CRITICAL REVIEW OF WHO-UNEP STATE OF THE SCIENCE OF ENDOCRINE DISRUPTING CHEMICALS – 2012

March 14, 2014

CRITICAL REVIEW OF

WHO-UNEP STATE OF THE SCIENCE OF

ENDOCRINE DISRUPTING CHEMICALS – 2012

Prepared by:

Exponent
Karyn L. Hentz, MSPH, RAC, DABT
James C. Lamb, IV, Ph.D., DABT, ATS
Jane Staveley, MSPH
Gerard Swaen, Ph.D.
Amy Lavin Williams, Ph.D., DABT

Gradient
Julie E. Goodman, Ph.D., DABT
Lorenz R. Rhomberg, Ph.D., ATS

Icahn School of Medicine at Mount Sinai
Paolo Boffetta, M.D., MPH

McMaster University
Warren G. Foster, Ph.D.

University of Guelph
Glen Van Der Kraak, Ph.D.

March 14, 2014

Disclaimer:
This independent review has been commissioned by joint funding from:
American Chemistry Council (ACC)
CropLife America (CLA), CropLife Canada (CLC) and Croplife International (CLI)
European Chemistry Industry Council (Cefic)
European Crop Protection Association (ECPA)

Contents

1. Executive Summary

The endocrine system is composed of several specialized glands and hormone producing tissues distributed throughout the body. Hormones produced by the endocrine glands and tissues regulate the body's internal environment and influence the orderly progression of development. By definition, endocrine disruption is an alteration or "disruption" in the endocrine system that results in an adverse health effect. Such disruption does not include normal hormonal fluctuations and responses and does not include changes that do not result in harm. By extension, an endocrine disruptor is a substance that is not part of the body's normal chemistry that causes harm through changes to the endocrine system. Three conditions must be met to consider a substance an endocrine disruptor:

1) The substance must cause harm or an adverse effect;
2) The substance must cause an alteration in the functioning of endocrine system; and
3) The harm or adverse effects must be the result of specific changes to hormones or the endocrine system and not due to other factors (for example, general systemic toxicity).

It should be further noted, that an observation of endocrine disruption in experimental animals, particularly at high doses, may not be relevant to much lower human or wildlife exposures.

Over the last twenty years there has been heightened public interest, media coverage, and an abundance of scientific research in the field of endocrine disruption. Studies have been conducted to identify and characterize chemicals that might be endocrine disruptors and to understand how adverse effects can occur by endocrine-mediated modes of action. In order to evaluate the wealth of existing data and understand the implications of these data for human health and wildlife, a systematic assessment is essential. Given recent advancements and interest in this topic, a current evaluation of what is known, what uncertainties exist, and areas of future research would be valuable.

Early in 2013, the World Health Organization (WHO) and the United Nations Environment Programme (UNEP) jointly published a *State of the Science of Endocrine Disrupting Chemicals - 2012* (WHO-UNEP 2012a). This WHO-UNEP (2012a) report claims to be an update to the 2002 state-of-the-science report prepared by the WHO with the International Programme on Chemical Safety (IPCS) (WHO-IPCS 2002). Although characterized as a state-of-the-science review that would meet the needs of a current systematic assessment of endocrine disruption, the WHO-UNEP 2012 report falls well short of this goal.

The present paper is a critical review of the strengths and weaknesses of the WHO-UNEP 2012 report. This paper necessarily includes a comparison with the earlier 2002 report, which was highly regarded and provided a framework for assessing the state of the science for endocrine disruptors. The aim of this critical review is not to re-evaluate each of the points made in the WHO-UNEP 2012 report and does not provide an alternative, comprehensive assessment of the data. Rather, a number of limitations of the WHO-UNEP 2012 report in this paper are identified and described with select examples that highlight concerns with the approach used and conclusions reached in the WHO-UNEP 2012 report.

The WHO-UNEP 2012 report is not technically a state-of-the-science review. A state-of-the-science review, by definition, involves a systematic assessment of evidence, taking into account the quality of the data and interpreting this information in context of the accepted definition of an endocrine

disruptor. A structured framework for evaluating the totality of the evidence ensures that all perspectives are represented and alternative explanations are considered. Instead of adopting the framework developed in the 2002 report, the 2012 report relied on "best professional judgment" for the review and evaluation of these data, which raises significant questions and concerns about conclusions reached in the report.

The 2012 report also does not function as an update to the 2002 report. In many cases, the WHO-UNEP 2012 report relies on much of the same information cited in the earlier report rather than focusing on what information was new or how it impacts conclusions reached in the earlier report. Any changes in opinion or differences seen in the evidence based on current data compared to the data from the 2002 report are not clearly presented or explained. In fact, there is no reference to the conclusions or research recommendations provided in the 2002 report. Thus, the WHO-UNEP 2012 report cannot be characterized as an update to the WHO-IPCS 2002 report.

Several shortcomings were consistently noted in the WHO-UNEP 2012 report that further call into question the reliability of the assessment of evidence and the statements made regarding endocrine disruptors:

- Endocrine disruption is implied rather than established. In many cases, only one or two of the three conditional elements of the definition of an endocrine disruptor are presented. Supporting data are often lacking to demonstrate that the chemical actually causes harm or an adverse effect; rather, the available studies may only show an association between the chemical and the effects, but not a causal relationship. Often, information is not available on mode of action. Without data to establish that the effects are being caused by hormonal changes or alterations in the endocrine system, endocrine disruption cannot be established.

  One of the examples discussed in this paper to highlight this deficiency is for prostate cancer. The WHO-UNEP 2012 report states that pesticides are associated with an increased risk for prostate cancer. However, data are very limited on specific pesticides, with only one or two studies suggesting an association and other studies that did not observe an association. Many studies that did not observe an increased risk of prostate cancer in association with pesticide exposure were not cited in WHO-UNEP 2012 report. Although the role normal hormones play in the development of the prostate are discussed, no mode-of-action information is provided for any of the pesticides discussed. Therefore, despite the implication presented in the WHO-UNEP 2012 report, insufficient data exist to link pesticides to an increase in prostate cancer risk, much less that any observed risk is a result of endocrine disruption.

- Disease trends are used inappropriately to raise concern about endocrine disruptors. Studying disease trends can be useful to explore new research ideas, but is not suitable for showing that a chemical causes a particular disease. One of the reasons why this kind of information cannot be used to establish a causal relationship is because it is not known who was exposed to the chemical of interest and whether or not these same people got the disease of interest. The WHO-UNEP 2012 report suggests that disease trends are associated with certain chemicals, but this cannot be proven. In addition, there are often many factors that are recognized to increase

a person's risk for disease, but the WHO-UNEP 2012 report often does not fully acknowledge these other factors as contributing to these disease trends.

One of the disease trends discussed in the WHO-UNEP 2012 report is an "indisputable" increase in autism spectrum disorders. As discussed in this paper, several factors are likely to have contributed to the observed increasing trend, including: changes in the diagnostic criteria for autism, greater awareness of the condition, the increased availability of medical and support services for those diagnosed with autism, and historically low rates of diagnosis in certain subpopulations due to the stigma associated with these types of behavioral conditions. Not only does the WHO-UNEP 2012 report minimize the role of these factors in the observed trend for autism spectrum disorders, but no information is provided on which chemicals or endocrine modes of action might be related to the development of autism. Therefore, it is not clear how this trend speaks to the issue of endocrine disruption.

- The WHO UNEP 2012 report failed to fully address a number of factors that must be considered when evaluating the relevance of endocrine disruption observed in the laboratory setting to real life conditions. When assessing the relevance of experimental animal data, it is important to consider the doses at which different responses occur and how those doses compare to the levels at which people or wildlife are exposed. In the WHO-UNEP 2012 report, there was no consideration of dose, exposure, or thresholds where effects might not be seen. Although the report emphasized the potential for low dose effects and non-monotonic dose responses, these issues do not exclude the need to consider dose and exposure. Finally, the potency of potential endocrine disruptors compared to naturally-occurring hormones in the body is important for understanding the ability of a substance to induce an adverse effect.

  Later in this paper, more detail is provided on the differences in exposures between experimental animals and humans, as well as the importance of thresholds, dose-response and potency. For example, the doses at which the naturally occurring phytoestrogen, genistein, has been shown to cause endometrial cancer in animal studies are at least 50-fold greater than the typical daily intake of humans. Even greater differences between experimental animal doses associated with effects and typical human exposures are seen for estrogen causing endometrial cancer in mice and polychlorinated biphenyls (PCBs) inducing thyroid hormone changes in rats. The failure to discuss and account for these differences in exposure is misleading.

In addition to these significant issues, the companion report *Summary for Decision-makers* (WHO-UNEP 2012b) is distinct from the state-of-the-science report and is not, in fact, a summary of the main report. In particular, there are many conclusions, including specific chemicals and allegations of associated adverse health effects, that are not supported by information or data provided in the main report (e.g., the role of endocrine disrupting chemicals (EDCs) in causing early puberty, association between dioxins and breast cancer). It would seem that in an effort to simplify information for decision makers, the *Summary for Decision-makers* goes well beyond the main report in drawing conclusions and includes discussions that are more advocacy-based than a reporting on the state of the science.

New data and new understanding of the endocrine activity of chemicals have been developed since 2002 and using this information to build upon the 2002 analysis is a worthwhile objective. However, the WHO-UNEP 2012 report does not achieve this goal. The WHO-UNEP 2012 report merely attempts to highlight a wide range of potential concerns associated with endocrine disruption, some of which are not justified given the evidence provided. In other cases, conclusions of endocrine disruption are reached without considering the role of other non-endocrine mechanisms, the influence of extraneous factors, the biological plausibility of mechanisms, or whether exposures are sufficient to warrant concern. These implications can be alarming to the uninformed reader and are misleading as they are based on only a partial presentation of the available evidence. Unfortunately, given the limitations identified in this critical review, the WHO-UNEP 2012 report is not a true assessment of the state of the science and should not be used to support evidence-based decisions regarding endocrine disruption.

2. Introduction

Endocrine disruption is currently a hot topic in the media, an issue of potential concern in the public, and a booming area of research. By definition, endocrine disruption is an alteration or "disruption" in the endocrine system that causes an adverse health effect. An endocrine disruptor, therefore, is a chemical that is not part of the body's hormonal system which adversely affects the functioning of the endocrine system. Endocrine disruption does not include the normal fluctuations in endogenous hormones that control homeostasis and does not include changes in the endocrine system that does not result in adverse effects.

In 2002, the World Health Organization (WHO), in collaboration with the International Programme on Chemical Safety (IPCS), produced a "state-of-the-science" assessment of endocrine disruptors (WHO-IPCS 2002). It was not an assessment of particular agents or risks, but set out to summarize the prevailing state of scientific knowledge – what was known, what was uncertain, and what the prospects were for resolving the uncertainties with further research and data collection regarding endocrine disruption. The report described patterns in natural human and animal populations that were considered possible manifestations of endocrine disruption and assessed the basis for evaluating whether these patterns should be regarded as real and robust, whether explanations for them other than endocrine disruption could be possible, and what might be the state of toxicological evidence for attributing them to an interference with endocrine-mediated control by exogenous chemicals at prevailing environmental concentrations. Issues under debate were described forthrightly along with the nature and extent of evidence available to support the differing points of view. Importantly, the assessment was notable not only for its product, but also its process. A large and widely representative set of international experts, including those with a variety of views, articulated and employed a weight-of-evidence methodology to integrate various kinds and lines of evidence and to gauge how and how well the collective evidence supported conclusions, which were then extensively reviewed. The aim was to be appropriately circumspect, yet earnestly probing – that is, neither to be alarmist, focusing only on feared possibilities, nor to be complacent and dismissive of concerns that had yet to be adequately supported scientifically. The WHO-IPCS 2002 assessment largely succeeded in these aims and it won wide acceptance and respect as an objective picture of what science had to say (and the limits as to what current knowledge allowed it to say) about the possibilities, prevalence, and magnitude of impacts of environmental chemicals on natural populations through interaction with endocrine systems.

Since 2002, the interest in the question of endocrine disruption as a possible environmental issue has only increased, and a good deal of further research has been done. In 2012, the WHO, in collaboration with the United Nations Environment Programme (UNEP), published what is presented as an "update" to the 2002 report (WHO-UNEP 2012a), which also presents itself as a review of the now-current state of the science. Unfortunately, the 2012 report falls well short of the standard set by the 2002 assessment, both in openness and objectiveness of process and as a substantial evaluation of current scientific knowledge and thinking on the issues. Whereas the 2002 assessment was produced by consensus among a large set of scientists spanning the range of views on the matter, the 2012 report was produced by a more limited set of authors who largely represent one part of the spectrum of opinion. The 2002 report articulated and used a weight-of-evidence evaluation process and while the

2012 report criticizes that process, it does not replace it with anything else, relying instead on unexplained "professional judgment." The 2002 report attempted to integrate information on exposure, toxicological testing (including dose-dependence of effects), the ability of putative disruptors to interfere with endocrine-mediated control, and patterns of appearance of possibly endocrine-related effects in populations. In contrast, the 2012 report discusses each of these independently and specifically declines to consider how these aspects can be brought together to assess whether there are real, current endocrine disruption problems or how well an integrated view of the scientific evidence can answer that question.

The present paper identifies several concerns about the WHO-UNEP 2012 report. Namely, the report fails to present an objective assessment of the current state of the science of endocrine disruption and does not serve to update the 2002 assessment. Instead, the 2012 report seeks to replace the earlier assessment with a much less thorough evaluation that stresses possibilities of concern rather than an assessment of evidence about whether those possibilities result in real human health or environmental problems. It is acknowledged that new data and new understanding of the endocrine activity of chemicals have been developed since 2002 and using this information to build on the 2002 analysis is worthwhile. However, the WHO-UNEP 2012 report fails to achieve this.

An underlying concern with the report is the presentation of evidence in a manner that infers that the information demonstrates endocrine disruption without full consideration of alternative explanations for the observed effects. This is partially achieved by the imprecise use of key terms or concepts. For example, throughout the report there are sections titled “Epidemiological evidence for EDCs [endocrine disrupting chemicals] causing [insert health effect under discussion, e.g., early puberty].” This title gives the reader an impression that evidence will be presented on chemicals that *cause* that particular effect, when these sections should have more appropriately been characterized as a discussion of EDCs *associated* with these effects. Section 6 provides other examples and more detail on the use of inference to imply rather than demonstrate that EDCs are causally associated with certain effects. The following table provides a summary of key terms, with their definitions, as used in this paper.

Table 1: Key Terms and Definitions

Key Terms	Definition
Endocrine disruptor	“An endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub) populations” (WHO-IPCS 2002)
Adverse effect	“[C]hange in morphology, physiology, growth, development, reproduction or life span of an organism, system, or (sub) population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influence.” (IPCS 2004)
Association	Statistical relationship between two or more events, characteristics, or other variables.
Causation	To say that an agent causes an adverse effect means that the agent interacts with an organism to produce changes that lead to adverse effects that would not have occurred had the agent not been present.

	Bradford Hill (1965) suggested several factors be considered to differentiate a causal association from a non-causal association. Based on a modification of the Bradford Hill (1965) criteria, WHO-IPCS (2002) developed causation criteria in a "structured format for assessing postulated relationships between altered health outcomes and exposure to EDCs [endocrine disrupting chemicals]."
Non-monotonic dose response	"[N]on-monotonic … dose response curves are often referred to as U-shaped (with maximal responses observed at low and high doses) or inverted U-shaped (with maximal responses observed at intermediate doses)." (p. 8, WHO-UNEP 2012a) "Non-Monotonic Dose Responses (NMDRs) – measured biological effects with dose response curves that contain a point of inflection where the slope of the curve changes sign at one or more points within the tested range." (USEPA 2013)
Threshold	Dose or exposure concentration of an agent below which a stated effect is not observed or expected to occur (EPA IRIS 2014, IPCS 2004)
Potency	The power of an agent to produce a desired effect (Dorlands Medical Dictionary) Together, affinity and efficacy determine the potency of a ligand [an agent] to activate specific hormone receptors and to elicit specific cellular responses in target tissues (Borgert et al. 2013)

To begin, it is unclear how the WHO-UNEP 2012 report is to be considered an assessment of the state of the science. It never defines what might be meant by "state of the science" nor discusses what such an assessment should cover and characterize. On this question, the aim of the 2002 report was to evaluate the potential magnitude and scope of endocrine disruption as a real public health problem and review the basis for those findings by considering the evidence at hand, issues arising during its interpretation, and uncertainties that remain. The existence of scientific uncertainties – and even controversies – and the different perspectives held among scientists in the field would seem to be a part of the state of the science, yet the 2012 report does not discuss or even acknowledge such topics generally. There is no discussion of how the literature was surveyed, what particular topics or depth of coverage was aimed at, or what was excluded. Indeed, very little discussion of methodology exists in the 2012 report.

