DRAFT

Framework for Incorporating Human Epidemiologic & Incident Data in Health Risk Assessment

January 7, 2010

Office of Pesticide Programs
US Environmental Protection Agency

Table of Contents

List of Tables

List of Figures

I. INTRODUCTION

EPA's Office of Pesticide Programs (OPP) is a licensing program regulating pesticide products in the U.S. As part of this, OPP evaluates the effects of pesticides on human health and the environment. Through the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and the Federal Food, Drug, and Cosmetic Act (FFDCA), OPP receives extensive hazard and exposure information to characterize the risks of pesticide products. Information on the toxic effects of pesticides is generally derived from studies with laboratory animals. In the past, information from well designed epidemiology studies on pesticides has not been typically available to inform evaluations of potential risks. This is changing, in large part due to increased availability of studies from the NIEHS/EPA Children's Centers (http://epa.gov/ncer/childrenscenters) and the Agricultural Health Study (http://aghealth.nci.nih.gov/). In addition, the National Children's Study (http://www.nationalchildrensstudy.gov) will, in time, provide valuable information on the effects of environmental chemicals on children. OPP intends to employ these prospective epidemiology studies, along with other sources of human effects information such as other kinds of epidemiology studies and human incident data, in its human health risk assessment. Consistent with Administrator Lisa Jackson's commitment to transparency and scientific integrity[1], OPP's goal is to use such information in the most scientifically robust and transparent way. To accomplish this, OPP is proposing a framework to describe the scientific considerations that EPA will weigh in evaluating how such studies and scientific information can be integrated into risk assessments of pesticide chemicals. This draft framework along with the draft case studies (Attachments A-C) will be reviewed by the FIFRA Scientific Advisory Panel (SAP) and will receive public comment in February, 2010. Subsequently, OPP will evaluate the comments from the Panel and public and make the appropriate revisions to the framework.

Two recent reports by the National Research Council (NRC) of the National Academy of Science (NAS), "Toxicity Testing in the 21st Century: A Vision and A Strategy (2007)" and "Science and Decisions (2009)", together provide new directions in toxicology and risk assessment. These two NRC reports advocate far reaching changes in how toxicity testing is performed, how such data are interpreted, and ultimately how regulatory decisions are made. Specifically, the 2007 report on 21st century toxicity testing advocates a shift away from the current focus of using apical toxicity endpoints to using toxicity pathways[2] to inform toxicity testing, risk assessment, and ultimately decision making. This bold, new approach is based on the rapidly evolving scientific understanding of how genes, proteins, and small molecules interact to form molecular pathways that maintain cell function in human cells. The goal for the new toxicity testing paradigm is to determine how exposure to environmental agents

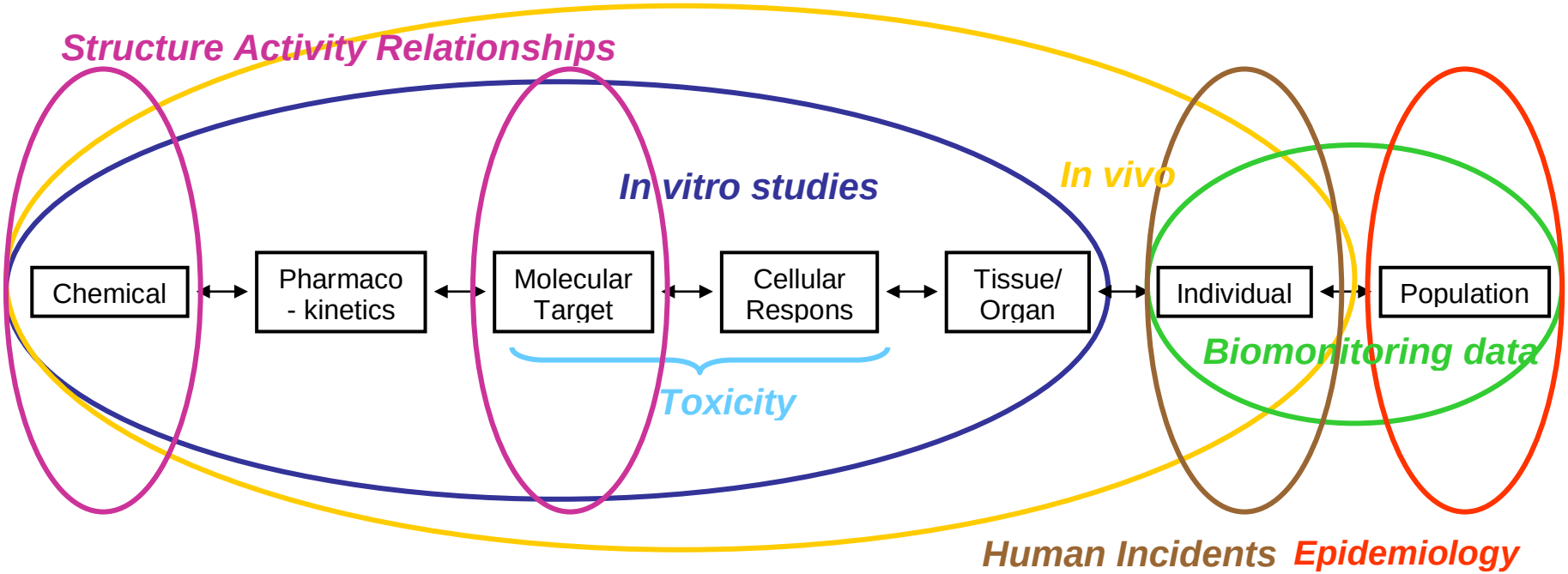
[1] April 23 and May 9, 2009 Memoranda to EPA Staff from Administrator Lisa Jackson

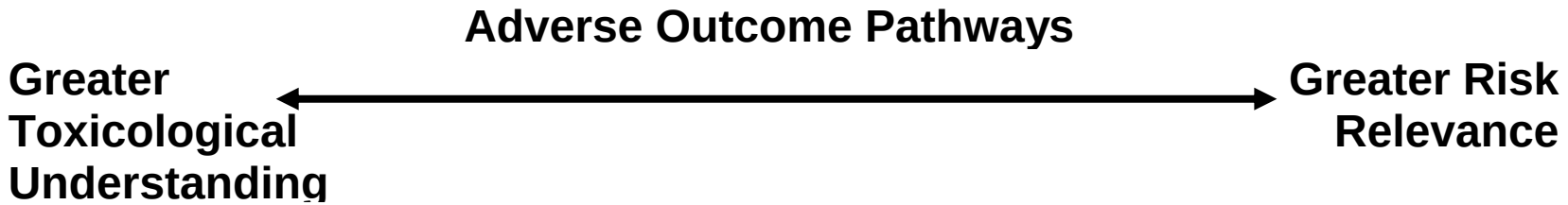
[2] Toxicity pathways are cellular response pathways that, when sufficiently perturbed, are expected to result in adverse health effects.

can perturb these pathways, thereby causing a cascade of subsequent key events leading to adverse health effects. Human information like that found in epidemiology studies, human incident databases, and biomonitoring studies, along with experimental toxicological information are expected to play a significant role in this new approach. Specifically, these types of human information provide insight into the effects caused by actual chemical exposures in humans and thus can contribute to problem formulation and hazard/risk characterization. In addition, epidemiologic and human incident data can guide additional analyses or data generations (e.g., dose and endpoint selection for use in *in vitro* and targeted *in vivo* experimental studies), identify potentially susceptible populations, identify new health effects or confirm the existing toxicological observations.

This new vision of toxicity testing and risk assessment will involve data from multiple levels of biological organization ranging from the molecular level up to population-based surveillance with a goal of considering chemical effects from their source to the ultimate health outcome and effects on populations. Figure 1 suggests how different types of information relate to each other. Such data will come from *in vitro* and *in vivo* experimental studies along with *in silico* and modeled data. OPP's proposed framework for incorporating epidemiology and incident data is conceptually consistent with the 2007 NRC report on 21st century toxicity testing in that both emphasize the use of the best available information from multiple data sources are compiled in a weight of the evidence (WOE) analysis. The approach described in this draft framework is desirable for on-going evaluations on a number of pesticide risk assessments. As the Agency moves forward in implementing the transformative approach in the 2007 and 2008 NRC reports, OPP will re-evaluate and update this draft framework as appropriate.

Figure 1. Source to Outcome Pathway: Chemical effects across levels of biological organization (adopted from NRC, 2007)





OPP's approach proposed in this draft framework relies on existing guidance documents and frameworks (Table 1) as the starting point for reviewing and evaluating epidemiology and human incident data for use in risk assessment. In brief, OPP's draft framework proposes to use the Bradford Hill Criteria as modified in the Mode of Action[3] (MOA) Framework as an organizational tool for describing and reviewing data from animals and humans. The International Programme for Chemical Safety (IPSC), International Life Sciences Institute (ILSI), and EPA introduced a MOA framework. This MOA framework begins with identifying the series of key events that are along the causal path, that are established on weight of evidence, using criteria based on those described by Bradford Hill, taking into account factors such as dose-response and temporal concordance, biological plausibility, coherence and consistency. Using this analytic approach, epidemiologic and human incident findings can be evaluated in the context of other human information and experimental studies to evaluate consistency, reproducibility, and biological plausibility of reported outcomes and to identify areas of uncertainty and future research.

Table 1. Key guidance documents and frameworks used by OPP

NAS	1983: Risk Assessment in the Federal Government: Managing the Process
	1994: Science and Judgment
	2007: Toxicity Testing in the 21st Century
	2009: Science and Decisions: Advancing Risk Assessment
WHO/IPCS	2001-2007: Mode of Action/Human Relevance Framework
	2005: Chemical Specific Adjustment Factors (CSAF)
EPA	1991-2005: Risk Assessment Forum Guidance for Risk Assessment (e.g., guidelines for carcinogen, reproductive, developmental, neurotoxicity, ecological, and exposure assessment, guidance for benchmark dose modeling, review of reference dose and reference concentration processes) http://epa.gov/raf/pubhumanhealth.htm
	2000: Science Policy Handbook on Risk Characterization http://www.epa.gov/spc/pdfs/rchandbk.pdf
OPP	2001: Aggregate risk assessment http://www.epa.gov/pesticides/trac/science/aggregate.pdf
	2002: Cumulative risk assessment http://www.epa.gov/pesticides/trac/science/
	Food Quality Protection Act 10X Safety Factor: http://www.epa.gov/pesticides/science/policies.htm,

[3] A mode of action is defined by the major steps (which may be involve PK or PD events) leading to an adverse health effect following interaction of a pesticide with biological target;.

Information on the human effects of pesticides can be found from different sources. Some sources include *in vitro* studies with human tissues; observational human studies like those performed for worker monitoring; epidemiology studies; human incident data; population-scale biomonitoring studies (e.g., National Health and Nutrition Examination Survey, NHANES) or human studies involving intentional exposure. The focus of this draft framework is on interpreting and using ***epidemiology*** and ***human incident data*** in human risk assessment; other sources of human information are not addressed in this document in any depth. Specifically, this draft does not extensively discuss research with pesticides involving intentional exposure of human subjects. Both the conduct of such research and EPA's reliance on data from such research are governed by EPA's Rule for the Protection of Human Subjects of Research (40 CFR Part 26.) Among other things, these rules forbid research involving intentional exposure of pregnant or nursing women or of children, require prior review of proposals for new research by EPA and by the EPA Human Studies Review Board (HSRB), and require further review by EPA and the HSRB of reports of completed research. In addition, this document does not discuss observational studies with agricultural workers at length. In the last several years, OPP has extensively evaluated existing worker monitoring studies in efforts to improve the data and approaches used in worker exposure assessment; those evaluations can be found elsewhere (http://www.epa.gov/scipoly/sap/meetings/2007/010907_mtg.htm)..

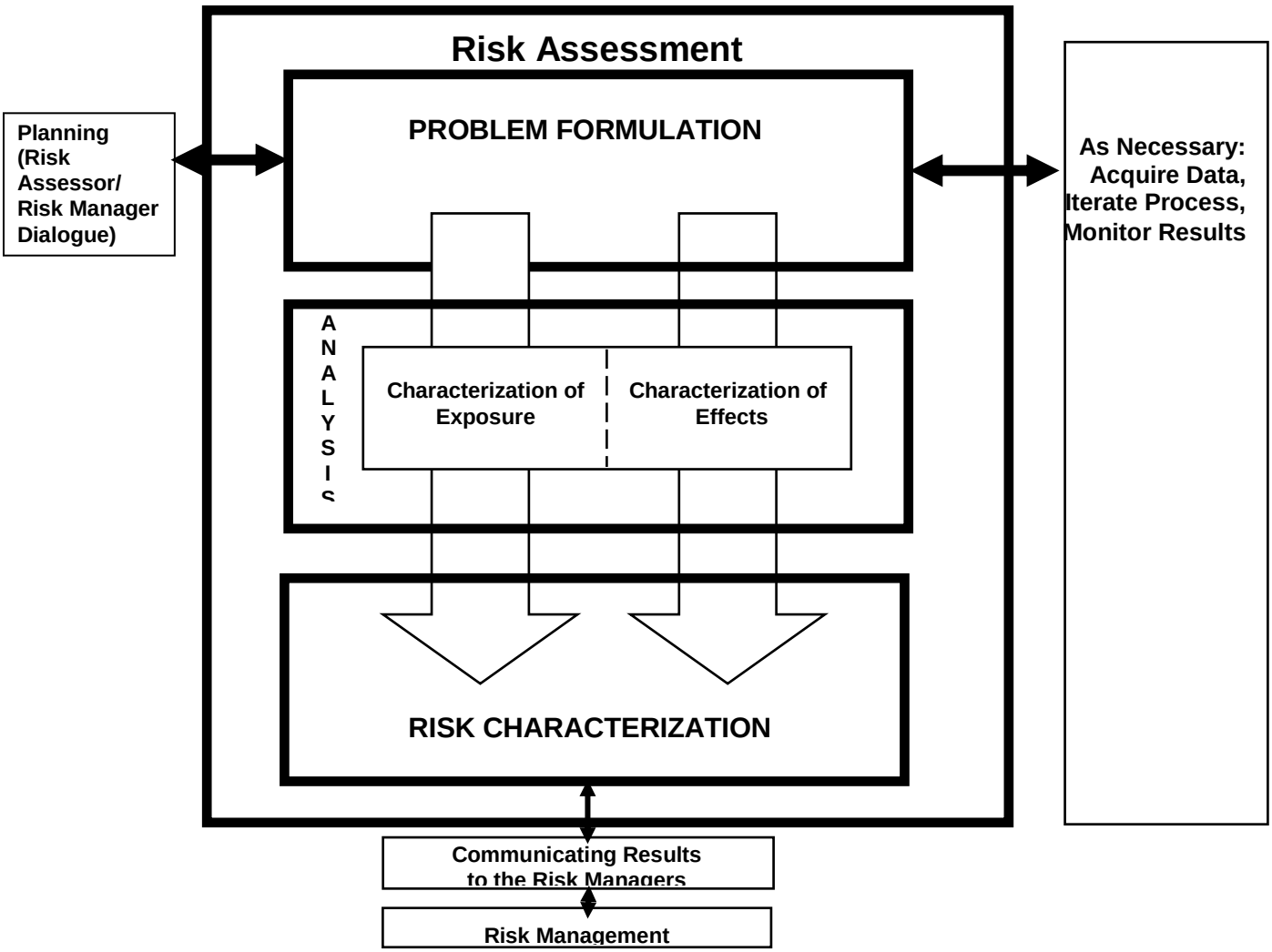
Like the other parts of EPA, OPP follows the NAS paradigm for conducting risk assessment. Risk is a function of both the hazard of a chemical and the levels of exposure to a chemical. See "Science and Decisions: Advancing Risk Assessment" NRC report (2009). In accordance with FIFRA and FFDCA, OPP assesses risk to the general population (including susceptible lifestages and/or many age groups ranging from infants to adults >50) and to agricultural and other workers. The general population can be exposed to pesticides via the oral route in food and water (i.e., from run-off or leaching). The general population can be exposed to pesticides from residential uses and potentially from inhalation or dermal exposure through volatization and/or spray drift (USEPA, 2009). Workers can be exposed during the mixing, loading, or application of a pesticide, as well as a result of entering and performing activities in a pesticide treated site. These exposures typically occur by the dermal and inhalation routes. Accordingly, OPP assesses risk to a multiple routes of exposure and to a variety of durations of exposure ranging from a single day up to chronic durations. Appendix 1 contains a brief summary of the types of risk assessments developed by OPP.

As recommended in "Science and Decisions: Advancing Risk Assessment (NRC, 2009)," OPP uses problem formulation as a tool to identify exposure pathways and potential health outcomes along with the appropriate methods and approaches for the scientific analysis. Before a risk assessment is conducted, the risk assessors and risk managers engage in a dialogue that considers stakeholder concerns and includes discussion of management goals as well as the scope and complexity of the assessment. During the problem formulation stage a plan for analyzing data and characterizing risk is developed (Figure 2). The plan identifies the appropriate

methods/models and data sources. If missing data are critical to the assessment, options are discussed as to how best to obtain that information (e.g., required testing, research). The peer review process is identified and the time line for completing the assessment is defined.

