects of chlorpyrifos in rodents indicate that it is usually a few, isolated markers of certain behaviors that were determined to be statistically significant in pair-wise comparisons with controls, but these were generally contradicted by other studies reporting no effects in similar dose ranges using similar routes of exposure, and sometimes showed effects in the other direction. In general, for each endpoint we did not find a pattern of an effect that was consistent enough over doses and time points within studies or consistent enough across studies to constitute a repeatable finding. This indicates that the reported effects may likely be due to chance, or, at the very least, need better corroboration before they can be considered compelling results.

3.3.3.3. Exposure-response It is important to keep in mind that many of the reported neurodevelopmental effects in the animal studies were observed at doses that are much lower than those at which the established neurotoxic effects of chlorpyrifos, acting through inhibition of AChE activity in the nervous system, have been reported to occur. If chlorpyrifos is a causal factor for neurodevelopmental effects at low doses, one would still expect to see a relationship between the dose and any reported effects, both within and among studies.

The studies examining the neurodevelopmental effects of chlorpyrifos in rodents do not demonstrate consistent exposure-response relationships either within or across studies. Ricceri *et al.* (2003) reported a larger effect on socioagonistic behavior in mice after exposure to 1 *vs.* 3 mg/kg-day chlorpyrifos, and Ricceri *et al.* (2006) reported some indication of socioagonistic behavior only after exposure to 3, but not 1 mg/kg-day chlorpyrifos. Braquenier *et al.* (2010) reported a chlorpyrifos-associated increase in anxiety in female mice that occurred at an exposure concentration of 1 mg/kg-day chlorpyrifos, but not at the lower or higher exposures in that study (0.2 and 5 mg/kg-day chlorpyrifos, respectively). Ricceri *et al.* (2006) reported increased locomotor activity associated with prenatal exposure to 6 mg/kg-day chlorpyrifos when followed by postnatal exposure to 1 mg/kg-day chlorpyrifos but not to 3 mg/kg-day chlorpyrifos. Levin *et al.* (2002) reported that female rats had a greater number of working and reference memory errors than controls in the radial arm maze after exposure to 1 mg/kg-day chlorpyrifos, but not 5 mg/kg-day. In the foraging maze, male mice showed increased food recognition and position learning after exposure to 1 mg/kg-day chlorpyrifos, but not at the higher dose of 5 mg/kg-day (Haviland *et al.*, 2010).

Across studies, anxiety was increased in female mice assessed in the elevated plus-maze after oral exposure of their dams to 1 mg/kg-day chlorpyrifos during GD

15– PND 14 (Braquenier *et al.*, 2010), but not in offspring of mouse dams exposed orally to doses of 3 or 6 mg/kg-day chlorpyrifos during GD 15–18 followed by subcutaneous exposure of the offspring to 1 or 3 mg/kg-day chlorpyrifos during PND 11–14 (Ricceri *et al.*, 2006). Ricceri *et al.* (2003) reported increased locomotor activity in mice exposed to 1 or 3 mg/kg-day by subcutaneous injection during PND 11–14, but both Dam *et al.* (2000) and Levin *et al.* (2001) reported no effects on locomotor activity in rats exposed to 5 mg/kg-day chlorpyrifos by the same route during the same postnatal period. Transient decreases in alternation latency in the T-maze test were reported in two studies after subcutaneous exposure to 1 or 5 mg/kg-day chlorpyrifos during prenatal periods (Icenogle *et al.*, 2004; Levin *et al.*, 2002), but another study reported no effects in the T-maze test after oral exposure to 0.3, 1, or 5 mg/kg-day chlorpyrifos during the perinatal period (Maurissen *et al.*, 2000). The studies that assessed neuromotor function and sensorimotor reflexes did not provide sufficient data to support an exposure-dependent effect of chlorpyrifos on these endpoints, except perhaps at exposures above 6 mg/kg-day chlorpyrifos. Together, the animal data do not demonstrate clear exposure-response relationships between chlorpyrifos exposure and neurodevelopmental effects.

3.3.3.4. Biological significance of responses Addressing the question of causation at doses well below those traditionally recognized as causing biological responses is a challenge because the magnitude of the putative responses may be only marginally detectable. Even if an effect is due to the treatment, one must evaluate the biological relevance of the effect. Statistical significance may be overruled by a lack of biological relevance, such as if the magnitude of response is small or the observed change is not interpretable as an adverse response (Goodman *et al.*, 2010). Many of the behavioral changes associated with chlorpyrifos treatment were mild and/or transient, and several were observed to be in the wrong direction for adversity (*e.g.*, fewer memory errors). Many could also not be replicated under the same or similar conditions in other studies and, thus, may be chance fluctuations or due to another factor. The biological significance of neurodevelopmental effects of chlorpyrifos in animals also depends on whether they are observed at doses above or below the threshold for inhibition of AChE activity in the brain. If chlorpyrifos causes neurodevelopmental effects at very low doses that are presumed to occur through an alternative mechanism besides AChE inhibition, they should be consistently observed at doses below this threshold. Although AChE activity was not assessed in many of the studies, particularly those with exposure during prenatal periods, often the reported effects of chlorpyrifos exposure were observed in conjunction with AChE inhibition or at doses shown to be associated with AChE inhibition in other studies (*i.e.*, ≥ 1 mg/kg-day). This indicates that even if the reported neurodevelopmental effects are real, they do not likely act through a mechanism that only operates at doses below the threshold for AChE inhibition in the brain.

3.3.4. Conclusions for animal data

As a whole, the studies examining potential associations between chlorpyrifos exposure and neurodevelopmental effects in rodents indicate that only a few, isolated markers of certain behaviors are associated with low exposures at a given developmental stage. The specific changes in these markers are not necessarily in agreement across studies, are often transient or in the wrong direction for an adverse effect, are often observed in conjunction with AChE inhibition or at doses shown to be associated with AChE inhibition in other studies, and do not demonstrate consistent exposure-response relationships, except perhaps at very high exposure levels. The majority of studies used subcutaneous injection as the exposure route, which is not as relevant to humans, and studies using the more relevant oral exposure route do not appear to be more likely to report associations with neurodevelopmental effects. In fact, studies with the most weight report largely null effects across various neurodevelopmental tests, and those that do report treatment-related effects often report inhibition of AChE activity in the brain at the same doses. Overall, the animal data are not sufficiently robust to support the hypothesis that chlorpyrifos exposure causes neurodevelopmental effects at exposure levels below those associated with systemic toxicity or AChE inhibition.

3.4. Evaluation of mechanistic data

3.4.1. Introduction

It is well-established that the MoA for acute neurotoxicity of chlorpyrifos is inhibition of AChE activity in the nervous system *via* chlorpyrifos-oxon, with high doses leading to cholinergic toxicity (as reviewed by Eaton *et al.*, 2008). As described earlier in Section 3.1, there is a threshold for this inhibition, as it requires chlorpyrifos exposures that are high enough to overwhelm detoxification pathways, allowing chlorpyrifos-oxon to reach the brain, and clinical symptoms are evident only when at least 70% or greater inhibition of AChE activity is reached (Clegg and van Gemert, 1999). The studies in rodents reviewed above indicate that the LOAEL for AChE inhibition in the brain is 1 mg/kg chlorpyrifos. Below this level, however, cell damage and loss in the developing brain have been reported (Qiao *et al.*, 2002).

Because AChE inhibition is not observed at low chlorpyrifos exposures in humans, other mechanisms have been proposed for potential neurodevelopmental effects of chlorpyrifos at exposures below the threshold for cholinesterase inhibition. The proposed mechanisms involve the action of chlorpyrifos itself, rather than chlorpyrifos-oxon. Evidence for these mechanisms comes mainly from *in vitro* studies, as there are little *in vivo* data available. Below, we review the studies assessing these mechanisms and discuss their relevance to the neurodevelopmental outcomes assessed in epidemiology and animal studies.

3.4.2. Analysis of mechanistic data

3.4.2.1. Neuronal differentiation Cholinesterases play an important role in early development of the nervous system. Several studies suggest that in addition to its enzymatic activity, AChE has morphogenic activity in the developing nervous system. If chlorpyrifos causes effects at low doses during neurodevelopment, perturbation of this morphogenic activity is one potential mechanism that has been proposed to lead to adverse effects on neuronal differentiation and synaptic function.

Several studies have demonstrated that in noncholinergic neuronal cells *in vitro*, chlorpyrifos inhibits neurite and axonal outgrowth and enhanced dendritic growth at concentrations lower than those which inhibit cholinesterase activity in the brain. Chlorpyrifos inhibited nerve growth factor (NGF)-induced neurite outgrowth in rat pheochromocytoma (PC12) cells at 1 μg/mL (2.85 μM) without inhibiting cholinesterase activity, which occurred at 10 μg/mL (28 μM) (Das and Barone, 1999). PC12 cells are immature neuronal precursors that differentiate into postganglionic sympathetic neuron-type cells with high AChE activity upon NGF stimulation. Chlorpyrifos inhibited axon outgrowth in primary cultures of embryonic rat sympathetic neurons at exposures ≥ 0.001 μM in the presence of a CYP450 inhibitor, indicating the effect is likely attributable to chlorpyrifos and not chlorpyrifos-oxon (Howard *et al.*, 2005). Chlorpyrifos also enhanced BMP-induced dendritic growth at concentrations between 1 and 10 μM. AChE was inhibited only at 1 μM and above. The authors noted that these effects on axonal and dendritic growth are likely both independent of AChE inhibition because, even though the dendritic effects were observed at same concentrations as AChE inhibition, TCPy (which does not inhibit AChE) also enhanced dendritic growth (Howard *et al.*, 2005). Yang *et al.* (2008) also reported decreased axonal length in primary sensory neurons from embryonic rat dorsal root ganglia at concentrations from 0.001 μM to 10 μM. Inhibition of AChE activity was observed at concentrations of 0.1 μM and above. These effects on axon length required the presence of AChE, however, as they were not observed in AChE knockout neurons.

Both *in vitro* and *in vivo* studies have reported effects of chlorpyrifos on cholinergic neuronal development and function in the absence of cholinesterase inhibition. Jameson *et al.* (2006) reported a reduction in the activity of choline acetyltransferase (ChAT), a marker for the cholinergic phenotype, in PC12 cells exposed to 5 μM chlorpyrifos at the beginning of differentiation. Reduced ChAT activity was not observed in undifferentiated cells or in cells exposed during mid-differentiation, whereas increased activity of tyrosine hydroxylase (TH), a marker for the catecholamine phenotype, was observed in both of these cell types. These results indicate that the start of differentiation is a critical period for chlorpyrifos to impair the development of the cholinergic phenotype, whereas promotion of the expression of the catecholaminergic phenotype occurs in both undifferentiated and differentiated cells.

Dam *et al.* (1999) exposed neonatal rats subcutaneously to 1 mg/kg-day chlorpyrifos on PND 1–4 or to 5 mg/kg-day chlorpyrifos on PND 11–14 and examined effects on the development of cholinergic neuronal function in the brainstem, forebrain, and cerebellum, using indices of synaptic proliferation (ChAT activity) and synaptic activity (hemicholinium-3 [HC-3] binding). The brainstem and forebrain develop prominent cholinergic inputs, whereas the cerebellum is sparse in cholinergic projections. Early treatment decreased synaptic proliferation without affecting synaptic activity in the forebrain. Neither measure was affected in the brainstem. Effects of chlorpyrifos were observed on the catecholamine pathways as well: early or late treatment increased the synaptic activity of the neurotransmitters norepinephrine and dopamine, with the greatest effects in the cerebellum. Effects on catecholamine systems were unrelated to the magnitude or temporal pattern of cholinesterase inhibition. The observed deficient cholinergic synaptogenesis and increased catecholaminergic synaptic activity are consistent with the results of the *in vitro* study by Jameson *et al.* (2006).

Subcutaneous exposure of rat dams on GD17–20 with 1 mg/kg-day chlorpyrifos induced small changes in synaptic proliferation and marked suppression of synaptic activity in the forebrain of neonatal offspring. This reduction in synaptic activity returned to normal by weaning, but deficits were again apparent in the regions of the forebrain involved in learning and memory (cerebral cortex, hippocampus, and striatum) in adolescence and adulthood (Qiao *et al.*, 2003). Similar exposure on GD 9–12 increased synaptic proliferation and decreased synaptic activity in the hippocampus and striatum in adolescence and adulthood (Qiao *et al.*, 2004).

Together, these studies indicate that chlorpyrifos exposure at concentrations as low as 1 nM *in vitro* or at doses of 1 or 5 mg/kg-day during the prenatal or neonatal period in rodents can induce effects on neuronal differentiation and function. Some of these effects may be independent of cholinesterase inhibition in neuronal cells *in vitro*, but *in vivo* they are observed at doses shown in other studies to inhibit AChE in the brain.

3.4.2.2. Oxidative stress Several studies have suggested that chlorpyrifos can induce oxidative stress in various neuronal cell types *in vitro* and *in vivo* at concentrations below the threshold for cholinesterase inhibition. If chlorpyrifos is a developmental neurotoxicant at these concentrations, production of oxidative stress in the developing nervous system, leading to oxidative neuronal cell damage, has been proposed as an underlying mechanism (Crumpton *et al.*, 2000).

Crumpton *et al.* (2000) reported a dose-dependent increase in the generation of reactive oxygen species (ROS) measured concurrently with acute treatment (10 minutes duration) of PC12 cell suspensions with 0.5–50 μg/mL (1.4–142 μM) chlorpyrifos. Acute treatment

with 10 μg/mL (28 μM) chlorpyrifos-oxon had no effect on ROS generation. The effects with chlorpyrifos were transient, however, as no increases in ROS generation were observed immediately following prolonged (24–72 hours) chlorpyrifos exposure in either undifferentiated or NGF-differentiated PC12 cells.

Qiao *et al*. (2005) reported increased lipid peroxidation, an effect of oxidative stress, in both undifferentiated and NGF-differentiated PC12 cells after exposure to chlorpyrifos at concentrations of 1 μM and above. The authors noted that the lack of enhancement of sensitivity to this effect by differentiation (which increases AChE activity) is consistent with a noncholinergic mechanism.

Oxidative stress was also examined in oligodendrocytes, which are glial cells that are essential to neuronal differentiation, myelination, impulse propagation, and homeostatic maintenance. Disruption of oligodendrocyte function can manifest as motor, cognitive, or behavioral dysfunction. Saulsbury *et al*. (2009) exposed CG-4 cells (oligodendrocyte progenitors) to chlorpyrifos at concentrations between 15 and 120 μM. Chlorpyrifos induced a dose-dependent increase in cell death at concentrations of 30 μM and above. Reactive oxygen species (ROS) and intermediates were observed in cells with 15, 60, and 120 μM chlorpyrifos, but not 30 μM chlorpyrifos. Superoxide generation was induced at 30 and 60 μM, and pretreatment with diethyl maleate (DEM), which reduces intracellular levels of the antioxidant glutathione, enhanced chlorpyrifos-induced cell toxicity at 15 and 30 μM. Addition of a nitric oxide synthase inhibitor did not fully reverse chlorpyrifos-induced toxicity, suggesting that the toxicity is partly caused by production of nitric oxide. Vitamin E, a nonspecific antioxidant, completely spared cells from the toxic effects of chlorpyrifos, further indicating that chlorpyrifos exposure leads to generation of ROS.