The WHO-UNEP 2012 report is characterized as an "update"; however, it falls well short of meeting this purported goal. A true update would cite the 2002 conclusions, articulate what data, findings, or new understanding since 2002 should be considered, and evaluate how and whether the 2002 conclusions need to be modified in light of the newer information. That is, one would expect an update to build on and modify the earlier analysis, giving reasons and support for the changes that seem in order. Although there is a section in each subchapter on specific health effects that is titled: *Scientific progress since 2002*, this information does not fulfill the requirements of an update.

Importantly, although other definitions are acknowledged, the 2012 report begins by noting explicitly that it uses the 2002 IPCS definition of an endocrine disruptor. This definition refers specifically to an agent or mixture that "alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations." Adversity is specifically indicated, as is the condition that such adversity be found in real populations. This suggests that the 2012 report would assess endocrine disruption based on whether *actual* exposures are indeed responsible for *actual* adverse effects in real populations. Instead, the focus of the 2012 report appears to be on the potential ability of agents to interact with the endocrine system, and then by implication, that these interactions are tied to health and environmental effects. That is, the report implicitly uses an alternate definition that has more to do with the potential for functional interactions than with whether such potential actually results in real and meaningful impacts. The WHO-UNEP 2012 report states its intent is to:

> "introduce the evidence that exogenous chemicals can interfere with hormone action and produce adverse effects. It is not our intention to develop a list of known EDCs [endocrine disrupting chemicals], or to identify the properties of EDCs. Rather, we provide a description of the logic that must be implemented to identify EDCs and their properties." (p. 11, WHO-UNEP 2012a)

This statement moves away from the adversity-in-real-populations definition toward an approach based solely on hazard and mode of action (MOA). It also sets as the report's objective the identification of possible effects rather than the evaluation of the actual extent of an endocrine disruption problem in the world, or the degree to which it is understood.

Additional concerns have been identified in the WHO-UNEP 2012 report, including the lack of a formal framework for assessing whether or not exposures are *causally* linked to adverse effects. Instead, the report adopted a narrative approach for the data review that does not represent a weight of evidence assessment. Rather than demonstrate causation, the report relies on inference to suggest that exposures to chemicals and adverse outcomes are related. For example, temporal trends in human diseases or wildlife populations are presented without an evaluation of evidence for actual causes or consideration of alternative explanations for these trends (especially diagnostic criteria and reporting changes). Exposures to chemicals considered as possible disruptors exist and are suggested as contributing to the observed trends, but there is little consideration of whether these exposures are sufficient to explain the alleged effects and whether the patterns of exposure are congruent with the trends. Finally, the report discusses, but does not give appropriate consideration to thresholds and the dose-response for adverse effects. Despite the chosen definition of endocrine disruption as producing adverse changes, the report treats all effects as evidence of disruption and makes an *a priori* rejection of thresholds. Furthermore, where examples of animal testing are discussed there is little consideration of dose-response, when, in fact, only some doses of some compounds can cause endocrine disruption in the laboratory.

In addition, it is important to call attention to another report, the "Summary for Decision-makers" which is a separate product of the WHO's undertaking. Readers should be aware that this "Summary" is not a summary of the main report, even though it includes as a (minor) part a brief recapitulation of principal

findings from main report. The Summary was produced by a subset of the authors of the main report, and the process for its review and relation to the larger assessment is not clear. In particular, there are many conclusions, including those about specific chemicals and their impact on human or animal health, that are not supported by any analysis or citation of data either in the Summary itself or in the main report (e.g., the role of EDCs in causing early puberty, association between dioxins and breast cancer). Indeed, the Summary goes well beyond the main report in drawing conclusions, and it includes discussion that is more advocacy than a reporting on the state of the science.

In the text that follows, observations and comments on the WHO-UNEP 2012 report are provided with further discussion and examples. The aim is not to reevaluate these points, nor to conduct a comprehensive assessment, but rather to illustrate with specific examples where the 2012 report has made many statements that claim or imply a finding about endocrine disruption as a cause of actual effects which are not supported by an unbiased and thorough evaluation of the pertinent evidence.

3. Summary for Decision-Makers is a Separate and Inconsistent Assessment

The relationship between the 2012 main report (WHO-UNEP 2012a) and the *Summary for Decision-Makers* (WHO-UNEP 2012b) is problematic. Based on the title of this document, one might presume that this document is a summary of – or at least based on the analysis of – the main report. But a close examination reveals that the Summary is actually characterized as "another product" of the process. In some of its sections, the Summary does summarize some of the findings of the main report, but there are many parts of the Summary, including many conclusions and assertions, that do not reflect analysis or discussion from the main report and include conclusions and findings on matters not mentioned at all in the main report. It seems very important to draw this distinction and make clear that the *Summary for Decision-Makers* is not a basis for discussing the contents of the main report. Some specific examples of the statements made in the *Summary for Decision-Makers* that are not supported by the main report are described below.

The *Summary for Decision-Makers* reportedly presents information on endocrine disruptors and human health in Section 4. However, no information is provided in this section on specific EDCs being associated with any particular human health outcomes. Only broad statements are made regarding EDCs and the role they may have in increasing the incidences of various diseases. The only definitive information provided on specific health effects is presented in Figure 5 (p. 7, WHO-UNEP 2012b). Figure 5 is titled “Diseases induced by exposure to EDCs during development in animal models and human studies” and lists a range of conditions from breast cancer to attention deficit hyperactivity disorder (ADHD). The word “induced” is a synonym for caused; and therefore, this figure title is stating that these diseases can be caused by exposure to EDCs during development. This is a significant overstatement of the evidence available for most of these diseases and is not supported by the data presented in the main report. For example, endometriosis is one of the reproductive/endocrine diseases mentioned in this table. However, in the main report in the discussion of a hormonal mechanism it is stated that “it is now hypothesized” that developmental exposure can contribute to the risk of endometriosis. A hypothesis falls well short of an actual demonstration that EDCs induce endometriosis. Similarly, autoimmune thyroid disease is discussed in the main report, but the information presented is only speculative in that “it is possible that EDCs play an important role in the development of these [autoimmune] diseases.” The only specific EDC mentioned in relation to autoimmune diseases is bisphenol A (BPA), which is characterized as “could be a factor in the development of inflammatory and autoimmune diseases.” Likewise, no data on specific EDCs are presented to demonstrate that these exposures induce heart disease and hypertension. Only intrauterine growth restriction (IUGR) is mentioned in the main report as a risk factor during development that can lead to hypertension and cardiovascular disease later in life; however, no data were provided to show that EDCs cause IUGR. Finally, Alzheimer disease and Parkinson disease are not even mentioned in the main report, so no data are provided to show that these diseases can be caused by EDCs, much less from exposure during development. The use of this title to present these data is misleading and suggests to the reader a level of confidence in the relationship between EDCs and specific diseases, particularly for exposure during development, that does not exist. It is interesting to note that later in the Summary several of these diseases are described less definitively as “potential diseases or dysfunctions originating from early exposure to EDCs” (p. 13, WHO-UNEP 2012b).

On page 9 of the *Summary for Decision-Makers* (WHO-UNEP 2012b), human disease trends are presented as evidence for concern regarding endocrine disruption. As described later in this paper, disease trends alone are not sufficient for connecting exposure to EDCs with an alleged increase in disease. However, the Summary specifically states that "EDC exposures have been linked with increased rates of neurobehavioural disorders, including dyslexia, mental retardation, ADHD and autism." In the main report, however, there is only a single use of the word "dyslexia," which occurs in the initial "overview" of Chapter 1, where the existence of trends in various public health endpoints are discussed (the causes of those trends, be they EDCs or otherwise, are not discussed). The main report states: "[n]eurobehavioural disorders, including dyslexia, mental retardation, attention deficit disorder, and autism affects nearly 20% of children in those countries where it has been evaluated" (p. 2, WHO-UNEP 2012a). This statement does not provide any evidence that EDC exposures are linked with the observation of dyslexia. Similarly, autism is discussed mostly as a public health trend in the main report, as in Chapter 2 where it states "[w]hilst the increase in autism spectrum disorders is indisputable, questions remain as to whether the increase in the incidence and prevalence of ADHD represent a true increase rather than an artifact" (p. 109, WHO-UNEP 2012a). Nowhere in the main report is autism linked to any EDC exposure, nor are data presented to support the statement given in the *Summary for Decision-Makers*. Discussion of possible causal agents in the development of autism is limited to two statements that suggest that thyroid hormone deficiencies may contribute to the development of this condition (p. 93, 119; WHO-UNEP 2012a), but no specific chemicals are identified. Thus, the *Summary for Decision-Makers* overstates the evidence by saying that there is a link between EDC exposures and autism and dyslexia, when no data as such are presented in the main report.

In the discussion of endocrine disruptors and wildlife health in the *Summary for Decision-Makers*, it is noted that there are important parallels between the increased incidences of human disorders and those observed in wildlife. An example is provided of cryptorchidism (failure of one or both testes to descend into the scrotum) in black-tailed deer in Alaska and white-tailed deer in Montana, which is compared to testicular dysgenesis syndrome (TDS) in humans. However, this is the only example of a parallel drawn between humans and wildlife in the main report. The statement and example in the Summary goes beyond the information provided in the main report by implying that multiple examples exist of similar observations in humans and wildlife when this is not the case.

Exposures to EDCs are discussed in the *Summary for Decision-Makers,* with an emphasis on the various routes of exposure and the fact that many chemicals have been measured in the serum, blood, or urine of humans and wildlife. However, as addressed in greater detail later in this paper, the mere presence of a chemical in the body does not mean that harm can occur. In the Summary (p. 15, WHO-UNEP 2012b), it is noted that “[a]t this time there are no data showing how exposure to mixtures of virtually hundreds of EDCs at low concentrations will affect human and wildlife health.” Yet, on page 19, the Summary states that “endocrine diseases and disorders are occurring at current exposure levels” and “there are situations in which individually safe exposures of EDCs have reached a collectively harmful level” (WHO-UNEP 2012b). While the main report discusses the potential for additive effects, data are not presented to show that the combination of current exposures is resulting in harm to humans or wildlife. Thus, the *Summary for Decision-Makers* not only overstates the issue of potential additivity as

presented in the main report, but is internally inconsistent regarding the effects of low level exposures to EDCs.

These few examples demonstrate a broader scope in discussions of the topics covered and more definitive statements made in the *Summary for Decision-Makers* compared to the main report. The lack of references in this report (or even references to particular sections in the main report) hinders the reader in finding the basis for many of these general statements. Efforts to simplify the information for decision makers or lay people have resulted in a failure to properly characterize the data gaps and therefore, the strength of the conclusions is overstated. In addition to some of the inconsistencies between the *Summary for Decision-Makers* and the main report, the Summary approached the state of the science from the same perspective and consequently has many of the same limitations as the main report. The rest of this paper focuses on these limitations and outlines specific concerns with the lack of a framework for identifying, reviewing, and evaluating data; the failure to update the 2002 WHO-IPCS report as stated; the informal approach to assessing causation from EDCs; the reliance on disease trends to suggest associations with EDCs; and the disregard for the role of dose and potency in endocrine disruption.

4. The 2012 Report is Not a State of the Science Review

Although the goal of the WHO-UNEP 2012 report was described as "providing a current state-of-science of endocrine disruptors as it relates to human and wildlife population health," (p. 3, WHO-UNEP 2012a) several elements were lacking to accomplish this goal. A state of the science review should have a defined scope with a systematic approach to the collection and review of data, and a clear methodology for the integration and assessment of these data. Several factors need to be considered in the process of integrating and interpreting data, particularly when evaluating the relevance of experimental animal studies to human health or wildlife. To that end, one of the key concerns is how the dose-responses observed in the experimental animal studies compare to the exposures potentially experienced by humans or wildlife. Other factors that also should be considered in the integration of data include: the quality of the data available, the consistency of the results, confounding factors that may influence the findings, and the identification of data gaps. When interpreting data, the complete spectrum of findings should be considered and any controversy or debate on the issues at hand should be presented. As discussed in greater detail below, given the undefined scope and lack of a structured methodology for integrating and assessing the weight of evidence in the WHO-UNEP 2012 report, the state of the science was not reviewed in a consistent, objective manner. In some cases, the weaknesses of the data are acknowledged, such as the lack of data on EDC exposure and ovarian cancer; in other cases, important studies are ignored that counter those cited in the report. Some examples of these weaknesses are provided in the sections below.

The most critical points of divergence between the 2002 and 2012 reports relate to the approach and methods used in evaluating the state of the science. In the 2002 report, the purpose and scope of the document describes not only what the review entailed, but also what the review was not intended to represent. Furthermore, in the WHO-IPCS 2002 report, a systematic method was used to integrate and interpret the data on endocrine disruption and a framework was adopted for:

> "evaluating the collective information from diverse data sets in a structured manner to provide objective assessments of the state-of-the-science of determining causality between exposures to EDCs and selected outcomes" (p. 1, WHO-IPCS 2002)

In contrast, other than stating that it is an update to the 2002 report, the WHO-UNEP 2012 report does not provide any information on what the scope of the review is perceived to include; nor are the limitations of the review described. Thus, the reader is left with the potential impression that this is a comprehensive review of the data on endocrine disruption. Most importantly, the 2012 report did not apply the framework developed in the 2002 review. In fact, this framework appears to have been overlooked based on the conclusions that:

> "[e]fforts are needed to develop systematic and transparent approaches to identify, evaluating and synthesizing the scientific evidence for endocrine disruptors that consider the science of endocrine action" (p. 20, WHO-UNEP 2012a).

Nor was an alternative approach proposed or applied to evaluate the weight of evidence. Instead, "best professional judgment" (p. 19, WHO-UNEP 2012a) is used without fully describing any objective criteria

by which such judgment was to be applied. Specific concerns about the lack of a structured approach for assessing causation in the 2012 report are further discussed later in this paper.

The literature published on the potential endocrine activity of specific chemicals is extensive and beyond the scope of either the 2002 or 2012 reports. Neither report could be expected to undertake complete reviews for even a small subset of chemicals, but a systematic methodology would have ensured that a representative spectrum of the available literature was captured in the review. However, it does not appear that such a systematic approach was utilized in the collection and review of data for the WHO-UNEP 2012 report. The title of the report "State of the Science" infers that it is minimally based on a systematic review of the available scientific literature. This, however, is not the case. In a systematic review, a fixed format is followed, which has been described and applied on many occasions (see for example, Khan et al 2003). In any systematic review, an organized and through search of the literature should be carried out, with a predefined set of key search items, using at least two, but preferably more, appropriate literature databases (e.g., PubMed, Toxline, Biosis, Scisearch). The literature search should be done in a transparent manner, with predefined inclusion and exclusion criteria and should be described in such a way that the search can be easily reconstructed and results verified. The review should also include a critical assessment of the quality of the individual studies. Studies should be summarized and interpreted in a manner that is clear to the reader. This approach was not followed in the WHO-UNEP report: no systematic search of the literature with predetermined key search terms was conducted. Furthermore, the handpicked publications were not assessed for their quality, nor were the results placed in a broader perspective. Since the report did not follow this widely accepted methodology it is difficult to assess whether the publications on which the report relied are representative of good quality research reported in the open scientific literature. The only description of the literature retrieval process indicates that emphasis was placed on literature published after 2000 through March 2012. Despite this claim that the focus was on literature available in the last ten years or so, a substantial number of citations relied upon in the 2012 report are ones that were previously cited in the 2002 report or were published prior to the turn of the century. For example, in the bone subchapter (p. 152-154), 19 of the 35 references (54%) predate 2002 and in the thyroid subchapter 66 of the 228 references (29%) were available prior to 2002. To be a robust state-of-the-science review, the report should have had clear parameters for the search criteria that excluded studies previously addressed in the 2002 report. The strong reliance on older citations further supports the conclusion that the WHO-UNEP 2012 report is not an actual update to WHO-IPCS 2002 report, but rather, a reworking of that earlier report. This is discussed more fully in Section 5 of this paper.