In problem formulation, hazard information from experimental toxicity studies and human studies are evaluated. Key scientific issues related to hazard assessment considered in problem formulation include: What are the effects associated with exposure? What are the modes/mechanisms associated with these effects? What are the temporal aspects of the effects? Are there susceptible populations and if so, who are they and what toxicological factors contribute to susceptibility? Exposure information is also evaluated in problem formulation. Key scientific issues related to exposure assessment considered in problem formulation include: How is the pesticide used? What are all of the relevant use sites of exposure? To what chemical substances will people be exposed? What are the routes, durations, and frequency of exposures? Who may be exposed? Does the exposure pose different risks to different groups (e.g., due age or activity patterns?)

Figure 2. Schematic of problem formulation (adapted from USEPA (1998)



The human health risk assessments of pesticide chemicals follow some key scientific principles for integrating effects and exposure information. First, OPP emphasizes using data on pharmacokinetics (PK) and modes/mechanisms of action to develop health protective and scientifically supportable risk assessments. Specifically, PK properties of a pesticide determine the dose available at the biological compartment and MOA describes the cascade of events leading from early precursors events to final toxic outcomes. PK properties and key events in a MOA provide a starting point for evaluating dose-response assessment, species differences, lifestage susceptibility, route to route differences, and biological plausibility of reported outcomes. Second, realistic and accurate exposure assessments strive for which reflect actual pesticide use/usage. Thus, OPP uses pesticide residue data from monitoring programs whenever available. Since pesticide exposure can vary greatly among different populations, it is important to maintain geographic, temporal and demographic specificity in risk assessments of pesticide chemicals. Furthermore, OPP strives to match the timeframe of the toxicity endpoint(s) to the timeframe of exposure

assessment in order to appropriately assess the temporal characteristics of risk. Lastly, OPP characterizes assumptions and remaining uncertainties used in risk assessment to identify areas where new research could best improve exiting and future risk assessments

II. REVIEWING EPIDEMIOLOGY STUDIES FOR USE IN PESTICIDE RISK ASSESSMENT

Epidemiology is an observational science that seeks to identify and evaluate relationships between exposure to chemical, physical or biological agents, and the health status of populations (Boyes et al, 2007). Calderon (2000) described four major uses of such studies: 1) describe the health status of a population and discover important time trends in disease and exposure frequency; 2) explain the occurrence of diseases by identifying factors that are associated with specific diseases or trends; 3) predict the number of disease occurrences and the distribution of health states in specific populations; and 4) improving the health status of the population by identifying factors that affect environmental or human health. In the case of pesticides, epidemiology focuses on the relation between exposure and adverse health effects in the general population and in specific sub-populations, such as occupationally exposed workers or applicators.

When considering the use of observational epidemiology studies in risk assessment, OPP will consider a number of characteristics of each study such as those provided here (USEPA, 2005). This list is not intended to be a check list. Instead, the degree of confidence one places on a study(ies) depends on how these factors are addressed.

- articulation of study objectives or hypothesis;
- selection and characterization of comparison groups (exposed and unexposed groups or case and control groups);
- characterization of exposure;
- length of follow-up for disease occurrence;
- methods for ascertainment of the causes of disease;
- consideration of bias and confounding/modification factors;
- the adequacy of the sample size to detect an effect;
- methodology for data collection and analysis;
- response rate and methodology for handling missing data; and
- complete and clear documentation of results.
- confirmation of study results by similar studies

A. Types of Epidemiology Studies

The major types of observational epidemiologic studies are described briefly below with consideration of their strengths and weaknesses (Lilienfeld and Lilienfeld, 1979;

Mausner and Kramer, 1985; Kelsey et al., 1996; Rothman and Greenland, 1998; Paddle and Harrington, 2000; USEPA, 2005; Purdue Pesticide Programs, PPP-43).

In **case-control studies**, groups of individuals with (cases) and without (controls) a particular disease are identified and compared with respect to exposure to determine whether those with disease are more likely to be exposed to the agent, i.e., whether there is an association. In case-control studies, first disease status (cases with disease; controls without) is determined and then exposure is measured. Because disease has already occurred at the time of selection into the case-control study, this study design is particularly useful in studying uncommon diseases, or diseases with long latency, and can be utilized to evaluate the relation between many different exposures and a disease outcome. Because case-control studies start with individuals who have the disease, compared to cohort studies, the studies can involve fewer subjects and can be completed in a relatively short time period. Challenges in case-control investigations include the selection of an appropriate control group, and the assessment of exposures which may have occurred long before the disease was diagnosed. Diseased individuals may remember exposures or events differently than those who remain healthy.

Population-based Cohort studies begin with a group of people that share common characteristics—the cohort—and evaluate their health over an extended time period. In cohort studies, a group of "exposed" and "nonexposed" individuals are identified and studied over time to determine differences in disease occurrence. All subjects are followed, over time, and their individual exposures and diseases documented. Disease occurrence is then analyzed to see if the rate of disease differs between the exposed and unexposed groups. In contrast to case-control studies, which are able to evaluate a single disease outcome, cohort studies can be utilized to evaluate multiple disease outcomes. Cohort studies can also be performed either prospectively, like the Agricultural Health Study (AHS, http://aghealth.nci.nih.gov/), or retrospectively from historical records. A prospective cohort design focuses on a group of people from a current point in time through a future point in time. A retrospective cohort design focuses on a group exposed at some point in the past, and compares disease rates after exposure occurred. Prospective cohort studies can be relatively lengthy and expensive to conduct for rare diseases. Specifically, significant resources and professional staff are required for a long period of time to collect high quality data.

Cross-sectional studies examine the relation between exposure and disease using information collected at the same point in time from individuals. Cross-section studies are generally used to identify patterns or trends in disease occurrence over time or in different geographical locations, and can be conducted quickly and relatively inexpensively. An important limitation of cross-sectional studies is they do not allow one to determine whether exposure precedes the disease. As such, cross-section studies are unable to establish temporal relationships between disease and exposure and typically require additional studies to confirm a hypothesized causal association suggested by a cross-sectional study.

Ecologic studies examine exposure and disease patterns using information reflecting group or population-level data. In an ecologic study, the unit of analysis is a group, e.g. geographic region, and not an individual. Using this design, it is not known whether all members of the exposed group are individually exposed, therefore, ecologic studies can suggest research hypotheses for studies and may contribute to problem formulation. Although they cannot in themselves establish a causal association, these studies should be noted in the hazard characterization.

B. Important Scientific Factors to Consider When Evaluating Epidemiologic Studies for Use in Risk Assessment

The following text describes some key scientific factors which are important to consider when evaluating epidemiologic data for use in pesticide risk assessment.

1. Exposure Assessment

Exposure assessment can be defined as the "process of estimating or measuring the magnitude, frequency and duration of exposure to an agent, along with the number and characteristics of the population exposed. Ideally, it describes the sources, pathways, routes, and the uncertainties in the assessment. (Zartarian et al, 2005)." In environmental epidemiology, exposure assessment poses a unique challenge, particularly for toxicants that are found in low concentrations in environmental media (NRC, 1991; NRC, 1997). Given the complexity of exposure pathways, researchers have developed a number of different approaches to assess exposure, which vary in accuracy, precision, and resource requirements (Niewenhuijsen, 2003). Some of these approaches are not specific to epidemiologic research but may be used to inform exposure assessment in a variety of scientific analyses. These approaches include indirect methods, based on historical records, questionnaires, and environmental monitoring, and direct methods, based on personal monitoring and biomonitoring (Table 2). A brief description of each method and its strengths and limitations is summarized below:

Table 2. Summary of indirect and direct exposure assessment methods.

Approach	Method/Tools	Example	Exposure Estimation
Indirect	Historical Records	Estimating proximity to agricultural crops using address information	Dichotomous or ordinal exposure
	Questionnaires	Determine potential for exposure based on pesticide-use responses	Dichotomous or ordinal exposure
	Environmental Monitoring	Measuring pesticide levels in community water drinking system	Dichotomous or ordinal exposure, although exposure can be estimated using modeling
Direct	Personal Monitoring	Measuring pesticide inhalation and dermal contact	Quantified exposure
	Biomonitoring	Measuring pesticide levels in blood and urine	Quantified internal dose

- **Historical records and questionnaires** are used to characterize key characteristics which may be associated with chemical exposure. When used in epidemiologic studies, historical records and questionnaires are not typically used to predict quantitative levels of exposure. Rather, historical record information or questionnaire responses are used to assign categorical levels of exposure. Examples of historical record information that can be used to assign exposure levels includes address in proximity to an agricultural crop and employment history information on job title and history. Similarly, questionnaires can be used to determine if individuals recall using pesticides or identify individuals that perform specific job functions that increase their potential for exposure. While historical records and questionnaires can be cost-effective sources of data on potential exposure, they do have limitations. Data collected from historical records and questionnaires is only a surrogate of exposure. As a result, these data sources may be an oversimplification of exposure and not accurately rank individuals exposure potential.

- **Environmental monitoring** is used to characterize the levels of contaminants in environmental media, including air, water, soil, food, and home and work environments. Many state and Federal programs collect environmental monitoring data that may be useful in epidemiologic studies. Environmental monitoring is particularly useful for exposure that can be defined geographic boundaries, such as air pollution and drinking water. As such, many epidemiologic studies have utilized ambient air monitoring data and community drinking water system data to characterize exposure to air pollution and drinking water contamination, respectively. While environmental monitoring data is useful for estimating exposures defined by geographic boundaries, it can be less reliable for the purposes of assigning individual-levels exposures, particularly when individuals live, work, and spend time in many different locations

- **Personal monitoring** is used to characterize exposure at the point of contact of a body boundary. Examples of personal monitoring include the use dosimeters to assess dermal contact with pesticides, personal air sampling devices to assess inhalation exposure, and collection of duplicate diet samples to determine pesticide levels in food. The advantage of personal monitoring is that it is likely to provide more accurate estimates of individual-level exposure than indirect methods. Personal monitoring also makes it possible to quantify exposure levels that can be useful for prioritizing the relevance of different routes of exposure. Additionally, personal monitoring can also be used to assess longitudinal exposure when repeated measurements are taken over time. While personal monitoring offers many advantages over indirect approaches, it also tends to be labor and resource intensive (Niewenhuijsen, 2003). As a result, it is not typically feasible to conduct large-scale epidemiologic studies that assess exposure using personal monitoring. Furthermore, personal monitoring is highly dependent on the measurement techniques and analytic tools used to obtain samples. As such, it is extremely important to consider the scientific rigor and reliably of personal monitoring methodologies that are used in epidemiologic studies.

- **Biomonitoring** is used to characterize exposure by measuring a chemical, its metabolite(s), or reactive product(s) in biological samples, such as blood, urine, saliva, milk, adipose, and other body tissues (Needham et al, 2007). Assessing exposure using biomonitoring has expanded rapidly as analytical tools have become more cost-effective and more biomarkers are identified. Compared with self-reported questionnaire or interview data, biomonitoring may reduce exposure misclassification and enhance the precision of the risk estimates. Similarly, biomonitoring integrates exposures from different routes and can be used to determine the amount of exposure that is absorbed into the body (Checkoway et al, 2004). Furthermore, in certain instances knowledge as to the role of the biomarker in the natural history of disease is known, such that biomarkers may help resolve temporality of exposure issues.

  While biomonitoring has many advantages over others exposure assessment methods, it also has its own limitations. In many studies, biological sample are only taken from a single point in time and may not reflect accurately reflect longitudinal patterns, particularly if exposures are highly variable. Furthermore, evaluation of biomarkers also requires an understanding of degradation and metabolism of chemicals in both the environment and human body. As such, biomarkers of exposure may differ between individuals for reasons other than exposure level. Differences in metabolism, co-morbidities such as kidney disease in relation to urinary measurements, uncertainty as to whether the biomarker measures exposure to the active ingredient or the environmental degradates may all account for apparent differences in biomarkers of exposure among individuals, and possibly between comparison groups.

Indirect exposure assessment methods are common in retrospective studies and based on factors that are surrogates of chemical exposure. As described above, indirect exposure data cannot generally be used to estimate quantitative exposure levels without additional modeling. For example, a questionnaire can be used to determine if an individual has ever used a pesticide, but can less reliably collect data on all the environmental and behavioral factors that are needed to calculate that individual's exposure. As such, indirect exposure data are often used to classify exposure using a dichotomous exposure variable (i.e. exposed/unexposed) or ordinal exposure scale. In contrast, direct exposure assessment methods are based on data on actual individual-level exposure through personal monitoring and biomonitoring. Thus, direct methods can be used to estimate individual exposure or internal dose levels. Direct methods are more common in prospective studies, but are also used in retrospective studies when existing biological samples are available from well-defined population groups.

Quantified personal measurements, such as personal monitoring and biomonitoring, are generally considered the best source of data for estimating actual exposure levels (NRC, 1991; NRC, 1997). While this is the case, accurate qualitative measures of exposure (e.g. dichotomous and ordinal exposure metrics) from indirect methods can be just as accurate for the purpose of epidemiology. Moreover, indirect methods are often easier to interpret and may require less additional research and development to demonstrate their utility in exposure assessment.

Regardless of the approach, exposure assessment methods should be able to provide exposure estimates that are reliable and valid. In the context of epidemiology, *reliability* general refers to the ability to reproduce results and *validity* generally refers to the extent that exposure estimates reflect true exposure levels (Checkoway et al, 2004). When evaluating a particular exposure assessment's reliability and validity, it is important to consider the exposure assessment's strengths and weaknesses in the context of the study's research objectives. Less refined exposure assessment may be suitable for exploratory studies. This is because exploratory studies help raise awareness about potential hazards that can encourage investment in more focused research. Conversely, studies with more focused hypotheses can be greatly strengthened through the use of more refined exposure assessment methods. Therefore, indirect and direct exposure assessment methods represent a spectrum of tools that are complimentary and can be used at different stages of research when exploring exposure-disease relationships.

2. Confounding Factors

A confounding factor is one which, if not properly measured and analyzed properly, will change the magnitude and direction of the estimated association between an exposure and health outcome. Confounding factors may include lifestyle exposures such as cigarette smoking, high energy diet and lack of physical activity all of which may adversely affect health and may be statistically associated with pesticide use. In

epidemiological analyses, confounding factors are measured in the study sample and typically included as “adjustments” to the final risk estimate through statistical analysis tools such as regression models. Depending upon the specific exposure-disease association under study, a factor may or may not be a confounding factor. Assessment of potential confounding is made on a study specific basis. When evaluating the quality of observational epidemiology studies, OPP will consider whether relevant confounding factors are properly identified, described, measured and analyzed such that an unbiased estimate of the specific association under study can be made.

3. Statistical Analysis

Epidemiologic studies are designed to measure an association between a specific exposure and a disease. When evaluating the quality of pesticide epidemiology studies, OPP will also consider the statistical methods used. Specifically, OPP will consider the extent to which the analytic methods described in the study are appropriate to the research question, the completeness of the description of the statistical methods utilized, the appropriateness of the methods for identification, assessment and adjustment of potentially confounding variables in the exposure-disease relation; and, the description and presentation of sub-group analyses performed.

Epidemiologic investigations typically utilize statistical modeling to measure risk. To do so, researchers must consider not only the relevant main exposure and outcome variables, but also consider relevant confounding factors, and whether the association under investigation may differ by level of these factors, i.e., effect modification (Szklo et al, 2004). Upon identification of a potentially confounding variable, one that changes the magnitude and/or direction of the association under study, adjustment through regression modeling can help to isolate the risk estimate of interest, i.e., the association under study. In addition, OPP will evaluate the stratification of the association by the level of the potential effect modifier under study or evaluation of statistical interaction. If the magnitude and direction of the association of interest differs greatly by level of a third variable, then the stratified results should be considered primary.

4. Potential Bias in Observational Research

Bias is a systematic error in the design or conduct of a study such that the results from a biased study will tend to be different from the true result. Studies may be biased in the way in which participants are selected into the study (selection bias), or the way in which information about exposure and disease status is collected (information bias). Bias reflects a problem in the design or conduct of the study, and should be addressed or discussed by researchers. However, no study is totally devoid of bias. One should consider the extent to which authors of published studies described potential bias in the study, and how, if at all, they attempted to address it in the study. Studies found to have significant potential to be biased will receive less weight in a WOE analysis than those without such significant bias.

5. Interpretation of Null studies

"Null" studies, or well-conducted studies which report no association between exposure to the pesticide and an adverse health outcome, will be evaluated carefully for their potential usefulness in human health risk assessment. The study may report a null result either because the association indeed does not exist in nature, or because the study as conducted fails to detect an association. To evaluate which of these two scenarios may be correct when reviewing "null" studies, one should consider other research reported concerning the same or similar research question, the manner in which exposure and outcome were assessed, the statistical methods used including the identification and analysis of confounding variables in the association, to interpret null studies. Statistical power refers to the probability that researchers may correctly identify that there is a difference between the two comparison groups, i.e., there is an association between exposure and disease, when in fact there is a true association in nature. Studies that are "low powered" may falsely conclude there is no association, when an association actually exists.

6. External Validity (Generalizalbity)

As noted above, *validity* generally refers to the extent that exposure estimates reflect true exposure levels (Checkoway et al, 2004). *External validity*, or *generalizability*, refers to the ability to extend the epidemiologic study results derived from a sample of the population (e.g., pesticide applicators) to the other populations (e.g., all agricultural workers). To assess external validity, comparison of characteristics in the sample to the larger population (if known) can be made. Generalizability is of particular importance because it is important to understand whether and how individual study results may be applied to the larger group or targeted sub-groups in regulatory risk assessment. For example, the AHS has reported statistical associations between some cancer and non-cancer health outcomes for some pesticide chemicals. OPP has an interest in evaluating the extent to which the reported findings may apply to pesticide applicators in states other than North Carolina and Iowa or to farm workers who primarily do post-application activities.