Slotkin *et al*. (2005) reported no increases in lipid peroxidation in the developing brain of rats exposed to chlorpyrifos by subcutaneous injection on GD 17–20 or PND 1–4, even at doses well above the threshold for cholinesterase inhibition. Lipid peroxidation increased in the forebrain and cerebellum of males, but not females, with exposure to 5 mg/kg-day on PND 11–14, but no increases were reported in the brainstem. The authors stated that cholinergic hyperstimulation is not responsible for the oxidative damage, as the cerebellum, which, as noted above, is sparse in cholinergic projections, was affected more than the brainstem, which has major cholinergic inputs. These data are not consistent with the *in vitro* data of Qiao *et al*. (2005), who reported lipid peroxidation in undifferentiated cells which correspond to earlier stages of neurodevelopment. By contrast, Slotkin and Seidler (2009) interpreted the Qiao *et al*. (2005) data as indicating that co-exposure to NGF enhanced the lipid peroxidation induced by chlorpyrifos, which is consistent with increased sensitivity to oxidative stress during neurodifferentiation, and that the study by Slotkin *et al*. (2005) confirms the Qiao *et al*. (2005) results by showing greater lipid peroxidation in vulnerable brain regions during peak periods of axonogenesis and synaptogenesis.

Slotkin and Seidler (2009) examined the effects of 30 μM chlorpyrifos on mRNA levels of genes involved in oxidative stress responses and genes encoding receptors of glutamate (a neurotransmitter) in PC12 cells. They reported larger and more widespread transcriptional changes in genes related to oxidative stress response in differentiating cells compared to undifferentiated cells. In undifferentiated cells, they reported more robust effects on the expression of genes for ionotropic glutamate receptors, which mediate excitotoxic cell death in the developing brain, than on genes for oxidative stress or for metabotropic glutamate receptors, which are not involved in excitotoxic cell death. These results suggest a greater role for excitotoxicity than oxidative stress in the undifferentiated state (earlier stages of neurodevelopment) and an increasing role for oxidative stress as cells undergo differentiation (later stages of neurodevelopment).

Together, these studies indicate that chlorpyrifos exposure induces oxidative stress in neuronal cells *in vitro* at concentrations of at least 1 μM, particularly when these cells undergo differentiation, and induction of oxidative stress is observed *in vivo* at doses shown in other studies to inhibit AChE in the brain.

3.4.2.3. cAMP-related cell signaling Perturbations of the adenylyl cyclase (AC) signal transduction pathway have been proposed as a mechanism for the potential neurodevelopmental toxicity of chlorpyrifos at exposures below the threshold for cholinesterase inhibition in several studies. Stimulation of the AC pathway catalyzes the synthesis of cyclic AMP (cAMP), which is involved in the control of cell replication and differentiation. Higher levels of cAMP increase cAMP-dependent kinase (PKA)-mediated phosphorylation of several proteins, including the transcription factor CREB. CREB is critical for synaptic plasticity and transcription-dependent forms of memory and has a role in cell survival and differentiation during brain development. Perturbation of this pathway during development would be expected to have an impact on brain cell development and cognitive function.

Schuh *et al*. (2002) reported that chlorpyrifos increased pCREB levels (suggesting a stimulation of the AC pathway) in primary rat cortical neurons in culture with an EC50 of 60 pM. AChE activity was not affected at concentrations up to 100 nM chlorpyrifos, but was observed at 1 and 10 μM. Chlorpyrifos also increased pCREB levels in rat hippocampal neurons with an EC50 in the range of 1–10 nM, but not in astrocytes at concentrations up to 10 μM. The CYP450 inhibitor SKF-525A did not alter the effects of chlorpyrifos, indicating that metabolism to chlorpyrifos-oxon is not necessary for these changes in pCREB levels.

Slotkin *et al*. (2007) reported that exposure to 30 μM chlorpyrifos decreased basal, fluoride-stimulated, and

forskolin-stimulated AC activity in differentiating PC12 cells. Acetylcholine receptor inhibitors did not protect cells from these effects, indicating no contribution of cholinesterase inhibition. Addition of vitamin E worsened the effects on AC signaling, indicating that oxidative stress is also not involved. Theophylline, which prevents breakdown of cAMP, restored AC activity to normal or supranormal levels.

Adigun *et al*. (2010) also examined effects of chlorpyrifos on AC signaling in PC12 cells. Treatment with 50 μM chlorpyrifos had no effect on AC signaling in undifferentiated PC12 cells, but treatment of differentiating cells produced deficits in all AC measures (basal activity and response to fluoride, forskolin, and manganese) when exposure started at the onset of differentiation. If chlorpyrifos exposure was continued for six days, or if cells were exposed for two days and then examined four days later, there was complete reversal of the inhibitory effects on AC signaling. Effects on cell signaling were distinct from those on indices of cell number and neurite outgrowth, which showed progressively greater effects at six days than at two days. This indicates that early exposure reprograms the function of the AC signal transduction pathway.

Song *et al*. (1997) subcutaneously injected neonatal rats with 1 or 5 mg/kg-day chlorpyrifos on PND 1–4 or with 5 mg/kg-day chlorpyrifos on PND 11–14 and examined effects on components of the AC cascade in brain regions enriched (forebrain) or sparse (cerebellum) in cholinergic innervation. Rats exposed to 5 mg/kg-day chlorpyrifos on PND 1–4 exhibited > 50% mortality during the exposure period. The authors measured inhibition of brainstem AChE activity 24 hours after the last dose and reported 25% inhibition for the 1 mg/kg-day group and 65% inhibition for the 5 mg/kg-day group exposed on PND 11–14, with substantial recovery from inhibition during the next five days. In the forebrain, 1 mg/kg chlorpyrifos induced deficits of 25–35% in basal and stimulated AC activities by PND 10 in rats exposed on PND 1–4. There were no effects on AC activity at PND 5, but effects worsened over the next five days after treatment cessation, the period in which recovery from cholinesterase inhibition occurred. The effects were not as pronounced (5–15% deficits) in animals exposed to 5 mg/kg chlorpyrifos on PND 11–14. The deficiencies in AC activity were also reported in the cerebellum and the heart, indicating that cholinergic overstimulation alone cannot account for these changes. Raising the dose to induce systemic toxicity (*i.e.*, 5 mg/kg-day on PND 1–4) did not further enhance the effects on AC activity, suggesting that these effects may also occur at lower concentrations in the absence of cholinesterase inhibition.

Meyer *et al*. (2004) subcutaneously injected rat dams with 1 or 5 mg/kg-day chlorpyrifos on GD 9–12 or GD 17–20, or rat pups with 1 mg/kg-day chlorpyrifos on PND 1–4 or 5 mg/kg-day on PND 11–14, and examined the function of the AC signaling pathway in several different brain regions during adulthood (PND 60). Effects on the AC pathway in rats exposed on GD 9–12 required the higher dose of 5 mg/kg-day chlorpyrifos. Exposures on GD 17–20 and later produced sex-specific alterations in the AC pathway. The effects were either stimulatory or inhibitory to the pathway, depending on the time of exposure, sex, and brain region. The authors noted that this rules out the possibility that chlorpyrifos interacts directly with the neurotransmitter receptors or proteins of the AC signaling cascade, because otherwise the alterations would have been similar in every region, for both sexes, and for each dosing regimen. Instead, they note that their results suggest that chlorpyrifos disrupts the program for development of cell signaling, with targeting of specific brain regions for each sex that depend upon the maturational phases of vulnerability of various neural cell populations.

Together, these studies indicate that chlorpyrifos exposure can lead to a disruption and reprogramming of signaling cascades related to the AC signal transduction pathway, but *in vivo*, these effects are only observed at doses shown in other studies to inhibit AChE activity in the brain.

3.4.2.4. Serotonergic dysfunction Neurodevelopmental alterations from chlorpyrifos are not confined to cholinergic systems and may involve a wide variety of neurotransmitters, such as serotonin (5HT) or dopamine (DA). Several studies have reported effects of chlorpyrifos on the functioning of 5HT synapses in rats after subcutaneous exposure during different stages of early development. These effects involved alterations in levels of the 5HT presynaptic transporter (5HTT), a biomarker for the concentration of 5HT nerve terminals that is responsible for regulating the concentration of 5HT in the synapse, levels of 5HT receptors that control cell signaling, or activities of 5HT or DA.

Gestational exposure to 1 or 5 mg/kg-day chlorpyrifos on GD 9–12 increased 5HT activity in the cerebral cortex, but not the midbrain or brainstem, in both sexes during adolescence (PND 30) (Slotkin and Seidler, 2007). Similar effects were observed for DA activity, but at lower magnitude. In adulthood, elevations in levels of 5HTT and the 5HT receptors, 5HT1A and 5HT2, were observed in the cerebral cortex, midbrain, and brainstem of both sexes (Aldridge *et al*., 2004).

Exposure to 1 or 5 mg/kg-day chlorpyrifos on GD 17–20 increased 5HT activity in brain regions with either 5HT projections or cell bodies in males at both doses and in females only at the higher dose (above the threshold for cholinesterase inhibition) during adolescence (Slotkin and Seidler, 2007). Similar effects were reported for DA activity, but at lower magnitude and with no sex preference (Slotkin and Seidler, 2007). This exposure scenario also induced larger effects on elevations of 5HT1A, 5HT2, and 5HTT during adulthood compared to those with exposure during GD 9–12, with selectivity for brain regions with 5HT nerve terminals and preferential effects in males (Aldridge *et al*., 2004). Also during adulthood,

decreased 5HT levels were reported in animals exposed to 5 mg/kg-day and a net increase in 5HT activity was reported with both doses, with no preference for these effects in either sex (Aldridge *et al.*, 2005b). DA content was unaffected in most brain regions, but large deficits were observed in the hippocampus with both doses of chlorpyrifos in both sexes (Aldridge *et al.*, 2005b). DA turnover was increased in the cerebral cortex, striatum, and midbrain in both sexes with the 5 mg/kg dose (Aldridge *et al.*, 2005b).

Exposure to 1 mg/kg-day chlorpyrifos on PND 1–4 resulted in an increase in levels of 5HT receptors during adulthood, with larger effects in males and in regions with 5HT cell bodies (Aldridge *et al.*, 2004). Levels of 5HTT were increased in both sexes in the brainstem and decreased in all other brain regions examined in females (Aldridge *et al.*, 2004). In behavioral tests conducted during adulthood, treated animals showed abnormalities related to 5HT deficiencies. The normal sex differences for the elevated plus maze and 16-arm radial maze tests were ablated, as the behavior of treated males was "feminized," resulting in similar scores to those observed for control females (Aldridge *et al.*, 2005a). The levels of 5HT were unchanged in males and slightly decreased in females, whereas 5HT activity was increased in both sexes (Aldridge *et al.*, 2005b). DA content and activity were decreased in the cerebrocortical area and increased in the striatum, and DA activity was increased in the midbrain (Slotkin *et al.*, 2002).

Following exposure to 5 mg/kg-day chlorpyrifos on PND 11–14, smaller increases in 5HT1A and 5HT2 were observed in adulthood, and 5HTT levels were decreased with the same regional and sex selectivity as observed in animals exposed on PND 1–4 (Aldridge *et al.*, 2004). There were no effects on the content or activity of either 5HT (Aldridge *et al.*, 2005b) or DA (Slotkin *et al.*, 2002).

Together, these *in vivo* data indicate that the immediate perinatal period has the greatest sensitivity and sex-selectivity to effects of chlorpyrifos on indices of serotonergic activity in rodents. The windows of GD 17–20 and PND 1–4 encompass the peak period of sexual differentiation in the brain (MacLusky and Naftolin, 1981), and exposures prior to this period did not produce male-female differences in outcomes. Aldridge *et al.* (2005b) suggested that the effects on 5HT indices observed during the perinatal period are indicative of deficient synaptic communication that is consistent with a "miswiring" of 5HT circuits and that effects on 5HT may be one component of a larger spectrum of chlorpyrifos-induced disruption of synaptic development and function that can ultimately contribute to behavioral anomalies. One hypothesized example could be through disruption of 5HT-mediated cell signaling, which includes AC signaling (Aldridge *et al.*, 2004); another is that deficiencies in 5HT systems could create a situation of increased reliance on 5HT mechanisms for cognitive function that aren't normally called into play (Aldridge *et al.*, 2005a). Alternatively, chlorpyrifos may not specifically target serotonergic systems, as effects could be secondary to those on neuronal differentiation (Aldridge *et al.*, 2004). It should be noted, however, that these effects are only observed with chlorpyrifos exposures that are associated with inhibition of AChE activity in the brain in other studies.

3.4.3. Conclusions for mechanistic data

Several mechanisms that presumably do not involve inhibition of AChE activity in the nervous system have been explored to determine whether chlorpyrifos can act as a neurodevelopmental toxicant *via* other mechanisms. The potential mechanisms include perturbation of the morphogenic, rather than enzymatic, activity of AChE; neuronal cell damage caused by induction of oxidative stress; disruption of the AC signal transduction pathway; and dysfunction of serotonergic systems. These mechanisms have been suggested to be involved in a large spectrum of effects such as chlorpyrifos-induced neuronal cell damage or disruption of systems controlling neuronal differentiation and synaptic function, although serotonergic dysfunction is involved in appetite and affective (depression) disorders and it is unclear how this would be relevant to the neurodevelopmental outcomes assessed in the epidemiology and animal studies described above. Evidence for the action of the proposed mechanisms at doses not affecting AChE activity comes mainly from *in vitro* studies, so their relevance to potential outcomes in children with very low exposures to chlorpyrifos is unclear. In addition, the chlorpyrifos concentrations used in the *in vitro* studies must be quantitatively considered for their relevance to human systemic concentrations. Chlorpyrifos concentrations in the blood of subjects in the Columbia cohort averaged 4 pg/g, which is equivalent to 0.01 nM, but were also measured as high as 0.1 nM (Eaton *et al.*, 2008). Using a blood/brain partition coefficient of 33 calculated by Timchalk *et al.* (2002) for chlorpyrifos in rats, brain concentrations would be estimated to range from 0.33–3.3 nM (Eaton *et al.*, 2008). Almost all of the *in vitro* studies reported potential mechanistic effects at micromolar concentrations, although three studies reported effects at concentrations ≤ 1 nM (Schuh *et al.*, 2002; Howard *et al.*, 2005; Yang *et al.*, 2008). AChE inhibition was also reported at micromolar concentrations in the *in vitro* studies (Das and Barone, 1999; Howard *et al.*, 2005; Schuh *et al.*, 2002). Overall, these data indicate that chlorpyrifos exposures at which potential mechanistic effects and AChE inhibition were observed in most of the *in vitro* studies are 1,000-fold higher than the estimated exposures to chlorpyrifos in the epidemiology studies.