An additional concern regarding the data collection process is that there seems to be a strong preference toward citing studies that report an association with exposure and omitting the studies that do not support such associations. For example, the studies that showed some association between levels of pesticides in the environment and adverse effects on frogs are cited (e.g., Hayes et al. 2003, McDaniel et al. 2008), while studies to the contrary (e.g., Murphy et al. 2006, Smith et al. 2005, Du Preez et al. 2009, Skelly et al. 2010, Spolyarich et al. 2011) are not mentioned. The WHO-UNEP report goes on to report that atrazine is associated with a suite of responses in amphibians related to testicular dysgenesis, yet the large number of studies that have failed to show such an association are not mentioned (e.g., LaFiendra et al. 2008, Oka et al. 2008, Storrs and Semlitsch 2008, Kloas et al. 2009).

The same is true of the experimental studies cited for thyroid effects from PCBs. While twelve studies were cited in the discussion, half of these studies were from one research laboratory (University of Massachusetts, Amherst), and thus, do not represent independent verification of findings in separate laboratories. Other studies investigating PCB effects on the thyroid were not included in the review (e.g., Martin and Klaassen 2010; Kato et al. 2004, 2010). Another example of the selective citation of literature in the 2012 report is failure to cite many of the experimental animal studies on BPA that were conducted according to standard guidelines under Good Laboratory Practices (GLP) (e.g., Stump 2010, Tyl et al. 2002, 2008). As recommended by Conrad and Becker (2011a, b), all well conducted laboratory studies, both GLP and non-GLP, should be considered in a review in order to provide a comprehensive understanding of the MOA, hazards and risks of a chemical. It is not expected that all of the available data could be critically reviewed, but some sense of balance would have helped the reader better understand the controversy that exists in the field of endocrine disruption.

A discussion of divergent results in the literature is essential to determine why different results were generated, particularly since one of the hallmarks of science is reproducibility. If only one laboratory is able to generate a particular outcome, it suggests that the outcome is not very robust or that there may be something unique in how the results were obtained, and thus, they are not reliable. An additional reason for discussing divergent results is to make the reader aware of potential controversy that is fundamental to the scientific process rather than to give a false sense of agreement. Only an informed individual can make decisions about the impact of discrepancies in the use of these data. Finally, identifying studies with divergent findings characterizes important gaps in understanding and helps to formulate potential research priorities. The failure to identify studies with different results is potentially misleading, contrary to best scientific practices, and does not facilitate sound policy decisions.

An additional concern regarding the specific citations referenced in the WHO-UNEP report are the discrepancies that exist between the findings reported in the publications and what is described in the WHO-UNEP report. For example, it is stated in the WHO-UNEP report "[a]rsenic exposure is strongly associated with prostate cancer" and two citations are provided (p. 131, WHO-UNEP 2012a). One of these citations, Benbrahim-Tallaa and Waalkes (2008), merely references the other. The statement in the WHO-UNEP 2012 report does not reflect the outstanding questions about potential prostate cancer risk at low environmental exposures. The association between arsenic in drinking water and prostate cancer is primarily based on the results of epidemiology studies of an ecological design conducted in Taiwan, which are not sufficient for assessing causation. The potential MOA is not defined and is only speculated to be endocrine-mediated. Furthermore, the International Agency for Research on Cancer (IARC 2004) reviewed the same evidence as Benbrahim-Tallaa and Waalkes (2008) and concluded that there was sufficient evidence for arsenic in drinking water causing skin cancer, lung cancer, and bladder cancer only; no conclusions were reached based on prostate cancer. Thus, the conclusion that arsenic is "strongly associated with prostate cancer" is not supported by the citations referenced and is not consistent with IARC's more comprehensive review of the literature.

Later in the discussion of mechanisms for prostate cancer, Soto et al. (1995) is cited as support for evidence that "organochlorine pesticides [are] shown to be associated with increased prostate cancer risks." In fact, the publication of Soto et al. (1995) relates to an *in vitro* assay and does not at all deal

with the risk of prostate cancer from pesticides. Similarly, Yolton et al (2011) is cited as showing that concentrations of BPA and phthalates in maternal urine during early pregnancy were associated with higher hyperactivity and aggression in 2-year old girls, but not in boys (p. 114-115, WHO-UNEP 2012a). However, Yolton et al. (2011) did not study hyperactivity or aggression and only measured early infant neurobehavior five weeks after delivery using the NICU Network Neurobehavioral Scale. An earlier publication from the same research group, Braun et al. (2009), did report an association between mean urinary BPA concentrations and behavior scores in female children, but there is no assessment of phthalates in this earlier study. Finally, Figure 2.31 (p. 182, WHO-UNEP 2012a) shows data on concentrations of DDE (dichlorodiphenyl-dichloroethylene, a metabolite of DDT or dichlorodiphenyltrichloroethane) concentrations in osprey eggs, with a straight line drawn through the points to depict a clear decline over time. This figure is reported to be "based on data from Henny et al. (2010)." The original paper, however, does not draw a best-fit line through these values, but rather, the data are presented in a bar graph with no statistically significant differences between the time periods: 1981-82, 1993, 1998, and 2000-01. In other words, there is no real trend for these four periods; only the values in 2006 and 2008 show a statistically significant decrease in egg concentrations. These are just a few examples of the discrepancies found, but do not reflect all of the citations that have been mischaracterized or mistakenly referenced. In some cases, these may just be mistakes in referencing the correct citation, but in others there appears to be a tendency to exaggerate the findings of the original authors.

Moreover, there is no description of a process applied for assessing the quality or reliability of the studies considered for review in the WHO-UNEP 2012 report. In fact, the quality of the underlying studies has not been evaluated at all. This is of particular concern because not all studies represent the same weight of evidence. For example, epidemiological studies that employ weak research designs (e.g., ecological and cross-sectional study designs) or include small sample sizes should not be given as much weight as studies with stronger designs (e.g., case-control and cohort studies) and larger sample sizes. Also, comparing *in vitro* exposures to relevant *in vivo* exposures is fraught with difficulties. *In vitro* studies can be relevant for investigating MOA and potential activity, but cannot provide useful information on dose-response, do not take into account the disposition of a chemical in the body (its absorption, distribution, metabolism and elimination), and fail to account for other pathways and processes that respond to certain modes of action. Various methods exist to evaluate and weigh the quality and reliability of studies included in a review, such as the systematic approach for evaluating toxicological and ecotoxicological data described by Klimisch et al. (1997), which is also employed by the Organization for Economic Cooperation and Development (OECD) and relied upon by the OECD's 34 member countries in the investigation of high production volume chemicals (OECD 2005). Yet inexplicably, neither this approach nor an alternative approach for evaluating the quality of the data reviewed was applied in the WHO-UNEP 2012 review.

In the WHO-UNEP 2012 report, the evidence for endocrine disruption in humans and wildlife is presented as narrative reviews of the data. These assessments were described as being founded on an aggregation of the information related to biological plausibility, relevant exposures, consistency of the

data across species, dose-response and temporality. Although this information may have been provided for specific health or environmental adverse outcomes[1], there was little, if any, integration of the data. For example, biological plausibility is often cited in the 2012 report as the basis for concern of causality, but the evidence generally is limited to data on the role of endogenous hormones and not based on mechanistic data for any of the chemicals of potential concern. The data on exposure are clearly not integrated with the rest of the data as exposure information is provided separately in Chapter 3. Species concordance is frequently used to try and bridge experimental animal or wildlife data with human data on particular observed effects, but there is a lack of integration of the data on exposures or mechanisms. Potential species differences do not seem to be considered, and in some cases, these differences may be critical to the interpretation of the data (this is discussed later in more detail regarding thyroid effects). Although the report purports to incorporate dose-response in the evaluation of the evidence, it tends to be ignored when specific chemicals or adverse outcomes are discussed; this issue is addressed in more detail later in this paper.

Other factors not specifically mentioned in the WHO-UNEP 2012 report that should be incorporated in the integration and interpretation of data in a state of the science review include the reproducibility of the data and consistency of data across different lines of evidence (epidemiological, *in vivo* and *in vitro* data), data gaps, and the existence of controversy or differences in interpretation of study findings. For example, the reproducibility of a finding in different studies, different research labs, or in different study population was not addressed. Regarding data gaps, the 2012 report notes in a number of places in the main text where significant gaps still exist in the data. However, these data gaps and their implications are not necessarily carried forward into the report conclusions. Additionally, the 2002 report made specific research recommendations based on such data gaps – an element missing from the 2012 report.

A final issue of concern with the WHO-UNEP 2012 report as a state of the science review relates to the definition of an endocrine disruptor and how this definition was applied to the interpretation of the data. The 2012 report (p. 11) clearly states that it is using the WHO-IPCS 2002 definition of an endocrine disruptor:

> "…an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub) populations. A potential endocrine disruptor is an exogenous substance or mixture that possesses properties that might be expected to lead to endocrine disruption in an intact organism, or its progeny, or (sub) populations." (p. 11, WHO-UNEP 2012a)

Yet, later in the same section of the report, it is stated that: "[e]ndocrine disruptors are exogenous chemicals or chemical mixtures that can interfere with any aspect of hormone action" (p. 19). These descriptions are not consistent and the use of the latter definition falls short of the formal definition of an endocrine disruptor. The 2012 report would be stronger if truly positive EDCs were better distinguished from those that are only *potential* EDCs as initially defined in the WHO-UNEP 2012 report

[1] When referring to adverse outcomes in this document, this includes adverse health effects in humans, experimental animals, and wildlife.

(i.e., chemicals that have shown activity on the endocrine system, but for which specific adverse effects or a causal linkage between the endocrine activity of the chemical and its adverse effects have not been specifically shown). Chemicals that have only been shown to have endocrine activity cannot be considered endocrine disruptors. For example, parathion is a potent organophosphate pesticide that can cause adverse effects, but the meaningful adverse effects are from cholinesterase inhibition, not endocrine disruption; nevertheless, parathion is repeatedly cited as an EDC in the 2012 report without evidence of an endocrine MOA (p. 13, 168, 172, WHO-UNEP 2012a). This issue underscores the need for providing the context in which effects are observed. The failure to differentiate between the potential for endocrine disruption and cases where endocrine disruption is clearly established leaves the reader with a misleading impression of the weight of evidence for particular effects.

Overall, the WHO-UNEP 2012 report falls short of what should be a state of the science review. Other recently published reports have done a better job of providing state of the science reviews on controversial issues in science. A few examples are presented below to highlight how such reviews can and should be conducted.

4.1. EPA (2013) Report on NMDR

The U.S. Environmental Protection Agency (EPA) recently published a thoughtful review of the scientific issues surrounding the phenomenon of a non-monotonic dose-response (NMDR) (USEPA 2013). This draft review is instructive not only because of its conclusions, but also the approach taken by the EPA in conducting this review. The report addresses a highly controversial issue that can affect both risk assessment and risk management. EPA took a methodical and even-handed approach to its review. The scope of the report was provided, clearly stating the scientific questions that were to be addressed. EPA also clearly described and respected the limits of the report; for example, they acknowledged that it was not a comprehensive treatise. For each of the topics reviewed, a description of the literature search used to identify the relevant articles for inclusion in the state of the science review was provided. In a succinct and well-written manner, EPA describes the evidence that both supports and conflicts with its conclusions in support of NMDRs. As part of this assessment the uncertainty associated with the interpretation of the data was also considered. Overall, the EPA report on NMDRs is an excellent example of a well-conducted state of the science review and illustrates how a controversial issue such as NMDRs and endocrine disruption can be approached in an objective manner.

4.2. EFSA (2010) Scientific Opinion on BPA

The European Food Safety Authority (EFSA) reviewed a developmental neurotoxicity (DNT) study in rats (Stump 2010) and other recent scientific literature in context of the risk assessment for BPA (EFSA 2010), a compound mentioned in the WHO-UNEP report as being an EDC. Specifically, EFSA considered the impact of these new data on the tolerable daily intake (TDI) established in an earlier review by EFSA (2006) was assessed. The EFSA 2010 review included a summary of the previous risk assessments that had been performed on BPA, as well as the key DNT studies that were highlighted in these earlier evaluations. In addition to the critical review of the DNT study by Stump (2010), the panel conducted a literature search to identify toxicological data published between 2007 and 2010. Specific parameters for the inclusion of studies were identified, including: peer-reviewed research papers reporting original

data and available in the public domain and all epidemiology studies, except those studies that only provided data on biomonitoring. EFSA described its reliance on clear "quality criteria" to assess the strengths and weaknesses of the studies they reviewed. Those criteria addressed issues of study design, conduct, recordkeeping, and interpretation.

BPA is a highly controversial example of an EDC, and although constantly described as such, other critical non-endocrine-related adverse effects for risk assessment were considered in EFSA's assessment. The EFSA review addressed the toxicokinetics and elimination of BPA as it pertains to the use of rat data in the risk assessment. They also discussed the lack of sufficient evidence to show that BPA operates *in vivo* to cause adverse effects through a NMDR. EFSA also addressed other alleged effects of BPA, such as immune effects and induction of breast tumors, and described the limitations associated with these studies. One important element in the EFSA report is the presence of a minority report which provided an opportunity for dissenting opinions on the weight of evidence to be presented. This approach clearly provides a balanced view of the scientific controversy.

The panel concluded that no new study could be identified that would call for a revision of the current TDI (0.05 mg BPA/kg body weight/day), which was established by EFSA in 2006 based on multi-generation reproductive toxicity studies in rodents. It is interesting to note that the critical effects that were selected as the basis for the TDI were changes in body and organ weights in adult and offspring rats and liver effects in adult mice – not endocrine-related effects. The EFSA study provides an example where a comprehensive evaluation of the literature considers explanations for toxicity beyond those related to endocrine disruption; a feature which is not generally considered in the WHO-UNEP 2012 report.

4.3. NRC (2005) Health Implications of Perchlorate Ingestion

As a result of the controversy associated with the derivation of a regulatory threshold for perchlorate, the National Research Council (NRC) was asked to independently review the adverse health effects associated with perchlorate ingestion from a clinical, toxicological, and public-health perspective. An expert panel was created and asked to:

> "[e]valuate the current state of the science regarding potential adverse effects of disruption of thyroid function in humans and laboratory animals at various stages of life."
> (p. 29, NRC 2005)

Specifically, the panel was requested to assess the potential for neurodevelopmental and cancer effects, the levels associated with iodine uptake and changes in thyroid hormones, and the relevance of animal studies to human health. Based on this review, the panel was also asked to review and comment on the U.S. EPA's risk assessment for perchlorate to determine if the findings were consistent with current scientific evidence. The scope of the panel's review was outlined with specific issues to be addressed. The identification and review of the literature was described to include materials provided by various agencies, the literature cited in the EPA risk assessments, and any publically available literature. The panel was specifically requested to determine if the EPA "considered all relevant literature (both supporting and non-supporting), consistently critiqued that literature, and then used appropriate

scientific studies to develop its health risk assessment" (p. 30). Thus, the full spectrum of evidence was considered in the review and a systematic approach used to evaluate the data. In addition, the panel was required to evaluate the evidence according to a specified set of categories and establish conclusions based on weight of evidence for causation. These categories ranged from "no evidence" to "evidence establishes a causal relationship" with several categories in between to reflect the range of evidence available.