C. Benefits & Uses of Epidemiologic Data in Human Health Risk Assessment

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, 2000). Alternatively, epidemiology studies may be used to

compare with points of departure and/or reference doses (RfD) derived from experimental animal studies to characterize assumptions used in deriving such values. In other cases, outcomes reported in epidemiologic studies may be compared qualitatively with those seen in *in vitro* and animal studies to evaluate biological plausibility or human relevance of animal findings (Hertz-Picciotto, 1995). As discussed in the Introduction of this draft framework, human information like that found in epidemiology studies are expected to play a significant role in the new vision of toxicity testing recommended by the NRC (2007). Specifically, epidemiology studies provide insight on actual chemical exposures in humans and thus can contribute to problem formulation and hazard/risk characterization. Human information may guide additional studies (e.g., dose and endpoint selection for use in *in vitro* and targeted *in vivo* experimental studies); and identify novel health effects or host susceptibilities which can be investigated with future research.

When laboratory data from animal studies provide the primary source of information for hazard characterization, one potential source of uncertainty is the relevance of animal models to humans. In the absence of data to support the contrary, animal findings are assumed to be relevant to humans. Furthermore, in the absence of data to support the contrary, EPA assumes that humans are more sensitive than laboratory animals. In actuality, humans may be more or less sensitive to pesticides than other animal species. Epidemiology and human incident data provide scientific to inform uncertainties associated with species extrapolation. With respect to population variability, epidemiology studies better characterize potential variability than do animal studies. Specifically, epidemiologic data include the genetic diversity, and variability inherent in human populations and thus better represent actual population response to environmental chemicals than laboratory animals(Caulderon, 2000).

With respect to dose-response characterization, animal toxicology studies have the benefit that studies can be designed to cover a broad range of exposure levels. However, animal toxicology studies generally use exposures which are much larger (sometimes orders of magnitude) than those found in the environment. These high exposure levels in animals studies dictate the need for extrapolation from high to low doses. This extrapolation introduces added uncertainty into the risk assessment. Epidemiology studies and human incident data involve actual real-world exposures and thus high dose extrapolation may not be needed.

Animal studies do not replicate the length, magnitude, duration, routes of exposure and variability in exposure experienced by humans (Caulderon, 2000). Human exposure often occurs through multimedia exposure pathways, including food, water, air, and indoor and outdoor environments. In contrast, controlled laboratory studies typically use a single route of exposure. In addition, humans may experience exposure to multiple chemicals and/or non-chemical stressors simultaneously, whereas most animal studies involve a single chemical stressor. On one hand, this multi-chemical exposure in epidemiology studies can provide a challenge when attempting to attribute epidemiologic outcomes to a single pesticide chemical. On the other hand, because epidemiologic research considers real-world exposure, along with experimental

approaches, it may help address questions associated with multiple chemical exposures which are difficult to evaluate.

III. HUMAN INCIDENT SURVEILLANCE DATA

Sources of human data include case reports and surveillance of acute pesticide poisoning incidents studies. Generally speaking, epidemiology studies on pesticides focus on lower levels of exposure that are less likely to result in acute clinical symptoms. In contrast, human incident information provides insight on potential acute or short-term, often reversible effects from single exposures. Thus, data from these sources can be combined and help aid acute and chronic hazard identification as part of the risk assessment process[4].

OPP considers several specific data sources in the review of acute pesticide poisoning incidents. Pesticide registrants are required under the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA) to notify EPA if and when they become aware of "factual information regarding unreasonable adverse effects on the environment of the pesticide." The Agency also obtains acute poisoning event data from a variety of sources like, poison control centers, states, and other organizations (Table 3). These data sources can all be considered case series or a compilation of case reports. Case reports are valuable because they can identify unusual or novel occurrences of an adverse health effects plausibly associated with use of a specific pesticide. While the data sources OPP utilizes in its evaluation of acute pesticide poisoning events vary with regard to plausibly associating pesticide exposure and adverse health outcome, taken together these reports can aid in identify severe adverse events, temporal trends or geographic or sub-group patterns in need of regulatory action. Synthesizing information from acute animal toxicity testing results, to documented human health effects of acute poisoning (Handbook of Pesticide Management), with case reports collected through various data sources provides OPP a "first-line" of evidence of adverse health effects plausibly associated with pesticide products (van den Brandt, 2001).

Case reports published in the peer reviewed literature describe a particular effect in an individual or group of individuals who were exposed to a substance. These reports are often anecdotal or highly selective in nature. They can, however, can be particularly valuable in identifying previously unidentified toxic effects. Case reports for pesticides typically describe the effects from an atypical (high exposure/dose, illegal, off-label) acute or short-term exposure. If similarities are detected across multiple medical case studies, more weight is given to the findings. Medical case studies that include quantitative exposure information can be compared to exposure estimates in the risk assessment (which are based on labeled application rates and surrogate exposure

[4] OPP is aware of efforts by IPSC to consider human incident data in risk assessment. http://www.who.int/ipcs/publications/methods/human_data/en/index.html

information) to characterize margins of exposure expected from typical use, when appropriate.

Human incident databases have different strengths and limitations based on how the information is collected and organized (Table 3). OPP has access to the following five human incident data sources: the OPP Incident Data System (IDS), the American Association of Poison Control Centers (PCC), National Pesticide Information Center (NPIC), the Centers for Disease Control and Prevention/National Institute for Occupational Safety and Health Sentinel Event Notification System for Occupational Risk (NIOSH SENSOR) and the California Pesticide Illness Surveillance Program (PISP).

- **OPP Incident Data System (IDS)** maintained by OPP and incorporates data submitted by registrants under FIFRA section 6(a)(2), as well as other incidents reported directly to EPA. IDS include information on incidents involving humans, plants, wild and domestic animals where there is a claim of an adverse effect, as well as detects of pesticides in water. The vast majority of reports are received in paper format. IDS entries act as a pointer to copies of original reports, retained on microfilm and scanned images in OPP's Information Service Center. Many companies use standardized, industry-developed Voluntary Incident Reporting Forms. While IDS reports are broad in scope, the system does not consistently capture detailed information about incident events, such as occupational exposure circumstances or medical outcome. In most cases data going into IDS is not validated or verified, though some reports are collected from calls to contract poison control centers.

  IDS include both occupational and non-occupational incidents. However, IDS contains more non-occupational incidents compared with occupational incidents. IDS also includes narrative information on exposure scenario and hazard information. As IDS is populated mostly by registrants, the Agency has relatively high confidence in the product information provided. The scope of IDS is national. Severity rankings are included for each incident (as specified by CFR §159.184), however, upon review of narratives, severity rankings often appear inflated compared to other databases (and subsequently, the Agency has less confidence in these classifications). Symptom information is sometimes included in the narrative portion of the incident; however this information is usually not validated/confirmed by a professional. Overall, IDS provides good information about national trends and frequency of incidents for pesticides; and can provide valuable insights into the hazard and/or exposure potential of a pesticide.

- The **National Poison Data System (NPDS)**, formerly called the Toxic Effects Surveillance System (TESS), is maintained by the American Association of Poison Control Centers (PCC) and is supported with funding from several federal agencies. NPDS is a computerized information system with geographically specific and near real-time reporting. Although the main mission of Poison Control Centers is in helping callers respond to emergencies, NPDS data can

help identify emerging problems in chemical product safety. Hotlines at 61 PCC's nationwide are open 24/7, 365 days a year. Hotlines are staffed by specially trained nurses to provide poisoning information. Using computer assisted data entry, standardized protocols, and strict data entry criteria, local callers report incidents. These reported incidents are retained locally and are updated in summary form to the national database. Information calls are tallied separately and not counted as incidents. The PCC system covers nearly all the US and its territories and is undergoing major computer enhancements since 2001. There are 1,546,503 records of "incidents" in the PCC database for pesticides, algicides and disinfectants. Not all of these records are complete

NPDS includes mainly non-occupational incidents. NPDS does not include narrative information and the (product information is not complete regarding pesticide product information. The Agency matches product identification numbers in NPDS to EPA product identification numbers (and subsequently, pesticide specific information). NPDS provides severity rankings and symptom information that are designated/recorded by trained specialists, and therefore the Agency has relatively high confidence in this information. NPDS also provides some information on the likelihood of the adverse effect being a result of the reported exposure. Overall, NPDS provides good information about national trends, frequency of incidents for pesticides (which is somewhat limited by NPDS's product database), as well as the hazard potential for particular pesticides.

- The **National Pesticide Information Center (NPIC**) is funded by EPA to serve as a source of objective, science-based pesticide information in response to inquiries and to respond to incidents. NPIC functions nationally during weekday business hours, under a cooperative agreement between Oregon State University and EPA. Similar to Poison Control Centers, NPIC's primary purpose is to provide information — not to collect incident data. NPIC does collect information about incidents from inquirers and reports that information to EPA (about 10% of NPIC's annual calls are considered "incident" related). The Center's main role is to provide information on a wide range of pesticide topics, and to direct callers for pesticide incident investigation and emergency treatment.

  Like IDS and PCC, the incidents in NPIC are mainly non-occupational. NPIC incidents include narratives and product information when the caller provides the information. The scope is national, however, there are significantly fewer incidents reported to NPIC than to NPDS or IDS. Hazard information includes severity rankings, route of exposure and symptoms – which are recorded by trained personnel. NPIC also provides information on how likely the link between exposure and adverse effect is (which they call a certainty index). NPIC also publishes annual reports and analyses in the open literature which are valuable resources.

- The **National Institute of Occupational Safety and Health's (Sentinel Event Notification System for Occupational Risks, SENSOR)** database contains data for occupational pesticide incidents that is collected from state health departments. It should be noted that not all states are represented in this data, and that the number of states reporting data to SENSOR program increased over time. Currently 12 states are included. Further, although there are data available back to 1974, a substantial number of cases were not recorded until 1998. Currently data up to and including the year 2006 is contained in the database. SENSOR reports include narratives and product information which are input by trained professionals. Hazard information includes severity rankings, and narrative information can include information on symptoms and treatment. Overall, NIOSH SENSOR provides good information on both occupational and non-occupational incidents, and sometimes valuable insights into the hazard and/or exposure potential of a pesticide. NIOSH SENSOR also conducts analyses of its own data and publishes these in the Morbidity and Mortality Weekly.

- The **California Pesticide Illness Surveillance Program (PISP)** is maintained by the State of California. This database documents pesticide-related illnesses and injuries. Case reports are received from physicians and via workers' compensation records. The local County Agricultural Commissioner investigates the circumstances of the exposure. Medical records and investigative findings are then evaluated by California's Department of Pesticide Regulation (DPR) technical experts and entered into an illness registry.

Table 3. Human Pesticide Incident Data Sources

Data Source	Years	Database Characteristics
OPP Incident Data System (IDS)	1992-present	• Centralized system • Incident reports from various sources • Case reports • Uneven level of detail • Labor intensive; not fully automated • Largely anecdotal reports/allegations
American Association of Poison Control Centers (PCC)	1993-2005	• National scope • Able to summarize organize information based on data fields that are systematically recorded by trained poison center professions • Clinically oriented • Over 1.5 million records • Provides greater information on incidents in residential settings
National Pesticide Information Center (NPIC)	1978- present	• National scope • Focus on incidents in residential settings • Limited scale;
Sentinel Event Notification System for Occupational Risk (SENSOR)	1998 - present	• Best available data for occupational incidents • Includes data from multiple sources • Provides detailed information • Standardized data fields • Covers 12 states • Level of reporting likely to vary from state to state • Focus on occupationally-related cases (although approximately 50% of cases are non-occupational)
California Pesticide Illness Surveillance Program (PISP)	Standard collections from 1982; Methods revised 1992	• Unique infrastructure for follow-up • Includes all types of pesticides • Provides detailed information • Standardized information • Limited to California • Occasional lag time between incident and report

When evaluating human incident data, OPP considers several general criteria. OPP considers the relative severity and frequency of symptoms. Additionally, OPP generally has greater confidence in reports in which temporal association can be verified or are at least plausible. Lastly, other factors that are used to evaluate human incident data include evidence of an exposure response association, consistency in reported health effects, biological plausibility of reported health effects, elimination of alternative causes of health effect such as pharmaceutical use, and, the specificity of the type of health effects seen in the exposed. Currently, OPP evaluates human

incident data on a chemical-specific basis. Incidents from each database are analyzed for hazard potential (deaths, frequency of more severe incidents, and patterns/trends of reported symptoms) and exposure potential (frequency of incidents/ trends over time, patterns/trends of exposure scenarios, of factors affecting exposure or of products). Additionally, narratives of more severe incidents are often evaluated for any temporal association between time-of-exposure and effects reported to determine whether an association is supported (*e.g.*, if a child's death is attributed to a pesticide used in the home, yet the death occurs in a locked tent outside of the home in severe weather conditions, that incident may not be attributed to the pesticide exposure; however, if a severe incident occurs at the time of exposure, or if symptoms and health conditions occur immediately after exposure without prior medical history and symptoms progress over time, or are reproduced upon re-exposure, then that severe incident may be more likely due to pesticide exposure). When preliminary analyses indicate a possible hazard/exposure potential of interest, additional databases are consulted for consistency and reproducibility, and considered in light of all other available information.

OPP uses human incident information for several purposes. Most broadly, the program uses incident data to inform risk assessment/risk management activities. To this end, OPP evaluates human incident data for trends over time and examines patterns in the severity and frequency of different pesticide exposures. In some cases, incident information can indicate need for a new risk assessment or new risk management measures. Incident information can also help assess the success of risk mitigation actions after they are implemented. Thus, incident information is an integral part of OPP's performance accountability system, to ensure the effectiveness of risk management actions that OPP has taken to protect human health and the environment. Lastly, incident information can be useful in targeting enforcement activities and can serve as a source for information on compliance with incident reporting regulations.

IV. PROPOSED WEIGHT OF THE EVIDENCE (WOE) ANALYSIS

OPP plans to use a WOE analysis for evaluating epidemiology and human incident data, such that all available data are evaluated and conclusions are made on the preponderance of the information rather than relying on any one study. As mentioned above, in the WOE analysis, OPP will use the best available data across multiple lines of evidence and from *in vitro*, *in vivo*, and *in silico* data sources to describe the cascade of events from the exposure source to the ultimate health outcome (Figure 1). Furthermore, OPP plans to use modified Bradford Hill criteria like those in the MOA framework as a tool for organizing and integrating information from different sources (U.S. EPA, 1999, 2005; Sonich-Mullin et al., 2001; Meek et al., 2003).

Data used in the WOE analyses will come from studies submitted for purposes of pesticide registration and from the scientific literature. In addition data may be from tools such as structure activity relationships (SAR), physiologically-based pharmacokinetic (PBPK) or biological dose-response models, biomonitoring or other exposure studies or analyses. Toxicity studies submitted for pesticide registration provide information on a wide range of adverse health outcomes, routes of exposure,

exposure durations, species, and lifestages. These studies typically used internationally accepted protocols and guidelines which ease comparisons across studies and chemicals. Literature studies provide additional information on toxic outcomes not reported in registration studies and are often important sources of information on exposure, PK, and MOA.

Practically speaking, epidemiology and human incident data are available on a pesticide only after its registration and widespread use has occurred. Epidemiology studies are generally conducted on the most widely used pesticides; these pesticides also tend to have to be well-studied in the scientific literature. Thus, OPP expects in many cases where epidemiologic data are available, a significant body of literature data on toxicology, exposure, PK, and MOA will also be available. Human incident data are available on a broader range of chemicals; some of which have robust databases and others which do not. In those cases where there are significant human incident cases and little is known about the MOA or PK of a particular pesticide, the WOE analysis can be used to identify areas of new research. The following text describes the aspects in the proposed WOE analysis:

1. *Review of the Epidemiologic and/or Human Incident Data:* The first step in a WOE analysis is to review all available studies. This review considers a variety of factors including, but not limited to, research hypothesis; study design (i.e., sample size, sufficient controls, quality of measurements, etc), exposure dose/concentration, statistical analysis, and conclusions. In the specific case of epidemiology and human incident data, important considerations used in reviewing such data are discussed in Sections II and III. A concise written review of the study is developed. This written review describes the study design, results, conclusions, and the strengths and weaknesses of the study.

   The quality of the epidemiologic exposure assessment is an important factor in determining what role epidemiologic data will play in the risk assessment. As such, it is important to fully characterize the assumptions used in the epidemiologic exposure assessment and the degree to which these assumptions affect the interpretation and generalizability of the epidemiologic findings. The evaluation of the epidemiologic exposure assessment may include a consideration of past and present exposure patterns (e.g., exposed populations, pathways, routes, and levels of exposure) and may include significant changes in use patterns (e.g., risk mitigation actions or new use patterns).