3.5. HBWoE evaluation of the potential neurodevelopmental toxicity of chlorpyrifos

A general hypothesis that has been put forth in the scientific literature is that chlorpyrifos causes adverse neurodevelopmental effects in humans at exposures below

the threshold for cholinesterase inhibition in the brain by chlorpyrifos-oxon, and that chlorpyrifos itself is acting as the neurotoxicant at low exposures by one or more proposed mechanisms during critical periods in the developing brain. The concern for chlorpyrifos-induced neurodevelopmental effects stems from a few epidemiology studies reporting potential associations between low-dose chlorpyrifos exposure and effects on infant neurobehavior, as well as cognitive and motor development and behavior outcomes in children, although other studies do not show these effects. The epidemiology studies have shortcomings, however, such as being subject to substantial confounding variables including low SES, maternal smoking and alcohol use, and exposures to other pesticides, as well as a lack of specific, reliable biomarkers of exposure. The studies with more robust factors, such as reliable exposure metrics or larger sample sizes, do not appear to be more likely to report associations with adverse neurodevelopmental effects.

In addition to the epidemiology studies, many rodent studies have been conducted to examine the potential neurodevelopmental effects of chlorpyrifos. These studies indicate that only a few isolated markers of certain behaviors are associated with low exposures (compared to those which cause systemic toxicity) at a given developmental stage. The specific changes in these markers are not necessarily in agreement across studies, are often transient or in the wrong direction for an adverse effect, are often observed in conjunction with AChE inhibition or at concentrations shown to be associated with AChE inhibition in other studies, and do not demonstrate consistent exposure-response relationships, except perhaps at very high exposure levels. While some investigators have proposed potential mechanisms for the effects of chlorpyrifos at doses below those associated with cholinesterase inhibition, the evidence for these comes mainly from *in vitro* studies, and the animal data for neurodevelopmental outcomes do not provide strong support for chlorpyrifos neurotoxicity at doses below this threshold.

Below, we evaluate the scientific data relevant to examining whether there is a causal association between exposure to chlorpyrifos and adverse neurodevelopmental effects in humans using the structured HBWoE approach (Rhomberg *et al*., 2010, 2011). This approach weighs all of the data from epidemiology, animal toxicity, and mechanistic studies in terms of quality and relevance to humans, allowing each of these data sets to inform one another. All of the data are then evaluated together to determine whether a causal relationship between chlorpyrifos at low exposures and neurodevelopmental effects in humans is plausible. This evaluation considers the uncertainties and inconsistencies in the data sets, as well as any *ad hoc* assumptions that may be required for some of the hypotheses put forth. The key outcome of this approach is an evaluation and comparison of alternative accounts (or hypotheses) of all the available data. If data of poor quality are used as a basis to support one of the accounts, the logic of how these data have been interpreted and the *ad hoc* assumptions needed to fit these data to the proposed hypothesis are discussed. Various competing accounts are weighed by comparing the *ad hoc* assumptions needed for each, with more credence given to the hypothesis that requires the least amount of assumptions.

We consider two hypotheses for MoAs that have been put forth in the literature to explain the existence of human risks for adverse neurodevelopmental effects of chlorpyrifos. These hypotheses are based on the human, animal, and mechanistic data from studies assessing whether there are adverse neurodevelopmental effects of chlorpyrifos at doses below those at which other effects have been observed, and rely on several lines of evidence regarding potential mechanisms for low-dose effects.

3.5.1. Hypotheses under consideration

We reiterate that stating the hypotheses in the HBWoE approach requires more than just putting forth the question of whether chlorpyrifos causes neurodevelopmental effects in human populations at the levels to which they are exposed. It is important to articulate the logical basis one is invoking to consider the available studies as evidence that is relevant to the potential for risk in the target human population. It is this articulation of a logical basis that identifies what is asserted as being in common between each studied situation and the others, and between each studied situation and the target human population, such that the relevance of findings and the ways of accounting for similarities and differences and for consistencies and inconsistencies can be considered as one brings the results to bear on the motivating question.

The first hypothesis is that chlorpyrifos induces adverse effects on the developing nervous system at doses below those which inhibit the activity of AChE in the brain. Moreover, any such effects are hypothesized to apply similarly across mammalian species, such that effects observed in animal experiments would be expected to apply across species, including in humans. Low-exposure human studies can be evaluated to determine whether they show indications of this presumed ability to affect neural development, which would be consistent with this hypothesis. Several specific mechanisms for effects at doses below those which inhibit brain AChE activity have been proposed by various research groups, each involving the action of chlorpyrifos itself rather than chlorpyrifos-oxon. These potential mechanisms include perturbation of the morphogenic, rather than the enzymatic, activity of AChE; neuronal cell damage caused by induction of oxidative stress; disruption of the AC signal transduction pathway; and dysfunction of serotonergic systems. It has been suggested that these mechanisms may be involved in a large spectrum of chlorpyrifos-induced neuronal cell damage or disruption of systems

controlling neuronal differentiation and synaptic function. Evidence for these mechanisms comes mainly from *in vitro* studies, as there are little *in vivo* data available, so their relevance to potential neurodevelopmental outcomes in children is unclear.

Under the first hypothesis, it would need to be presumed that the various kinds of measured outcomes across studies are all manifestations of a common underlying neurotoxic mechanism that is being hypothesized to apply across settings. Thus, it would be presumed that the observation of somewhat different particular results in each study and in each species can be taken as evidence of the operation of the same mechanism of action. Under this view, differences in particular outcomes across studies are attributed to the way investigators have chosen to measure the observable manifestation of this common underlying mechanism of toxicity.

The second hypothesis is that chlorpyrifos induces neurodevelopmental toxicity only through the established mechanism of inhibition of AChE activity in the nervous system by its metabolite, chlorpyrifos-oxon. This mechanism requires sufficient doses of chlorpyrifos for chlorpyrifos-oxon to reach the brain. If neurodevelopmental effects are observed in animal studies, this mechanism potentially underlies these effects if observed with doses that inhibit AChE activity to some extent, even if this inhibition is not high enough for systemic toxicity (*i.e.*, > 70% inhibition). This hypothesis is consistent with a lack of neurodevelopmental effects in humans because exposures in humans are far below those that are associated with AChE inhibition.

3.5.2. Evaluation of hypotheses for each line of evidence

We next considered the two hypotheses in the context of each line of evidence (epidemiology, animal, and mechanistic) and evaluated how well the hypotheses are in agreement with the available data, how well they would explain patterns in the data if they were true, what other events or processes should be observed if they are true, and whether these processes, in fact, are observed. For each hypothesis, the following questions become evident:

1. Are the data from the epidemiology studies compelling? Are there alternative explanations for the few positive associations observed in these studies?
2. What is the evidence that chlorpyrifos is associated with neurodevelopmental effects in animals, and are these effects observed in the absence of cholinesterase inhibition in the brain?
3. What is the evidence that the candidate mechanisms act only through chlorpyrifos and not chlorpyrifos-oxon?

1. Are the data from the epidemiology studies compelling? Are there alternative explanations for the few positive associations observed in these studies?

There are few studies of each specific neurodevelopmental outcome examined in each cohort, limiting the ability to look for consistency of outcomes across studies or cohorts. The studies that carry more weight, such as those with more reliable exposure metrics or larger sample sizes, do not appear to be more likely to report associations with adverse neurodevelopmental effects.

The outcome examined in the largest number of studies was newborn head circumference (Perera *et al.*, 2003; Whyatt *et al.*, 2004; Barr *et al.*, 2010; Eskenazi *et al.*, 2004; Berkowitz *et al.*, 2004; Wolff *et al.*, 2007). Regardless of the weight of each study, null results were reported across all studies of this outcome, increasing the likelihood that the overall findings are robust and that there is no association between chlorpyrifos exposure and decreased newborn head circumference.

The two studies that examined infant neurobehavior (Engel *et al.*, 2007; Young *et al.*, 2005) reported associations between chlorpyrifos exposure and abnormal reflexes in the BNBAS, but these studies carry less weight because of their potential for exposure measurement error from the use of urinary DEPs as the exposure metric and misclassification of outcome from a single assessment of neurobehavior.

Four studies examined cognitive and motor development, and three of these studies used the BSID-II and reported no associations between chlorpyrifos exposure and MDI or PDI scores up to 24 months of age in the Columbia (Rauh *et al.*, 2006), CHAMACOS (Eskenazi *et al.*, 2007), and Mount Sinai (Engel *et al.*, 2011) cohorts. Only Rauh *et al.* (2006) examined these outcomes in children at 36 months of age, using a single measure of chlorpyrifos in cord blood as the exposure metric. At this age, associations with lower PDI scores and with mental and motor delays were reported, although it is unclear whether these effects are clinically significant, as the mean scores for children in the highest exposure group were well within the normal range. Of the two studies that assessed scores on the WISC-IV, one reported an association between cord blood chlorpyrifos and a decrement of 0.35 to 0.81 points in working memory scores at age seven in the Columbia cohort (Rauh *et al.*, 2011), whereas the other reported no association with changes in scores on the WISC-IV in children between seven and 9 years of age in the Mount Sinai cohort (Engel *et al.*, 2011). Rauh *et al.* (2011) did not report the mean or range of WISC-IV scores, so it is unclear whether the modest decrement reported for the working memory index is clinically significant.

Two studies conducted a single assessment of behavioral outcomes at one age per study using the CBCL. This may lead to outcome misclassification, particularly when the measure is based on reporting of behavior by mothers, which is subject to reporting bias. The study using the more robust exposure metric at only one time point (Rauh *et al.*, 2006) reported associations with attention problems, ADHD, and PDD, whereas the study with the larger sample size but a less-reliable exposure metric at two time points (Eskenazi *et al.*, 2007) only reported an association with PDD. In both studies, very few children

scored in the clinical ranges of the CBCL, limiting the clinical significance of the results. The lack of confirmation of the associations, and the methodological issues in these studies increases the likelihood that there are alternative explanations for the observed outcomes other than exposure to chlorpyrifos.

There are many alternative explanations for the few positive associations with neurodevelopmental outcomes reported in some of the cohort studies. One explanation is that exposure measurement error increases the uncertainty that any associations are specifically attributable to chlorpyrifos exposure. Direct measurement of the parent compound more accurately reflects the chlorpyrifos dose in the brain than do measurements of its metabolites in urine. Urinary TCPy originates from exposure to not only chlorpyrifos, but to chlorpyrifos-methyl and to TCPy itself, and urinary DEPs originate from diazinon and disulfoton in addition to chlorpyrifos, so measurements of these biomarkers are not specific to chlorpyrifos and can overestimate exposure to chlorpyrifos. Most of the exposure metrics used in the cohort studies were measured at only one point in time, but there is large intraindividual variability in exposures to chlorpyrifos, so a single measurement may not represent average exposure over time or exposure at some earlier or later time. In addition, because of the rapid elimination of chlorpyrifos and its metabolites from the body, any measure of chlorpyrifos or its metabolites in blood or urine at a single time point reflects exposure during the brief period of time prior to measurement and may not accurately reflect exposure throughout the entire critical period of neurodevelopment.

Another explanation for the positive findings in some of the cohort studies stems from the lack of adequate adjustment for several confounding factors, decreasing the likelihood that any observed effects are attributable to chlorpyrifos exposure. Each cohort was exposed to multiple types of pesticides besides chlorpyrifos, but these exposures were not controlled for in any of the studies reporting associations. Maternal smoking and alcohol use have been associated with adverse neurodevelopmental outcomes, but these factors were likely underestimated in the few cohorts in which they were considered. The cohorts in which positive associations were reported all come from populations with low SES, which has also been associated with effects on neurodevelopment. Of the cohort studies reporting associations with neurodevelopmental outcomes, all examined at least some confounding factors related to SES (*e.g.*, mother's education; household income; quality of home environment) and those that were associated with outcomes were included as covariates in the final models. Although there was adjustment for several different confounders in each cohort, other factors that could affect the results may not have been accounted for, which increases the likelihood that there are alternative explanations for the observed outcomes other than exposure to chlorpyrifos.

Finally, the positive associations reported in some of the cohort studies could actually be statistical anomalies rather than actual associations with chlorpyrifos exposure. This is because in each study that reported associations between chlorpyrifos exposure and neurodevelopmental outcomes, many different analyses were conducted and there was no adjustment for multiple comparisons. This increases the likelihood that several results were statistically significant by chance. In addition, if studies with smaller sample sizes report associations but studies with larger sample sizes do not, it is more likely that the reported associations are statistical anomalies or that publication bias occurred.

Overall, we found that the epidemiology data are not sufficiently robust to support the hypothesis that chlorpyrifos is a causal factor for neurodevelopmental effects. The cohort studies do not report consistent results, and there is a lack of clear exposure-response information. Because of the many uncertainties in these studies, the few positive results may also support alternative explanations that factors other than chlorpyrifos are causal for the reported outcomes, or that the observed associations are statistical anomalies.

2. What is the evidence that chlorpyrifos is associated with neurodevelopmental effects in animals, and are these effects observed in the absence of cholinesterase inhibition in the brain?

The studies assessing potential neurodevelopmental effects of chlorpyrifos in rodents indicate that usually a few, isolated markers of certain behaviors were determined to be statistically significant in pair-wise comparisons with controls, but these were generally contradicted by other studies reporting no effects in similar dose ranges using similar routes of exposure, and sometimes showed effects in the other direction. In addition, many of the behaviors were observed in conjunction with AChE inhibition or at concentrations shown to be associated with AChE inhibition in other studies, and they did not demonstrate consistent exposure-response relationships, except perhaps at very high exposure levels. For each endpoint, we did not find a pattern of an effect that was consistent over doses and time points within studies or consistent across studies to constitute a repeatable finding, increasing the likelihood that the reported effects may be due to chance. This was not only true for the animal data as a whole, but also for the most rigorous animal studies, which should carry the most weight. When we separately examined the studies that used dose groups with a relatively high number of animals of each sex, considering both oral and subcutaneous exposures and including the one study that complied with US EPA Pesticide Assessment Guidelines and GLP regulations, we found that these studies reported largely null effects across various neurodevelopmental tests, and those that did report treatment-related effects often reported inhibition of AChE activity in the brain at the same doses, suggesting that effects occur *via* this pathway.

Several animal studies examined AChE activity in the brain after both oral and subcutaneous chlorpyrifos exposures during postnatal periods, and inhibition of AChE was observed in these studies at doses as low as 1 mg/kg-day when assessed within one day of exposure cessation. Very few of the neurodevelopmental studies assessed AChE inhibition after prenatal chlorpyrifos exposure, however, and only at doses of 3 mg/kg-day and greater, with the exception of one study with dermal exposure to 1 mg/kg-day that reported *increased* AChE activity 90 days after exposure cessation (Abou-Donia *et al.*, 2006). Given the rapid recovery from AChE inhibition after exposure to chlorpyrifos, it is expected that inhibition would have been observed shortly after this exposure. In addition, a study by Qiao *et al.* (2002) reported a no observed adverse effect level (NOAEL) of 1 mg/kg-day and a LOAEL of 2 mg/kg-day for AChE inhibition in the brain when assessed 24 hours after exposure of fetal rats *via* subcutaneous injection of dams on GD17–20. It remains possible that the 1 mg/kg-day dose could have inhibited brain AChE activity if assessed within a few hours of exposure cessation. Regardless of the limited data for AChE inhibition at doses below 3 mg/kg-day, none of the studies with prenatal oral exposures reported neurodevelopmental effects at doses below 6 mg/kg-day, with the exception of two studies reporting different markers of increased anxiety in female mice. In addition, in almost all studies that reported effects with prenatal subcutaneous exposures, they were observed at doses of at least 5 mg/kg-day, with the exception of a few effects observed at a dose of 1 mg/kg-day. These effects included a slower trend of habituation 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 transiently observed in the opposite direction.