Perchlorate has been described as an EDC. Drinking water ingestion is a biologically plausible cause of potential adverse effects on the thyroid because perchlorate can inhibit iodine uptake through inhibition of the sodium-iodide symporter (NIS), and active transport protein, in rats. However, perchlorate has the benefit of showing a clear dose-response relationship and adverse effects appear to only happen when a threshold is exceeded (NRC 2005). The panel concluded that on the basis of the biology of human and rat thyroid tumors, "it is unlikely that perchlorate poses a risk of thyroid cancer in humans" (p. 111). The MOA is biologically plausible, but the likelihood of harm is low under typical exposures. The NRC report is an example of a solid, systematic review utilizing a weight of evidence approach to establish a causal relationship for perchlorate and effects on the thyroid. Furthermore, this review considered dose-response in evaluating potential human health effects, which the WHO-UNEP report failed to consider.

4.4. Conclusions

As seen by comparison with the above examples of state-of-the-science assessments conducted based on objective reviews of the literature, using sound methodology, and defined goals, the WHO-UNEP 2012 report did not provide a summary of the state of the science for endocrine disruption. A number of limitations regarding the lack of a defined scope for the review, the absence of a process for identification, integration, and interpretation of data, and the lack of a structure for evaluating the weight of the evidence calls into question the conclusions reached in the report. The WHO-UNEP 2012 report can be more appropriately characterized as a summary of proposed aspects of science that should be considered when discussing endocrine disruption, but it is not a summary of the state of the science.

5. WHO-UNEP 2012 Is Not an Update of the WHO-IPCS 2002 Report

Although the WHO-UNEP 2012 report is self-described as an update of the WHO-IPCS 2002 *Global Assessment of the State-of-the-Science of Endocrine Disruptors,* it falls short in this regard. An update relies and builds upon the information reviewed and assessments of data from an earlier document. The 2012 report reviews much of the same information cited in the 2002 report by frequently citing literature from the year 2001 or earlier. Although most sections in Chapter 2 of the 2012 report contain a small subsection titled, *Scientific Progress since 2002*, these are only brief sketches that are not sufficiently scientifically robust to achieve the objective of fully revisiting the state of scientific understanding. Often, the information contained therein is limited to a series of bullets that lack citations in support of the statements made. In some cases, the support for these bullets in the text of the chapter is scant, and it generally does not reflect the scientific consensus or weight of evidence on the topic at hand (examples are provided below). Importantly, there is no consideration of whether new information has changed the state of understanding since 2002, nor whether the research needs identified in 2002 have been addressed. The WHO-UNEP 2012 report does not specifically provide any follow-up on the research recommendations presented in the WHO-IPCS 2002 report; nor is there any reference to the conclusions reached in the WHO-IPCS 2002 report, much less any support or explanation when conclusions differ from the 2002 report. Thus, it is unclear in what way the WHO-UNEP 2012 report can be considered an "update."

Finally, and most significantly, the WHO-UNEP 2012 and WHO-IPCS 2002 reports differ in terms of how the data on endocrine disruption were reviewed and evaluated. In the 2002 report, a framework was developed and proposed for assessing a causal association for EDCs. As previously discussed, this framework was based on established and broadly accepted scientific methods for the evaluation of data to determine whether associations could be considered causally related – not just chance findings or the result of bias. In addition to the reasons provided above, the WHO-UNEP 2012 report cannot be considered an update since it does not rely on the weight-of-evidence framework developed in the 2002 report. Furthermore, the WHO-UNEP 2012 report failed to provide an appropriate alternative framework by which the data could be assessed to determine causal relationships. Consequently, the WHO-UNEP 2012 report is not an update but, rather, a selective re-evaluation of information largely included in the 2002 report that describes the potential for endocrine disruption.

In order to better understand why and how the WHO-UNEP 2012 report is not an update of the WHO-IPCS 2002 report, selected examples are provided below of adverse outcomes reviewed in both documents and compared with respect to the data relied upon, characterization of the evidence and state of the science, and conclusions reached about the weight of evidence.

5.1. Semen/Sperm Quality

Sperm or semen quality was evaluated in both state of the science reviews. The WHO-IPCS 2002 report concluded that a global trend for declining semen quality was not supported by the existing data. This conclusion was based on a broad review of studies investigating sperm counts, covering the first study suggesting a decline (Nelson and Bunge 1974), the first meta-analysis (Carlsen et al. 1992), longitudinal retrospective studies in single centers (e.g., Auger et al. 1995), and broader investigations around the

world (e.g., Auger and Jouannet 1997, YoungLai et al. 1998, Swan et al. 1997, Jørgensen et al. 2001). The review in the 2002 report described the limitations and biases of the various studies; concerns included the use of a retrospective study design, evaluation of semen or sperm samples from men that may not have been representative of the general population (e.g., patients at infertility clinics), differences in methods for recruiting study subjects, variability in the analytical methods, and lack of control or consideration of the other factors that are known to impact sperm quality (e.g., age, sexual abstinence). The lack of data on specific chemical exposures raised questions about assessing the strength of the association. Both epidemiological and experimental data were discussed and found to be conflicting. The WHO-IPCS 2002 report acknowledged that, while it was biologically plausible and some experimental evidence was available to support EDCs as affecting sperm quality, the "lack of any demonstration to date of an endocrine-disrupting mechanism for other chemical exposures indicates the need for more studies before firm conclusions can be drawn" (p. 56, WHO-IPCS 2002). In Chapter 7 of the WHO-IPCS 2002 report, an evaluation of the strength of evidence was conducted for several examples based on the proposed framework of causal criteria to assess endocrine disruptors. In *Semen Quality and Testis Function in Humans* (Section 7.4.1, p. 124), it was concluded that the overall strength of evidence was weak based on an assessment of the temporality, strength of the association, consistency, and biological plausibility of the association.

The WHO-UNEP 2012 report describes various studies of declining semen quality starting with the first meta-analysis by Carlsen et al. (1992) and many of the same references relied on in the WHO-IPCS 2002 report. The discussion is not limited to the newest studies on sperm or semen quality, nor is it a comprehensive review of all of the data. Although several prospective studies of the general population were noted to have been conducted by Nordic, Baltic, German, Spanish and Japanese researchers, the citations provided are two studies of Finnish and Danish men (Jørgensen et al. 2011, 2012). The WHO-UNEP 2012 report also discusses another recent study investigating a large number of French men; this was a retrospective analysis of men with total infertile partners that were subjects in assisted reproductive technology. Several other studies published since the WHO-IPCS 2002 report that do not show a decrease or an increase in sperm counts are not cited in the WHO-UNEP 2012 report (e.g., Costello et al. 2002, Marimuthu et al. 2003, Pal et al. 2006, Axelsson et al. 2011, Elia et al. 2012). In a recent commentary from Bonde et al. (2011), it is shown that sperm cell counts in Danish military draftees have remained stable and, in fact, suggest higher counts in the last four years of the study between 2007 and 2010. Again, this contribution in the extensive and ongoing discussion about sperm cell count trends is not cited at all. The selective citation of literature and the failure to include many studies that do not support a decline in sperm counts suggests an unbalanced review of the literature.

Unlike the WHO-IPCS 2002 report, there is no mention of the limitations and potential biases of any of the studies reviewed in the WHO-UNEP 2012 report despite the fact that all of the limitations noted in the 2002 report continue to apply (since, as mentioned above, the 2012 report relies heavily on the same studies). The prospective studies highlighted in the 2012 report have the additional problem of a low participation rate, which raises concerns about whether the men in the studies were representative of the general population. For example, the participation rate in the Finnish study was 13.4% (Jørgensen et al. 2011) and overall participation was 24% in the Danish study (Jørgensen et al. 2012). Furthermore, there is only a brief mention that the issue of declining sperm counts remains controversial. The

questions about the implications of the various findings related to sperm quality are not carried over into the conclusions[2] on male reproductive health. Recent reviews of this issue continue to characterize the reports of declining sperm counts as controversial because of the differences seen geographically and temporally (Sharpe 2010, Fisch and Braun 2013). These reviews identify a number of factors that may account for these differences that are not related to endocrine disruption, such as lab techniques, sexual behavior resulting in differences in abstinence, lifestyle factors (e.g., obesity, drug use), and genetic variations.

The WHO-UNEP 2012 report presents limited data regarding specific EDCs and the potential for affecting sperm quality. Epidemiology studies are characterized as showing weak associations with EDCs; this conclusion is based on single citations for most of the chemicals of possible concern identified in the report. Single publications are not sufficient to assess the weight of evidence (or to provide one's "professional judgment") of a potential association between an exposure and an adverse effect. Furthermore, the 2012 report does not evaluate limitations of these studies, nor does it consider experimental animal studies in conjunction with epidemiology to assess the hypothesis that EDCs could cause male reproductive disorders, including effects on sperm or semen quality. The 2012 report relies only on "suboptimal or poor semen quality in large proportions (20-40%) of men in countries in which this has been studied" and that there is "some evidence for a declining semen quality," thereby suggesting that these perceived trends are the consequence of exposure to endocrine disruptors with no strong evidence to support this claim.

Despite acknowledging that the epidemiological data only show weak associations for a decline in sperm quality related to specific EDCs, the conclusions in the 2012 report for male reproductive health focus on the observation of decreased sperm counts, ignoring the variability in sperm quality reported around the world and the questions that have been raised about this issue. Based on the evidence presented in the WHO-UNEP 2012 report, it does not appear that the evidence for changes in sperm quality differ from that reported in the WHO-IPCS 2002 report. Therefore, the difference in conclusions between the two reports is difficult to explain. When an objective, structured, and transparent weight-of-the-evidence analysis reaches one conclusion and a subjective analysis concludes the opposite, logic dictates the driving force for such a difference stems from the lack of systematic methodology and the bias inherent in a subjective analysis.

5.2. Adrenal Disorders

As noted in the WHO-UNEP 2012 report, adrenal dysfunction was not discussed in detail in the WHO-IPCS 2002 report. Adrenal disorders were not reviewed in detail in the 2002 report because the "available research to date is very limited" (p. 86, WHO-IPCS 2002). Although, the observation of severe adrenocortical hyperplasia in the Baltic ringed and gray seal were mentioned. Furthermore, it was stated that although adrenal effects in wildlife were associated with dichlorodiphenyldichloroethane

[2] Conclusions for each subchapter in Chapter 2: *Evidence for endocrine disruption in humans and wildlife* include the sections titles: *Main messages*, *Scientific progress since 2002*, and *Strength of evidence*.

(DDD), DDT, and PCBs, the involvement of these compounds in the cause of these disorders was uncertain.

The WHO-UNEP 2012 report includes a discussion of adrenal disorders and this review seems appropriate given the lack of systematic review in 2002. The information presented in the WHO-UNEP 2012 report focuses heavily on data and literature from 2002 or earlier, particularly regarding effects in wildlife. This suggests that the state of knowledge regarding endocrine disruption related to perturbations of the adrenal gland has not changed significantly since 2002. No new data are provided on the observation of adrenocortical hyperplasia in Baltic seals – the one new article, Lind et al. (2003), addresses bone mineral density in Baltic grey seals, not adrenocortical hyperplasia. Significantly, there is a change in the conclusions about the weight of evidence regarding adrenocortical hyperplasia in wildlife and various compounds, despite the lack of new information. In 2002, this was considered an uncertain association, but in 2012, it was concluded that this was a causal relationship. This change in the interpretation of the data should have been highlighted and the evidence to support this change should have been stated explicitly. If the WHO-UNEP 2012 were an update to the 2002 report, this change in opinion regarding the evidence for adrenocortical hyperplasia in wildlife and various compounds should have been highlighted. Regardless, it is unclear how this more definitive conclusion regarding adrenocortical hyperplasia in wildlife species was reached in 2012 based on essentially the same data reviewed in 2002. Additional concerns about the strength of evidence regarding the causal relationship between exposure to EDCs and the observation of adrenocortical hyperplasia in Baltic seals are discussed later in this paper.

5.3. Endometriosis

In the WHO-IPCS 2002 report, endometriosis was reviewed as a reproductive outcome possibly associated with EDCs. Evidence from epidemiology studies was reviewed, and, while PCBs and dioxins were noted to have been associated with endometriosis in some studies, other studies failed to observe an association. Experimental animal evidence was presented for 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and the data were judged to be conflicting – some studies indicated that TCDD may have a role in the development of endometriosis, but other studies did not support an association. The 2002 report mentioned the criticisms of the studies of TCDD in rhesus monkeys, including factors potentially confounding the results. The question of relevance of these data to humans was also raised given the high doses required to induce endometriosis in rodents. Regarding possible MOAs, the role of estrogen and progesterone were discussed in the context of known disease development and the aryl hydrocarbon receptor (AhR) was reviewed in relation to TCDD exposure and endometriosis. Endometriosis was one of the illustrative examples for assessing causation; specifically, PCBs and TCDD were evaluated for a causal role in the development of endometriosis. The conclusion was:

> "Relative to the hypothesis of an association between a stressor and an outcome, evidence is judged to be weak because of conflicting data from humans and animals, lack of association in women exposed to high amounts of TCDD, and antiestrogenic effects of TCDD. In humans, occurrence of endometriosis shows dependency on estrogen–progesterone balance, suggesting that an EDC-related mechanism may be possible." (p. 125, WHO-IPCS 2002)

The WHO-UNEP 2012 discussion on endometriosis is not an update of the data from the WHO-IPCS 2002 report as it mainly consists of a re-review of the information evaluated in 2002. The literature cited linking dioxin to endometriosis in the WHO-UNEP 2012 report was either cited in the WHO-IPCS 2002 report or is research from the same lab that provided studies considered in the 2002 review. For example, a fairly recent publication by Bruner-Tran et al. (2010) on dioxin and the role of progesterone in endometriosis is referenced in the 2012 report; however, a publication by Bruner-Tran (1999) discussing the same issue was cited in the 2002 report. The 2012 report also failed to note that the research conducted by Bruner-Tran et al. employs high doses of TCDD. These doses are at the highest end of the range highlighted in the 2002 report as being of questionable relevance to endometriosis in humans. Much of the PCB literature was also reviewed and cited in the earlier WHO-IPCS 2002 report.

Although the WHO-UNEP 2012 report cites nine publications as finding a "relationship between circulating phthalate (and phthalate esters) and endometriosis" (p. 44) – none of these publications mention phthalates. Only four publications cited in the 2012 report refer to studies on phthalates, and, as noted in the WHO-UNEP 2012 report, these studies are inconsistent because different studies have observed associations for only some, but not other phthalate esters. It is interesting to note that, in the discussion of phthalate esters and endometriosis, there is no mention of a study that did not observe an association between measures of phthalate metabolites in urine and endometriosis (Itoh et al. 2009). Regarding MOA for endometriosis, a discussion on the general hormonal mechanism of the disease is provided, but little information is reviewed on specific EDCs. The WHO-UNEP 2012 report state that epigenetic changes have been reported to be involved in endometriosis – particularly those induced by *in utero* exposure. However, only two review articles are cited: Guo et al. (2009) only mention dioxin and state that "so far there is no solid evidence linking dioxin exposure to endometriosis." Cakmak and Taylor (2010) do not mention any of the environmental chemicals highlighted as potentially associated with endometriosis (e.g., dioxin, PCBs, phthalates). Although the WHO-UNEP 2012 report acknowledges that there is limited and conflicting experimental and epidemiological evidence to support a role for EDCs in causing endometriosis, it is not clear what, if anything, has changed regarding the state of the science for endocrine disruption and endometriosis.

5.4. Conclusions

These examples illustrate how the WHO-UNEP 2012 report is not an update of the WHO-IPCS 2002 report. In some cases, the updated report reaches conclusions that conflict with those reached in the earlier report, despite the lack of new information to support a change in the weight of evidence. The fact that the 2012 report reaches more definitive conclusions based on the same data emphasizes a reliance on less stringent criteria for evaluating potential causal relationships compared to the earlier 2002 report. In light of this, although the 2012 report is stated to be an update, the 2012 report in actuality is a revised review of the state of the science because it does not build upon what was previously done in 2002 and is based on “best professional judgment” on these matters.