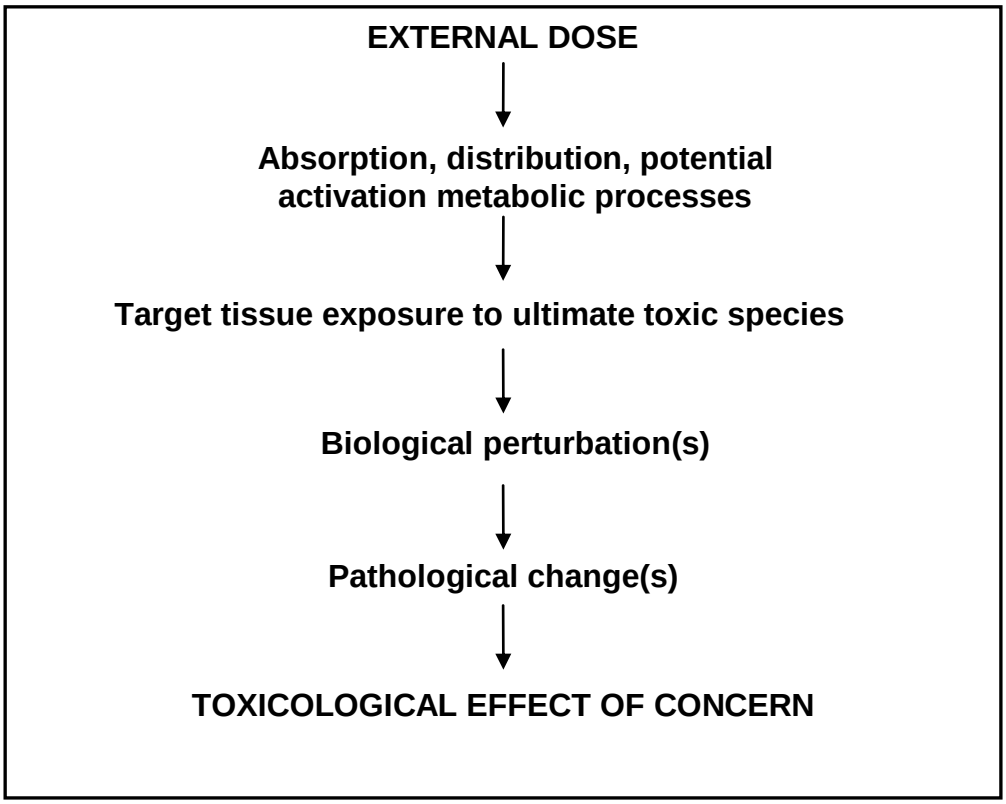
2. *Modified Bradford Hill Criteria & The Mode of Action Framework:* Effects, PK and MOA information provides the scientific foundation for evaluating dose-response, species differences, lifestage sensitivity, route to route differences, and biological plausibility of reported outcomes. Accordingly, OPP's proposed WOE analysis relies heavily on modified Bradford Hill Criteria as used in the MOA Framework (U.S. EPA,1999, 2005; Sonich-Mullin et al., 2001; Meek et al., 2003). The modified Bradford Hill criteria used in the MOA framework are described by the steps described below (adopted from Seed et al, 2005). Briefly, the concept of

MOA involves the identification of a biologically plausible sequence of key events that are obligatory and quantifiable steps leading to an adverse outcome. The MOA analysis takes into account factors such as dose-response and temporal concordance, biological plausibility, coherence and consistency. In this analysis, epidemiologic findings and human incident data can be evaluated in the context of other human information and experimental studies to evaluate biological plausibility, to identify areas of uncertainty and areas of further research.

- Postulated MOA. A brief description of the sequence of measured events, starting with chemical administration, to the formation of the toxic effect at a given site is first provided.

- Key events. Clear descriptions of each of the key events (i.e., measurable parameters) that are thought to underlie the MOA are given. Figure 3 provides a schematic of general key events that can describe a MOA from external dose to ultimate toxicological effect of concern. Data to inform the key events may come from a combination of *in vitro* or *in vivo* data sources (human or animal) and include information on toxicity pathways like that shown previously in Figure 1. Moreover, as shown in Figure 1 and Figure 3, these key events can be a combination of PK and pharmacodynamic (PD) events.

Figure 3. Schematic of general key events that can describe a MOA.



- Dose-response relationships. Dose-response relationships are identified for each key event; such data can be presented in figures or tables for ease of evaluation. Dose-response relationships are compared among key events. In some cases, the earlier key events may be more sensitive than later key events. In other cases, key events may share similar dose-response curves. In the risk assessment, a point of departure (PoD) may be derived from data on a key event when the dose-response relationships between the key event and ultimate health outcome has been established.

- Temporal association. The temporal sequence of key events over time that lead to the toxic effect are described. Temporal data can be presented in figures or tables for each of evaluation. In this part of the analysis, data are evaluated to ensure that the temporal sequence of events is supported. For example, this analysis considers key events which occur rapidly (e.g., metabolism to an active metabolite which could occur within minutes of exposure) and those which occur after longer durations (e.g., development of a tumor) to ensure coherence of the effects

- Strength, consistency, and specificity of association of key events and the toxic effect. Complete assessment and presentation of the relationships among key events, precursor lesions, and the toxic effect is provided. In this analysis, the consistency of observations across studies of different designs is described. OPP will place higher confidence in results which are replicated or reproduced from multiple studies. In this evaluation, a comparison of animal and human findings is provided. There are a number of possible outcomes to this species comparison. When animal and human data show a similar toxic profile, both quantitatively and qualitatively, there is high confidence in the human health risk assessment. In some cases, animal and human data show a qualitatively similar toxic profile but quantitative differences are observed. In other words, a particular chemical exhibits the same MOA in animals and humans but there may be species differences in dose-response characteristics. These dose-response differences could be due to tissue dosimetry (i.e., PK) or from different response characteristics (i.e., PD). Scientifically robust approaches for using PK or PD data to perform species extrapolation are described in IPSC (2005) and USEPA (2006b) In these cases, there is a high degree of knowledge about the PK and PD characteristics of animals and humans. Consequently, the human health risk assessment is based on strong scientific evidence and thus, there is high confidence in the human health risk assessment.

  In other cases, animal and human data show qualitatively dissimilar outcomes. This situation highlights the need to fully and objectively evaluate all available information in a transparent and comprehensive

manner. These situations also highlight the value of understanding MOA. Specifically, by knowing the key events leading to toxicity, one will better understand why species respond differently to pesticide exposure. When animal and epidemiologic data do not provide a consistent toxicological picture of a particular pesticide, more weight would likely be given to those studies with robust study design and availability of replication or confirmatory data. Dissimilar animal and human findings do not necessarily mean that animals are poor models of human effects. In some cases, a possible toxic effect has simply not been evaluated in animals and further animal research may be needed; this is particularly important for pesticides whose PK characteristics and tissue dosimetry vary significantly among routes of exposure and/or among lifestages or for pesticides which exhibit multiple MOAs.

In other cases, humans may express a broader range of outcomes for a particular toxicity than can be measured in animals, thus making experimental studies with animals difficult to design and conduct. In these situations, epidemiology studies with high quality study designs and strong exposure assessment may provide the most appropriate data for characterizing human risks. However, in most situations, the epidemiological study may not be sufficiently robust for deriving quantitative risk assessment values. In these situations, it is important to ensure that the risk assessment is protective of human health by selecting a sensitive endpoint from the animal studies as the basis for risk extrapolation. It is important to note that even though the adverse effect observed in two species may differ, the dose-response curves for those effects may not differ. Thus, by selecting a biologically plausible and sensitive endpoint from the animal studies, the risk assessment is protective of human health.

- Biological plausibility and coherence. Determination of whether key events and the sequence of events are consistent with current biological thinking, regarding both the specific toxic effect in general and the specific chemical under review. Specific to epidemiologic and human incident data, the degree to which reported outcomes compare with those expected from a known MOA and/or with health outcomes for other chemicals in the same chemical class are discussed.

- Other MOAs. Alternative MOAs that may be applicable for the chemical under review. Comparison of their likelihood vis-à-vis the proposed MOA. There may be cases where epidemiologic findings differ from those in animals or differ from expected outcomes in a proposed MOA, In these situations, the epidemiology data be used to suggest new hypotheses for alternative MOA(s) which can be further investigated with new research.

3. *Overall conclusions, statement of areas of confidence and uncertainty, and recommendations for risk assessment:* In the final portion of the proposed WOE analysis, the overall conclusions along with statement of areas of confidence and uncertainty. This section also identifies areas of additional research. This section recommends the source of data for regulatory values and the appropriate approach for extrapolating between species (if necessary) and among humans.

V. CASE STUDIES

To gain experience with this draft framework, OPP is presenting several case studies. These case studies are described in Attachments A-C.

- Case study 1 (Attachment A) describes several ecologic and retrospective epidemiology studies describing the statistical association between atrazine and birth outcomes.

- Case study 2 (Attachment B) describes on-going collaborative work between OPP, ORD, NCI, and NIEHS to compare the exposure assessment approaches used by OPP and the AHS.

- Case study 3 (Attachment C) will provide a retrospective analysis of reported human incident data for diazinon. This case study is not included in this draft. This document will be transmitted to the SAP in preparation for review at the February, 2010 meeting.

For each case study, work is on-going. OPP has **not** completed a WOE analysis for them. The atrazine and diazinon case studies describe the status of the data evaluation phase of the analysis. The atrazine studies included in Case study 1 will be incorporated in a WOE analysis scheduled for review by the FIFRA SAP in September, 2010. Similar to many other pesticides undergoing registration review, the human incident data for diazinon will be integrated with other *in vitro* and *in vivo* data along with epidemiology studies as part of the human health risk assessment.

Case study 2 evaluates exposure assessment method in a prospective epidemiology study (i.e., the AHS). Specifically, the case study involves a comparative analysis evaluating exposure assessments used by OPP and the AHS. An overview of this case study is provided in Attachment B. The Agency is developing an addendum for submission to the FIFRA SAP for review at the February, 2010 meeting. This forthcoming addendum will provide a more detailed discussion of the key science issues and project plan for this comparative analysis. Conceptually, the addendum will present this case study in terms of a problem formulation exercise. Specifically, the problem formulation will identify key differences in the exposure assessment methodologies used by AHS investigators and the Agency in an attempt to illustrate how these differences affect risk conclusions.

VI. SUMMARY & NEXT STEPS

This draft framework describes important factors in reviewing epidemiology and human incident data and describes a proposed WOE analysis for incorporating such data in pesticide human health risk assessment. This proposed WOE integrates data from a variety of *in vitro*, *in vivo*, and *in silico* sources using modified Bradford Hill criteria as an organizational tool. OPP is seeking peer review and public comment on this draft document and on-going case studies at a meeting of the FIFRA SAP in February, 2010. OPP acknowledges that toxicology and risk assessment are currently undergoing transformational changes towards implementing the new vision of 21st century toxicity testing. As these transformation changes occur, OPP will update this approach as appropriate.

VII. REFERENCES

Blair A, Tarone R, Sandler D, Lynch C, Rowland A, Wintersteen W, Steen W, Dosemeci M, Alavanja M. 2002 Reliability of reporting on lifestyle and agricultural factors by a sample of participants in the agricultural health study from iowa. Ann Epidemiol. Oct 1;10(7):478

Boyes W.K., Moser V.C., Geller A.M., Benignus V.A., Bushnell P.J., Kamel F. 2007. Integrating epidemiology and toxicology in neurotoxicity risk assessment. Hum Exp Toxicol. 26(4):283-93.

Calderon R.L. 2000. Measuring risks in humans: the promise and practice of epidemiology. Food and Chemical Toxicology. 38:S59-S63.

Carlile, D.J., K. Zomorodi and J.B. Houston. 1997. Scaling factors to relate drug metabolic clearance in hepatic microsomes, isolated hepatocytes, and the intact liver: studies with induced livers involving diazepam. Drug Metab. Dispos. 25(8):903-911.

Checkoway H., Pearce N., and Kriebel D. Research Methods in Occupational Epidemiology, 2nd Edition. Oxford University Press, New York, 2004.

Clark, L.H., R.W. Setzer and H.A. Barton. 2004. Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. Risk Anal. 24(6):1697-1717.

Hertz-Picciotto I. 1995. Epidemiology and quantitative risk assessment: a bridge from science to policy. American Journal of Public Health. 85(4): 484-491.

Hoppin, J.A., F. Yucel, M. Dosemeci, and D.P. Sandler. 2002. Accuracy of self-reported pesticide use duration information for licensed pesticide applicators in the Agricultural Health Study. Journal of Exposure Analysis and Environmental Epidemiology, 12: 313-318.

IPCS (2005). Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration response assessment. Harmonization Project Document 2. World Health Organisation, International Programme on Chemical Safety, Geneva, Switzerland.

Kelsey J.L., Whittemore A.S., Evans A.S., Thompson W.D.. Methods in Observational Epidemiology. 2nd ed. New York, NY: Oxford University Press; 1996.

Lilienfeld, AM; Lilienfeld, D. (1979) Foundations of epidemiology, 2nd ed. New York: Oxford University Press.

Mausner, JS; Kramer, S. (1985) Epidemiology, 2nd ed. Philadelphia: W.B. Saunders.

Meek, M.E., J.R. Bucher, S.M. Cohen et al. 2003. A framework for human relevance analysis of information on carcinogenic modes of action. Crit. Rev. Toxicol. 33:591-653.

Needham L.L., Calafat A.M., and Barr D.B. 2007. Uses and issues of biomonitoring. Int. J. Hyg. Environ. Health. 210: 229-238.

Nieuwenhuijsen M.J. Exposure Assessment in Occupational and Environmental Epidemiology. Oxford University Press, New York, 2003.

NRC (National Research Council). 1983. Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington, DC.

NRC (National Research Council). 1991. Environmental Epidemiology, Volume 1: Public Health and Hazardous Wastes. National Academy Press, Washington, DC.

NRC (National Research Council). 1994. Science and Judgment in Risk Assessment. National Academy Press, Washington, DC.

NRC (National Research Council). 1997. Environmental Epidemiology, Volume 2: Use of the Gray Literature and Other Data in Environmental Epidemiology. National Academy Press, Washington, DC.

Paddle G.M., Harrington J.M. 2000. Environmental epidemiology--strengths and weaknesses. Int Arch Occup Environ Health. 73:7-14.

Purdue Pesticides Programs. 2003. Pesticides and Epidemiology: Unraveling Disease Patterns. Purdue University Cooperative Extension Service. http://www.btny.purdue.edu/Pubs/PPP/PPP-43.pdf.

Rothman KJ, Greenland S. Modern epidemiology. 2nd ed. Lippincott Williams & Wilkins, Philadelphia, 1998.

Seed, J., E.W. Carney, R.A. Corley et al. 2005. Overview: Using mode of action and lifestage information to evaluate the human relevance of animal toxicity data. Crit. Rev. Toxicol. 35(8-9):664-672.

Sonich-Mullin, C., R. Fielder, J. Wiltse et al. 2001. IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. Regul Toxicol Pharmacol. 34:146-152.

Szklo, M. and Nieto, F.J. 2004. Epidemiology: Beyond the Basics. Jones and Bartlett Publishers: Boston, MA.

U.S. Environmental Protection Agency. 1998. Guidelines for Ecological Risk Assessment. Risk Assessment Forum, Office of Research and Development, Washington, D.C. EPA/630/R-95/002F. April 1998.

U.S. Environmental Protection Agency. (1999). Guidelines for carcinogen risk assessment. Risk Assessment Forum. SAB review draft. Washington, DC: U.S. Environmental Protection Agency. www.epa.gov/ncea/raf/crasab.htm.

U.S. EPA (U.S. Environmental Protection Agency). 2000. Science Policy Council Handbook: Risk Characterization. U.S. Environmental Protection Agency, Office of Research and Development, Office of Health and Environmental Assessment, Washington, DC. EPA/100/B-00/002. Available at http://www.epa.gov/iris/backgr-d.htm.

U.S. EPA (U.S. Environmental Protection Agency). 2001. Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment. Office of Emergency and Remedial Response. Washington, DC. EPA-540-R-02-002. Available at http://www.epa.gov/superfund/RAGS3A/index.htm.

U.S. EPA (U.S. Environmental Protection Agency). 2002a. Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility and Integrity for Information Disseminated by the Environmental Protection Agency. Office of Environmental Information, Washington, DC. EPA/260/R-02/008. Available at http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf.

U.S. EPA (U.S. Environmental Protection Agency). 2002b. A Review of the Reference Dose and Reference Concentration Processes. December. Risk Assessment Forum. Washington, DC. EPA/630/P-02/002F.

U.S. EPA (U.S. Environmental Protection Agency). 2004. An Examination of EPA Risk Assessment Principles & Practices. Staff Paper Prepared for the U.S. Environmental Protection Agency by members of the Risk Assessment Task Force. Office of the Science Advisor. U.S. Environmental Protection Agency, Washington, DC. EPA/100/B-04/001.

U.S. EPA (U.S. Environmental Protection Agency). 2005. Guidelines for Carcinogen Risk Assessment. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, DC. EPA/630/P-03/001F. Federal Register 70(66):17765-17817. Available at http://www.epa.gov/raf.

U.S. EPA. (U.S. Environmental Protection Agency). 2006a. Harmonization in Interspecies Extrapolation: Use of $BW^{3/4}$ as Default Method in Derivation of the Oral RfD (External Review Draft). U.S. Environmental Protection Agency, Washington, DC. EPA/630/R-06/001. Available at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=148525.

U.S. EPA (U.S. Environmental Protection Agency). 2006b. Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment (Final Report). U.S. Environmental Protection Agency, Washington, DC. EPA/600/R-05/043F.