Together, the animal data do not provide clear evidence that chlorpyrifos is associated with neurodevelopmental effects at doses that are below the threshold for inhibition of AChE in the brain. Most of the observed effects occurred at doses high enough to inhibit AChE, but not to induce systemic toxicity, which is hypothesized to occur when the extent of AChE inhibition is above 70% (Clegg and van Gemert, 1999). Because of this, it cannot be ruled out that exposures to concentrations of chlorpyrifos that induce a smaller degree of AChE inhibition are associated with certain neurodevelopmental effects in rodents.

3. What is the evidence that the candidate mechanisms act only through chlorpyrifos and not chlorpyrifos-oxon?

The specific mechanisms proposed to support the hypothesis that chlorpyrifos induces adverse neurodevelopmental effects at doses below those which inhibit the activity of AChE in the brain involve the action of chlorpyrifos itself rather than chlorpyrifos-oxon. In some of the *in vitro* studies, evidence for these mechanisms was observed in cells that were exposed to chlorpyrifos in the presence of a CYP450 inhibitor, which prevented metabolism of chlorpyrifos to chlorpyrifos-oxon (Howard *et al.*, 2005; Schuh *et al.*, 2002). Because of this, it is unlikely that any of the observed effects were attributable to chlorpyrifos-oxon unless it was present in the experiment through contamination. In addition, one of the *in vitro* studies reported ROS generation in cell suspensions after exposure to chlorpyrifos but not chlorpyrifos-oxon (Crumpton *et al.*, 2000).

The *in vivo* evidence for the proposed mechanisms is less clear, as some studies reported effects in both cholinergic and noncholinergic brain regions, as well as in serotonergic systems. These effects were all observed in the presence of AChE inhibition or at doses associated with AChE inhibition in other studies, with the possible exception of effects on synaptic proliferation and activity observed in rats exposed to 1 mg/kg-day chlorpyrifos during GD 17–20 (Qiao *et al.*, 2003), a treatment that was not associated with AChE inhibition in the brain in another study, at least when assessed 24 hours after exposure cessation (Qiao *et al.*, 2002). Although all of the *in vivo* mechanistic studies used subcutaneous injection as the exposure route to avoid first-pass metabolism of chlorpyrifos to chlorpyrifos-oxon, there is evidence for extrahepatic metabolism to chlorpyrifos-oxon in the brain (Chambers and Chambers, 1989). Because the *in vivo* studies do not provide evidence for an absence of AChE inhibition, chlorpyrifos-oxon is presumed to be present in the nervous system in sufficient amounts to inhibit AChE in these studies and could be driving the specific mechanisms, regardless of whether they act through noncholinergic processes. Although chlorpyrifos itself can also inhibit AChE in the brain, this requires much higher concentrations compared to chlorpyrifos-oxon. If the proposed mechanisms involve only the action of chlorpyrifos because they are presumed to occur at chlorpyrifos exposures below those that induce AChE inhibition in the brain from chlorpyrifos-oxon, then these mechanisms should also operate at exposures below those that allow chlorpyrifos itself to inhibit AChE. This is clearly not the case, however, because AChE activity is inhibited at the chlorpyrifos exposures used in the *in vivo* mechanistic studies.

3.5.3. Evaluation of hypotheses for all lines of evidence together

Now that we have considered the two hypotheses in the context of each line of evidence, we now consider all lines of evidence as a whole and how they inform the interpretation of each other. For each hypothesis, we ask the following questions:

1. How do the neurodevelopmental effects examined in the animal studies relate to those examined in the cohort studies? Are the rodent behavioral tests sufficient to detect the subtle effects asserted to be caused by chlorpyrifos exposure in humans?

2. How do the doses used in the animal studies compare to human exposures from the cohort studies and to the doses used in the mechanistic studies?
3. What is the evidence for the *in vivo* operation of the candidate mechanisms for neurodevelopmental effects at doses below those which induce brain cholinesterase inhibition? What consequences would these mechanisms have in humans, and is there any evidence for this in the literature?

1. How do the neurodevelopmental effects examined in the animal studies relate to those examined in the cohort studies? Are the rodent behavioral tests sufficient to detect the subtle effects asserted to be caused by chlorpyrifos exposure in humans?

The rodent tests for neurobehavioral assessment used in the animal studies were designed to measure similar functions as neurobehavioral tests conducted for humans but they cannot match the complexity of human behavior (Ulbrich and Palmer, 1996). The test batteries include tests of locomotor activity and crude assessments of learning, sensory, and motor integration, and these are typically parallel tests, in that they are conducted in a different manner in humans but it is believed that the same functions are being measured (Sharbaugh *et al*., 2003). Ideally, homologous tests, which follow the same procedure in both animals and humans, would provide a more accurate measure of the same cognitive function, but such tests are not available (Sharbaugh *et al*., 2003). For example, although a common endpoint for children is a score on a standardized test of intelligence or IQ, there are no standardized intelligence tests for animals that can identify subtle cognitive dysfunction (Rice, 2005; Winneke, 1992). The rodent test batteries are not capable of thoroughly characterizing the types of impairment produced in specific domains or in identifying the domains affected, so extrapolation from their results to specific deficits in children is problematic (Rice, 2005). In addition, humans are often not evaluated to the same extent as rodents after exposures to potential neurotoxicants, so generalizations from studies with rodent tests may be unfounded (Ulbrich and Palmer, 1996). Thus, current rodent models are only conceptual analogs to human studies, and results in animals models can be used to qualitatively characterize neurotoxic effects in humans (Winneke, 1992; Rice, 2005; Bellinger; 2005).

2. How do the doses used in the animal studies compare to human exposures from the cohort studies and to the doses used in the mechanistic studies?

Regarding the hypothesis that chlorpyrifos induces adverse effects on the developing nervous system in humans at doses below those associated with inhibition of AChE activity in the brain, if the animal studies report neurodevelopmental effects in conjunction with AChE inhibition or at doses in the range of those associated with AChE inhibition in other studies, then they should not be relevant to the human situation.

As noted in section 3.2.1.2, Eaton *et al*. (2008) estimated the daily intake of chlorpyrifos for mothers in the Columbia cohort as 0.008 μg/kg-day and for the CHAMACOS cohort as 0.007 μg/kg-day. These exposures are four orders of magnitude lower than the dose (100 μg/kg-day) that has been shown to significantly inhibit BuChE in plasma, which is the most sensitive *in vivo* biological effect of chlorpyrifos (Coulston *et al*., 1972), and five orders of magnitude lower than the LOAEL for AChE inhibition in the brain in the animal studies reviewed here. Because of this, it seems highly unlikely that exposures in these cohorts could produce inhibition of AChE in the maternal or fetal brain.

The chlorpyrifos exposures in the epidemiology studies are also much lower than those used in the *in vivo* and *in vitro* mechanistic studies. Chlorpyrifos concentrations in the brain of subjects in the Columbia cohort were estimated to range from 0.33 to 3.3 nM (Eaton *et al*., 2008). By contrast, Marty *et al*. (2007) reported that five-day-old rat pups exposed to 1 mg/kg chlorpyrifos had maximum blood concentrations (at two hours post-dosing) ranging from 16 to 140 nM, depending on the exposure route, which resulted in estimated brain concentrations of 0.5–4.6 μM. Almost all of the *in vitro* studies reported potential mechanistic effects at micromolar concentrations, although three studies reported effects at concentrations ≤ 1 nM (Schuh *et al*., 2002; Howard *et al*., 2005; Yang *et al*., 2008). AChE inhibition was also reported at micromolar concentrations in the *in vitro* studies (Das and Barone, 1999; Howard *et al*., 2005; Schuh *et al*., 2002). These data indicate that estimated exposures to chlorpyrifos in the cohort studies are 1,000-fold lower than those used in the animal studies or those at which effects, including AChE inhibition, were observed in most of the *in vitro* mechanistic studies.

3. What is the evidence for the in vivo operation of the candidate mechanisms for neurodevelopmental effects at doses below those which induce brain cholinesterase inhibition? What consequences would these mechanisms have in humans, and is there any evidence for this in the literature?

The underlying data used to support the various candidate mechanisms were mainly examined in animal cell lines *in vitro*. Herein, we summarize the limited *in vivo* evidence for these mechanisms and discuss whether there is evidence for their consequences in humans.

One proposed mechanism is the perturbation of the morphogenic, rather than enzymatic, activity of AChE in the developing nervous system. A few *in vivo* studies reported certain effects on synaptic activity and/or proliferation in brain regions that are either dense or sparse with cholinergic inputs after exposure to 1 mg/kg-day chlorpyrifos. It is not clear that these effects occurred in absence of AChE inhibition, however. There is also no

evidence that they are attributable only to chlorpyrifos and not chlorpyrifos-oxon, although some of the *in vitro* studies provide evidence to rule out chlorpyrifos-oxon, such as the presence of effects after co-exposure to CYP450 inhibitors and an absence of AChE inhibition. The inhibition of neurite outgrowth observed in the *in vitro* studies is a plausible mechanism for neurological deficits, as there is evidence in animals that neurite growth during brain development is essential for neuronal connectivity, and disruption of this process can lead to cognitive deficits (Berger-Sweeney and Hohmann, 1997; Webb *et al.*, 2001); however, it has not been shown that chlorpyrifos can induce such effects *in vivo*.

For the proposed mechanism involving neuronal cell damage caused by induction of oxidative stress, there is only one *in vivo* study available. This study reported increased lipid peroxidation only with postnatal exposure occurring after the period in which neuronal cell differentiation takes place, which is not consistent with results from *in vitro* studies showing similar amounts of lipid perioxidation in both undifferentiated as well as differentiated neurons. The effects were observed at a dose that has been shown to inhibit brain AChE in other studies (5 mg/kg-day), but the authors noted that lipid peroxidation was greater in the cerebellum, a brain region with sparse cholinergic innervations, suggesting that these effects do not involve cholinergic hyperstimulation. Although oxidative stress is a known mechanism for neuronal cell damage, including during human development (Ikonomidou and Kaindl, 2011), it has not been clearly shown that chlorpyrifos induces oxidative stress in the developing brain *in vivo* at concentrations that do not inhibit the activity of AChE.

Two *in vivo* studies examined the proposed mechanism of disruption of the AC signal transduction pathway. One study examined postnatal exposures and reported effects on this pathway only at doses that were also associated with AChE inhibition in the brain (Song *et al.*, 1997). The authors stated that because the effects were observed in several brain regions, including the cerebellum, and also in the heart, the effects cannot be attributable to cholinergic hyperstimulation alone. The other *in vivo* study used prenatal and postnatal exposures at doses that cause AChE inhibition and reported sex-specific alterations in the AC pathway (Meyer *et al.*, 2004). The effects were either stimulatory or inhibitory to the pathway, depending on the exposure period, sex, and brain region, and the authors hypothesized that they are indicative of a disruption and reprogramming of the AC signaling cascade during neurodevelopment. Perturbation of this pathway during development would be expected to have an impact on brain cell development and cognitive function, but the available evidence does not support effects on this pathway *in vivo* after exposure to chlorpyrifos at doses below the threshold for AChE inhibition in the brain.

The final proposed mechanism is dysfunction of serotonergic systems. All studies examining this mechanism were conducted *in vivo* using chlorpyrifos doses that have been shown to inhibit AChE in the brain in other studies. The effects are assumed to be noncholinergic because serotonin is not a neurotransmitter for cholinergic systems. Serotonergic dysfunction is involved in appetite and affective (depression) disorders, so it is unclear how this mechanism is relevant to the neurodevelopmental outcomes assessed in the epidemiology and animal studies.

3.5.4. Evaluation of alternative accounts

An HBWoE evaluation comes down to an evaluation of alternative "accounts," which are proposed sets of explanations for the observed phenomena across the body of relevant lines of evidence. These competing accounts should be evaluated to determine how the evidence supports them, what is necessary to assume for their support, and how the overall weight of the evidence for each suggests how compelling the account is. An account is most compelling when it is not only supported by the factual record, but also helps explain the data by finding common reasons for sets of observations and, moreover, achieves this ability much more readily than any competing account.

For chlorpyrifos, there are two competing accounts that need consideration:

1. The epidemiology evidence is sufficiently compelling that, even in the face of inconsistent evidence in animals with much higher doses, one or more of the proposed mechanisms for low-dose neurodevelopmental effects of chlorpyrifos must be right, and the biological implausibility of these mechanisms is mistaken (*i.e.*, because an effect appears, it must have a causal explanation).
2. Doubts about the potential mechanisms have merit, and the few apparent associations from epidemiology studies do not indicate a causal connection (*i.e.*, the appearance of some associations is due to chance or to shortcomings of the studies and should be deemed false positive results).

Acceptance of the first account is associated with many unanswered questions and *ad hoc* explanations for how the data should be interpreted as supporting it. This account requires that one dismiss the many alternative explanations for the few positive associations observed in the epidemiology studies, despite the plausibility of these explanations. Exposure measurement error from the use of unreliable exposure metrics and the presence of a wide variety of confounding factors that were not adequately adjusted for are important considerations when evaluating the epidemiology data for chlorpyrifos. In addition, this account requires that, although there were many statistical analyses in the epidemiology studies, one chooses to focus only on the few statistically significant findings, regardless of their clinical significance or lack of confirmation in other studies. This account also

requires that one accepts the existence of an exposure-response relationship, despite the lack of consistently observed exposure-response relationship within and among the epidemiology and animal studies. It requires that the human studies that failed to show an increase in risk for neurodevelopmental effects did so for plausible reasons, such that the lack of effects in these studies does not contradict the asserted neurodevelopmental toxicity of chlorpyrifos.

This account also requires the inclusion of an explanation for why the animal studies failed to show consistent neurodevelopmental effects in the absence of AChE inhibition. That is, what is being argued to be happening in humans must for some reason not be happening in experimental animals, or that a sufficient number of studies with concentrations low enough to not perturb AChE activity have not been conducted, such that further research in this area is needed. It requires that one rely heavily on the truth of the mechanistic hypotheses that permit a biologically plausible MoA in the absence of AChE inhibition in the brain, even though effects at doses below the threshold for AChE inhibition were only observed in a few *in vitro* studies and have not been validated *in vivo*. These mechanisms are *ad hoc* rather than *a priori*, making it necessary to find independent, positive evidence of their operation in humans.

To accept this account as true, one must accept that somehow chlorpyrifos can enter the brain of the fetus or child after exposure to doses in the range of background levels in the general population and induce neurotoxicity in the developing brain *via* mechanisms that are independent of AChE inhibition, even though there is a large body of evidence that does not support this *in vivo*. The proposed mechanisms appear to have been chosen to fit the low-dose hypothesis already put forth in the epidemiology studies, and they are not based purely on an evaluation of the WoE as a whole. Because of this, the alternative account should be considered.

The alternative, competing account is that the few apparent associations from epidemiology studies do not indicate a causal connection and there is not adequate support for a biologically plausible mechanism for neurodevelopmental effects of chlorpyrifos in the absence of AChE inhibition in the brain. This account is supported by the totality of the data, which provides plausibility for the few associations observed in the epidemiology studies to be deemed false positive results. This account requires that one accept that the animal data indicate that a few neurodevelopmental effects are observed in the presence of some degree of AChE inhibition in the brain, but not necessarily at the level required for systemic toxicity, as numerous animal studies reported neurodevelopmental effects only in the presence of AChE inhibition or at concentrations shown to be associated with AChE inhibition in other studies. If this account is true, a causal relationship between chlorpyrifos exposure and neurodevelopmental effects in the absence of AChE inhibition in the brain would be understood as not plausible for humans, and the few positive associations observed in epidemiology studies would be attributed to alternative explanations.