6. Causation Assessment

Causation is a critical element in the definition of an endocrine disruptor, and consideration of causal relationships should have been given more attention in the WHO-UNEP 2012 report. Instead, the 2012 report tended to focus on only part of the definition of an endocrine disruptor – that is, the potential for altered endocrine function. Consequently, information is presented or discussed on possible or potential endocrine disruption without considering whether actions on the endocrine system caused any adverse effects. The failure to differentiate between the *potential* for endocrine system interaction of some sort and actual disruption of the physiology or of development as a result of such interactions misleads the reader regarding the weight of evidence for particular disruptive effects. The impression that the strength of evidence is greater than warranted could result in inappropriate regulatory actions and research priorities when additional research might be needed to establish these causal relationships. Any state-of-the science review should be a balanced and objective appraisal of all of the available literature with identification of data gaps, along with clear conclusions supported by the data and an explanation of the inferences that cannot be supported. Only through such an approach can all parties (in particular, stakeholders without a strong scientific background) be confident in the evidence-based decisions that arise from such a report.

There have been many definitions for an endocrine disruptor over the years and as discussed above, the WHO-UNEP 2012 report states that the definition in general use today was relied on; this definition was also adopted in the WHO-IPCS 2002 report. Integral to this definition are three important components:

- The substance must act through altered function of the endocrine system;
- The substance must cause an adverse health effect; and
- That adverse effect must be causally related to and occur as a consequence of the altered endocrine function.

The first of the components in the definition of an endocrine disruptor is the demonstration of an endocrine-mediated MOA. This requires the differentiation between endocrine-mediated effects from other known modes of action for a particular chemical. For example, lead is mentioned as an example of an environmental exposure that causes cognitive and behavioral deficits. However, a large number of mechanisms, both direct and indirect, have been established for lead toxicity; most of these do not involve the endocrine system (Lidsky and Schneider 2003). Therefore, to demonstrate that a toxic effect of lead is due to endocrine disruption, it would be necessary to show that the effect is, for example, the result of disruption of thyroid hormone transport in the brain (endocrine-mediated) and not the result of interfering with calcium transport, storage or homeostasis (non-endocrine mediated). Further, just because a chemical is associated with changes in an endocrine organ does not necessarily make it an endocrine disruptor. If the process by which a chemical causes its effects is general cytotoxicity, oxidative stress, or other forms of systemic toxicity, then it should not be considered an endocrine disruptor. Only those chemicals that have *specific* interactions with endocrine pathways should be considered potential endocrine disruptors. Other chemicals or classes of chemicals, not normally considered potential endocrine disruptors, such as tobacco smoke, alcohol, or organophosphate pesticides, were presented in the WHO-UNEP 2012 report as though they were EDCs, but the report does not provide support that the effects discussed are the result of alterations in endocrine function.

In the 2012 report, interference with endocrine function by a chemical is often considered sufficient evidence for endocrine disruption, when in fact this only demonstrates that a chemical has the ability to interact with the endocrine system (Tinwell et al. 2013), and thus, has the *potential* to result in endocrine disruption.

The second component in the definition of an endocrine disruptor is that the substance must cause an adverse effect. Adverse effects do not include the normal fluctuations in homeostasis or adaptive responses (Goodman et al. 2010). Thus, the mere presence of a change does not necessarily mean that the outcome is adverse. This concept is discussed in more detail below in the section on dose-response.

The WHO-UNEP 2012 report frequently relies on the observation of adverse effects in endocrine organs or the existence of a possible endocrine MOAs to identify and discuss health outcomes in relationship to various chemicals and exposures. In some cases, the main text of the report reflects the uncertain nature of the association between an exposure and an effect or the limitations in the knowledge of an endocrine-mediated MOA. However, the characterization of these chemicals as only having the *potential* for endocrine interaction is sometimes lost in the discussion of the strength of evidence or in the companion report (*Summary for Decision-makers*). For example, as discussed later in this paper, the 2012 report notes an increasing trend in autism and proposes thyroid hormone deficiencies as possibly contributing to this disease. No environmental chemicals or exposures are mentioned in association with autism; yet, when discussing the strength of evidence, it states: "[m]oreover, there is sufficient evidence to conclude that a number of factors, including environmental, contribute to the increases in autism spectrum disorders" (p. 119, WHO-UNEP 2012a). Given the lack of data provided to support this statement, autism may be considered a disease of potential interest, but further research is needed to determine whether the trends in autism are even associated with endocrine disruption.

The final component of the definition of an endocrine disruptor is that the adverse effect must be *causally* related to exposure. To be able to reach this determination, one needs to have a systematic method for assessing causation. Evidence must be reviewed to determine whether the relationship between the exposure and the effect is real and that exposure actually causes that effect. Furthermore, for EDCs, the effects must be demonstrated to be the result of an endocrine MOA. Various methods have been developed to assess causal relationships; the most commonly referenced approach relates to the criteria outlined by Sir Austin Bradford Hill. In the WHO-IPCS 2002 report, a framework was based on Bradford Hill's criteria to assess relationships between exposures to potential EDCs and altered health outcomes (see WHO-IPCS 2002, Chapter 7). Several factors were specifically considered important in establishing the overall weight of evidence for a causal relationship: temporality, strength of the association, consistency of the observations, biological plausibility, and evidence of recovery. A number of illustrative examples were presented to demonstrate how the framework would work to evaluate hypotheses that particular EDCs cause specific adverse outcomes.

The WHO-UNEP 2012 report does not adopt this framework for assessing causation, claiming that this approach failed to distinguish between the quality of evidence and strength of the recommendations, as

recommended in the GRADE[3] scheme in clinical medicine (GRADE 2011). Despite reference to this alternative approach, the 2012 report did not apply the GRADE scheme in its review (which, because it was derived for clinical medicine, would not be appropriate for the evaluation of toxicology data and studies in wildlife anyway). Although there is substantial discussion of weight of evidence, no systematic approach is described or adopted to assess the weight of evidence for causation. Instead, "best professional judgment" is the basis used for making expert assessments of the data (p. 19, WHO-UNEP 2012a). Aggregated data are purported to be presented on trends, biological plausibility, relevant exposures, consistency across species, dose-response, and temporality. However, not all of these categories are captured in the discussion of all endpoints of concern; in particular, few data are presented on dose-response. Although a section titled *Strength of evidence* is included in each subchapter and conclusion-like statements are provided for some adverse health outcomes, in many cases, these generic statements do not reflect the uncertainties or limitations that were described in the main text. More importantly, these statements do not represent the totality of the evidence, often ignoring contrary study results that were inappropriately excluded. Furthermore, little effort was made to synthesize the findings across the categories in the report – from temporal trends to exposure and biological plausibility – in order to provide a complete picture of the state of science for MOA, dose response, and adverse health effects.

The use of the narrative approach in WHO-UNEP 2012 report allowed for a selective presentation of information without a critical review of the quality of the data. In addition, this approach did not result in a balanced assessment of the overall consistency of the data. While data limitations might have been mentioned in the main text, when the *Strength of Evidence* is presented at the end of each subchapter, it often does not reflect many of the uncertainties and limitations described earlier. Rather than conducting a formal assessment of causation, inference is relied on to suggest causation.

For each of the specific health outcomes, information is presented in the 2012 report for a sequence of topics in such a way to suggest that they were related when, often, there is no connection at all. Each subchapter discusses the trends in the subject disease or health outcome first, followed by a suggestion that exposure to environmental chemicals contributes to these trends. In some cases, limited or no data are provided to support that environmental exposures are contributing to the trends and often other causes of the trends are ignored. For example, the report states that there is a rising trend for breast cancer and notes that this trend cannot be explained by improved diagnosis or changes in risk factors, including genetic factors. The statement about the trends is followed by the comment that twin studies have highlighted the importance of environmental factors. The juxtaposition of these two sentences gives the reader the impression that the rising trend in breast cancer must be a consequence of these environmental exposures. However, the term "environment" as used in twin studies encompasses all modifiable (i.e., non-genetic) factors, not just environmental chemical exposures. Several publications have concluded that the observed increase in breast cancer incidence in some countries can be explained by the introduction and promotion of mammography and breast cancer screening (Glass et al. 2007, Séradour et al. 2009, Weedon-Fekjær et al. 2012). In addition, in several countries around the

[3] Grades of Recommendation Assessment, Development and Evaluation

world, the trend for breast cancer has been in decline since 2002, which has been attributed to the dramatic reduction in hormone replacement therapy in post-menopausal women (Glass et al. 2007, IOM 2012, LeClère et al. 2013, Séradour et al. 2009, Weedon-Fekjær et al. 2012). Therefore, the implication that the trends are due to environmental chemical exposures is misleading when evidence for some of the changes in breast cancer trends point to improved diagnostic tests or to changes in recommended post-menopausal therapies. In the IOM (2012) review of the state of the science regarding environmental risk factors for breast cancer, the factors with the clearest evidence included: hormone therapy products, oral contraceptives, being overweight or obese, alcohol consumption and ionizing radiation; the evidence for exposure to industrial chemicals was considered to be limited and in some cases only suggestive of a possible association for an increased risk of breast cancer. Further discussion of the issues associated with disease trends as presented in the WHO-UNEP 2012 report is described in more detail later in this paper.

Another part of the discussion for each adverse outcome includes a description of the normal role of endogenous hormones. This information is useful in understanding basic mechanisms and showing the potential biological plausibility of endocrine disruption; however, it does not demonstrate that environmental chemicals are acting in the same manner as endogenous hormones. Evidence is presented from epidemiology and experimental animal studies for particular chemicals associated with various health outcomes, but rarely is there any discussion of specific mechanisms for the highlighted chemicals. For example, androgens and estrogens are mentioned as playing a role in normal prostate development; however, an androgen- or estrogen-mediated MOA cannot be described for any chemical mentioned in this section. In fact, the report notes that "the precise mechanisms by which the chemicals related to prostate cancer induce the carcinogenic process remain to be resolved" (p. 131, WHO-UNEP 2012a). Overall, the reader is left with the impression that environmental chemicals could cause disease via the same mechanisms as endogenous hormones without any understanding of the normal feedback mechanisms that exist for homeostasis as well as differences in potency and dose-response (discussed in more detail below). Moreover, a failure to recognize the complexity of hormone-receptor interaction and activation is lacking. Specifically, despite a short discussion at the beginning of the WHO-UNEP 2012 report, the role of co-activators, repressors and transcription factor interactions and receptor cross-talk is completely ignored. Recent studies also point to competition for transcription factors (Kollara and Brown 2006); yet these issues are completely ignored.

Another shortcoming of the WHO-UNEP 2012 report that affects the cohesive evaluation of causation is the fact that little, if any, discussion of exposure occurred in the report in the context of specific effects or related to specific hormonal MOAs. Generally, when exposures were mentioned, a reference is made to Chapter 3 of the 2012 report. In Chapter 3, only general information is presented on major classes of chemicals that the WHO-UNEP 2012 report describes as known or potential endocrine disruptors. No quantitative information on exposure is presented for individual chemicals, and potential human exposures are not considered in context of dose-response data. The segregation of exposure data from the information on chemicals of concern, potential adverse effects associated with that exposure, and biological plausibility or possible modes of action for endocrine disruption makes the assessment of the causal relationship between an exposure and an effect impossible.

The lack of a systematic approach to assess causation for specific chemicals and associated health outcomes resulted in conclusions that were predisposed to the identification of potential EDCs. The selective citation of literature without discussion of contradictory studies and the failure to consider alternative causes of reported effects gives the reader the impression that the weight of evidence is stronger than is justified by the available scientific data. This calls into question the integrity of the decisions at all levels of the 2012 report. Specific examples are provided below that further demonstrate the issues with the evaluation of causation in the WHO-UNEP 2012 report, with an emphasis on highlighting the key factors that are typically used in a causation assessment.

6.1. Adrenocortical hyperplasia in seals

The WHO-UNEP 2012 report concluded:

> "There is sufficient evidence to show that adrenocortical hyperplasia and a suite of pathological changes characteristic of Cushing disease in Baltic seals were caused by exposure to a mixture of DDT, PCBs and their methylsulfonyl metabolites; along with the drastic reduction of DDT and PCBs in Baltic biota, the seal populations have gradually recovered" (p. 149, WHO-UNEP 2012a)

This conclusion is not supported by the discussion in the main text of the report; at best, it is based on limited data. Additionally, the report ignores conflicting data and fails to consider alternative causes for the adrenocortical hyperplasia.

The data on adrenocortical hyperplasia in seals or other aquatic mammals are limited and inconsistent. Although the WHO-UNEP 2012 report mentions that adrenocortical hyperplasia has not been reported in seal populations outside the Baltic Sea, it does not discuss specific evidence from Great Britain that reported contrasting results for other marine mammals. For example, Kuiken et al. (1993) measured the concentrations of chlorinated hydrocarbons in the carcasses of harbor porpoises and found that adrenocortical hyperplasia was not associated with increased levels of these chemicals, but rather associated with exposure to chronic stressors causing their death. Another study that is not mentioned, Clark et al. (2006), also found a significantly higher mass of the adrenal glands in Atlantic bottlenose dolphins that were chronically stressed compared to those that were acutely stressed, suggesting that other factors may be involved in the observation of adrenocortical hyperplasia in seals. The WHO-UNEP 2012 report also does not address alternative causes for these observations; Lair et al. (1997) suggest that the adrenal hyperplasia seen in beluga whales may be part of the normal aging process. Although these latter studies are with cetaceans rather than pinnipeds, they do point to alternative causes of adrenocortical hyperplasia in marine mammals and should have been considered in the data review and assessment of causation.

The report also fails to address the data inconsistencies for the specific persistent organic pollutants mentioned, including the differences observed in experimental animal studies and reports in wildlife. For example, two-year chronic toxicity studies of PCB exposures in female Sprague-Dawley rats (NTP 2009, 2006) found evidence of increased adrenocortical atrophy, which is in contrast to the findings of adrenocortical hyperplasia found in Baltic seals. While the degree of similarity between adrenal glands

of seals and rats is unknown, it would be more likely to observe similar rather than opposite effects from the same chemical. The cancer bioassays for several Aroclors did not report adrenal effects in rats exposed to these PCB commercial mixtures in the diet for two years (Mayes et al. 1998), which further calls into question the identification of PCBs as the cause of adrenocortical hyperplasia in seals. The WHO-UNEP 2012 report acknowledges in the introduction to the section on adrenocortical hyperplasia that a DDT metabolite showed degeneration and necrosis in the adrenal cortex of laboratory mice – a finding inconsistent with hyperplasia – but this is not mentioned further.

No information on plasma cortisol levels is available for the Baltic Sea seals, as the WHO-UNEP 2012 report notes, and it is critical to differentiate a stress response from a direct toxic effect on the hypothalamus-pituitary-adrenal axis as the cause of the observed hyperplasia (Harvey and Sutcliffe 2010). Although the WHO-UNEP 2012 report acknowledges the possible role of stress and aging of wildlife in the development of adrenal hyperplasia in the main text, when the strength of evidence is described, these factors are ignored. The 2012 report does not consider alternative causes for the species recovery, and stress can just as plausibly explain these observations given other changes (such as reductions in nutrient inputs, eutrophication, oxygen deficiency, and oil discharges) occurring during the same time period in the Baltic Sea (HELCOM 2012) which could have led to recovery.