U.S. EPA (U.S. Environmental Protection Agency). 2009. Scientific Issues Associated with Field Volatilization of Conventional PesticidesU.S. Environmental Protection Agency, Washington, DC. OPP Regulatory Public Docket EPA-HQ-OPP-2009-0687.

van den Brandt P, Voorrips L, Hertz-Picciotto I, Shuker D, Boeing H, Speijers G, Guittard C, Kleiner J, Knowles M, Wolk A, Goldbohm A. 2002 The contribution of epidemiology Food Chem Toxicol. 2002 Feb-Mar;40(2-3):387-424.

Zartarian V., Bahadori T., McKone, T. 2005. Adoption of an official ISEA glossary. Journal of Exposure Analysis and Environmental Epidemiology. 15:1-5.

Attachment A

CASE STUDY on Recent Epidemiologic Findings on the Association between Atrazine and Birth Outcomes

Office of Pesticide Programs
Office of Pollution Prevention and Toxic Substances
U.S. Environmental Protection Agency

FIFRA Scientific Advisory Panel
February 2-5, 2010

A.I. INTRODUCTION

A. Background

Atrazine, 2-chloro-4-(ethylamine)-6-(isopropylamine)-s-triazine, is currently one of the most widely used agricultural herbicides in the United States. During 2010, the Office of Pesticide Programs (OPP) is launching a comprehensive re-evaluation of atrazine. OPP's re-evaluation will be supported by two meetings of the FIFRA Scientific Advisory Panel (SAP) in April and September 2010. The April 2010 meeting will review findings from experimental studies, particularly those in laboratory animals and *in vitro* assays, and discuss sampling approaches to monitor atrazine levels in drinking water. The September 2010 meeting will build on the topics discussed during the April SAP and include a weight of the evidence of atrazine's health effects based on both toxicological and epidemiological findings.

OPP's evaluation of epidemiologic findings will be guided by the scientific principles outlined in its draft Framework for Incorporating Epidemiologic and Human Incident Data in Human Health Risk Assessment. As discussed in the draft Framework, there are several important factors to consider when evaluating epidemiologic studies including their goals, study population, characterization of exposure, ascertainment of disease, consideration of bias and confounding, and data collection, analysis, and documentation. Due to these factors, the size, scope, and quality of epidemiology studies can vary significantly. Ultimately these factors will determine the impact of epidemiologic findings on pesticide risk assessment and risk management decisions. Therefore, it is important to evaluate the strengths and limitations of epidemiologic studies when incorporating epidemiologic findings into risk assessment.

B. Case Study Objectives

The objective of the case study is to illustrate the types of strengths and limitations that should be considered when evaluating epidemiologic findings in risk assessment. While a broader epidemiological literature exists on atrazine, the case study focuses on six epidemiologic studies (Winchester, Huskins, and Ying, 2009; Mattix, Winchester, and Scherer, 2007; Ochoa-Acuña and Carbajo, 2009; Villanueva et al., 2005; Ochoa-Acuña et al., 2009; Mohanty and Zhang, 2009) which were conducted after OPP's 2003 Atrazine Interim Re-registration Eligibility Decision (IRED) (U.S. EPA, 2003). These six studies used either ecologic or retrospective cohort designs to examine the association between maternal atrazine exposure and several adverse birth outcomes, including birth defects, low birth weight (LBW), small-for-gestational age (SGA), and preterm delivery.

Although these six studies share many similarities in both design and research goals, they also have some important differences. These differences help illustrate the range of factors that must be considered when evaluating ecologic and retrospective studies, as well as epidemiology more generally. The remainder of the case study is organized into the following sections:

- ***Section A.II*** reviews the six epidemiologic studies that have been conducted since the 2003 Atrazine IRED and highlights their strengths and limitations.
- ***Section A.III*** then discusses the challenges of studying birth outcomes and general issues that should be considered when evaluating the six studies, as well as epidemiologic findings more generally.
- ***Section A.IV*** concludes with a summary of the studies' strengths and limitations and outlines OPP's future directions regarding its re-evaluation of atrazine.

During the February SAP, OPP will solicit comments from the SAP and public on the strengths and limitations of these studies, as well as the use of epidemiologic study findings in risk assessment more generally. Comments received on this case study will help advance both the re-evaluation of atrazine and broader needs of OPP as it expands its use of epidemiologic studies in risk assessment.

A.II. RECENT EPIDEMIOLOGIC STUDY FINDINGS ON ATRAZINE

Since OPP's 2003 Atrazine IRED, six epidemiologic studies have examined the relationship between maternal atrazine exposure and several different adverse birth outcomes, including birth defects, LBW, SGA, and preterm delivery. A summary of each study's research question(s), study design, metrics for exposure and adverse birth outcome, and findings is provided in Table A- 1**Error! Reference source not found.** below. Following this table, the remainder of the section provides more detailed reviews of each study. The reviews describe each study's objectives, design, and results and then discuss strengths and limitations that should be considered when incorporating the findings in risk assessment.

Table A- 1. Summary of recent studies evaluating the relationship between atrazine and adverse birth outcomes.

Review ID/ Study Title	Research Question(s)	Study Design	Exposure Assessment	Outcome Assessment	Findings
Review A: Agrichemicals in surface water and birth defects in the United States (Winchester, Huskins, and Ying, 2009)	Are annual peaks in pesticides and nitrates (April-July) correlated with an increase in birth defects?	Ecologic	USGS National Water Quality Assessment, 1996-2002	CDC National Natality Dataset, 1996-2002	A significant association was found between the season of elevated agrichemicals and birth defects. Specifically, the rates of 11 of 22 birth defect categories were significantly greater when conception occurred in months with the highest levels of atrazine, nitrates, and "pesticides."
Review B: Incidence of AWD is related to surface water atrazine and nitrate levels (Mattix, Winchester, and Scherer, 2007)	Are congenital AWD rates higher in Indiana than the U.S. and positively associated with surface water levels of atrazine and nitrates	Ecologic	Midwest surface water levels measure in USGS National Water Quality Assessment,	Indiana Birth Records Database and CDC National Natality Dataset, 1990-2002	The rate of birth defects in Indiana was significantly higher than the national rate in 1996, 1998, and 2001. Additionally, the Indiana monthly birth defect rate was positively correlated with the mean surface water atrazine levels in the Midwest.
Review C: Risk of limb birth defects and mother's home proximity to cornfields (Ochoa-Acuña and Carbajo, 2009)	What is the association between home proximity to corn and soybeans and specific birth defects during the crop growing season?	Retrospective Cohort	Proximity to corn and soy fields using spatial analysis of maternal address	Indiana Birth Records Database, 2000-2004	Risk of limb birth defects increased significantly in relation to proximity to corn fields (Adjusted OR=1.22; 95% CI=1.01, 1.47 per additional 10 ha planted with corn within 500 m), but not soy fields
Review D: Atrazine in municipal drinking water and risk of LBW, preterm delivery, and SGA status (Villanueva *et al.*, 2005)	Are LBW, preterm deliver, or SGA related to atrazine concentrations in drinking water?	Ecologic	Atrazine drinking water measurements taken from 112 water distribution systems in Finisteŕe, France, 1990-1998	Finisteŕe, France Birth Records, 1997-1998	Atrazine levels were not associated with an increased risk of LBW or SGA status. There was an increased risk of SGA status in cases in which the third trimester overlapped in whole or in part with the May–September period when atrazine levels were highest, compared with October to April (OR= 1.37, 95% CI: 1.04, 1.81).
Review E: Drinking water herbicide exposure in Indiana and prevalence of SGA and preterm delivery (Ochoa-Acuña *et al.*, 2009)	Is either preterm delivery or SGA related to atrazine concentrations in drinking water?	Retrospective Cohort	Atrazine time-series data from 19 drinking water systems in Indiana, 1993-2007	Indiana Birth Records Database, 1997-2007	Atrazine in drinking water during the third trimester and the entire pregnancy was associated with a significant increase in the prevalence of SGA. Atrazine in drinking water > 0.1 µg/L during the third trimester resulted in a 17–19% increase in the prevalence of SGA compared with the control group (< 0.1 µg/L). No significant association was found for preterm delivery.
Review F: Correlations of agrochemicals residues in drinking water and birth defects in Illinois (Mohanty and Zhang, 2009) [a]	What is the association between birth defects, adverse pregnancy outcomes, and preterm births with several drinking water contaminants?	Ecologic	Drinking water data from Illinois Community Water Systems, 1998-2002	Illinois Birth-Registry Database, 1998-2002	*Not Yet Published*

[a] OPP received an initial study report about this unpublished study and plans to conduct a review of the study when OPP is provided a report or when the results are published in the scientific literature. *Abbreviations*: **AWD** – Abdominal Wall Defects; **CDC** – Centers for Disease Control and Prevention; **CI** – Confidence Interval**; LBW** – Low birth weight; **OR** – Odds Ration; **SGA** – Small-for-gestational age; and **USGS** – U.S. Geologic Survey.

A. Review 1: "Agrichemicals in surface water and birth defects in the United States"

Winchester, Huskins, and Ying (2009) conducted an ecologic study to investigate if the rate of birth defects is greater in months with the highest concentrations of several surface water agrichemicals, including atrazine. National pregnancy and birth outcome data from 1996-2002 were obtained through the national natality database maintained by the Centers for Disease Control and Prevention (CDC).[5] Monthly surface water concentrations of atrazine, nitrates, and "other pesticides" were then estimated using data collected through USGS's National Water-Quality Assessment Program (NAWQA), which has collected water quality data from several major U.S. river basins and aquifers since 1991.[6] Based on these health outcome and drinking water monitoring sources, the investigators presented average monthly birth defect rates and geometric mean atrazine concentrations by aggregating monthly data for the year 1996- 2002, as shown in Figure A- 1.

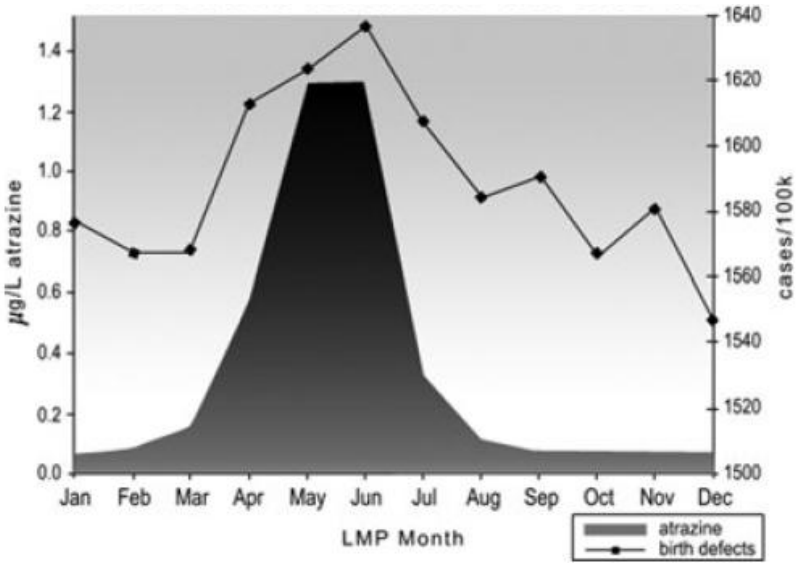


Figure A- 1. The United States mean birth defect rates by month of last menstrual period versus geometric mean atrazine concentrations. (Excerpted from Mattix, Winchester, and Scherer, 2007)

The investigators performed two analyses to explore the relationship between birth defects and agrichemical surface water concentrations. First, they compared the risk of 22 different birth defect types during April through July – the months with the highest levels of atrazine – with the rest of the year using a logistic regression model that adjusted for potential confounders, including race/ethnicity, metropolitan residence,

[5] CDC's natality database uses the United States birth-registration system and includes data on maternal risk factors and demographics and birth characteristics. The database is assessable at: http://wonder.cdc.gov/natality.html.

[6] The category "other pesticide" was not described in the study, so it is unclear what chemicals it represents.

alcohol use, tobacco use, maternal diabetes, and year. Second, they assessed the relationship between birth defects and atrazine, nitrates, and "other pesticides" over the entire calendar year using a logistic regression model that adjusted for the same potential confounders. The results of the first analysis are provided in Table A- 2. As shown, the rates of 11 different birth defect types were significantly higher in April-July when compared with all other months. Similarly, several positive associations were observed to be statistically significant in their second analysis (Table A- 3). In particular, the odds ratios for 9 of 11 birth defect types were associated with atrazine when their model treated atrazine separately from the other agrichemicals. Similarly, when they considered all three groups of agrichemicals in their logistic regression, 5 of 6 birth defects were positively associated with atrazine.

Table A- 2. Individual birth defects by month of last menstrual period (time of conception) (Excerpted from Winchester, Huskins, and Ying, 2009)

Birth Defect Type [a]	Mean (SE) Birth Defects per 100,000, Adjusted for Year [b]		
	April-July	Other Months	p-value
Spina	21.93 (0.5)	20.31 (0.34)	<0.01
Circul	134.99 (1.25)	131.09 (0.85)	<0.01
Cleft lip	83.09 (0.98)	79.07 (0.66)	<0.01
Adactyly	85.65 (1.00)	81.88 (0.68)	<0.01
Musculo	223.49 (1.61)	217.36 (1.11)	<0.01
Down	46.23 (0.74)	43.22 (-0.49)	<0.01
Other con	455.89 (2.33)	443.89 (1.59)	<0.01
Tracheo	13.33 (0.39)	12.32 (0.26)	<0.05
Gastro	32.22 (0.61)	30.82 (0.42)	<0.05
Urogen	105.37 (1.11)	102.54 (0.76)	<0.05
Clubfoot	58.17 (0.82)	56.23 (0.56)	<0.05
Anen	10.7 (0.35)	10.67 (0.24)	NS
Hydro	23.24 (0.52)	22.8 (0.36)	NS
Micro	6.3 (0.27)	6.08 (0.18)	NS
Nervous	20.97 (0.49)	20.98 (0.34)	NS
Heart	118.85 (1.18)	117.44 (0.82)	NS
Rectal	9.05 (0.32)	8.39 (0.22)	NS
Omphalo	28.89 (0.58)	28.1 (0.4)	NS
Genital	79.36 (0.96)	79.09 (0.66)	NS
Renalage	13.99 (0.4)	13.52 (0.28)	NS
Hernia	12.11 (0.38)	11.83 (0.26)	NS
Chromo	36.24 (0.65)	34.91 (0.44)	NS

[a] Complete names of birth defect types were not reported by authors.
[b] Values in cells are mean (standard error) adjusting for years.
NS = not significant with $p > 0.05$. p-values are obtained from logistic regression models.

Table A- 3. Odds ratio (OR) of selected individual birth defects in relation to atrazine, nitrates, and "other pesticides." (Excerpted from Winchester, Huskins, and Ying, 2009)

Birth Defect Type [a]	Simple Model [b]			Multiple Model [c]		
	Atrazine	Nitrate	Other Pesticides	Atrazine	Nitrate	Other Pesticides
Spina	1.023 (1.000,1.047)*	1.016 (0.903,1.143)	0.988 (0.949,1.029)	1.018 (0.988,1.050)	1.012 (0.883,1.160)	0.973 (0.928,1.020)
Circul	1.004 (0.995,1.013)	1.068 (0.893,1.151)	1.006 (0.990,1.023)	1.006 (0.994,1.018)	0.932 (0.882,0.986)	1.007 (0.988,1.027)
Tracheo	1.030 (1.001,1.061)*	0.959 (0.825,1.115)	1.069 (0.986,1.113)**	1.016 (0.978,1.056)	0.941 (0.790,1.122)	1.060 (1.001,1.094)*
Gastro	1.021 (1.003,1.041)*	0.974 (0.884,1.074)	0.985 (0.951,1.019)	1.024 (0.999,1.051)*	0.926 (0.825,1.040)	0.972 (0.933,1.012)
Urogen	1.007 (0.997,1.017)	0.735 (0.613,1.015)	1.021 (1.004,1.040)*	1.007 (0.994,1.021)	0.982 (0.923,1.044)	1.018 (0.957,1.038)
Cleft lip	1.021 (1.009,1.033)**	0.991 (0.933,1.053)	0.999 (0.978,1.020)	1.024 (1.009,1.040)**	0.960 (0.895,1.031)	0.983 (0.959,1.008)
Adactyly	1.022 (1.011,1.034)**	1.024 (0.965,1.087)	1.023 (1.003,1.045)*	1.023 (1.007,1.039)**	0.971 (0.906,1.042)	1.008 (0.984,1.032)
Clubfoot	1.016 (0.996,1.028)**	0.993 (0.924,1.067)	1.005 (0.980,1.031)	1.014 (0.995,1.033)	0.983 (0.903,1.071)	0.996 (0.967,1.025)
Musculo	1.015 (1.008,1.022)**	1.025 (0.988,1.064)	1.031 (1.018,1.045)**	1.008 (0.999,1.018)	1.004 (0.961,1.049)	1.024 (1.009,1.040)*
Down	1.021 (1.005,1.037)**	1.009 (0.930,1.096)	0.999 (0.971,1.028)	1.027 (1.005,1.049)*	0.982 (0.891,1.082)	0.980 (0.947,1.013)
Other con	1.010 (1.005,1.015)**	1.149 (1.120,1.178)**	1.031 (1.022,1.040)**	1.011 (1.002,1.025)**	1.177 (1.143,1.212)**	1.027 (1.016,1.037)**

[a] Complete names of birth defect types were not reported by authors.
[b] Values in cells are mean (95% confidence interval) of odds ratio (OR) in response to one unit increase of each agrichemical predictor (in log). The 'simple' logistic regression models use only one agrichemical predictor and are adjusted for maternal risk factors, maternal demographics and years.
[c] Values in cells are mean (95% confidence interval) of odds ratio (OR) in response to one unit increase of each agrichemical predictor (in log). The 'multiple' logistic regression models use all three agrichemical predictors and are adjusted for maternal risk factors, maternal demographics and years.
* Indicates $p < 0.05$; ** Indicates $p < 0.01$.