When assessing the weight of the available evidence in support of the competing accounts, it is clear that the first account requires far more *ad hoc* assumptions and is not adequately supported by the data as a whole. Because of this, the WoE for this account is weak in comparison to the more substantial WoE supporting the lack of a causal association at chlorpyrifos doses below the threshold for inhibition of AChE activity in the brain.

4. US EPA, ECETOC, and HBWoE frameworks

US EPA and ECETOC have proposed frameworks specifically as guidance for weighing evidence in the context of evaluating potential human disease causation (US EPA, 2010; ECETOC, 2009). Below, we provide perspective on our approach compared with these frameworks, by describing and evaluating the US EPA and ECETOC frameworks and contrasting their rationales with that of the HBWoE approach.

4.1. US EPA framework for incorporating human epidemiology and incident data in health risk assessment

US EPA's OPP has proposed a Draft "Framework for Incorporating Epidemiologic and Incident Data in Health Risk Assessment" (US EPA, 2010). The framework is designed to incorporate epidemiology and human incident data into human health risk assessments specifically for pesticides and to be consistent with the NRC report of 21st Century Toxicity Testing (NRC, 2007). The NRC report promotes shifting toxicity testing away from apical toxicity endpoints to toxicity pathways (cellular response pathways) to inform potential adverse effects in humans and, ultimately, risk decision making. The US EPA framework proposes to use the Bradford Hill Criteria as modified in the MoA framework (US EPA, 1999, 2005; Sonich-Mullin *et al.*, 2001; Meek *et al.*, 2003; Seed *et al.*, 2005).

The US EPA framework first describes strengths and weaknesses of different types of epidemiology studies (*e.g.*, case-control, cohort, longitudinal, cross-sectional, ecologic) and important factors to consider when evaluating epidemiology data (*e.g.*, exposure assessment, confounding factors, statistical analysis, potential bias in observational research, interpretation of null studies, external validity). The framework describes the benefits and uses of epidemiology data, stating, "Epidemiology studies have the potential to help inform multiple components of the risk assessment in a variety of ways. High quality studies with robust exposure assessment may be used to estimate risk quantitatively. However, often due to resource constraints, most epidemiology studies suffer some limitations in size, scope, exposure assessment, or

data analysis which prevent their use in quantitative risk assessment (Caulderon [sic], 2000)." The framework describes how human studies are expected to play a significant role in the new vision of toxicity testing in the 21st century (NRC, 2007) in that human chemical exposure information can help guide *in vitro* and *in vivo* studies that are focused on investigating toxicity pathway *vs*. apical effect dose-response endpoints. Further, the framework describes how potential sources of uncertainty in animal data can be informed by human studies, emphasizing species extrapolation and population variability and "real-world exposures" *vs*. high-dose animal studies. US EPA notes that, while epidemiology studies can pose a challenge in interpretation, the evaluation of multiple routes and multiple chemical exposures may be very informative.

The US EPA framework next describes the utility of human incident data, including case reports and surveillance studies of acute pesticide poisoning incidents. These studies are often high dose exposures of short-term duration (frequently one-time exposures), and effects are often reversible. The framework indicates that human incident studies are often valuable because they can identify novel health effects potentially associated with a specific chemical (*e.g.*, pesticide) exposure, and can be compared to effects from acute animal studies. The US EPA framework describes how human incident data are used broadly to evaluate trends over time and patterns of severity and frequency of pesticide exposure, and further to inform risk assessment and risk management decisions. The framework describes the strengths and weaknesses of sources of human incident data: OPP Incident Data System (IDS); American Association of Poison Control Centers (PCC); National Pesticide Information Center (NPIC); Sentinel Event Notification System for Occupational Risk (SENSOR); California Pesticide Illness Surveillance Program (PISP).

The US EPA framework then describes how to conduct a WoE evaluation of the epidemiology and human incident data so that the conclusions are based on all of the available data rather than on any one study, and introduces the idea that multiple lines of evidence should be considered in addition to the epidemiology and human incident data (*i.e.*, *in vitro*, *in vivo*, and *in silico* data). The framework describes the specific steps involved in the WoE analysis, including, as a first step, a written review of each epidemiology and human incident study that describes the study design, results, conclusions, strengths and weaknesses, and the quality of the exposure assessment. The second step of the US EPA framework relies on the modified Bradford Hill Criteria as used in the MoA framework (US EPA, 1999, 2005; Sonich-Mullin *et al.*, 2001; Meek *et al.*, 2003; Seed *et al.*, 2005) which includes the following steps for organizing and integrating information: postulated MoA; key events; dose-response relationships; temporal association; strength, consistency, and specificity of association of key events and the toxic effect; biological plausibility and coherence; other potential MoAs. The framework briefly describes each step and the important factors to consider in the MoA analysis so that areas of uncertainty and areas of future research may be identified. The final step of the US EPA framework is the "Overall conclusions, statement of areas of confidence and uncertainty, and recommendation for risk assessment." The framework briefly states that this section should discuss the overall conclusions based on the WoE; that is, identify new areas of research, provide recommendations for source data for regulatory values, and extrapolate from animals to humans, if necessary.

The US EPA framework lays out the necessary elements that are important to consider for determining the strength and limitations of each epidemiologic or human incident data set, and suggests a written summary of each study. Although the framework provides a small discussion regarding how the overall conclusions and WoE should be presented, this discussion simply states that this should be done, but provides no guidance on how one should actually weigh the evidence. The framework is put forth as being designed to incorporate epidemiology and human incident data specifically into human health risk assessments for pesticides, but there is no discussion about how to actually *integrate* the epidemiologic data into a risk assessment or how to weigh this evidence with other types of data (*e.g.*, animal, pharmacokinetic, exposure, and MoA studies). The framework generally discusses how potential sources of uncertainty in animal data can be informed by human studies, and states that the framework is designed to include multiple lines of evidence (*i.e.*, epidemiology, toxicology, exposure, pharmacokinetic, and MoA data), but provides no real framework for how one should systematically consider all of the evidence.

As shown in Table 1, the US EPA framework provides guidance on key WoE aspects 1 and 2, as described above in Section 2, with a focus on epidemiology data. That is, the framework focuses mostly on the intrinsic quality of individual epidemiology studies and how to evaluate the body of epidemiology data across studies, but provides little guidance on how to integrate these data with animal studies or MoA data. Therefore, the framework is not really a framework *per se* for *integrating* epidemiology data with other data, but more a conceptual guidance on how to *evaluate* epidemiology studies. Although guidance on human data evaluation and interpretation is necessary, it is only part of what is required for integrating these data with other relevant data in evaluating human health risk and disease causation.

4.2. ECETOC framework for integration of human and animal data in chemical risk assessment

ECETOC has proposed a Framework for Integration of Human and Animal Data in Chemical Risk Assessments (ECETOC, 2009). The framework addresses quality aspects of both animal and human data, strongly encourages the use of both types of data in a combined approach, and suggests that human and animal data ideally should be "complimentary and should confirm each other (*i.e.*,

both indicate excess risk, or both indicate the absence of risk)." The framework indicates that when there are apparent contradictions, efforts should be made to better understand the biological basis of the contradictory evidence which will often further inform the risk assessment process.

Overall, the ECETOC framework involves three steps, which are discussed in more detail in the next sections:

1. Assessment of collective WoE of human data with regard to quality;
2. Assessment of collective WoE of animal data with regard to quality and relevance to humans;
3. Integration of the available evidence.

4.2.1. Human data quality in the ECETOC framework

The ECETOC framework provides an overview of different types of human data, including observational epidemiology (cross-sectional, cohort, and case-control) and controlled experimental studies, describing the quality aspects of human data (study design, exposure information, health outcome data, and other data quality issues similar to those discussed above in the US EPA framework), and how to evaluate the strengths and limitations of a single human study. The framework acknowledges that all data quality requirements are rarely met in epidemiology studies, which complicates their interpretation. The framework describes the criteria for evaluating the quality of human and animal data, the relevance of the animal data in evaluating human risk, and the integration of the two data sets.

The ECETOC framework describes Human Data Quality Criteria (based on the Bradford Hill Criteria), emphasizing that the criteria should be applied to human data in an integrated fashion and should generally consider the stage of the risk assessment. The framework provides techniques that can be used to arrive at a judgment about causal interpretations of each risk assessment stage. That is, the framework describes how Bradford Hill guidelines or meta-analyses can be used for hazard identification; how comparisons of LOAELs and NOAELs of different effects can be used to determine the critical or lead effect; and how different exposure levels should be considered to determine the appropriate dose-response curve to arrive at a LOAEL or NOAEL. The ECETOC framework emphasizes that determining whether an effect is critical depends on "the severity of the effect, its reversibility or whether it is deemed to be 'adverse.' "

The ECETOC framework provides a scheme for scoring the quality of human data. The framework first describes two prerequisites for human data that must be satisfied: (1) exposures must have occurred, and (2) the health effect should be determined adequately. If these criteria are not met, then the study is considered to provide "no information" and is not considered further in the evaluation. According to the framework, if a study meets these prerequisites, then the intrinsic quality of the human data are assessed as "High," "Good," "Compromised," "Poor," or "No Information." The framework provides fairly prescriptive guidance (checklists) on how to assess the quality of the study in terms of these five categories. Once the quality has been determined, the next step in the framework is to determine the nature of the health effect in the study (*i.e.*, chronic, sub-chronic, or acute, and whether the effect is specific or non-specific). The basis of the scoring considers that an effect that is chronic requires a stronger data set than an acute effect, with the order being (from less to more data required): acute specific effect < acute non-specific effect < sub-chronic or chronic specific effect < sub-chronic or chronic non-specific effect. Based on the combined nature and intrinsic quality of a study, the framework assigns a human data quality score of A, B, C, D, or X (*e.g.*, a high quality study with an acute specific effect would have a score of "A," and a poor quality study with a non-specific sub-chronic or chronic effect would have a score of "D," and various combinations would fall in between). The framework discusses how adjustments to the scoring scheme are possible, noting that identical considerations should go into evaluating both positive and null data, but that there are some exceptions (*e.g.*, size of population and confidence intervals of null studies) that should be considered that could change the scoring for null studies. The ECETOC framework states that, "A small positive study may be all that is needed to unequivocally suggest causality or, perhaps potency (*e.g.*, consider a lethal concentration of a substance); in contrast, the absence of effects usually requires larger population sizes." We address this point further in Section 4.2.5. The framework discusses that the stage in the risk assessment for which the data are applicable should affect the scoring because each stage has different requirements (*e.g.*, observation of an effect is important for hazard identification, but at this point in the assessment, exposure level is not important because dose-response is not being assessed).

4.2.2. Animal data quality in the ECETOC framework

The ECETOC framework describes quality aspects of the animal data, emphasizing that animal data need to be integrated with the human data, where good human and animal data are available. The three criteria ECETOC uses to describe the quality of animal data in human risk assessment are reliability, relevance, and adequacy, based on the criteria put forth by Klimisch *et al.* (1997) to determine whether data are sufficient or if more studies are needed in the context of the Organization for Economic Co-operation and Development (OECD) Existing Chemicals Program that is intended to ensure sufficient quality data for high production chemicals.

1. **Reliability** refers to the quality of the test, and takes into account whether standardized methodologies are used. The framework applies the four reliability

categories proposed by Klimisch *et al*. (1997) to determine reliability of the animal studies.

2. **Relevance** is defined as the extent to which data and tests are appropriate for a particular hazard, and how closely toxicity in a test species predicts toxicity in humans. Studies that may lack relevance are atypical species, *in vitro* studies with no *in vivo* confirmation of effect, or exposure routes that may not be relevant. The framework provides general guidelines for determining relevance.
3. **Adequacy** is defined as the usefulness of the data for hazard and/or risk assessment purposes. The framework discusses the importance of considering statistical significance, the types of effects observed (*i.e.*, adaptive *vs*. adverse; reversibility and severity as discussed by Lewis *et al*., (2002)), whether effects could be due to chance, and consideration of concordance in deciding whether a study is adequate or not. The framework provides general guidelines for determining adequacy and guidance on what action to take when the data are discordant.

In reference to adequacy of an animal study, the ECETOC framework states that if there is discordance between animal species then the study of higher quality should be used, and if the studies are of equal quality, then the worst-case study should be used. The framework further states with regard to this point that, "If the discordance between animal studies cannot be rationalized in terms of MoA and the animal studies are both Klimisch category 1 or 2 (considered reliable without or with restrictions, respectively), then the worst-case data should be used in the assessment, while also taking quality aspects into account." We discuss this point further below.

4.2.3. Relevance of animal data in human risk assessment

The ECETOC framework next describes the relevance of the body of animal data to human risk assessment. The steps involved in interpreting the body of animal studies in the context of human risk, within the framework, are as follows: forming a MoA hypothesis; dosimetry; relevance to humans; and dose-route extrapolation.

The framework briefly discusses that the MoA hypothesis should be based on considering possible key events reviewed against the modified Bradford Hill Criteria (Seed *et al*., 2005; Boobis *et al*., 2006) to determine whether the WoE is sufficient to derive a MoA. The framework further describes how sometimes a MoA hypothesis is not possible based on a lack of data or fundamental understanding of the biology, and sometimes more than one plausible MoA is possible, but as discussed below, provides little guidance on what to do in these cases. The dosimetry step is briefly described within the framework as identification of toxicant, target organ, dose-response, and temporal aspects of dose-response, and that animal data and kinetic modeling techniques should be applied to animal data to analyze dosimetry.

The framework provides a scheme for determining relevance of animal data to human risk, based on considering each MoA key event in animals and its plausibility in humans (Boobis *et al*., 2008; Seed *et al*., 2005; Boobis *et al*., 2006). The framework recommends an extended version of the Human Relevance Framework developed by the International Programme on Chemical Safety (IPCS), including the following scoring criteria for data quality and availability:

A. Reliable animal data directly relevant to humans (may also include null findings with confidence that they are applicable to humans)
B. Reliable animal data relevant to humans (less confident in how reliable null findings apply to humans)
C. Reliable animal data, with uncertain but assumed relevance to humans
D. Unreliable animal data
X. No relevant animal data, or data not relevant to humans

A flow diagram is provided in the ECETOC framework that incorporates reliability and relevance (in the context of MoA as outlined in the IPCS framework) into evaluating human relevance of animal data, with descriptions of how to determine the various categories A–X. The framework then describes dose-route extrapolation as the next step when the route of human exposure being considered is different from the critical animal study, briefly describing how dosimetry considerations and exposure modeling can assist in this step.