In conclusion, the WHO-UNEP 2012 report states that there is sufficient evidence that adrenocortical hyperplasia was caused by exposure to various persistent organic pollutants (POPs), including DDT, PCBs, and their methylsulfonyl metabolites. However, this conclusion is based on limited data from one region, the Baltic Sea, and is inconsistent with findings from other studies (Kuiken et al. 1993, Clark et al. 2006). Furthermore, the WHO-UNEP 2012 report fails to consider alternative factors for the recovery of these species in the Baltic Sea. Additionally, it is presumed that because the effect occurred in an endocrine organ, it must be the result of endocrine disruption, but no data are provided to show that these effects are the result of an endocrine MOA. Finally, the report does not evaluate stress as a factor in the development of adrenocortical hyperplasia. Given the limited data available on the observation of adrenocortical hyperplasia, inconsistent findings in experimental animal studies, conflicting data in other wildlife species, and stress as a plausible alternative cause for these observations, it is questionable that there is sufficient evidence to demonstrate that these compounds caused the adrenocortical hyperplasia observed in the Baltic seals.

6.2. Prostate cancer

Prostate cancer is included in the discussion of various hormonal cancers in Section 2.7.2.4 of the WHO-UNEP 2012 report and it is stated that there is sufficient evidence for a link between pesticide exposures and prostate cancer. However, the report does not reach any conclusions or make any statements about the evidence for an endocrine-mediated MOA. The mere association between pesticides and prostate cancer is insufficient to demonstrate causation. More importantly, this link has not been shown to be attributable to an alteration in endocrine function. Moreover, based on the data presented in main text of the report, from an objective view, there is not sufficient evidence to conclude that a link between pesticide exposures in general and prostate cancer even exists.

The report mentions that individual pesticides have been reported to be associated with prostate cancer; however, these data are not consistent. For example, the report mentions that oxychlordane was linked with an increased risk for prostate cancer based on Ritchie et al. (2003). However, the fact that another study cited in the report (Hardell et al. 2006) did not observe an association between oxychlordane and prostate cancer is ignored. Other biomonitoring studies that examined oxychlordane levels that are not cited in the 2012 report include two studies that failed to observe an association with prostate cancer (Aronson et al. 2010, Sawada et al. 2010) and one that did (Xu et al. 2010). Thus, the evidence from biomonitoring studies for an association between oxychlordane and prostate cancer is inconsistent. Similar results were seen for other organochlorine pesticides where only one or two statistically significant associations were reported for any individual pesticide. Although other epidemiology studies are cited in the WHO-UNEP 2012 report, these studies did not directly measure exposure through analysis of blood, fat, or urine, and therefore, are considered to be more susceptible to bias and should be given less weight in an overall assessment of the strength of evidence. For example, the Agricultural Health Study (AHS) (Alavanja et al. 2003, Koutros et al. 2010) obtained exposure information based on a questionnaire that collected data on duration and frequency of pesticide use, which is less precise when compared to biomonitoring studies.

It is interesting to note that the two meta-analyses of pesticide applicators (van Maele-Fabry and Willemas 2004) and pesticide manufacturers (van Maele-Fabry et al. 2006) characterized the strength of evidence for pesticide exposure for these workers as weak (rate ratios of less than two). A general limitation of the epidemiology studies on pesticide exposures and cancer is the use of multiple comparisons to assess risks for many different types of pesticides and various cancer endpoints, which can lead to false positive findings. In addition, the long latency between initiation and detection of the cancer make it very difficult to identify relevant exposures. Overall, the data for pesticides in general are weak and those for individual pesticides are limited with regard to an association with prostate cancer.

While it has been speculated that hormones play a role in the development of prostate cancer given the involvement of sex steroids in the development of the prostate, it is recognized that many other factors may be involved in the etiology of prostate cancer. These factors include age, family history (genetics), race, dietary fat, and other dietary factors. As the WHO-UNEP 2012 report notes, the mechanism by which pesticides could induce prostate cancer is currently unknown. Therefore, while it may be biologically plausible for endocrine disruption to be contributing to prostate cancer, insufficient data are available to show that pesticides are involved in the induction of prostate cancer by an endocrine-mediated MOA.

In conclusion, while the WHO-UNEP 2012 report states that there is sufficient evidence for an association between pesticides and prostate cancer, this evidence is weak and generally limited to single observations for individual chemicals – not all pesticides. Furthermore, it is acknowledged that the mechanism by which prostate cancer is induced has not been established. Consequently, there is a lack of evidence to demonstrate that this association is due to endocrine disruption. Based on these issues regarding the strength of the association, lack of consistency among studies of pesticide workers and manufacturers and lack of evidence for an endocrine MOA, the overall strength of evidence for

pesticides causing prostate cancer through endocrine disruption is weak as concluded in the WHO-IPCS 2002 report.

6.3. Conclusion

As the above examples illustrate, the WHO-UNEP 2012 report presented information on chemicals and various adverse outcomes, but whether the exposure causes these effects was not determined. Several factors, such as demonstrated exposure to the chemical, dose-response, and consistency in the data were frequently ignored. Above all, the lack of a framework to collectively evaluate the data on specific chemicals and the alleged adverse outcomes is a significant shortcoming to the WHO-UNEP 2012 review.

7. Temporal Trends in Health Outcomes

The WHO-UNEP 2012 report indicates that the high incidence and increasing trends of many endocrine-related disorders in humans is one of "three strands of evidence [that] fuel concerns over endocrine disruptors" (p. vii). The report indicates that, "worldwide, there has been a failure to adequately address the underlying environmental causes of trends in endocrine diseases and disorders" (p. ix, WHO-UNEP 2012a).

The report concludes that, because the increase in disease trends has occurred primarily over the last few decades, it cannot be entirely attributable to genetic causes. It is implied that if these trends are not the result of genetic heritability, then the only other explanation is environmental exposure. However, environmental factors traditionally cover an endless array of factors, including diet, exercise, lifestyle factors, infectious agents, and even drug use; for wildlife, it includes factors related to habitat, food supply, disease, predation, and competition – factors which are not related to environmental chemical exposures. Yet, the WHO-UNEP 2012 report does not go out of its way to point out that, much of the time, the environmental causes of the diseases being discussed are not chemical exposures.

Perhaps most importantly, there are many other factors that can influence the appearance of an increasing trend (either temporally or geographically) in disease incidence. For human health considerations, these include changes in diagnostic criteria, screening, medical interventions, and treatment. Other trends in lifestyle can have substantial effects, for example, giving birth to children at an older age impacts the incidence of birth defects and congenital abnormalities. Another significant trend is the obesity endemic and being overweight has been found to increase the risk for male infertility (Hammoud et al. 2008). In addition, trends sometimes are assessed by compiling data from different sources, so what appears to be a trend actually may be a reflection of different data collection methods.

By selectively referencing publications, an impression is created for the reader that certain diseases have an increasing prevalence. However, publications that are not cited in the WHO-UNEP 2012 report provide opposing evidence that the disease prevalence of interest is not on the rise at all. For example, several papers are cited to point out that the incidence of hypospadias is increasing. However, Fisch et al. (2010) state that “[a] review of the epidemiologic data on this issue amassed to date clearly demonstrates that the bulk of evidence refutes claims for an increase in hypospadias rates.” The postulated decreasing trends in semen quality have also been highly contested by other scientists.

In the end, even if an exposure and a health outcome trend are related spatially or temporally, all one can be sure of is that a statistical correlation exists. Because it generally cannot be known if the exposure caused the health outcome (or vice versa), whether they each have a common cause, or whether they are completely independent, these analyses are the weakest form of scientific evidence for evaluating causation. It also cannot be known whether people with health effects are the same people with a particular exposure.

The reasoning in the WHO-UNEP report that the rising disease trends must be associated with exposure to EDCs becomes even more questionable given the fact that exposure to most of the compounds

named in the report have not increased over the last twenty or thirty years, but rather, have decreased. Concentrations of DDE in human milk have gone down from 1µg/l in 1984 to levels close to 0.1 µg/l in 2001 (Wilhelm et al. 2007). Another biomonitoring study clearly shows that human exposure to phthalates has declined over time with the exception of some new compounds that were recently introduced to the market (Wittasek et al. 2007). Historical biomonitoring data from a number of countries including the Czech Republic, Norway, and the U.S. indicate that human exposure to compounds like persistent chlorinated pollutants has decreased in the last two decades (Cerná et al. 2012, Ferriby et al. 2007, Nøst et al. 2013). These declining trends in human exposure over the last two decades in the Western world run counter to rising disease trends.

In some cases, the WHO-UNEP 2012 report puts trends in perspective. For example, the report indicates that elevated BPA levels could be a result of polycystic ovarian syndrome and not the other way around. In other cases, however, alternative explanations are discussed, but dismissed in favor of endocrine disruption as an explanation without a sufficient evaluation of the science. For example, as noted earlier, the report briefly acknowledges, but appears to dismiss, the role of alternative factors in the observed tends for breast cancer. The introduction and promotion of mammography and breast cancer screening are associated with the increasing trend in breast cancer incidence since the 1980s and in some parts of the world, declines in breast cancer incidence since 2002 are associated with changes in the use of hormone replacement therapy (Glass et al. 2007, Seradour et al. 2009, Weedon-Fekjaer et al. 2012).

Overall, any claims of trends indicating an endocrine cause must be supported by systematic review of all relevant data regarding disease trends, exposure trends, and alternative explanations for observed statistical correlations. Below, two examples are discussed in which the report describes trends of increasing endocrine-related disorders and concludes they are due to environmental EDCs, without considering whether the weight of evidence supports such conclusions or alternative explanations. Note that these examples are not weight-of-evidence analyses, but the identification of factors and limitations related to the discussion of health trends in the WHO-UNEP 2012 report and how this may lead to erroneous conclusions.

7.1. Autism spectrum disorders (ASDs)

Autism spectrum disorders (ASDs) – which include autistic disorder, Asperger's syndrome, and pervasive developmental disorders not otherwise specified- – are developmental disabilities that are diagnosed based on behavioral symptoms and failure to reach developmental milestones (CDC 2012). ASDs are characterized by communication and socialization problems, as well as atypical behaviors and interests. ASD symptoms are usually apparent before the age of three and can vary in severity and presentation. While IQ decrement can co-occur with ASDs, they are not associated with ASDs *per se* (CDC 2012).

The WHO-UNEP 2012 report claims that "the increase in autism spectrum disorders is indisputable" (p. 109) and there is "sufficient evidence to conclude that a number of factors, including environmental, contribute to the increases in autism spectrum disorders" (p. 119). The first claim is based on two studies, one published in 1976 and the other in 2007 (Wing et al. 1976, Rice 2000). As indicated in the title of the latter article ("Prevalence of autism spectrum disorders--autism and developmental

disabilities monitoring network, six sites, United States, 2000"), Rice (2000) did not evaluate current prevalence of ASDs. In addition, there is a considerable body of other literature evaluating ASD prevalence that the WHO-UNEP 2012 report did not consider.

At least some of the increase in ASD prevalence is due to better diagnostic techniques and increased case ascertainment (CDC 2012). For example, when the Centers for Disease Control and Prevention (CDC) created new diagnostic standards for its Autism and Developmental Disabilities Monitoring Network (ADDM), there was a significant increase in ASD prevalence compared to the prevalence estimated using older standards (Rice et al. 2012). Also, a recent reanalysis of older studies found ASD prevalence to be consistent with current reports when the data were analyzed according to contemporary diagnostic criteria, indicating that ASD was likely underestimated in earlier studies (Duchan and Patel 2012). Increases in ASD prevalence may also be partly attributed to diagnostic substitution, as children who would have been diagnosed with learning disabilities or mental retardation in the past are currently diagnosed with ASD (reviewed by Fombonne et al. 2009). ASD diagnosis relies on behavioral identification, which leaves room for wide variation in clinical judgment and is influenced by differing cultural and social norms worldwide (Elsabbagh et al. 2012). Even within a culture, there is evidence of low inter-evaluator agreement about diagnoses, although it is not clear in which direction it would influence prevalence measures (Duchan and Patel 2012). Finally, the success of national awareness efforts may also contribute to a perceived increase in ASD prevalence, with more children being tested, diagnosed, and treated (Duchan and Patel 2012). Therefore, it is not clear that there is a true increase in ASDs.

Regardless of temporal trends, the WHO-UNEP 2012 report presents no evidence that environmental factors, much less EDCs, contribute to ASDs. The report summary states that "insufficiency of thyroxine during pregnancy is also associated with reduced intelligence quotient, ADHD and even autism in children" (p. xii, WHO-UNEP 2012a). Yet, the WHO-UNEP 2012 report does not provide a reference to support this claim nor discuss any other factor(s) that may contribute to ASDs. Many potential non-EDC risk factors have been studied, including genetics, older paternal age, sex, prenatal nutrition, and *in utero* exposure to antidepressants and pain killers (e.g., Guinchat et al. 2012, Duchan and Patel 2012, Schmidt et al. 2011, Gentile et al. 2013, Kinast et al. 2013). Although the strength of the evidence supporting these associations varies, the available data indicate that a good deal of research on possible causes of ASDs is not considered in the WHO-UNEP 2012 report.

Overall, the two references (Rice 2000, Wing et al. 1976), cited regarding endocrine disruption and trends reported for ASDs, on which the WHO-UNEP 2012 report relies are not representative of the literature as a whole. Thus, the report's conclusions that there are actual increases in the spectrum of autism-related disorders and that these increases are due to endocrine disruption are not supported based on the current state of the science.

7.2. Wildlife Population Declines

The WHO-UNEP 2012 report states that the evidence for "endocrine disrupting POPs such as PCBs and organochlorines" (p. 186) as causes of wildlife population declines has increased since 2002 due to observed increases in the populations since restrictions on the use of these chemicals. The logic

presented – that as chemical exposures increased, populations declined and, conversely, as chemicals were removed from the market and exposures declined, populations recovered – is logical insofar as 1) the chemical exposures are documented; 2) the levels of exposure occurring are sufficient to impact the organisms; 3) the organism-level impacts are manifested in population-level impacts; and 4) other possible causes for population changes are adequately considered. The WHO-UNEP 2012 report falls short in demonstrating the linkages that would be required to make a case based on all of these points.

The two most prominent examples cited are links between DDT and bird populations and between tributyl tin (TBT) and snail populations. For the latter example, the WHO-UNEP 2012 report cites publications by Jörundsdóttir et al. (2005) and Morton (2009). Jörundsdóttir et al. (2005) observed reductions in the levels of imposex in the dogwhelk (*Nucilla lapillus*) in Iceland, mainly near small harbors with no change in larger harbors. As no measurements were made of TBT concentrations, Jörundsdóttir et al. (2005) stated that the continued impacts in the large harbors are "presumably associated" with continued use of TBT paints on larger vessels. Morton (2009) documented a 20-fold increase in the population of *N. lapillus* on the southeastern coast of England during the period May 2004 to August 2008, which coincided with the period over which TBT was banned as an anti-foulant paint globally. Morton (2009) stated that, "due to the lack of confirmatory chemical data, the changes in population size, structure, and reproduction herein reported upon for *N. lapillus* cannot be correlated positively with changes in ambient TBT levels." The WHO-UNEP 2012 report discusses recovery in the abundance of North Sea brown shrimp, although there is no known mechanism of endocrine disruption by TBT in crustaceans. Verhaegen et al. (2012), as cited in the WHO-UNEP report, state that the inability to demonstrate an "unarguable causative link" between decreased organotin concentrations and recovery of the shrimp stock is due to the lack of data on both exposure and effects in these organisms. None of these weaknesses in the conclusions of the cited studies are mentioned in the WHO-UNEP 2012 report, and therefore, the observed trends of declines in imposex cannot be attributed to reductions in TBT, much less endocrine disruption.

Despite the fact that the ability of organotins to cause masculinization of female gastropods (including the development of imposex) is probably the most-recognized EDC effect in wildlife over the past 30+ years, the vast majority of field studies on this phenomenon do not include chemical analyses of body burdens (Titley-O'Neal et al. 2011). Although this article by Titely-O'Neal et al. (2011) is cited in the WHO-UNEP 2012 report, a number of interesting points from the review were not mentioned. For example, the WHO-UNEP report does not note the lack of agreement among researchers on the mechanism of induction of effects, the observation of imposex prior to the use of TBT, the natural occurrence of imposex in some species, the lack of sensitivity of a number of species, and the fact that female masculinization by TBT or triphenyl tin (TPT) has been confirmed in the laboratory in only a small fraction of species affected (7.5% or 20 species confirmed out of 268 total species). Thus, the statement in the WHO-UNEP 2012 report that the "temporal relationship between a measure of exposure and population parameters" for TBT is an example of the "best evidence of a relationship between EDCs and wildlife populations" does not reflect the uncertainties in the available information, including studies cited in the report.