Overall, the study provides useful descriptive information on the seasonal patterns of both birth defects and the levels of several agrichemicals, including atrazine, in surface water. Specifically, the study demonstrated that monthly U.S. birth defect rates generally follow a seasonal pattern that peaks during late spring to early summer. Similarly, it also highlighted that the use of pesticides and fertilizers can lead to higher levels of residues in surface waters, some of which may subsequently be sources of public drinking water supplies. These findings help raise awareness in the public health community, help generate research hypotheses, and inform future birth defect research on atrazine and other risk factors that exhibit similar seasonal patterns.

Although the study reported positive associations between several types of birth defects and atrazine levels in surface water, it has limitations to consider when evaluating the results in the context of risk assessment. The most important limitations of the study relate to its ecologic design and use of USGS surface water monitoring data as an

exposure surrogate. In ecologic designs, the unit of analysis is at a group level, rather than individual level. Ecologic studies are frequently used to generate hypotheses in epidemiologic research, since they utilize existing data sources and are not resource intensive in terms of time and cost. A limitation of ecologic studies, however, is that findings at a group-level may not translate to individuals. For example, Winchester, Huskins, and Ying (2009) characterized exposure levels using monthly geometric mean estimates of the concentration of atrazine in U.S. surface water. As this measure of exposure is based on regional surface water data, rather than drinking water data, it cannot provide direct evidence that mothers who had children with birth defects were actually exposed to elevated levels of atrazine. Due to this limitation, the validity of the study must be confirmed through additional research that uses individual-level exposure data. The investigators do emphasize the limitations of their ecologic study design and indicate, "While a causal link between agrichemicals and birth defects cannot be proven from this study an association might provide clues to common factors shared by both variables."

B. Review 2: "Incidence of abdominal wall defects is related to surface water atrazine and nitrate levels"

Mattix, Winchester, and Scherer (2007) conducted an ecologic study first to determine if the rate of congenital abdominal wall defects (AWD) has been higher in Indiana than in the entire U.S. and then to characterize the association between the AWD rates and surface water atrazine and nitrate levels in the Midwest. The national incidence of AWD was determined using the Centers for Disease Control and Prevention's (CDC) National Natality Dataset for the years 1990 and 1995-2002. Similarly, the incidence of AWD in Indiana was determined using birth data collected by the Indiana State Department of Health on patients diagnosed with omphalocele or gastroshisis during 1990-2002. The investigators then characterized potential population-level exposure to atrazine and nitrates during 1990-2002 using surface water monitoring data from U.S. Geologic Surface (USGS). The USGS surface was not specifically collected from Indiana and represented the Midwest region of the U.S., which was defined by USGS as Indiana, Illinois, Iowa, and Nebraska.

The investigators performed a descriptive analysis comparing the rate of AWD in Indiana, the Midwest, and the U.S. As shown in Figure A-2, the annual rate of AWD in Indiana was generally higher than the entire Midwest and U.S. A chi-squared test was then used to statistically compare the annual rate of AWD in Indiana with the Midwest and U.S. rates. Based on this approach, the annual rate of AWD in Indiana was reported to be significantly higher than the U.S. rate in 1996, 1998, and 2001, but not for others years examined. Similarly, it was also reported that the annual rate of AWD in Indiana was significantly higher that the rate in the entire Midwest in 1998.

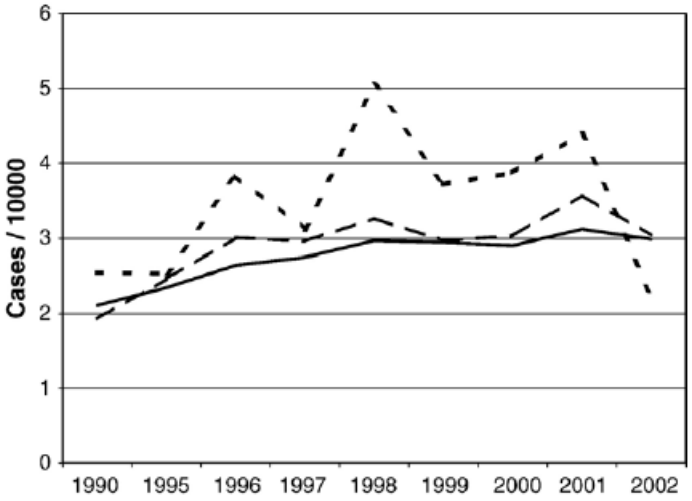


Figure A-2. : Incidence of AWD. Data obtained from the CDC natality set comparing the birth rate of AWD in Indiana (dotted line), in the Midwest (dashed line), and nationally (solid line) (Excerpted from Mattix, Winchester, and Scherer, 2007).

The investigators next evaluated the association between the monthly rate of AWD with monthly mean estimates of surface water levels of atrazine and nitrate in the Midwest. A descriptive analysis was performed and Pearson correlation coefficients were calculated to evaluate the linear association between the monthly AWD rate and mean monthly atrazine and nitrate levels. As shown in Figure A-3**Error! Reference source not found.** below, the monthly AWD rate in Indiana and surface levels of atrazine and nitrate were reported to follow similar seasonal patterns and exhibited peak values during the month of June. While both atrazine and nitrate surface water levels followed somewhat similar visual patterns, only the associations between the monthly AWD rate and atrazine levels was reported to be statistically significant. This correlation was observed when using both the CDC national natality data (0.60 Pearson correlation coefficient, p-value = 0.0392) and Indiana state registry data (0.69 Pearson correlation coefficient, p-value = 0.0125).

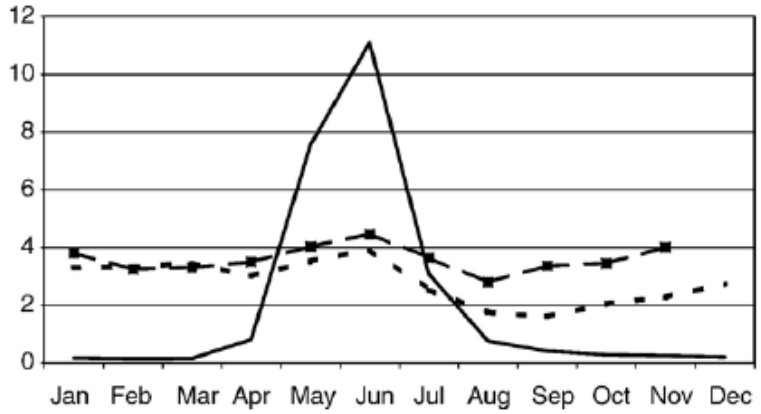


Figure A-3 Abdominal wall defect rates from Indiana birth certificate registry data compared with surface water nitrate and atrazine levels from 1990-2001. The AWD incidence by month of last menstrual period/estimated date of conception is represented in cases per 10,000 live births by the dotted line with square data points. Mean nitrate levels (in milligrams per liter) are shown by the dotted line, and mean atrazine levels (in micrograms per liter) are shown by the solid line. Peak incidence of each is seen in June (Excerpted from Mattix, Winchester, and Scherer, 2007)

Mattix, Winchester, and Scherer (2007) performed descriptive analyses of ADW trends in both Indiana and the U.S. These descriptive analyses provide information on the burden of AWD in Indiana relative to the Midwest and to the entire U.S., as well as evidence that both ADW rates and atrazine levels in Midwest surface waters follow the same seasonal trend. While the study provides useful descriptive information, the study did not consider any potential population-level confounders (e.g. demographic and behavioral risk factors). In addition, another important limitation of the study relates ecologic design. More detailed discussion of ecologic study limitations has been discussed in the review of Winchester, Huskins, and Ying (2009) which is provided in the previous section (*Section A*). Because both studies shared many similarities, including their coauthors, ecologic design and data sources, the general limitations discussed in *Section A* are relevant to both studies and not repeated in the review of Mattix, Winchester, and Scherer (2007).

C. Review 3: "Risk of limb birth defects and mother's home proximity to cornfields"

Ochoa-Acuña and Carbajo (2009) conducted a retrospective cohort study to assess the relationship between birth defects and maternal home proximity to corn and soy fields. The study population was identified using the Indiana Birth Records Database, which includes data on maternal demographics, behavioral risk factors, and health outcomes of babies delivered in Indiana. Using the Indiana Birth Records Database, the investigators defined their study population as (1) mothers living in rural geographic

boundaries based on the 2000 U.S. Census who had (2) singleton births conceived between May 1 and August 31 of the years 2000 through 2004. Based on these selection criteria, 48,216 births were included in the study population.

Births defect cases were then identified using standardized birth anomaly disease codes and exposure levels were assigned to each pregnancy using maternal home proximity to land with corn and soy fields as an indicator of agrichemical exposure. In order to estimate maternal home proximity to corn and soy fields, the investigators performed a spatial analysis of maternal address data and land crop cover data collected by the U.S. Department of Agriculture. Using this spatial analysis, the investigators classified the exposure of each birth record in their study population two ways. First, they created dichotomous variables in which births were defined as "exposed" to a given crop (corn or soy) if the planted crop area near a maternal address exceeded the median planted crop area for the entire study population. Second, they created continuous exposure variables by utilizing their spatial analysis to develop estimates of exposure expressed in terms of areas of corn and soy that were within 500 meters of each maternal address.

After assigning "exposure levels," the investigators performed two separate statistical analyses to assess whether there were associations between birth defects and their dichotomous and continuous exposure variables, respectively. In both statistical analyses, multivariate logistic regression models were used to calculate birth defect odds ratios, after controlling for potential demographic and maternal risk factors, including age, race, education, prenatal care, smoking, use of alcohol during pregnancy, and health status during pregnancy. The results of the analysis that utilized dichotomous exposure variables for corn and soy fields are provided in Table A- 4**Error! Reference source not found.** below. A higher number of birth defect cases were generally observed in the "exposed" groups for both corn and soybeans. However, the only statistically significant association was between limb birth defects and exposure to corn fields (Odds Ratio = 1.76, 95% CI: 1.12, 2.78). When maternal home proximity to corn and soy fields were treated as continuous exposure variables, a similar pattern of associations was also identified. In particular, it was reported that there was a 22% increase in the risk of limb birth defects per every 10 ha increase in planted area with corn (Odds Ratio = 1.22, 95% CI: 1.01, 1.47) (*See Table A- 5).*

Table A- 4. Number of birth defect cases and adjusted odds ratio estimates of living within 500m of more than 3.4 ha of cornfields; and of living within 500m of more than 2.4 ha of soybeans in Indiana (Excerpted from Ochoa-Acuña and Carbajo, 2009)

	Corn			Soybeans		
Birth Defect Type	**<3.4ha.**	**>3.4ha**	**Odds ratio (95% CI)**	**<3.4ha**	**>3.4ha**	**Odds ratio (95% CI)**
Abdominal cavity	16	23	1.50 (0.72,3.10)	19	20	1.09 (0.50,2.39)
Craniofacial	46	64	1.36 (0.88,2.10)	46	64	1.54 (0.96,2.45)
Heart	49	65	1.26 (0.82,1.93)	55	59	0.91 (0.58,1.45)
Limb	38	67	1.76 (1.12,2.78)	49	56	1.14 (0.71,1.82)
Neural tube	2	7	3.57 (0.62,20.5)	3	6	1.75 (0.33,9.27)
Other nervous system	16	15	0.91 (0.40,2.08)	15	16	1.49 (0.62,3.62)
Respiratory	41	66	1.40 0.88,2.23)	37	70	1.35 (0.82,2.21)
Urogenital	78	68	0.98 (0.67,1.44)	70	76	1.16 (0.77,1.74)

Note: Only singleton infants born outside city limits that were conceived during May–August of each year (2000 through 2004) in Indiana were included in this study. Sample size for each exposure/crop group was 24,108.

Table A- 5. Adjusted odds ratio estimates for specific birth defects in relation to the area within 500 m of home planted with corn or soybeans (Excerpted from Ochoa-Acuña and Carbajo, 2009)

	Corn		Soybeans	
Birth Defect Type	**Odds ratio**	**95% CI**	**Odds ratio**	**95% CI**
Abdominal cavity	0.84	0.57,1.25	1.08	0.77,1,52
Craniofacial	1.1	0.89,1.33	0.98	0.79,1.20
Heart	1.11	0.91,1.35	1.11	0.92,1.34
Limb	1.22	1.01,1.47	1.04	0.85,1.28
Neural tube	1.02	0.51,2.03	1.72	0.98,3.02
Other nervous system	1.03	0.69,1.54	1.11	0.75,1.63
Respiratory	1.06	0.87,1.30	1.17	0.97,1.40
Urogenital	0.83	0.67,1.03	1.15	0.96,1.38

Note: The odds ratios represent the change in odds of birth defects per a 10-ha increase in the area planted with either crop around the home. Only singleton infants born outside city limits conceived during May–August of each year (2000 through 2004) in Indiana were included in this study.

Compared with the previous two studies, Ochoa-Acuña and Carbajo (2009) conducted a more focused study that was able to estimate each birth's exposure potential. The ability to estimate individual-level exposure was a strength of the study; however, the use of a proximity to corn fields is only a surrogate of potential chemical exposure. Without actual exposure information, it is difficult to link the potential increase in limb birth defects to atrazine or any other specific agrichemical. The investigators address this issue in their discussion and suggest that exposure to agrichemicals applied preferentially to corn, including higher quantities of the fertilizer nutrients nitrogen, phosphorous, potassium and atrazine, may be alternative explanations of their findings. However, it cannot be determined if pregnant mothers were actually exposed to elevated levels of these chemicals without confirmatory exposure information. Therefore, it would be beneficial to collect environmental or human exposure data in follow-up research to verify that proximity to corn fields is associated with increased chemical exposure.

An additional limitation that makes it difficult to evaluate the study is its inclusion of only births that were conceived during May 1 to August 31. Focusing on the period of peak atrazine exposure allowed the investigators to enhance their ability to detect an association, as well as control for potential seasonal confounders. However, this approach also introduces uncertainty in the results because information is not provided on the association between birth defects and proximity to corn fields during the remainder of the year when exposure to atrazine and other agrichemicals may be lower. If the association between birth defects and proximity to corn fields is due to higher levels of maternal atrazine exposure, then the association between birth defects and proximity to corn fields should be weaker when there may be less potential for exposure during the remainder of the year. Therefore, it would also be beneficial to examine whether there is a relationship between birth defects and proximity to corn fields when there may be less potential for exposure to atrazine and other agrichemicals.

D. Review 4: "Atrazine in municipal drinking water and risk of low birth weight, preterm delivery, and small-for-gestational-age status"

Villanueva *et al.* (2005) conducted an ecologic study to evaluate the association between atrazine levels in municipal drinking water and the risk of preterm delivery, LBW, and SGA. The study focused on Finisteŕe, a region of France that is highly agricultural and uses a large amount of pesticides in corn production. Birth record data were identified for 9,721 live births in Finisteŕe during October 1, 1997 to September 30, 1998. All birth records data were obtained before infants were discharged from hospitals and included infant health information, maternal demographic characteristics, and municipality of maternal residence prior to birth.

Preterm delivery, LBW, and SGA were determined by comparing infant health information with a standard definition of preterm delivery and French population growth curves. Atrazine exposure levels for each birth were then classified using drinking water monitoring data from 2,661 atrazine measurements that were sampled from 112 water distribution units during 1990 to 1998. Using this atrazine monitoring data, an exposure index was created by calculating the geometric mean atrazine levels for each water distribution unit over the entire eight-year sampling period. Based on distribution of their exposure index, the investigators then created low, medium, and high exposure groups using tertiles as cut-offs for both raw (≤ 0.05, > 0.0.05– 0.075, > 0.0.075 μg/l) and treated water(<0.029, 0.029– 0.036, > 0.036 μg/l). After creating this exposure index, the individual birth record data were linked to the exposure index data by determining the water distribution unit that serves the municipality of each maternal residence.

The investigators performed two statistical analyses to evaluate the association between their measure of maternal atrazine exposure and risk of preterm delivery, LBW, and SGA. First, logistic regression was used to calculate odds ratios for each health outcome relative to the low exposure category, adjusting for several potential confounders, including infant sex, number of prenatal consultations, and parent occupational information. The results of the investigators study are presented in Table A- 6 below. As shown, no statistically significant associations were reported between atrazine and preterm delivery, LBW, or SGA. In their second analysis, the investigators considered the timing of potential exposure relative to each trimester of pregnancy because their descriptive analysis of the atrazine monitoring data showed a distinct peak during the late spring through summer (i.e. May to September). Logistic regression was used to calculate odds ratios of preterm delivery, LBW, and SGA for each trimester of pregnancy by comparing trimester periods that overlapped with any part of May-September with October-April. Based on this analysis, a significant association was identified between SGA and atrazine. In particular, as shown in Table A- 7, the investigators reported that births with third trimesters in May-September had a 37% increase in the risk of SGA when compared with births that had their third trimester in October-April (Odd Ratio = 1.37, 95% CI: 1.04, 1.81).