4.2.4. Integration of animal and human data

Finally, the ECETOC framework applies a matrix for placing the body of human and animal data into categories A-D so that it is clear what data (human or animal) are more relevant for the risk assessment. That is, the quality of the human data and the quality and relevance of the animal data are considered together and scored, for example, as A/A, A/B, C/C, *etc.* As such, the outcome will indicate which data should take precedence in the human risk assessment. The framework discusses that when the human and animal data have equivalent scores, the data needing fewer adjustments should be used, and that is typically the human data. Further, when the scoring is equivalent but the data are not concordant, the data suggesting a hazard should generally take precedence. If both suggest a hazard, the one with the lower level should take precedence, considering an upper bound for the other data source. The framework indicates that when the human data are scored as "A," then these data should take precedence regardless of animal data. We discuss this point further below.

ECETOC discusses that it is not possible to construct a matrix that can easily be applied to all situations, and that it is acceptable to deviate from the procedures outlined in the framework as long as they can be scientifically justified. The framework further indicates that if both human

and animal data are category "C" (*i.e.*, poor or compromised studies in humans, and reliable animal data but with uncertain relevance to humans), hazard and risk assessments should proceed with caution, particularly if the data are discordant.

The ECETOC framework provides 15 case study examples so that application and implementation of the framework can be more clearly understood.

4.2.5. Practical application of WoE in the ECETOC framework

The ECETOC framework defines the various categories of the animal data and human data and provides some guidance (mostly by example in case studies within which there are brief descriptions of categorization) about how to weigh all of the animal or human evidence in order to assign a category. The flowchart (Figure 3 of the ECETOC framework) and table (Table 1 of the ECETOC framework) of the framework provide some general guidance for categorization. The framework discusses the importance of human relevance when interpreting the animal data, and the category choice relies on what level of relevance is determined, but the framework could provide more guidance (even if by example) on exactly how to weigh all of the animal evidence so that one can assign the appropriate category of relevance to humans. The case studies include very brief descriptions of each step, providing little discussion of the actual process of weighing the evidence.

The scoring for determining human data quality within the framework consists of a checklist of requirements for each category (High, Good, Compromised, Poor, or No Information), with a prescribed number of requirements that need to be met for a study to fall into a given category. To determine whether each requirement is met, however, involves some judgment on the part of the risk assessor because the requirements are very qualitative. "Quality" is judged on intrinsic properties of the study conduct relative to prevailing standards for studies of that type, rather than on how study strengths and shortcomings affect the application of results to the causality questions at hand. Therefore, although some structure is useful, determining data quality should also consist of a narrative discussion of the logic for how the quality was determined. Further, the ECETOC framework discusses that there are some exceptions that should be considered that could change the scoring for null studies that may be different from positive studies (*e.g.*, size of population and confidence intervals of null studies). Identical considerations should go into evaluating both positive and null data, however, and the determination of study quality should be made based on methodology and not study results.

One question presented in the framework flowchart regarding relevance of animal data to humans (Figure 3 of the ECETOC framework) is whether the MoA is established in animals, with a yes or no answer guiding how to categorize the data. The ECETOC framework provides little guidance on how to actually weigh the MoA evidence in animals in order to determine whether an MoA has been established. The answer may not be a simple yes or no; if MoAs are suggested but not yet "established," it is not clear how the unproven but perhaps informative possibilities affect the interpretation of available results. The framework suggests that if the MoA has not been established in animals that it should be assumed that it is relevant to humans and given a data category of "C." There should be an option, however, to modify that categorization based on MoA data that may or may not be relevant to humans. That is, the animal MoA may not be definitively established, but there may be enough data to hypothesize several MoAs, and if so, it would be important to consider how plausible those proposed MoAs are in humans. This would involve considering animal, human, pharmacokinetic, exposure, and MoA data together, and allowing the data sets to inform one another in weighing all of the evidence to determine plausibility. This is often the case for the available data (*i.e.*, proposed MoAs that are being tested for a given causal question, as opposed to definitive MoAs that are already established).

Further, the framework discusses, with regard to discordance within the animal data, that "it should be noted that a lack of concordance between sexes or species, or even between strains of the same species, could provide invaluable information about the mode of action (MoA) of the substance.…If the discordance between animal studies cannot be rationalized in terms of MoA and the animal studies are both Klimisch category 1 or 2 [considered reliable], *then the worst-case data should be used in the assessment*, while also taking quality aspects into account" [emphasis added]. The first part of this statement is true, in that lack of concordance between animal data could provide useful information with regard to mode of action in animals and humans, and this point should be emphasized more within the WoE for the framework. The framework does not elaborate on this point, however, but instead indicates that the worst-case data should be used. One should consider that, in a case where there is uncertainty in the MoA because of discordance in animals, particularly if the studies are reliable, this uncertainty implies that relevance to humans is uncertain and proceeding with the risk assessment may not be appropriate. These data should perhaps suggest more studies as the recommended next step (rather than assuming worst-case) to determine what is causing the lack of concordance. Understanding the lack of concordance in animals would provide useful information for potential human relevance. Further, although not mentioned in the ECETOC framework, different animal species may be more or less relevant to humans (*e.g.*, monkey data *vs*. rodent data) and this should be considered in the context of the question of the particular human disease causation (Gray *et al.*, 2005).

The ECETOC framework provides a methodology for integrating the animal and human effect data based on the scoring of each body of data (animal or human), with the outcome being whether precedence should be

given to human or animal data for the final risk assessment. The framework, however, does not explicitly consider how all of the data together (*i.e.*, in weighing all of the evidence), both negative and positive, can help inform interpretation of one another (animal, human, exposure, and MoA studies). Instead, it emphasizes choosing one basis for inference (the "best" according to the scoring scheme), taking that study's results as most indicative of the true situation, and not dealing with how discrepancies between this single basis and the remaining body of evidence is to be accounted for. Although it does discuss forming a MoA hypothesis and considering dosimetry as part of determining whether the animal MoA data are relevant to humans (based on the IPCS human relevance framework), it does not explicitly consider what to do when there is more than one plausible MoA, or the importance of how these MoA data may inform interpretation of the epidemiology or animal toxicology data.

The framework discusses, with regard to integrating the animal and human data, that when the scoring is equivalent for humans and animals, but the data are not concordant, the data suggesting a hazard should generally take precedence (apart from a category of A for human data, which should generally be given precedence regardless of whether an effect was observed), and if both suggest a hazard, the one with the lowest level should take precedence, considering an upper bound for the other data source. In this case, it would seem more appropriate to carefully analyze all of the data (animal, human, and MoA) to see how all of the data inform interpretation of each other, so that the logic can be clearly traced with regard to the choice of the dose-response information to use in the assessment. For example, in some cases it may seem more logical to accept category B human data with a less sensitive effect than the more sensitive category B animal data because there may be MoA data (in animals and/or humans) that suggest the proposed animal MoA is not plausible in humans. If nothing else, such a case should suggest further study to determine the plausibility of the animal MoA in humans.

Although intended to be flexible, the ECETOC framework perhaps provides too many steps and checklists so that each piece of the puzzle is dealt with separately, and perhaps eliminated before other data are considered that might have informed interpretation of that particular piece of data. Parts of the framework could be very useful, however, perhaps as tools applied to a more holistic approach to weighing all of the evidence. As noted in the conclusions of the ECETOC framework: "any attempt to systematize reporting, conduct, or classification of data is likely to be criticized. The objection is well-founded; if a classification framework is too rigid it can stifle creativity and if it is too lax, it may only provide the veneer of an evaluation. The Task Force believes that the primary benefit of the proposed ECETOC framework will be an evolving improvement towards the transparent evaluation and integration of human and animal data in the risk assessment process."

Overall, the ECETOC framework provides a very useful first step in ranking human and animal data quality and relevance of animal data to humans. The framework, however, would benefit from more discussion of: (i) importance of carefully weighing all the animal and human data, and guidance on how one should go about this, so that the appropriate data quality category can be determined; (ii) going beyond the categorization of animal and human effect data to integrate other important data sets (exposure, MoA, and pharmacokinetic); and (iii) the importance of considering all of the data together and how all of the data can inform each other (both negative and positive data, and of varying quality) so that the evidence as a whole can be truly weighed. As presented in Table 1, the ECETOC framework provides guidance for key aspects 1 and 2 of a WoE evaluation, and provides some guidance on key aspect 5 with regard to integrating human and animal data.

4.3. Comparison of HBWoE, US EPA, and ECETOC frameworks

Although the methodology varies, the three WoE frameworks described here each include a systematic review of the quality of the individual studies relevant to the question of human disease causation, and each examine the data within a particular line of investigation (*i.e.*, epidemiology, animal toxicology, or MoA studies) for particular endpoints, evaluating consistency, specificity, and reproducibility of outcomes (described in steps 1 and 2 in Table 1). One difference between the frameworks is that the US EPA framework focuses on epidemiology data, the ECETOC framework focuses on epidemiology and animal toxicology data, and the HBWoE framework evaluates all relevant data (*i.e.*, epidemiology, animal toxicology, and MoA data) individually and then within each line of investigation. Within these steps, the frameworks vary in the degree to which the steps are explicitly described, with the US EPA framework providing little prescriptive guidance and the ECETOC framework perhaps providing too much. We think the HBWoE framework falls somewhere in between, incorporating the key aspects of each, with more flexibility than the ECETOC framework, but a bit more guidance than the US EPA framework. Because our framework is intended to be flexible, various aspects of the US EPA and ECETOC frameworks could be applied within the HBWoE framework for the first two steps of the evaluation.

As summarized in Table 1, the US EPA and ECETOC frameworks provide guidance for only the first two key aspects of a WoE evaluation. Although the US EPA and ECETOC frameworks discuss the importance of integrating all of the relevant data, there is little guidance on how to actually do that. The HBWoE framework goes beyond these steps in that it integrates all of the relevant data, within the context of proposed hypotheses, so that each line of evidence can inform interpretation of one

another (steps 3–5 in Table 1). The integration includes, but goes well beyond, simply noting the patterns and degrees of concordance and discordance, among studies within a realm (*i.e.*, across human studies or in different sexes and species of animals); it takes the stance that what makes data relevant to human hazard identification is the hypothesized commonality that the source data and the target human population have in how the agent acts to provoke observed effects in the source and, presumably, similar or at least mechanistically related effects in the target population. Under this view, discordant results are possible but require their own (at least tentative) explanations for why the causal process proposed to be common to the key studies and the target human population do not apply to the discordant studies, and the WoE for human hazard is judged by the success and biological plausibility of the set of explanations of concordant as well as discordant results. That is, the HBWoE framework compares the various accounts of the observations at hand, discussing consistencies and inconsistencies within the data and the *ad hoc* assumptions required to support each account, and tracing the logic and reasoning for how the data support (or do not support) each account's hypotheses (step 6). In this way, the HBWoE framework does not seek to prove or disprove any one hypothesis, nor to definitively choose one and reject the others; rather, it seeks to present the lines of reasoning for each account of the observations so that the data will speak for themselves.

As part of comparing various accounts, the HBWoE framework will often require tracing the logic and reasoning for how a poor quality study is used to support a particular line of argument (or hypothesis). In contrast to the ECETOC framework that explicitly describes how poor quality studies should be eliminated early on in the evaluation, a key aspect of the HBWoE framework is that all data, positive and negative, and of varying quality (even poor quality) are maintained and carried through the evaluation. Poor quality studies may have some useful information, and it should not be taken as self-evident that their results are false; rather, such outcomes should have lesser weight. In HBWoE, this lesser weight arises naturally from the consideration of the comparative inability of poor studies (compared to more robust ones) to provide outcomes that differentiate between the generally operating causal factors being evaluated and extraneous, study-specific explanations that could produce spurious outcomes. In a HBWoE evaluation, the logic for how a poor quality study fits (or does not fit) with the available data needs to be considered and articulated as part of one account of the observations at hand, so that it can be compared to other accounts of the available data.

In tracing the logic and reasoning for how certain studies or lines of evidence fit (or do not fit) with the available data, the HBWoE framework necessitates inclusion of all data relevant to the various hypotheses that have been put forth. By contrast, checklists do not work in weighing all of the evidence if they lead one to make certain assumptions about a given study or data set without consideration of all of the data when enough features of the array of criteria seem to fit. In fact, the criteria developed by Bradford Hill (which he called "postulates") were designed to articulate the basis for judgments and facilitate the integration of evaluations across criteria, and were not intended simply as checklists from which causality could be concluded. Hill saw the postulates as guides to thinking rather than as measures of evidence. The HBWoE framework emphasizes the importance of how each piece of information (positive or negative) might inform interpretation of one another, or how studies of varying quality (even poor quality) need to be considered insofar as they have bearing on distinguishing between alternative explanations, particularly if the study is the basis for a particular line of argument that needs to be articulated as part of one account, and in this way integrates all data relevant to questions of potential human disease causation.

In comparison to the other frameworks, there are aspects of more traditional WoE approaches that the HBWoE framework does not have. The outcome of the HBWoE framework is complex and not easily summarized succinctly. The HBWoE framework does not arrive at decisions, but it is used to inform decision makers by characterizing uncertainty and plausibility of alternative conclusions. The HBWoE framework is not readily codified, so can be quite complicated in practice, requiring deep and broad expertise. Although judgments will still be needed, and these judgments will instill scientific debate, the debate can be more clearly focused on the scientific bases of the various lines of argument.

4.4. Comparison of frameworks in the context of chlorpyrifos

Applying the HBWoE framework to evaluate the health effects of chlorpyrifos, we concluded that the most likely account of the epidemiology, animal toxicology, and mechanistic data is that the few apparent associations from epidemiology studies do not indicate a causal connection and there is not adequate support for a biologically plausible mechanism for neurodevelopmental effects in the absence of AChE inhibition in the brain.

The US EPA and ECETOC frameworks provide guidance for reviewing studies systematically and examining the consistency, specificity, and relevance of outcomes across studies. They do not, however, provide guidance on how to integrate all of the relevant data or how to use each line of evidence to inform the integration of other kinds of data. For example, these frameworks do not provide information regarding how the results of chlorpyrifos toxicology and mechanistic studies should inform the interpretation of the epidemiology studies. In this case, the toxicology and mechanistic data

indicate a lack of effects at exposures below those which cause AChE inhibition, casting doubt on the validity of positive associations in epidemiology studies. The US EPA and ECETOC frameworks also do not compare the different accounts. That is, while they aim to determine how well the data support a specific hypothesis, they do not consider how well the data support alternative hypotheses, nor do they explicitly address the question of why discordant results exist and how these should be accounted for. By using the HBWoE framework, it is clear that the WoE supports the account of no causation much more than the account of causation, and that more *ad hoc* assumptions are required to support the causation account. This is not as evident using the other frameworks.

Both the US EPA and ECETOC frameworks focus on the "best" studies. That is, while they discuss reviewing all the studies, they both come down to focusing on a specific set of data, sometimes ignoring other data. For example, US EPA might choose the most robust epidemiology study, but not consider whether results from this study are consistent with those of less robust studies, or whether this study is robust enough to draw conclusions. Similarly, ECETOC states that poor quality studies should be eliminated early on. In a HBWoE evaluation, the logic for how a poor quality study fits (or does not fit) with the available data is considered and articulated as part of one account of the observations at hand; poor studies are poor because they fail to discriminate between the causal hypothesis being evaluated and other extraneous explanations of their outcomes, so they do not help to differentiate the relative plausibility of competing accounts. Thus, using either of these frameworks would have resulted in some or several of the chlorpyrifos studies being ignored, and their role in each account would have been missed.