The WHO-UNEP 2012 report acknowledges the difficulty in making the link between declines/recoveries in wildlife populations and EDCs, stating that many factors may be responsible. These factors may include food, habitat, competition, predation, overall environmental quality, climate change, and human activities (e.g., harvesting, traffic, noise). Regardless, the report emphasizes chemicals, specifically EDCs, as the causative factor. Even in the case of TBT and *N. lapillus*, which is arguably the best known example of EDC effects on wildlife, there are other factors that impact the distribution and abundance of this gastropod. This species is sensitive to changes in nutrient levels, substrate loss, toxic algal blooms, and oil spills (Bryan 1968, Gibbs et al. 1999, Robertson 1991,). For the DDT example, a number of confounding factors are likely to have affected the recovery of osprey populations, as discussed by Henny et al. (2010), who stated that "expansion of suitable habitat (reservoirs) and enhanced use of artificial nest sites confounds a simple conclusion that recent population increases were solely a recovery from earlier contaminant exposure," especially in the western United States. The WHO-UNEP 2012 report concludes that the "strength of the evidence linking EDC exposure to most wildlife population declines is insufficient," then goes on to make the statement that "an endocrine mechanism for wildlife declines is probable but not conclusive" (p. 186, WHO-UNEP 2012a). It would be more appropriate to conclude that the strength of the evidence that EDC exposures are solely responsible for wildlife population declines is insufficient, even for the two best known examples, DDT and TBT, and an endocrine mechanism for wildlife declines is possible, but only one of many potential factors.

7.3. Conclusion

The WHO-UNEP 2012 report does not provide sufficient evidence that trends in health outcomes are correlated with trends in exposures to EDCs, much less that the associations are causal. There are many assertions made without references to support them or, at best, a limited number of references are given when many more are available and may reflect an opposing view. Even for the references that are cited, in some instances, the study authors' own conclusions about the weaknesses of the significance of the findings are ignored. As demonstrated with the two examples above (ASD and wildlife population declines), when a more thorough review is conducted it is clear that the state of the science does not support environmental EDCs as responsible for observed trends in adverse outcomes as a general rule.

8. Dose-response and Potency

In Chapter 1 of the WHO-UNEP 2012 report, an effort is made to describe the endocrine system – the glands involved, hormones produced, molecular mechanisms involved in mediating responses, and physiological processes that are regulated by this system. A few of the feedback mechanisms that are an integral part of this system (e.g., how insulin secretion is affected by changes in blood glucose levels) are also mentioned. These various negative feedback loops are important in regulating the production and release of hormones, the expression of various hormone receptors, and generally maintaining homeostasis (i.e., a stable internal environment). What the WHO-UNEP 2012 report fails to fully discuss, however, is the fact that the endocrine system is specifically designed to respond to environmental fluctuations and such homeostatic responses which generally are considered normal and adaptive as long as they are transient and within the normal homeostatic range (Rhomberg et al. 2012, Goodman et al. 2010). In fact, the responsive nature of the endocrine system is essential to health as seen in the hormonal changes that occur when a woman becomes pregnant. In other words, not all modulations of endocrine function are necessarily adverse. Based on this fact, it can be generally accepted that endocrine activity observed through *in vitro* testing (or even some *in vivo* assays) is not sufficient to classify a substance as an endocrine disruptor if these tests do not address whether the alterations cause actual harm in a whole organism or its offspring. Rather, such a substance may be considered endocrine-active (EFSA 2013a); without a clear indication of consequent adversity in a living organism, the substance does not reach the level of an endocrine "disruptor." Nevertheless, the WHO-UNEP 2012 report often presented evidence of *in vitro* or *in vivo* endocrine modulation (rather than adversity) as support for certain substances being classified as EDCs. For example, in the section on adrenal disorders, only *in vitro* data are presented as evidence of potential endocrine disruption in humans; no epidemiologic evidence of adrenal effects in people is available. Given the fact that a large proportion of the cited *in vitro* studies relate to alterations in gene expression, these data fall short of demonstrating an adverse effect and alone do not show endocrine disruption. Despite the lack of robust evidence for adrenal disorders in humans as a result of exposure to environmental chemicals, the WHO-UNEP 2012 report identifies the adrenal cortex as "the most commonly affected and vulnerable endocrine organ in toxicology" (p. 148, WHO-UNEP 2012a).

Although endocrine disruption is generally posited throughout the WHO-UNEP 2012 report in terms of "adverse" outcomes, the report fails to provide a concrete definition for what may be considered an adverse response. In particular, the 2012 report did not adopt the IPCS (2004) definition of an adverse health effect: “change in morphology, physiology, growth, development, reproduction or life span of an organism, system, or (sub) population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influence.” The need to clearly delineate adverse from adaptive responses, particularly when considering results from *in vitro* assays, was addressed and possible definitions for these terms proposed in a recent workshop (Keller et al. 2012). Certain endocrine responses – such as cancer or reproductive disorders – clearly can be judged as detrimental. For other endpoints – such as alterations in hormone levels – it is more difficult to delineate an adaptive response that is within the limits of homeostasis from one that has gone beyond those limits and is therefore considered adverse. This is particularly true when the response is only measured at a single time point shortly after exposure with

no indications of whether the response is transient or more permanent. Thus, the mere presence of a change does not necessarily mean that the outcome is adverse. In the subchapter on thyroid disorders (p. 97, WHO-UNEP 2012a), the report suggests that chemicals can interfere with thyroid hormone signaling without affecting serum hormone levels, but it is stated that methods to evaluate this are not yet available. Changes in thyroid hormone signaling alone cannot be characterized as adverse without evidence to show that these changes lead to impairment in function (Bianco and Kim 2006). Furthermore, the failure to observe changes in serum thyroid hormone levels would indicate a lack of consequence from the change in signaling.

In its opinion on the scientific criteria for the identification of endocrine disruptors, EFSA specifically discuss a "threshold of adversity," noting that toxicologically relevant responses occur only when the degree of endocrine modulation elicited is beyond that which could be counteracted through homeostatic mechanisms (EFSA 2013a). These thresholds are not unlike those that exist for responses measured from other physiological systems operating within the body. Further, their importance in characterizing endocrine disruption has been emphasized by toxicologists concerned about the European Commission's recommendations for the regulation of EDCs (Dietrich et al. 2013). In contrast, the WHO-UNEP 2012 report proposes that thresholds for endocrine disruption "should not be assumed" (p. 19, WHO-UNEP 2012a) and exposures to endocrine-active substances – no matter the level – will add to the already present hormone levels in the body and thus alter endocrine function in a threshold-independent manner. However, using simple mathematical calculations within a systems biology construct, Borgert et al. (2013) have shown that the endocrine system is able to discriminate potent hormonal signals from the "background noise" of other endogenous molecules, thus making the system relatively robust in its responses and resistant to spurious interferences by substances with lower potency.

Because thresholds exist, not only for inducing an endocrine response but also for moving beyond adaptive modulation into the realm of adversity, it is important to understand at what doses the observed responses occur and how these doses compare to the levels at which people or wildlife are typically exposed. In the Chapter 2 discussions, regarding various adverse outcomes, however, the WHO-UNEP 2012 report often fails to mention the doses at which findings are observed. It is important to note that doses administered in experimental animal studies are often above those to which people or wildlife are generally exposed. Frequently, if one delves deeper into the literature, it is discovered that the doses associated with the reported findings are extremely high, well above what people may be typically exposed. To illustrate, a few selected examples are discussed below. These are not isolated examples but, rather, representative of how the lack of the consideration of dose in the WHO-UNEP 2012 report leads to a false impression that humans are at risk of endocrine effects from their daily exposures to chemicals.

8.1. DES or Genistein and Endometrial Cancer

In the discussion of animal studies of EDCs and endometrial cancer, the WHO-UNEP 2012 report cites the study of Kabbrarah et al. (2005) as showing "(G)reater than 90% of CD-1 pups neonatally exposed to DES or the phytoestrogen genistein develop endometrial cancer by 18 months of age whilst C57Bl/6 mice are resistant" (p. 130, WHO-UNEP 2012a). In this study, both DES and genistein were injected

subcutaneously into the pups (a route of administration not relevant to environmental exposures) at doses of 1 mg/kg/day and 50 mg/kg/day, respectively, on postnatal days 1-5. The DES dose is over 1,000-fold higher than the typical estrogen dose that women receive from low-dose daily oral contraceptive pills (Kripke 2005). Further, the daily intake of genistein in Japanese subjects has been shown to be <1 mg/kg/day (Wakai et al. 1999, Nakamura et al. 2000), at least 50-fold lower than the dose administered to mice in the above study. Thus, the doses used in the study by Kabbrarah et al. (2005) are well beyond those to which people would be typically exposed. Another issue is that this study was actually conducted in knockout mice with a genetic predisposition for DNA repair errors – a fact that is not mentioned in the WHO-UNEP 2012 report discussion and that brings the human relevance of the findings into question.

8.2. PCBS and Neurodevelopment

In another example, in the discussion on neurodevelopmental disorders, the WHO-UNEP 2012 report cites three studies as consistent evidence that "PCB exposures decrease serum thyroid hormone levels," (p. 113) with no mention of study details. However, the doses at which effects were observed in those studies are extremely high. In Goldey et al. (1995) and Zoeller et al. (2000), rats were exposed to 1, 4, or 8 mg/kg/day of Arochlor 1254 on gestational days 6-21. Although circulating T_4 levels were reduced on postnatal days 1-30, they recovered by postnatal day 45, indicating a transient effect. More importantly, Goldey et al. (1995) reported that pup mortality was 20% and 50% in the 4 mg/kg/day and 8 mg/kg/day dose groups, respectively, indicating that these doses were extremely high. In the third study (Bastomsky, 1974), adult rats were injected with an even higher dose of 25 mg/kg/day of Arochlor 1254 for 4 days. In contrast, the mean intake of PCBs from consumption of the French diet was recently estimated at 2.71 ng/kg/day for adults and 3.77 ng/kg/day for children (Sirot et al. 2012), while that from consumption of the Japanese diet was estimated at 1.45-2.08 pg/kg/day (Nakatani et al. 2011). These levels are 1,000,000-1,000,000,000 times lower than those used in the studies cited in the WHO-UNEP 2012 report as consistent evidence of effects of PCBs on thyroid function.

It should be further noted that, although the thyroid develops and functions in a manner generally similar between rodents and humans, differences exist that make neonatal rats more susceptible to alterations and less capable of compensating for possible alterations in function than humans. For example, the human fetal pituitary can respond to thyroxin-releasing hormone (TRH) as early as gestation week 25 and thyroid stimulating hormone reaches peak serum levels sometime around then, while the hypothalamic-pituitary-thyroid axis in rats does not response to TRH signals until a couple of weeks after birth (Howdeshell et al. 2002). Further, free thyroid hormone levels can be maintained during pregnancy in humans via increased peripheral metabolism and enhanced thyroid hormone binding to serum proteins (Howdeshell et al. 2002, Ahmed et al. 2008). Consequently, children with congenital hypothyroidism may be born with low-normal concentrations of thyroid hormone due to compensation by the maternal system (Ahmed et al. 2008).

8.3. BPA and Adverse Effects

Throughout the WHO-UNEP 2012 report, BPA is mentioned as being responsible for a variety of adverse findings in rodents, including fibroids development in mice and rats (p. 42); defeminization and other

alterations in social behaviors in female rats(p. 115); altered mammary gland development leading to increased tumor induction(p. 128); endometriosis in offspring of exposed mice(p. 130); and modified immune responses in mice(p. 169), to name a few. In all of these cases, the doses of BPA associated with these findings are not reported. Further, in Chapter 3 of the WHO-UNEP 2012 report, the various ways in which people may be exposed to BPA is emphasized (e.g., in the call-out box on page 196 under the heading called "origin and use"). The report also notes that BPA is found in virtually all people (p. 225), but no information is provided on the magnitude of exposure, the biological concentrations that have been measured in people, or whether these would be sufficient to cause effects. The implication of these data is that people are at risk of adverse health effects because they are exposed to BPA. However, a number of recent weight-of-evidence evaluations have been conducted to assess the potential risks to humans from BPA exposure (Goodman et al. 2006, 2009, Hengstler et al. 2011, Teeguarden and Hanson-Drury 2013, EFSA 2013b). These reviews find that some BPA results reported in investigatory experiments have not been replicated in subsequent studies; many studies have used non-oral exposure routes that bypass first-pass liver metabolism and, thus, are not relevant to human oral exposures; the majority of BPA studies have been conducted at doses well above those to which humans are generally exposed; and human exposures are generally well below the current BPA TDI of 0.05 mg/kg/day derived from two- and three-generation reproductive studies in rodents. More specifically, daily BPA exposures were recently estimated by EFSA (2013b) to be ≤857 ng/kg/day for toddlers and ≤495 ng/kg/day for infants 1-5 days of age; these values are 50- to 100-fold lower, respectively, than the BPA TDI value of 0.05 mg/kg/day. Daily BPA exposures were also estimated by the U.S. Food and Drug Administration (FDA) to be 100-200 ng/kg/day for children and adults and 200-400 ng/kg/day for infants (FDA 2009); these values are even lower than those estimated by EFSA. In other words, the implication of human health risks from BPA exposure raised in the WHO-UNEP 2012 report is unfounded when the data are considered in the context of actual doses administered and concentrations to which people are typically exposed.

It is unclear why the WHO-UNEP 2012 report fails to consider dose in its discussions of evidence for endocrine disruption in humans and wildlife. In the beginning of Chapter 2, it is noted that the focus was on the "identification of the characteristics of the hazards posed by endocrine disruptors rather than risk assessment," (p. 23) because accurate risk assessments are difficult in light of limited human exposure data and the combined effects of mixtures. However, the WHO-UNEP 2012 report often draws conclusions that appear to go beyond a simple assessment of potential hazard. For example, the report concludes that environmental exposures *play a role* in the observed increased incidences of hormonal cancers (rather than saying that these exposures *have been associated with* the cancers) and (p. 137) the adrenal changes seen in Baltic seals *were caused by* exposure to DDT, PCBs, and their metabolites (instead of saying that they *have been associated with* these exposures) (p. 147). Further, at the end of Chapter 1, the WHO-UNEP 2012 report states that "[b]est professional judgment was used to make expert assessments of the data linking exposure to chemicals with each disease/dysfunction," (p. 19) and relevant exposures and dose-responses were considered. Therefore, dose and exposure were specified as important factors in the evaluation, although this does not appear to be the case.

The WHO-UNEP 2012 report contends that hormonal dose-response curves are non-monotonic (i.e., have an inflection point at which the sign of the dose-response curve (+ or -) changes) and endocrine

disruptors can act at very low doses (i.e., doses below the no observed adverse effect level (NOAEL) or a dose that is environmentally relevant to humans). The implication is that there is no threshold for adverse effects; therefore, dose is not relevant. Perhaps this is one of the reasons that the WHO-UNEP 2012 report generally fails to fully consider the doses at which effects are observed in experimental studies. As support for NMDRs and low dose effects, the report cited a recent review by Vandenberg et al. (2012) in which numerous examples of endocrine-disrupting chemicals that exhibit these types of behaviors were illustrated. This review has been duly criticized, however, for its selective dismissal of studies that do not show these effects and the general acceptance of those studies that do show these types of responses without any type of critical evaluation of study quality; the inclusion of studies that do not address adverse effects, but rather, transient, adaptive responses; and a failure to consider whether the doses examined in these studies are of any relevance to human exposure levels (Rhomberg and Goodman 2012). Further, the EFSA noted that studies of low-dose effects often suffer from various methodological shortcomings (including the use of small numbers of animals and single doses) and their findings are of questionable toxicological relevance and frequently cannot be replicated in subsequent, more robust studies (EFSA 2010).