Table A- 6. Distribution of pregnancy outcomes by atrazine level in municipal drinking water, odds ratios (OR) and 95% confidence intervals (CI) from a logistic regression adjusted for maternal age, sex of the newborn, and percentage of samples below the detection limit from May to September (dichotomized at the median, 30% in raw and 60% in treated water) (Excerpted from Villanueva *et al.*, 2005)

Atrazine level from May to Sept. (μg/l)	Births	Preterm Delivery		Low Birth Weight		Small-for-gestational Age	
		Cases	OR (95% CI)	Cases	OR (95% CI)	Cases	OR (95% CI)
Raw Water [a]							
≤ 0.05	1262	51(4.0%)	1.00	71(5.6%)	1.00	101(8.0%)	1.00
> 0.0.05– 0.075	1050	47(4.5%)	1.34 (0.84, 2.14)	41(3.9%)	0.70 (0.44, 1.12)	72(6.9%)	0.81 (0.56, 1.18)
> 0.0.075	1198	39(3.3%)	1.15 (0.62, 2.13)	51(4.3%)	0.78 (0.45, 1.36)	68(5.7%)	0.65 (0.41, 1.03)
p-trend			0.638		0.397		0.066
Treated Water [a]							
<0.029	884	36(4.1%)	1.00	38(4.3%)	1.00	65(7.4%)	1.00
0.029– 0.036	1079	37(3.4%)	1.22 (0.73, 2.06)	53(4.9%)	1.29 (0.77, 2.15)	65(6.0%)	0.96 (0.62, 1.48)
> 0.036	1144	50(4.4%)	1.93 (0.85, 4.35)	52(4.5%)	0.92 (0.45, 1.86)	70(6.1%)	0.97 (0.53, 1.79)
p-trend	0.126		0.729		0.937		

[a] Total numbers do not add up to 3510 due to missing values. Availability of atrazine levels was lower for treated than raw water.

Table A- 7. Adjusted OR for the outcomes studied, by trimesters of pregnancy that overlapped with any part of May-September compared with October-April (Excerpted from Villanueva *et al.*, 2005).[a]

	Preterm Delivery		Low birth weight		Small-for-gestational-age	
Timing	**OR (95%CI)**	**Cases**	**OR (95%CI)**	**Cases**	**OR (95%CI)**	**Cases**
First trimester						
October–April	1.00	47	1.00	66	1.00	118
May–September	1.36 (0.95, 1.95)	89	0.95 (0.68, 1.32)	87	0.73 (0.56, 0.96)	121
Second trimester						
October–April	1.00	57	1.00	72	1.00	117
May–September	1.09 (0.77, 1.54)	79	0.88 (0.63, 1.21)	81	0.82 (0.63, 1.06)	120
Third trimester						
October–April	1.00	61	1.00	55	1.00	83
May–September	0.83 (0.58, 1.18)	69	1.24 (0.88, 1.74)	93	1.37 (1.04, 1.81)	155

[a] Obtained from logistic regression adjusting for sex of the newborn, maternal age, and geometric mean atrazine levels.

In contrast to the previous studies, Villanueva *et al.* (2005) utilized atrazine measurements from drinking water sampling to estimate exposure. Although the use of drinking water data would appear to be a higher quality measure of exposure, the number of measurements was not sufficient to characterize the seasonal concentration of atrazine in each water distribution systems during 1997-1998. Because there was not sufficient monitoring data from 1997-1998 alone, the investigators created an index of exposure based on sampling data from 1990-1998 even though their study population only included infants born in 1997-1998. This introduces considerable uncertainty into the study, as it is likely that their index of exposure may not reliably estimate all individual's level of atrazine exposure. As a result, the investigators measure of exposure may potentially make it more difficult to observe associations between atrazine and preterm delivery, LBW and SGA. In addition, other studies have highlighted that Villanueva *et al.* (2005) observed relatively small differences between the tertiles of exposure that were used to define their low, medium, and high "exposure" categories (Ochoa-Acuña *et al.*, 2009). As a result, there may actually be minimal biological differences in the different exposure groups. If this is the case, it may be difficult to observe significant differences, because the low, medium, and high exposure groups may actually represent relatively similar exposed populations.

E. Review 5: "Drinking water herbicide exposure in Indiana and prevalence of small-for-gestational-age and preterm delivery"

Ochoa-Acuña *et al.* (2009) conducted a retrospective study to assess the association between atrazine exposure and the prevalence of preterm delivery and SGA births among women in Indiana. A key feature of their study was the use of drinking water

monitoring data to estimate exposure to atrazine. In Indiana, drinking water monitoring data were available through four data sources: (1) Safe Drinking Water Act IDEM system quarterly measurements; (2) Acetochlor Registration Partnership (Hackett *et al.* 2005); (3) Novartis Atrazine Public Water System Voluntary Monitoring (Tierney *et al.* 1999); and (4) Atrazine and Simazine Re-registration Program (U.S. EPA, 2006). Using these data sources, the investigators re-constructed average (monthly) atrazine concentrations in finished drinking water by interpolating atrazine concentrations between the 7-14 day sampling intervals in the summer months and 1-6 month sampling intervals in the winter months. These re-constructed atrazine concentration profiles were then averaged over the exposure period of interest to estimate exposure (i.e., first month of pregnancy, last month of pregnancy, third trimester, or entire pregnancy).

The investigators restricted their study area to community water systems in urban areas that had well defined service boundaries and sufficient drinking water data to re-construct atrazine concentrations over the exposure periods of interest. Based on these restrictions, the investigators identified 19 community water systems with sufficient atrazine monitoring data. Infant and maternal characteristics and health outcome information was then identified using the Indiana Birth Record Database. Mother's residence at the time of delivery was used to match each mother to a specific community water system within the state. A spatial analysis was then performed to link birth records to the specific community water systems that supplied water to the maternal address. Based on this approach, a total of 25,154 births were linked to one of the 19 community water systems with atrazine concentration data. The number of births within each community water system varied considerably based on the populations served and number of years of atrazine data, although roughly 70% of the birth records came from the Ft. Wayne community water system.

The prevalence of preterm delivery and SGA were compared between mothers who resided in community water systems with low, medium or high concentration of atrazine in drinking water; categories were defined using the $<25^{th}$, $25\text{-}75^{th}$, $>75^{th}$ percentiles of the continuous distribution of atrazine concentration distribution during either the first month, last month, last trimester, or entire pregnancy period, depending on the research question. The then investigators then used a log-binomial model to calculate prevalence ratios after adjusting for several potential confounders, including various maternal and infant demographics characteristics, behavioral risk factors, and potential seasonal patterns in pregnancy outcomes. Based on this approach, the investigators reported that there was an association between atrazine concentration in drinking water and the prevalence of small-for-gestation age infants, but not for preterm delivery

(*See* Table A- 8). When the third trimester was used as the period of exposure, the reported prevalence of SGA in the medium and high exposure groups were 19% and 17% higher than the prevalence reported for the low exposure group, respectively. Similarly, when entire pregnancy period was used as the exposure period, the reported prevalence of SGA in the high exposure group was 14% higher than the low exposure group (Adjusted prevalence ratio = 1.14, 95% CI: 1.03, 1.24).

Table A- 8. Prevalence of preterm delivery and SGA in relation to mean level of atrazine in drinking water (µg/L) and adjusted PRs (95% CI) for comparisons between medium (≥ 25th, ≤ 75th percentiles), and high (> 75th percentile) and the control exposure group (< 25th percentile) (Excerpted from Ochoa-Acuña *et al.*, 2009).

Response	Atrazine exposure group (µg/L)[a]	Within-group percentiles			No. of births	Gestation length (weeks ± SD) or birth weight (g ± SD)	No. of preterm/ SFA cases	Preterm/SGA prevalence (CI)	Preterm/SGA adjusted PR (CI)[b]
		25th	50th	75th					
Preterm delivery									
First month	<0.057	0.001	0.020	0.050	4,995	38.9 ± 2.02	358	7.17 (7.07, 7.26)	
	0.057 -0.435	0.100	0.165	0.256	10,072	38.8 ± 1.95	736	7.31 (7.24, 7.37)	0.98 (0.87, 1.11)
	>0.435	0.655	1.121	1.781	5,034	38.8 ± 1.96	402	7.99 (7.88, 8.09)	1.07 (0.93, 1.22)
Last month	<0.057	0.001	0.037	0.050	5,407	38.9 ± 2.00	393	7.27 (7.18, 7.36)	
	0.057–0.507	0.100	0.180	0.281	10,889	38.8 ± 1.96	818	7.51 (7.45, 7.58)	1.04 (0.93, 1.18)
	>0.507	0.768	1.227	1.884	5,443	38.8 ± 1.95	409	7.51 (7.42, 7.61)	0.87 (0.72, 1.04)
SGA									
Third trimester[c]	<0.103	0.001	0.045	0.050	4,363	3,309 ± 528	479	11.0 (10.8, 11.1)	
	0.103–0.835	0.117	0.210	0.326	8,747	3,268 ± 526	1,251	14.3 (14.2, 14.4)	1.19 (1.08, 1.32)
	>0.835	0.872	1.116	1.482	4,373	3,276 ± 514	575	13.1 (13.0, 13.3)	1.17 (1.03, 1.34)
Entire pregnancy	<0.179	0.001	0.047	0.107	6,038	3,284 ± 568	723	12.0 (11.8, 12.1)	
	0.179–0.644	0.277	0.363	0.491	12,078	3,273 ± 534	1,609	13.3 (13.2, 13.4)	1.06 (0.98, 1.15)
	>0.644	0.740	0.822	0.950	6,038	3,237 ± 543	840	13.9 (13.8, 14.1)	1.14 (1.03, 1.24)

[a] Values listed correspond to < 25th, 25th–75th, and > 75th percentiles.
[b] Adjusted for mother's ethnicity, level of education, month prenatal care began, smoking status, and quarter of the year in which baby was conceived.
[c] Excludes records from preterm deliveries (< 37 weeks' gestation).

The investigators conducted a well-designed epidemiologic investigation on atrazine and the birth outcomes preterm delivery and SGA. The study shares many similarities with the study reviewed previously (Villanueva *et al.*, 2005), but utilized higher quality atrazine monitoring data that allowed the investigators to perform a more robust statistical analysis. Additionally, the study adjusted for a greater number of potential

confounders and considered seasonal trends in adverse birth outcomes. An additional strength of the study was that many of the risk factors that were treated as confounders had associations with adverse birth outcomes that were consistent with the scientific literature. This increases the reliability of the investigators' approach, since their results identify many factors that are commonly accepted as risk factors of SGA and preterm delivery.

Although the study used a more refined measure of exposure than the previous studies, additional information regarding the method of interpolation of the atrazine drinking water concentrations could be helpful to fully understand the study's results. For example, the statistical error involved in interpolation of average monthly concentrations of atrazine in drinking water using measurements 7-14 days apart in the spring and summer months is likely different than that involved in attempting to interpolate monthly atrazine concentrations in drinking water between measurements taken 1-6 months apart in the winter months. While the possible bias due to misclassification due to interpolation may be low, especially in the assignment of low, medium, high exposure categories, the investigators did not comment as to the impact of the varying data quality over the course of the exposure periods. In addition, investigators did not comment on the quality and comparability of the four drinking water monitoring programs utilized to re-construct atrazine drinking water concentration over the study period. Ideally, a comparison of detection methods, percent coefficient of variability between and within the separate programs, limit of detection, and other measures of quality control would be summarized.

F. Review 6: "Correlations of agrochemicals residues in drinking water and birth defects in Illinois"

The Agency's Office of Pesticide Programs has reviewed an initial study report and PowerPoint presentation from the 2009 Illinois Sustainable Technology Center Research Symposium (Mohanty and Zhang, 2009). Based on this available information, it appears that the study investigators conducted an ecologic study to compare the correlation between several drinking water contaminants and adverse birth outcomes at the county-level in Illinois. The water contaminants evaluated in the study included nitrate, nitrite, atrazine, and several trihalomethane and haloacetic acid disinfectant byproducts. The level of these water contaminants at the county-level were compared county-level rates of several adverse birth outcomes, including birth defects, LBW, and adverse pregnancy outcomes A more substantive review of the study's design, methods, and results will be prepared when the investigators complete a final study report.

A.III. DISCUSSION

This section discusses more general methodological issues that are important to consider when evaluating epidemiologic studies in risk assessment. Although the section focuses on issues that are related to the six atrazine studies reviewed in the previous section, many of the issues discussed below are relevant to the OPP's evaluation of epidemiologic studies more generally. Thus, guidance from the SAP is sought on the proper use of epidemiologic studies in general, as well as with respect to the atrazine re-evaluation.

A. Study Design Considerations

The epidemiologic studies reviewed in this case study focused on adverse birth outcomes, including birth defects, preterm delivery, LBW, and SGA. Many of these birth outcomes are relatively rare health outcomes. Studying rare health outcomes can be challenging in epidemiologic research because large study populations are often needed to identify a sufficient number of diseased individuals to be able to perform comparisons with sufficient statistical power. Due to the logistics and costs of recruiting and monitoring a large cohort, however, it is not typically feasible to routinely conduct prospective epidemiologic studies. As a result, epidemiologists often use retrospective study designs to study adverse birth outcomes and other rare health outcomes.

Retrospective studies enable researchers to identify large enough populations to study rare events, but are often considered weaker than prospective studies because they are limited to existing data sources. For example, the U.S. studies reviewed in this case study are based on birth outcome data from state-based birth registries. Because the U.S. does not have a uniform, nationwide birth registry, state-based registries collect birth outcome data differently. This makes it difficult to perform state-by-state comparisons (National Birth Defects Prevention Network, 2008). Similarly, birth registries often only include data that is collected immediately following delivery and may provide incomplete. As a result, it is possible that some adverse birth outcomes, particularly birth defects, may not be diagnosed until after neonates are discharged from the hospital after delivery (Weinhold, 2009), resulting in underreporting of adverse birth outcomes.

With regard to retrospective designs, the studies reviewed in this case study used either ecologic or retrospective cohort designs. As previously discussed in *Section A.II.A*, epidemiologic studies are considered ecologic if they evaluate disease-exposure relationships at the population-level. Since population-level relationships may not be valid at the individual-level, ecologic studies are generally considered one of the weaker designs in epidemiology. However, it should be noted that ecologic studies can help generate hypotheses that may be shown to be valid after further research. The

remaining studies used retrospective cohort study designs and were able to evaluate disease-exposure relationships at the individual-level. Retrospective cohort studies and other individual-level study types are generally stronger than the ecologic studies and should be given greater relative importance when evaluated in risk assessment.

B. Retrospective Exposure Assessment and Bias

Exposure assessment can be an important source of uncertainty in retrospective epidemiologic studies. As discussed in the draft Framework, exposure in retrospective studies is often estimated indirectly using surrogates of exposure, including historical information, questionnaires, and existing environmental monitoring data. The use of these indirect methods, rather than more direct measures of exposure, can lead to exposure misclassification – meaning that individuals are assigned an exposure classification that does not represent their true exposure.

In many retrospective studies, there may be no information to suggest that exposure misclassification is systematically different in diseased individuals compared to non-diseased individuals (*i.e.*, non-differential exposure misclassification). When this is the case, exposure misclassification makes it more difficult to observe associations between exposure and disease (*i.e.*, weaken the strength of association by biasing the results towards the null value of 1), but should not affect the validity of positive findings. Alternatively, exposure misclassification may be systematically different between diseased and non-diseased individuals (*i.e.*, differential exposure misclassification). Differential misclassification is often a greater concern than non-differential misclassification because it can affect the validity of study results by increasing the likelihood of observing false positive or false negative associations. With regard to the studies reviewed in *Section A.II*, there is no strong evidence to suggest that they were subject to differential misclassification. This is because the studies estimates exposure using monitoring data which are generally less prone to differential bias than other indirect approaches that require study participants to recall past exposures. Therefore, it is unlikely that the exposure assessment approaches used in these studies increased their likelihood of observing false associations.

In addition to introducing bias in epidemiologic studies, surrogates of exposure are potentially more difficult to quantitatively integrate into risk assessment than direct measures of exposure. In Winchester, Huskins, and Ying (2009), for example, the investigators used two different surrogates of exposure to atrazine, nitrates, and other agrichemicals. These exposure surrogates included (1) months with potential peak exposure and (2) geometric mean atrazine, nitrates, and other agrichemical levels in U.S. surface waters. Both of these measures of exposure cannot easily be evaluated in human health risk assessment, as neither can be used to directly estimate daily

chemical intake. Similarly, Ochoa-Acuña and Carbajo (2009) used living in proximity to corn and soy fields as surrogates for potential agrichemical exposure. In order to estimate exposure based on proximity to corn or soy fields, additional data would likely be needed on pesticide applications rates, meteorological conditions, and other environmental factors that influence the fate and transport of pesticides. The one study that used exposure data that could be more directly incorporated into human health risk assessment was Ochoa-Acuña *et al.* (2009). In this study, the investigators developed atrazine time series profiles for several community drinking water systems using drinking water monitoring data. This type of data can potentially be combined with water consumption data to estimate daily atrazine intake over different periods of time, depending on the health effect of concern. Therefore, of the studies reviewed, it may be the most easily integrated into quantitative risk assessment.