5. Conclusions

As regulatory agencies make greater use of human data in chemical risk assessments, it will be a challenge to determine how to assess all of the data that are relevant to the question of human disease causation. We compared three frameworks that have been proposed to guide risk assessors in this endeavor and assessed how well each framework incorporates key aspects of WoE.

While the three WoE frameworks each include a systematic review of the quality of the individual studies and examine the data within a particular line of evidence (*i.e.*, epidemiology, animal toxicology, or MoA studies), the US EPA and ECETOC frameworks provide little guidance for integrating all of the relevant data. By contrast, the HBWoE framework integrates all of the relevant data within the context of proposed hypotheses, so that each line of evidence can inform the interpretation of one another. Further, the HBWoE framework compares the various accounts of the observations at hand, discussing consistencies and inconsistencies within the data and the *ad hoc* assumptions required to support each account, tracing the logic and reasoning for how the data support (or do not support) each account's hypotheses. The HBWoE framework emphasizes the importance of how each piece of information (positive or negative) might inform interpretation of one another, and how studies of varying quality (even poor quality) need to be considered, characterizing the uncertainty and plausibility of alternative conclusions while integrating all of the data relevant to potential human disease causation questions.

In our application of the HBWoE framework to evaluate the data relevant to examining whether there is a causal association between exposure to chlorpyrifos and adverse neurodevelopmental effects in humans, we found that the epidemiology data are not sufficiently robust to support the hypothesis that chlorpyrifos is a causal factor for neurodevelopmental effects. The available studies do not report consistent results, and there is a lack of clear exposure-response information. Because of the many uncertainties in these studies, the few positive results may also support alternative explanations that other factors are causal for the reported outcomes, or that the observed associations are statistical anomalies. In addition, the animal toxicity data do not provide clear evidence that chlorpyrifos is associated with neurodevelopmental effects at doses that are below the threshold for inhibition of AChE in the brain; this would be relevant to exposures in the epidemiology studies, which are at least 1000-fold lower than those used in the animal studies. Further, the mechanisms proposed to underlie potential neurodevelopmental effects in humans at doses below those associated with inhibition of AChE activity in the brain have not been shown to operate in the developing brain *in vivo* at concentrations that do not inhibit the activity of AChE.

For chlorpyrifos to act as a neurodevelopmental toxicant at the near-background exposure levels in the epidemiology studies, it must be accepted that chlorpyrifos can enter the brain of the fetus or child after exposure to doses in the range of background levels in the general population and induce neurotoxicity in the developing brain *via* mechanisms that are independent of AChE inhibition, even though there is a large body of evidence that does not support this *in vivo*. Rather, the few apparent associations from epidemiology studies are not indicative of a causal connection, and there is not adequate support for a biologically plausible mechanism for neurodevelopmental effects of chlorpyrifos in the absence of AChE inhibition in the brain. The weight of the available evidence more strongly indicates that a causal association between chlorpyrifos exposure and neurodevelopmental effects in the absence of AChE inhibition in the brain is not plausible for humans, and the few positive associations observed in epidemiology studies would be attributed to alternative explanations.

Declaration of interest

This paper was prepared with financial support to Gradient, a private environmental consulting firm, by Dow AgroSciences. The work reported in this paper was conducted during the normal course of employment. The authors have the sole responsibility for the writing and contents of this paper.

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.

Achenbach T, Rescorla L. (2000). Manual for the ASEBA Preschool Forms and Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Achenbach TM, Dumenci L, Rescorla LA. (2003). DSM-oriented and empirically based approaches to constructing scales from the same item pools. J Clin Child Adolesc Psychol, 32:328–340.

Adigun AA, Ryde IT, Seidler FJ, Slotkin TA. (2010). Organophosphate exposure during a critical developmental stage reprograms adenylyl cyclase signaling in PC12 cells. Brain Res, 1329:36–44.

Agency for Toxic Substances and Disease Registry (ATSDR). (1997). Toxicological profile for chlorpyrifos (Update - Draft for public comment). Research Triangle Park, NC.

Ahdaya SM, Monroe RJ, Guthrie FE. (1981). Absorption and distribution of intubated insecticides in fasted mice. Pestic Biochem Physiol 16:38–46.

Aiuto LA, Pavlakis SG, Boxer RA. (1993). Life-threatening organophosphate-induced delayed polyneuropathy in a child after accidental chlorpyrifos ingestion. J Pediatr, 122:658–660.

Albers JW, Cole P, Greenberg RS, Mandel JS, Monson RR, Ross JH, Snodgrass WR, Spurgeon A, van Gemert M. (1999). Analysis of chlorpyrifos exposure and human health: Expert panel report. J Toxicol Environ Health B Crit Rev, 2:301–324.

Albers JW, Berent S, Garabrant DH, Giordani B, Schweitzer SJ, Garrison RP, Richardson RJ. (2004a). The effects of occupational exposure to chlorpyrifos on the neurologic examination of central nervous system function: A prospective cohort study. J Occup Environ Med, 46:367–378.

Albers JW, Garabrant DH, Schweitzer SJ, Garrison RP, Richardson RJ, Berent S. (2004b). The effects of occupational exposure to chlorpyrifos on the peripheral nervous system: A prospective cohort study. Occup Environ Med, 61:201–211.

Albers JW, Garabrant DH, Schweitzer S, Garrison RP, Richardson RJ, Berent S. (2004c). Absence of sensory neuropathy among workers with occupational exposure to chlorpyrifos. Muscle Nerve, 29:677–686.

Albers JW, Garabrant DH, Mattsson JL, Burns CJ, Cohen SS, Sima C, Garrison RP, Richardson RJ, Berent S. (2007). Dose-effect analyses of occupational chlorpyrifos exposure and peripheral nerve electrophysiology. Toxicol Sci, 97:196–204.

Aldridge JE, Levin ED, Seidler FJ, Slotkin TA. (2005a). Developmental exposure of rats to chlorpyrifos leads to behavioral alterations in adulthood, involving serotonergic mechanisms and resembling animal models of depression. Environ Health Perspect, 113:527–531.

Aldridge JE, Meyer A, Seidler FJ, Slotkin TA. (2005b). 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 JE, Seidler FJ, Slotkin TA. (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, 112:148–155.

American Psychiatric Association (APA). (2000). Diagnostic and Statistical Manual of Mental Disorders, 4th Edition. Washington, DC: American Psychiatric Association.

Aprea C, Sciarra G, Sartorelli P, Desideri E, Amati R, Sartorelli E. (1994). Biological monitoring of exposure to organophosphorus insecticides by assay of urinary alkylphosphates: Influence of protective measures during manual operations with treated plants. Int Arch Occup Environ Health, 66:333–338.

Bakke JE, Feil VJ, Price CE. (1976). Rat urinary metabolites from O,O-diethyl-O-(3,5,6-trichloro-2-pyridyl) phosphorothioate. J Environ Sci Health B, 11:225–230.

Barr DB, Angerer J. (2006). Potential uses of biomonitoring data: A case study using the organophosphorus pesticides chlorpyrifos and malathion. Environ Health Perspect, 114:1763–1769.

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

Barr DB, Barr JR, Maggio VL, Whitehead RD Jr, Sadowski MA, Whyatt RM, Needham LL. (2002). A multi-analyte method for the quantification of contemporary pesticides in human serum and plasma using high-resolution mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci, 778:99–111.

Bellinger DC. (2005). Neurobehavioral assessment in studies of exposures to neurotoxicants. Int Rev Res Ment Retard 30:263–300.

Berger-Sweeney J, Hohmann CF. (1997). Behavioral consequences of abnormal cortical development: Insights into developmental disabilities. Behav Brain Res, 86:121–142.

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, 112:388–391.

Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D, Farland W. (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. Crit Rev Toxicol, 36:781–792.

Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J, Vickers C. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. Crit Rev Toxicol, 38:87–96.

Bradman A, Whitaker D, Quirós L, Castorina R, Claus Henn B, Nishioka M, Morgan J, Barr DB, Harnly M, Brisbin JA, Sheldon LS, McKone TE, Eskenazi B. (2007). Pesticides and their metabolites in the homes and urine of farmworker children living in the Salinas Valley, CA. J Expo Sci Environ Epidemiol, 17:331–349.

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

Brazelton TB, Nugent JK. (1995). Neonatal Behavioral Assessment Scale (Third Edition). London, United Kingdom: Mac Keith Press.

Calderon RL. (2000). Measuring risks in humans: The promise and practice of epidemiology. Food Chem Toxicol, 38:S59–S63.

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, McKone TE, Barr DB, Harnly ME, Eskenazi B. (2003). Cumulative organophosphate pesticide exposure and risk assessment among pregnant women living in an agricultural community: A case study from the CHAMACOS cohort. Environ Health Perspect, 111:1640–1648.

Centers for Disease Control and Prevention (CDC). (2009). Fourth National Report on Human Exposure to Environmental Chemicals. (Online) Available at: http://www.cdc.gov/exposurereport/pdf/FourthReport.pdf. Accessed on 11 December 2009.

Chakrabarti TK, Farrar JD, Pope CN. (1993). Comparative neurochemical and neurobehavioral effects of repeated chlorpyrifos exposures in young and adult rats. Pharmacol Biochem Behav, 46:219–224.

Chambers JE, Chambers HW. (1989). Oxidative desulfuration of chlorpyrifos, chlorpyrifos-methyl, and leptophos by rat brain and liver. J Biochem Toxicol, 4:201–203.

Chanda SM, Mortensen SR, Moser VC, Padilla S. (1997). Tissue-specific effects of chlorpyrifos on carboxylesterase and cholinesterase activity in adult rats: An *in vitro* and *in vivo* comparison. Fundam Appl Toxicol, 38:148–157.

Chanda SM, Pope CN. (1996). Neurochemical and neurobehavioral effects of repeated gestational exposure to chlorpyrifos in maternal and developing rats. Pharmacol Biochem Behav, 53:771–776.

Clegg DJ, van Gemert M. (1999). Determination of the reference dose for chlorpyrifos: Proceedings of an expert panel. J Toxicol Environ Health B Crit Rev, 2:211–255.

Cole TB, Jampsa RL, Walter BJ, Arndt TL, Richter RJ, Shih DM, Tward A, Lusis AJ, Jack RM, Costa LG, Furlong CE. (2003). Expression of human paraoxonase (PON1) during development. Pharmacogenetics, 13:357–364.

Coulston F, Griffin T, Golberg L. (1972). Safety Evaluation of Dowco 179 in Human Volunteers. Albany, NY: Institute of Experimental Pathology and Toxicology, Albany Medical College.

Crumpton TL, Seidler FJ, Slotkin TA. (2000). Is oxidative stress involved in the developmental neurotoxicity of chlorpyrifos? Brain Res Dev Brain Res, 121:189–195.

Dam K, Garcia SJ, Seidler FJ, Slotkin TA. (1999). Neonatal chlorpyrifos exposure alters synaptic development and neuronal activity in cholinergic and catecholaminergic pathways. Brain Res Dev Brain Res, 116:9–20.

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 KP, 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, 160:217–230.

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

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. Am J Epidemiol, 165:1397–1404.

Engel SM, Wetmur K, 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. doi:10.1289/ehp.1003183.

Eskenazi B, Harley K, Bradman A, Weltzien E, Jewell NP, Barr DB, Furlong CE, Holland NT. (2004). Association of in utero organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population. Environ Health Perspect, 112:1116–1124.

Eskenazi B, Marks AR, Bradman A, Harley K, Barr DB, Johnson C, Morga N, Jewell NP. (2007). Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children. Environ Health Perspect, 115:792–798.

European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC). (2009). Framework of the Integration of Human and Animal Data in Chemical Risk Assessment. ECETOC Technical Report No. 104. p. 130.

Eyer P. (2003). The role of oximes in the management of organophosphorus pesticide poisoning. Toxicol Rev, 22:165–190.

Goodman JE, Dodge DG, Bailey LA. (2010). A framework for assessing causality and adverse effects in humans with a case study of sulfur dioxide. Regul Toxicol Pharmacol, 58:308–322.

Hack M, Breslau N, Weissman B, Aram D, Klein N, Borawski E. (1991). Effect of very low birth weight and subnormal head size on cognitive abilities at school age. N Engl J Med, 325:231–237.

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.

Hill AB. (1965). The environment and disease: Association or causation? Proc R Soc Med, 58:295–300.

Howard AS, Bucelli R, Jett DA, Bruun D, Yang D, Lein PJ. (2005). Chlorpyrifos exerts opposing effects on axonal and dendritic growth in primary neuronal cultures. Toxicol Appl Pharmacol, 207:112–124.

Huff RA, Corcoran JJ, Anderson JK, Abou-Donia MB. (1994). Chlorpyrifos oxon binds directly to muscarinic receptors and inhibits cAMP accumulation in rat striatum. J Pharmacol Exp Ther, 269:329–335.

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.

Ikonomidou C, Kaindl AM. (2011). Neuronal death and oxidative stress in the developing brain. Antioxid Redox Signal, 14:1535–1550.

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.

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.

Karanth S, Pope C. (2000). Carboxylesterase and A-esterase activities during maturation and aging: Relationship to the toxicity of chlorpyrifos and parathion in rats. Toxicol Sci, 58:282–289.

Kisicki JC, Seip CW, Combs ML. (1999). A Rising Dose Toxicology Study to Determine the No-Observable-Effect-Levels (NOEL) for Erythrocyte Acetylcholinesterase (AChE) Inhibition and Cholinergic Signs and Symptoms of Chlorpyrifos at Three Dose Levels. Report to Dow AgroSciences LLC. p. 40.

Klimisch HJ, Andreae M, Tillmann U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharmacol, 25:1–5.

Krimsky S. (2005). The weight of scientific evidence in policy and law. Am J Public Health, 95 Suppl 1:S129–S136.

Landrigan PJ, Claudio L, Markowitz SB, Berkowitz GS, Brenner BL, Romero H, Wetmur JG, Matte TD, Gore AC, Godbold JH, Wolff MS. (1999). Pesticides and inner-city children: Exposures, risks, and prevention. Environ Health Perspect, 107 Suppl 3:431–437.

Lasky RE, Klein RE, Yarbrough C, Engle PL, Lechtig A, Martorell R. (1981). The relationship between physical growth and infant behavioral development in rural Guatemala. Child Dev, 52:219–226.

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.

Lemons JA, Schreiner RL, Gresham EL. (1981). Relationship of brain weight to head circumference in early infancy. Hum Biol, 53:351–354.

Lester BM, Als H, Brazelton TB. (1982). Regional obstetric anesthesia and newborn behavior: A reanalysis toward synergistic effects. Child Dev, 53:687–692.

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.

Lewis RW, Billington R, Debryune E, Gamer A, Lang B, Carpanini F. (2002). Recognition of adverse and nonadverse effects in toxicity studies. Toxicol Pathol, 30:66–74.

Li WF, Matthews C, Disteche CM, Costa LG, Furlong CE. (1997). Paraoxonase (PON1) gene in mice: Sequencing, chromosomal

localization and developmental expression. Pharmacogenetics, 7:137–144.

Linkov I, Loney D, Cormier S, Satterstrom FK, Bridges T. (2009). Weight-of-evidence evaluation in environmental assessment: Review of qualitative and quantitative approaches. Sci Total Environ, 407:5199–5205.