C. Confounding and Seasonal Patterns

An important consideration in any epidemiological study is the influence of factors that may act as confounders by obscuring disease-exposure relationships. Most epidemiologic studies are able to address many common factors that may act as confounders, including demographic and behavioral risk factors. The majority of studies reviewed previously in *Section A.II*, for example, controlled for a wide range of demographic and maternal risk factors, including age, race/ethnicity, tobacco and alcohol use, and health status during pregnancy.

While most epidemiologic studies consider many common confounders, less conventional confounders can be more difficult to address in epidemiologic studies. When studying adverse birth outcomes, one potential source of confounding is seasonal factors that may be associated with the rate of both birth and adverse birth outcomes. The annual pattern of birth in the U.S. has been shown to be elevated during August-September and depressed during April-May. Similarly, it has also been shown that different demographic and socioeconomic subpopulations have distinct seasonal birth patterns. Darrow *et al.* (2009), in particular, recently found that college-educated mothers were more likely to give birth in the spring than less-educated mothers and the largest depressions in birth rates during April-May were observed in mothers who were unmarried, hispanic or non-hispanic black, and had less than a high school education. Based on their findings, Darrow *et al.* (2009) suggest that seasonal birth patterns amongst different subpopulations can be a source of confounding that has interpretational implications.

Although not reviewed extensively for this case study, seasonal trends in adverse birth outcomes have also been reported in other epidemiologic studies. Moreover, the first two studies reviewed (Mattix, Winchester, and Scherer, 2007; Winchester, Huskins, and

Ying, 2009) also suggest that birth defect rates in the U.S. follow a seasonal pattern where the rate peaks for births conceived late-spring to early-summer. This peak in the birth defect rate may be attributable to increased atrazine exposure, but it may also be associated with other covariates that follow the same seasonal pattern. Addressing potential seasonal confounders may be challenging because it is unlikely that data on other seasonal covariates are readily available, particularly in retrospective studies. An alternative approach is to address seasonal confounding in a study's analysis phase. In the study reviewed in *Section A.II.E* (Ochoa-Acuña *et al.*, 2009), for example, the investigators included the quarter of the year in which conception occurred in their statistical model to account for potential seasonal patterns. This approach may also be challenging, given that seasonal confounder may not follow an obvious pattern, such as the quarterly pattern considered in the analysis performed by Ochoa-Acuña *et al.* (2009).

A.IV. CONCLUSION

A. Summary of Findings

The epidemiologic studies reviewed in this case study used existing birth registry data and different types of environmental data to examine the potential associations between atrazine exposure and several adverse birth outcomes. The studies varied in overall quality and had different strengths and limitations, as summarized in the Table A- 9 below. The first two studies (Mattix, Winchester, and Scherer, 2007; Winchester, Huskins, and Ying, 2009) provided a snapshot of overall trends in both birth defect rates and atrazine surface water concentrations the U.S. Due to their ecologic measure of exposure, however, they could not provide direct evidence that that mother's who had infants with birth defects were more likely to be exposed to elevated levels of atrazine due to their ecologic design. These studies were followed by Ochoa-Acuña and Carbajo (2009), which used living in proximity to corn fields as potential surrogate of exposure to atrazine and other agricultural chemicals. The use of this surrogate exposure data enabled the investigators to classify exposure at the individual-level, but it also had limited specificity to atrazine or other risk factors that may be associated with proximity to corn fields.

The final studies reviewed in the case study used actual drinking water monitoring data to characterize potential atrazine exposure. Villanueva *et al.*, 2005 had less detailed monitoring data and could only assign a single exposure classification to each municipality that was included in the study. Additionally, atrazine usage and exposure patterns in France may differ from patterns in the U.S. This and other study population differences between Villaneuva *et al.* (2005) and the U.S. studies may be an important issue to consider. Finally, Ochoa-Acuña *et al.* (2009) utilized more robust monitoring data from four separate sources to develop atrazine time series profiles for 19

municipalities in Indiana. The incorporation of atrazine drinking water monitoring data enabled the researchers to assign individual-level exposure levels. Therefore, Ochoa-Acuña *et al.* (2009) appears to be the strongest of the studies for the purposes of informing the atrazine re-evaluation.

Table A- 9. Summary of strengths and limitations of each atrazine study

Study	Strengths	Limitations
Study A: Incidence of abdominal wall defects is related to surface water atrazine and nitrate levels (Mattix, Winchester, and Scherer, 2007)	▪ Described seasonal AWD trends in Indiana and the U.S. ▪ Highlighted that peak concentrations of atrazine correlated closely with seasonal trends in AWD ▪ Useful in hypothesis generation	▪ Ecologic study design can only demonstrate correlation ▪ Measure of exposure based on surface water measurements that may not reflect actual exposure
Study B: Agrichemicals in surface water and birth defects in the United States (Winchester, Huskins, and Ying, 2009)	▪ Described seasonal birth defect trends in the U.S. ▪ Highlighted that peak concentrations of atrazine correlated closely with seasonal trends ▪ Useful in hypothesis generation	▪ Ecologic study design can only demonstrate correlation ▪ Measure of exposure based on surface water measurements that may not reflect actual exposure
Study C: Risk of limb birth defects and mother's home proximity to cornfields (Ochoa-Acuña and Carbajo, 2009)	▪ Used proximity to corn fields as a novel measure of individual exposure ▪ Helped control for potential seasonal confounders by focusing on period of peak exposure	▪ Surrogate exposure measure cannot be directly linked to atrazine, although may be sufficient to accurately rank study participants ▪ Results may only be generalizable to rural births conceived in Spring/Summer period
Study D: Atrazine in municipal drinking water and risk of low birth weight, preterm delivery, and small-for-gestational-age status (Villanueva *et al.*, 2005)	▪ Utilized drinking water monitoring atrazine data to assign exposure levels ▪ Evaluated risk of birth outcomes during different trimesters of pregnancy	▪ Relied on a single estimate of exposure for each municipality because there was insufficient monitoring data to develop time series exposure estimates ▪ Relatively small differences in exposure between low, medium, and high exposure groups ▪ Because the study was conducted in France, there may be population differences to consider when generalizing the results to the U.S
Study E: Drinking water herbicide exposure in Indiana and prevalence of small-for-gestational-age and preterm delivery (Ochoa-Acuña *et al.*, 2009)	▪ Utilized multiple sources of drinking water monitoring data to assign exposure using atrazine time-series profiles ▪ Controlled for multiple confounders, including seasonality ▪ Findings on other risk factors agreed with published literature	▪ No comparison of different drinking water data sources is provided ▪ Study results are driven by data from a single community water systems that represented 70% of all births

The strengths and limitations of these studies, as well as broader methodological issues described in *Section A.III*, are intended to illustrate challenges that the Agency must consider when evaluating the epidemiologic findings in the context of risk assessment. Key challenges illustrated by these studies include:

- Ecologic studies generally cannot provide strong etiologic information on disease-exposure relationships in the context of human health risk assessment. However, ecologic studies are generally less resource intensive than other types of studies and can help raise awareness in the public health community and generate research hypotheses. Therefore, ecologic studies can help identify potential hazards to consider in the risk assessment process and encourage future research.

- Epidemiologic studies on rare health outcomes often utilize retrospective study designs. Because retrospective studies rely on existing data sources to estimate disease burden and exposure levels, they may not be able to address confounding and bias as well as prospective studies. For this reason, OPP closely follows prospective studies, including the Agricultural Health Study and cohorts being studied by the Children's Environmental Health Centers which are supported by U.S. EPA and the National Institute of Environmental Health Science.

- Exposure assessment can be an important area of uncertainty in epidemiologic studies. The use of surrogates of exposure, including questionnaires and proximity to potential sources of exposure, may be subject to exposure misclassification and may be more difficult to integrate into quantitative risk assessment. On the other hand, the use of drinking water monitoring data and other sources of more direct exposure data can strengthen epidemiologic studies and be more readily integrated in risk assessment.

B. Future Directions

In general, epidemiologic findings can be used qualitatively in the risk assessment process to evaluate concordance and help identify toxicological effects that are relevant to humans. While the atrazine studies reviewed in this case study highlight some general challenges, they also provide a number of findings that are relevant to the Agency's 2010 re-evaluation of atrazine. As such, the reported findings will be considered when evaluating the scientific literature on the health effects of atrazine. Moving forward, the Agency has planned additional meetings of the FIFRA SAP in April and September 2010. The April 2010 meeting will include a review on the findings of

experimental laboratory studies and the September 2010 meeting will integrate findings from the laboratory studies and epidemiologic studies using a weight-of-the-evidence approach which incorporates the best available science.

A.V. REFERENCES

Darrow L.A., Strickland M.J., Klein M., Waller L.A., Flanders W.D., Correa A., Marcus M., and Tolbert P.E. 2009. Seasonality of Birth and Implications for Temporal Studies of Preterm Birth. Epidemiology. 20(5): 699-706.

Hackett A.G., Gustafson D.I., Moran S.J., Hendley P., Van Wesenbeeck I., Simmons N.D., et al. 2005. The acetochlor registration partnership surface water monitoring program for four corn herbicides. Journal of Environmental Quality. 34:877-889.

Mattix K.D., Winchester P.D., and Scherer L.R. 2007. Incidence of abdominal wall defects is related to surface water atrazine and nitrate levels. Journal of Pediatric Surgery. 42, 947-949.

Mohanty M.K. and Zhang B. 2009. Correlations of agrochemicals residues in drinking water and birth defects in Illinois. Presentation at the Illinois Sustainable Technology Center Research Symposium; September 9, 2009. Available: http://www.istc.illinois.edu/research/symposium.cfm [Accessed 7 December 2009].

Ochoa- Acuña H., Carbajo C. 2009. Risk of limb birth defects and mother's home proximity to cornfields. Science of the Total Environment. 407: 4447-4451.

Ochoa-Acuña H., Frankenberger J., Hahn L., and Carbajo C. 2009. Drinking water herbicide exposure in Indiana and prevalence of small-for-gestational-age and preterm delivery. Environmental Health Perspectives. 117(10): 1619-1624.

Tierney D.P., Clarkson J.R., Christensen B.R., Hines N.A. 1999. Exposure to the herbicides atrazine and simazine in drinking water. Abstr Pap Am Chem Soc. 217:080-AGRO.

U.S. EPA (U.S. Environmental Protection Agency) (2006). 2003 Revised Atrazine Interim Re-registration Eligibility Decisions. Available: http://www.epa.gov/opp00001/reregistration/atrazine/ [Accessed 7 December 2009].

U.S. EPA (U.S. Environmental Protection Agency). 2006. Triazine Cumulative Risk Assessment. Available: http://www.epa.gov/oppsrrd1/REDs/triazine_cumulative_risk.pdf [Accessed 7 December 2009].

Villanueva C., Durand G., Coutte M., Chevrier C., and Cordier S. 2005. Atrazine in municipal drinking water and risk of low birth weight, preterm delivery, and small-for-gestational-age Status. Environmental Health Perspectives. 2005, 62(6): 400-405.

Winchester P.D., Huskins J., and Ting J. 2009. Agrichemicals in surface water and birth defects in the United States. Acta Paediatrica. 98(4): 664-669.

Attachment B

Office of Pesticide Programs

& The Agricultural Health Study

Comparison of Exposure Assessment Approaches

Office of Pesticide Programs
Office of Pollution Prevention and Toxic Substances
U.S. Environmental Protection Agency

FIFRA Scientific Advisory Panel
February 2-5, 2010

The Environmental Protection Agency's (the Agency) Office of Pesticide Programs (OPP) has developed a draft framework to incorporate the results of epidemiological research and human incident data into the risk assessment process. As part of this effort, the Agency is developing three case studies. Attachment A illustrates the application of the science considerations pertinent to evaluating retrospective and ecological study designs in the context of human health risk assessment. Attachment C will be submitted to the Panel as an addendum and will provide an analysis of human incident data.

This document describes a proposed case study to use results from a prospective epidemiologic study to inform the risk assessment process. The case study will use information from the Agricultural Health Study (AHS) (http://aghealth.nci.nih.gov/). The AHS is an extensive, high quality epidemiologic study which collects information on pesticide exposure. The AHS cohort includes approximately 90,000 pesticide applicators and their spouses in two major areas of pesticide use (i.e., Iowa and North Carolina). To date, AHS investigators have authored over 100 articles in the peer reviewed literature on many topics including ground-breaking work on cancer and non-cancer disease outcomes and exposure metrics. The AHS is therefore an excellent basis for a case study because of the robust methods used, the size of the selected cohort, the quality of the data, and the direct applicability of those studied in the cohort to routine risk assessments for pesticide users developed by the Agency[7].

This specific case study analysis will focus on exposure assessments for people who mix, load, or apply pesticides (handlers). The Agency has well-developed, transparent, and peer reviewed methods for estimating occupational handler exposures. (http://www.epa.gov/scipoly/sap/meetings/2007/010907_mtg.htm) . These methods, however, differ from the approaches used in the AHS and in epidemiological analysis, in general, to evaluate cohort participants' exposures. The AHS has developed peer review methods for exposure classification in the epidemiological study. The general differences in scope and purpose of exposure assessment in observational epidemiologic investigation and quantitative risk assessment have been well articulated in the scientific literature (e.g., van den Brandt et al.; Food and Chem. Tox.; 40:387-424, 2002). In brief, the primary goal of the Agency's occupational handler assessments is to evaluate the high end of exposure distributions associated with specific application scenarios for a particular pesticide. EPA methods result in a single point estimate of anticipated exposure, based upon certain occupational handler activities. In these evaluations, the Agency assesses the impact on the potential exposure of different variables such as, using various kinds of personal protective equipment and/or engineering controls, as well as applicable label rates and acres treated per day. In contrast, the goal of epidemiologic exposure assessment within the AHS is to develop a

[7] The AHS has several inherent mechanisms in place to protect the quality of the data from this study and the privacy/rights of the participants. The Agency will abide by all of these criteria to ensure full compliance with these processes.

relative exposure ranking of individuals who are actual pesticide users within a cohort. It is not practicable or feasible to directly measure actual exposure in observational analyses such as the AHS. The AHS exposure information is ascertained from questionnaires completed out by individual cohort members. The AHS has documented the quality of self-reported information and has performed field measurement sub-studies to assess the exposure intensity algorithm.

Because the AHS and the Agency have different purposes for evaluating pesticide applicator exposure, there are inherent differences in the occupational handler exposure methodologies between the AHS and Agency. Whether or not these differences could lead to differences in conclusions regarding the risk of adverse health effects of exposure is under investigation by a collaborative effort between EPA's OPP, Office of Research and Development (ORD), the National Cancer Institute (NCI), and the National Institute of Environmental Health Sciences (NIEHS). The case study will compare the exposure algorithms and input data used by both AHS and also the Agency in their respective occupational handler exposure assessments. The multi-agency collaborative effort is proposed for three steps.

1. The first step involves a side by side evaluation of each exposure determinant used by the AHS and Agency approaches to pesticide applicator exposure assessment to inform where differences may be present.

2. The second step will build on the initial side by side comparison by evaluating exposure and biomonitoring data for a subset of pesticide applicators from the AHS cohort (Bakke et al, 2008; Thomas et al, 2009). The Bakke et al (2008) and Thomas et al (2009) studies are unique for AHS in that they provide exposure information on individual applicators in that cohort. These data will allow the comparison of exposure estimates developed for individuals by the Agency and AHS approaches with calculations of actual exposure derived from the biomonitoring data from the same individuals. This three-way comparison will then be considered in context with Agency's previous analysis of biomonitoring data (mostly from agricultural workers) which was presented to FIFRA SAP in 2007 (http://www.epa.gov/scipoly/sap/meetings/2007/010907_mtg.htm).

3. The third step in the exposure assessment comparison between the AHS and the Agency's approaches may involve a large-scale comparative analysis of atrazine and alachlor users from AHS cohort to those predicted by the Agency. The feasibility of this large-scale analysis will be assessed after the completion of steps 1 and 2. This step will proceed only if it is determined that the input the data are sufficiently similar to allow for an appropriate comparison. Following the feasibility analysis, the methodology for the large-scale analysis will be determined.

The Agency is developing an addendum for submission to the FIFRA SAP for review at the February, 2010 meeting. This forthcoming addendum will provide a more detailed discussion of the science issues and project plan discussed above. The Agency will solicit comments from the Panel on this plan along with suggestions for additional or alternative analyses. Conceptually, the addendum will present this case study in terms of a problem formulation exercise, like that discussed in the draft Framework and the 2009 National Research Council document on advancing risk assessment science and decision making (http://dels.nas.edu/dels/rpt_briefs/IRA_brief_final.pdf). Specifically, the problem formulation will identify key differences in the exposure assessment methodologies used by AHS investigators and the Agency in an attempt to illustrate how these differences can elucidate potential differences.

.

Attachment C

Draft Case Study

Integrating Human Incident Data into Regulatory Risk Assessment

Diazinon Human Incident Data

Office of Pesticide Programs
Office of Pollution Prevention and Toxic Substances
U.S. Environmental Protection Agency

FIFRA Scientific Advisory Panel
February 2-5, 2010

Note to reader: Attachment C will be provided to the SAP in a separate transmission.