Lotti M, Moretto A, Zoppellari R, Dainese R, Rizzuto N, Barusco G. (1986). Inhibition of lymphocytic neuropathy target esterase predicts the development of organophosphate-induced delayed polyneuropathy. Arch Toxicol, 59:176–179.

MacIntosh DL, Needham LL, Hammerstrom KA, Ryan PB. (1999). A longitudinal investigation of selected pesticide metabolites in urine. J Expo Anal Environ Epidemiol, 9:494–501.

MacLusky NJ, Naftolin F. (1981). Sexual differentiation of the central nervous system. Science, 211:1294–1302.

Marty MS, Domoradzki JY, Hansen SC, Timchalk C, Bartels MJ, Mattsson JL. (2007). The effect of route, vehicle, and divided doses on the pharmacokinetics of chlorpyrifos and its metabolite trichloropyridinol in neonatal Sprague-Dawley rats. Toxicol Sci, 100:360–373.

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.

McMullin TS, Lowe ER, Bartels MJ, Marty MS. (2008). Dynamic changes in lipids and proteins of maternal, fetal, and pup blood and milk during perinatal development in CD and Wistar rats. Toxicol Sci, 105:260–274.

Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J, Patton DE. (2003). A framework for human relevance analysis of information on carcinogenic modes of action. Crit Rev Toxicol, 33:591–653.

Meyer A, Seidler FJ, Aldridge JE, Tate CA, Cousins MM, Slotkin TA. (2004). Critical periods for chlorpyrifos-induced developmental neurotoxicity: Alterations in adenylyl cyclase signaling in adult rat brain regions after gestational or neonatal exposure. Environ Health Perspect, 112:295–301.

Massoulié J, Sussman J, Bon S, Silman I. (1993). Structure and functions of acetylcholinesterase and butyrylcholinesterase. Prog Brain Res, 98:139–146.

Morgan MK, Sheldon LS, Croghan CW, Jones PA, Robertson GL, Chuang JC, Wilson NK, Lyu CW. (2005). Exposures of preschool children to chlorpyrifos and its degradation product 3,5,6-trichloro-2-pyridinol in their everyday environments. J Expo Anal Environ Epidemiol, 15:297–309.

Moretto A, Lotti M. (1998). Poisoning by organophosphorus insecticides and sensory neuropathy. J Neurol Neurosurg Psychiatr, 64:463–468.

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

Mueller RF, Hornung S, Furlong CE, Anderson J, Giblett ER, Motulsky AG. (1983). Plasma paraoxonase polymorphism: A new enzyme assay, population, family, biochemical, and linkage studies. Am J Hum Genet, 35:393–408.

Muto MA, Lobelle F Jr, Bidanset JH, Wurpel JN. (1992). Embryotoxicity and neurotoxicity in rats associated with prenatal exposure to DURSBAN. Vet Hum Toxicol, 34:498–501.

Namba T, Nolte CT, Jackrel J, Grob D. (1971). Poisoning due to organophosphate insecticides. Acute and chronic manifestations. Am J Med, 50:475–492.

National Research Council (NRC) of the National Academies. (2007). Toxicity Testing in the Twenty-first Century: A Vision and a Strategy. Washington, DC: National Academies Press.

Needham LL. (2005). Assessing exposure to organophosphorus pesticides by biomonitoring in epidemiologic studies of birth outcomes. Environ Health Perspect, 113:494–498.

Nolan RJ, Rick DL, Freshour NL, Saunders JH. (1984). Chlorpyrifos: Pharmacokinetics in human volunteers. Toxicol Appl Pharmacol, 73:8–15.

Ounsted M, Moar VA, Scott A. (1988). Head circumference and developmental ability at the age of seven years. Acta Paediatr Scand, 77:374–379.

Palmer JS, Rowe LD, Crookshank HR. (1980). Effect of age on tolerance of calves to chlorpyrifos. Am J Vet Res, 41:1323–1325.

Perera FP, Rauh V, Tsai WY, Kinney P, Camann D, Barr D, Bernert T, Garfinkel R, Tu YH, Diaz D, Dietrich J, Whyatt RM. (2003). Effects of transplacental exposure to environmental pollutants on birth outcomes in a multiethnic population. Environ Health Perspect, 111:201–205.

Pope CN, Chakraborti TK. (1992). Dose-related inhibition of brain and plasma cholinesterase in neonatal and adult rats following sublethal organophosphate exposures. Toxicology, 73:35–43.

Pope CN, Chakraborti TK, Chapman ML, Farrar JD, Arthun D. (1991). Comparison of *in vivo* cholinesterase inhibition in neonatal and adult rats by three organophosphorothioate insecticides. Toxicology, 68:51–61.

Qiao D, Seidler FJ, Padilla S, Slotkin TA. (2002). Developmental neurotoxicity of chlorpyrifos: What is the vulnerable period? Environ Health Perspect, 110:1097–1103.

Qiao D, Seidler FJ, Abreu-Villaça Y, Tate CA, Cousins MM, Slotkin TA. (2004). Chlorpyrifos exposure during neurulation: Cholinergic synaptic dysfunction and cellular alterations in brain regions at adolescence and adulthood. Brain Res Dev Brain Res, 148:43–52.

Qiao D, Seidler FJ, Slotkin TA. (2005). Oxidative mechanisms contributing to the developmental neurotoxicity of nicotine and chlorpyrifos. Toxicol Appl Pharmacol, 206:17–26.

Qiao D, Seidler FJ, Tate CA, Cousins MM, Slotkin TA. (2003). Fetal chlorpyrifos exposure: Adverse effects on brain cell development and cholinergic biomarkers emerge postnatally and continue into adolescence and adulthood. Environ Health Perspect, 111:536–544.

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, 119:1196–1201.

Rauh VA, Garfinkel R, Perera FP, Andrews HF, Hoepner L, Barr DB, Whitehead R, Tang D, Whyatt RW. (2006). Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. Pediatrics, 118:e1845–e1859.

Rauh VA, Whyatt RM, Garfinkel R, Andrews H, Hoepner L, Reyes A, Diaz D, Camann D, Perera FP. (2004). Developmental effects of exposure to environmental tobacco smoke and material hardship among inner-city children. Neurotoxicol Teratol, 26:373–385.

Rhomberg LR, Bailey LA, Goodman JE, Hamade AK, Mayfield D. (2011). Is exposure to formaldehyde in air causally associated with leukemia? – A hypothesis-based weight-of-evidence analysis. Crit Rev Toxicol (in press).

Rhomberg LR, Bailey LA, Goodman JE. (2010). Hypothesis-based weight of evidence: A tool for evaluating and communicating uncertainties and inconsistencies in the large body of evidence in proposing a carcinogenic mode of action–naphthalene as an example. Crit Rev Toxicol, 40:671–696.

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.

Rice DC. (2005). From animals to humans: Models and constructs. Int Rev Res Ment Retard 30:301–337.

Richardson RJ. (1995). Assessment of the neurotoxic potential of chlorpyrifos relative to other organophosphorus compounds: A critical review of the literature. J Toxicol Environ Health, 44:135–165.

Rushton JP, Ankney CD. (2009). Whole brain size and general mental ability: A review. Int J Neurosci, 119:691–731.

Saulsbury MD, Heyliger SO, Wang K, Johnson DJ. (2009). Chlorpyrifos induces oxidative stress in oligodendrocyte progenitor cells. Toxicology, 259:1–9.

Schoenemann PT, Budinger TF, Sarich VM, Wang WS. (2000). Brain size does not predict general cognitive ability within families. Proc Natl Acad Sci USA, 97:4932–4937.

Schuh RA, Lein PJ, Beckles RA, Jett DA. (2002). Noncholinesterase mechanisms of chlorpyrifos neurotoxicity: Altered phosphorylation of Ca2+/cAMP response element binding protein in cultured neurons. Toxicol Appl Pharmacol, 182:176–185.

Seed J, Carney EW, Corley RA, Crofton KM, DeSesso JM, Foster PM, Kavlock R, Kimmel G, Klaunig J, Meek ME, Preston RJ, Slikker W Jr, Tabacova S, Williams GM, Wiltse J, Zoeller RT, Fenner-Crisp P, Patton DE. (2005). Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. Crit Rev Toxicol, 35:664–672.

Shah PV, Fisher HL, Sumler MR, Monroe RJ, Chernoff N, Hall LL. (1987). Comparison of the penetration of 14 pesticides through the skin of young and adult rats. J Toxicol Environ Health, 21:353–366.

Sharbaugh C, Viet SM, Fraser A, McMaster SB. (2003). Comparable measures of cognitive function in human infants and laboratory animals to identify environmental health risks to children. Environ Health Perspect, 111:1630–1639.

Slotkin TA. (2004). Cholinergic systems in brain development and disruption by neurotoxicants: Nicotine, environmental tobacco smoke, organophosphates. Toxicol Appl Pharmacol, 198:132–151.

Slotkin TA, MacKillop EA, Ryde IT, Seidler FJ. (2007). Ameliorating the developmental neurotoxicity of chlorpyrifos: A mechanisms-based approach in PC12 cells. Environ Health Perspect, 115:1306–1313.

Slotkin TA, Oliver CA, Seidler FJ. (2005). Critical periods for the role of oxidative stress in the developmental neurotoxicity of chlorpyrifos and terbutaline, alone or in combination. Brain Res Dev Brain Res, 157:172–180.

Slotkin TA, Seidler FJ. (2007). Prenatal chlorpyrifos exposure elicits presynaptic serotonergic and dopaminergic hyperactivity at adolescence: Critical periods for regional and sex-selective effects. Reprod Toxicol, 23:421–427.

Slotkin TA, Seidler FJ. (2009). Oxidative and excitatory mechanisms of developmental neurotoxicity: Transcriptional profiles for chlorpyrifos, diazinon, dieldrin, and divalent nickel in PC12 cells. Environ Health Perspect, 117:587–596.

Slotkin TA, Tate CA, Cousins MM, Seidler FJ. (2002). Functional alterations in CNS catecholamine systems in adolescence and adulthood after neonatal chlorpyrifos exposure. Brain Res Dev Brain Res, 133:163–173.

Smith GN, Watson BS, Fischer FS. (1967). Investigations on dursban insecticide: Metabolism of [36Cl] O,O-diethyl-O- 3,5,6-trichloro-2-pyridyl phosphorothioate in rats. J Agric Food Chem 15:132–138.

Song X, Seidler FJ, Saleh JL, Zhang J, Padilla S, Slotkin TA. (1997). Cellular mechanisms for developmental toxicity of chlorpyrifos: Targeting the adenylyl cyclase signaling cascade. Toxicol Appl Pharmacol, 145:158–174.

Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, Grant D, Hartley M, Knaap A, Kroese D, Mangelsdorf I, Meek E, Rice JM, Younes M; International Programme on Chemical Safety. (2001). IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. Regul Toxicol Pharmacol, 34:146–152.

Sood B, Delaney-Black V, Covington C, Nordstrom-Klee B, Ager J, Templin T, Janisse J, Martier S, Sokol RJ. (2001). Prenatal alcohol exposure and childhood behavior at age 6 to 7 years: I. dose-response effect. Pediatrics, 108:E34.

Sultatos LG, Murphy SD. (1983). Kinetic analyses of the microsomal biotransformation of the phosphorothioate insecticides chlorpyrifos and parathion. Fundam Appl Toxicol, 3:16–21.

Sultatos LG, Murphy SD. (1983b). Hepatic microsomal detoxification of the organophosphates paraoxon and chlorpyrifos oxon in the mouse. Drug Metab Dispos, 11:232–238.

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

Ulbrich B, Palmer AK. (1996). Neurobehavioral aspects of developmental toxicity testing. Environ Health Perspect, 104 Suppl 2:407–412.

US EPA. (1999). Guidelines for carcinogen risk assessment (SAB Review Draft). Risk Assessment Forum, NCEA-F-0644. [Online] Available at: www.epa.gov/ncea/raf/crasab.htm. Accessed on 27 June 2001.

US EPA. (2002). Interim reregistration eligibility decision for chlorpyrifos. EPA 738-R-01-007. [Online] Available at: http://www.epa.gov/oppsrrd1/REDs/chlorpyrifos_ired.pdf.

US EPA. (2005). Guidelines for Carcinogen Risk Assessment. Washington, DC: Risk Assessment Forum. EPA/630/P-03/001F. p. 166.

US EPA. (2010). Framework for Incorporating Human Epidemiologic & Incident Data in Health Risk Assessment (Draft). Office of Pesticide Programs. p. 68.

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.

Venerosi A, Ricceri L, Scattoni ML, Calamandrei G. (2009). Prenatal chlorpyrifos exposure alters motor behavior and ultrasonic vocalization in CD-1 mouse pups. Environ Health, 8:12.

Webb SJ, Monk CS, Nelson CA. (2001). Mechanisms of postnatal neurobiological development: Implications for human development. Dev Neuropsychol, 19:147–171.

Weed DL. (2005). Weight of evidence: A review of concept and methods. Risk Anal, 25:1545–1557.

Wessels D, Barr DB, Mendola P. (2003). Use of biomarkers to indicate exposure of children to organophosphate pesticides: Implications for a longitudinal study of children's environmental health. Environ Health Perspect, 111:1939–1946.

Whitney KD, Seidler FJ, Slotkin TA. (1995). Developmental neurotoxicity of chlorpyrifos: Cellular mechanisms. Toxicol Appl Pharmacol, 134:53–62.

Whyatt RM, Camann DE, Kinney PL, Reyes A, Ramirez J, Dietrich J, Diaz D, Holmes D, Perera FP. (2002). Residential pesticide use during pregnancy among a cohort of urban minority women. Environ Health Perspect, 110:507–514.

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 chlorpyrifos exposure within an inner-city cohort during pregnancy. Environ Health Perspect, 117:559–567.

Whyatt RM, Rauh V, Barr DB, Camann DE, Andrews HF, Garfinkel R, Hoepner LA, Diaz D, Dietrich J, Reyes A, Tang D, Kinney PL, Perera FP. (2004). Prenatal insecticide exposures and birth weight and length among an urban minority cohort. Environ Health Perspect, 112:1125–1132.

Winneke G. (1992). Cross species extrapolation in neurotoxicology: Neurophysiological and neurobehavioral aspects. Neurotoxicology, 13:15–25.

Wolff MS, Engel S, Berkowitz G, Teitelbaum S, Siskind J, Barr DB, Wetmur J. (2007). Prenatal pesticide and PCB exposures and birth outcomes. Pediatr Res, 61:243–250.

Yan X, Lashley S, Smulian JC, Ananth CV, Barr DB, Ledoux TA, Hore P, Robson MG. (2009). Pesticide concentrations in matrices collected in the perinatal period in a population of pregnant women and newborns in New Jersey, USA. Hum Ecol Risk Assess 15:948–967.

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

Young JG, Eskenazi B, Gladstone EA, Bradman A, Pedersen L, Johnson C, Barr DB, Furlong CE, Holland NT. (2005). Association between

in utero organophosphate pesticide exposure and abnormal reflexes in neonates. Neurotoxicology, 26:199–209.

Youngstrom E, LaKind JS, Kenworthy L, Lipkin PH, Goodman M, Squibb K, Mattison DR, Anthony BJ, Anthony LG. (2010). Advancing the selection of neurodevelopmental measures in epidemiological studies of environmental chemical exposure and health effects. Int J Environ Res Public Health, 7:229–268.