The Danish Centre on Endocrine Disrupters,[4] in its examination of the evidence presented by Vandenberg et al. (2012) for NMDRs, noted that the majority of these data were from *in vitro* studies and inappropriately included findings for which the U-shaped curves were the product of general toxicity (DTU Food 2013). The IDEC concluded that 45% of the *in vitro* examples cited by Vandenberg et al. (2012) were the result of cytotoxicity and, thus, were not examples of true NMDRs. Of the remaining examples cited, appropriately one-third were judged to be false and another one-third were considered questionable. Further, only 5 of the 34 *in vivo* examples cited by Vandenberg et al. (2012) were considered to show "clear evidence" of NMDRs. In other words, while examples of NMDRs do exist, they are not as common as Vandenberg et al. (2012) would lead readers to believe. Recently, the U.S. Environmental Protection Agency (EPA) also conducted an expert review of the experimental evidence for NMDRs (USEPA 2013). In this draft report, EPA noted that such responses are not uncommon in *in vitro* studies and often relate to "lower-order biological endpoints" rather than apical endpoints. However, "[t]here is currently no reproducible evidence that the early key events involved in the expression of NMDRs that are identified at low doses are predictive of adverse outcomes that may be seen in humans or wildlife populations for estrogen, androgen or thyroid endpoints" (p. 8, USEPA 2013). The EPA concluded that, while NMDRs for adverse effects have been occasionally seen in intact organisms, NMDRs are relatively uncommon. Further, such dose-response curves – when observed – typically occur at high doses, well above the NOAELs identified in standard testing paradigms. In summary, the limited available evidence for low dose effects and NMDRs do not preclude the need to consider dose in assessing the potential hazards of chemicals to the endocrine system.

Finally, the WHO-UNEP 2012 report generally does not address the fact that most EDCs have much lower potency than endogenous hormones (Nohynek et al. 2013, Sharpe 2003). In fact, the report claimed that in the DES case study, while other chemicals may be less potent than DES, these effects are

[4] The common European spelling of "disruptor" is "disrupter."

"equally undesirable when the exposure occurs in early development where potency seems less important" (p. 25, WHO-UNEP 2012a). In other words, the report suggests that at vulnerable developmental stages, potency may not be very relevant. This statement confuses the issue of potency in different life stages. Although certain substances may be more or less potent depending on the particular life stage at which exposure occurs, potency is always important – no matter the developmental window. The WHO-UNEP 2012 report rightly stated that hormone potency and receptor affinity are not the same things (p. 12, WHO-UNEP 2012a). It is further suggested that potency depends on many different factors, but it is vague as to what those factors might be (other than receptor abundance). At the receptor level, potency is determined by both the affinity of a substance to bind to a receptor site as well the efficacy with which that substance activates the receptor (Borgert et al. 2013). Endogenous hormones have both strong affinity for their receptor sites as well as high efficacy for activation of these receptors; thus, hormones generally are highly potent for modulating endocrine function. Exogenous chemicals, on the other hand, are rarely as potent as hormones, either due to reduced affinity, reduced efficacy, or both (e.g., Gaido et al. 1997, Nilsson 2000). An example of this comparing DES and BPA was previously presented. Although the WHO-UNEP 2012 report claims that "very low concentrations of environmental endocrine disruptors could add to the endogenous hormone effect to produce a response that is much greater than would be predicted based on the hormone alone" (p. 8, WHO-UNEP 2012a), this theory fails to consider the existence of biological thresholds. As described by Borgert et al. (2013), given the lower potency of most exogenous chemicals, the additional presence of these chemicals will not significantly alter hormone receptor occupancy; thus, a biological threshold for potency exists.

At the whole organism level, potency relates to the ability of a substance to produce a biological effect and may be substantially different from the potency measured with *in vitro* assays (EFSA 2013a). In considering potential EDCs, therefore, potency should not be determined based on the results of *in vitro* studies. Rather, EFSA recommends that potency should be based on the ability of a substance to produce an adverse health effect *in vivo* (EFSA 2013a). This ability will depend on not just a substance's potency at its receptor site, but also on the timing of exposure (i.e., the particular life stage of development) and the dose and duration of exposure. Therefore, dose remains an important factor in assessing the potency of potential EDCs to cause adverse health effects.

8.4. Conclusion

In summary, the WHO UNEP 2012 report fails to fully address a number of factors related to dose-response and potency that must be considered when defining EDCs. First, the substance must be shown to cause an adverse effect in an intact organism, their progeny or (sub)populations; therefore, *in vitro* data alone are insufficient for classifying a compound as an endocrine disruptor. Further, the observed effect must be shown to go beyond simple modulation of endocrine function; that is, it must result in an adverse outcome. Second, thresholds exist for inducing such adverse effects. Therefore, the doses at which different responses are observed to occur and how those doses compare to the levels at which people or wildlife are exposed must be considered. Further, the potential for low dose effects and NMDRs do not exclude the need to consider dose and exposure. Finally, the potency of potential endocrine disruptors compared to endogenous hormones relates to the ability of a substance to induce

an adverse effect in a living organism and is important regardless of the life stage at which exposure occurs.

9. Conclusions

It must be acknowledged that creating a true "state of the science" overview is a large and complex task. It cannot be expected that such an overview examine each putative case and question in detail. This critique is not based on the claim that full assessments need to be done on any and every substance or observed population endpoint for which endocrine disruption is a question. That said, however, when such assessments have not been done, it is not supportable to draw conclusions about such particular instances. The WHO-UNEP 2012 report has reached conclusions in many instances, either explicitly by naming endocrine disruption causes for particular endpoints in cases where the needed assessment has not been made, or, more often implicitly, by appearing to accept putative examples at face value without considering whether other, non-endocrine mechanisms or extraneous factors might also explain the patterns, posited mechanisms are biologically plausible, exposure levels are sufficient, or the whole inferential case is consistent. It is not that the studies and phenomena referenced in the report are not part of the scientific evaluation of endocrine disruption; it is that these studies only represent a part of the totality of the evidence, and no single part by itself constitutes a sufficient basis to draw conclusions.

The WHO-IPCS 2002 report, which the WHO-UNEP 2012 report seeks to update, saw its task of characterizing the state of the science to mean that it should evaluate the potential scope, magnitude, and nature of endocrine disruption as a public health issue, as well as evaluate the ways in which then-current scientific understanding and pertinent evidence could be brought to bear on this question. It recognized that this inherently entails evaluation of possible alternative explanations for phenomena of potential concern. One must integrate the evidence, which includes an assessment of plausible explanations and characterizes the data limits to unambiguously answer the questions at hand. These elements are central to a well-crafted scientific evaluation. The aim of the 2002 report was not to identify definitive answers nor to prepare a mere list of kinds of data to consider, but rather an assessment of what science could and could not say about the issues at hand, and the prospects for new research to improve on this.

The WHO-UNEP 2012 report takes a much narrower view of the meaning of "state of the science." It does not take on the task of evaluating the actual scope and magnitude of endocrine disruption as a real and operating public health problem. Instead, it highlights temporal trends for endpoints that are plausibly affected by endocrine control and then implies, without doing any evaluation or critical examination, that these might be indications of such a problem. It names some basic tenets of endocrinology and notes the ways in which exogenous agents might in principle interfere with endocrine-mediated control, but it makes little assessment of whether these effects operate at significant levels in real populations at current or foreseeable levels of exposure. Although the 2012 report claims the intent to show the "logic" of evaluating evidence on endocrine disruption (as it simultaneously declines even in its stated intent to apply this logic to the evaluation of actual cases, even as illustrative examples), it does not do even this in practice. Instead, the report merely reviews some selected kinds and examples of possibly relevant evidence, naming the ways in which they might contribute to a finding of endocrine disruption as a problem, without considering contrary evidence on the same systems, countervailing arguments, or possible limits on the applicability of the elements to

the larger assessment question. It does not attempt to tie evidence together, consider how it should be tied together (and what other factors might limit the ability to draw firm conclusions), or assess whether bringing the named data to bear would enable a clear enough picture of the issues to serve as the basis for regulatory or other actions. This is a significant shortcoming of the 2012 report and the failure to address the methodology that would be used to confirm that endocrine disruption is occurring is an opportunity lost.

Neither can the WHO-UNEP 2012 report be considered an update to the WHO-IPCS 2002 report because it does not build on and modify the earlier analysis, giving reasons and support for changes in the state of the science. A true update to the earlier report would cite the 2002 conclusions, articulate what data, findings, or new understanding since 2002 should be considered and evaluate how and whether the 2002 conclusions need to be modified in light of the newer information. In addition, the WHO-UNEP 2012 report does not address research recommendations from the earlier report. In some cases, the 2012 report reaches conclusions that conflict with those of the earlier 2002 report, despite the lack of new information to support a change in the weight of evidence. The fact that the 2012 report reaches more definitive conclusions based on the same data emphasizes a reliance on subjective decision-making and less stringent criteria for evaluating potential causal relationships compared to the earlier 2002 report. Although the WHO-UNEP 2012 report is stated to be an update, this report in actuality is a revised review of the state of the science that does not build upon what was previously done and disregards the WHO-IPCS 2002 proposed framework for causation in favor of "best professional judgment" on these matters.

A key concern with the WHO-UNEP 2012 report is the use of subjective inference instead of a formal framework to assess the potential role of causation for endocrine disruption. The report adopted a narrative approach for the data review that does not represent a weight-of-evidence assessment. Rather than demonstrate causation, the report relies on inference to suggest that exposures to chemicals and adverse outcomes are related. The WHO-UNEP 2012 report presented information on chemicals and various adverse outcomes, but whether the exposure causes these effects was not determined or adequately considered in an objective, transparent, and scientific manner. Several key factors for establishing causation, such as a demonstrated exposure to the chemical, dose-response, and consistency in the data were frequently ignored. For example, temporal trends in human diseases or wildlife populations are presented without consideration of alternative explanations for these trends (especially diagnostic criteria and reporting changes). Exposures to chemicals considered to have the potential for endocrine disruption exist and are suggested as contributing to the observed trends, but there is little consideration of whether these exposures are sufficient to explain the alleged effects and whether the patterns of exposure are congruent with the trends. Above all, the lack of a framework to collectively evaluate, in an objective and comprehensive manner, the data on specific chemicals and the alleged adverse outcomes is a significant shortcoming in the WHO-UNEP 2012 review.

The WHO UNEP 2012 report fails to fully address a number of critical factors that must be considered when defining EDCs; specifically, dose, dose-response, potency, and adversity. First, the substance must be shown to cause an adverse effect in an intact organism, their progeny or (sub)populations;

therefore, *in vitro* data alone are insufficient for classifying a compound as an endocrine disruptor. Further, the observed effect must be shown to go beyond adaptive modulation of endocrine function; that is, it must result in an adverse outcome. Second, thresholds exist for inducing such adverse effects. The 2012 report does not give appropriate consideration to thresholds and dose-response for adverse effects. Despite the chosen definition of endocrine disruption as producing adverse changes, the report treats all effects as evidence of disruption and makes an *a priori* rejection of thresholds. Where examples from animal testing are discussed, there is little consideration of dose-response, when, in fact, only some doses of some compounds can cause endocrine disruption in the laboratory. The WHO-UNEP 2012 report also does not address the fact that most EDCs have much lower potency than endogenous hormones and potency is important regardless of the life-stage at which exposure occurs. Finally, consideration of low dose effects and NMDRs do not exclude the need to consider potency and exposure.

It is also important to emphasize that the *Summary for Decision-Makers*, while implied by the title to be a synopsis of the main report, it is not truly representative of the main report. In many cases, the *Summary for Decision-Makers* compounds the limitations of the main report by making statements without supporting references and providing more definitive conclusions. Consequently, this companion report cannot be considered a summary, nor should it be relied on to make decisions regarding the regulation of endocrine disruptors.

In short, when compared to the 2002 report, the WHO-UNEP 2012 report is not a true assessment of the state of the science, and neither is it really an update. It does not review the findings of the earlier 2002 effort, name new research that ought to bear on those earlier findings or challenge or affirm the 2002 document's analysis for any stated reasons having to do with updated scientific findings.

To reiterate, the above discussion of particular cases above is not aimed at doing full assessments of putative cases, but it is provided to illustrate with selected examples the ways in which the 2012 state of the art assessment has not considered the full complexity of data that should be considered when making full and scientifically sound evaluations. This paper has been prepared to document why simply naming possible and unevaluated implications of various carefully chosen studies does not constitute evaluating the state of the science on the questions at hand.

The WHO-IPCS 2002 report was largely successful in its evaluation, despite its much broader ambitions than those of the WHO-UNEP 2012 report. That said, there has indeed been a considerable development of pertinent scientific evidence since 2002, and a true update – in the sense of a revisiting of the WHO-IPCS 2002 conclusions and arguments to see how new information might alter them – is desirable. This has not yet been accomplished and it is hoped that a renewed exercise could be undertaken using the 2002 report as a starting point and an expert panel tasked with conducting a true update aimed at a sufficiently encompassing view of what it means to assess the state of the science.

Acronyms and Abbreviations

ADDM	Autism and Developmental Disabilities Monitoring Network
ADHD	Attention deficit hyperactivity disorder
AhR	Aryl hydrocarbon receptor
AHS	Agricultural Health Study
ASD	Autism spectrum disorder
ATSDR	Agency for Toxic Substances and Disease Registry
BPA	Bisphenol A
CDC	Centers for Disease Control and Prevention
DDE	Dichlorodiphenyldichloroethylene
DDD	Dichlorodiphenyldichloroethane
DDT	Dichlorodiphenyltrichloroethane
DES	Diethylstilbestrol
DNA	Deoxyribonucleic acid
DNT	Developmental neurotoxicity
EDC	Endocrine disrupting chemicals
EFSA	European Food Safety Authority
GRADE	Grades of Recommendation Assessment, Development and Evaluation
IPCS	International Programme on Chemical Safety
MOA	Mode of action
NAS	National Academy of Sciences
NIS	Sodium-iodide symporter
NMDR	Non-monotonic dose response
NOAEL	No-observed-adverse-effect-lLevel
PCB	Polychlorinated biphenyls
PCDF	Polychlorinated dibenzofurans
POP	Persistent organic pollutant
TBT	Tributyl tin
TCDD	Tetrachlorodibenzo-p-dioxin
TDI	Tolerable Daily Intake
TPT	Triphenyl tin
TRH	Thyroxin-releasing hormone
UNEP	United Nations Environment Programme
USEPA or EPA	United States Environmental Protection Agency
USFDA or FDA	United States Food and Drug Administration
WHO	World Health Organization

References

Ahmed OM, El-Gareib AW, El-Bakry AM, Abd El-Tawab SM, Ahmed RG. Thyroid hormones states and brain development interactions. Int J Dev Neurosci 2008 Apr; 26(2):147-209.

Alavanja MC, Samanic C, Dosemeci M, Lubin J, Tarone R, Lynch CF, Knott C, Thomas K, Hoppin JA, Barker J, Coble J, Sandler DP, Blair A. Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. Am J Epidemiol 2003 May 1; 157(9):800-14.

Aronson KJ, Wilson JW, Hamel M, Diarsvitri W, Fan W, Woolcott C, Heaton JP, Nickel JC, Macneily A, Morales A. Plasma organochlorine levels and prostate cancer risk. J Expo Sci Environ Epidemiol 2010 Jul; 20(5):434-45.

Auger J, Jouannet P. Evidence for regional differences of semen quality among fertile French men. Fédération Francaise des Centres d'Etude et de Conservation des Oeufs et du Sperme humains. Hum Reprod 1997 Apr; 12(4):740-5.

Auger J, Kunstmann JM, Czyglik F, Jouannet P. Decline in semen quality among fertile men in Paris during the past 20 years. N Engl J Med. 1995 Feb 2;332(5):281-5.

Axelsson J, Rylander L, Rignell-Hydbom A, Giwercman A. No secular trend over the last decade in sperm counts among Swedish men from the general population. Hum Reprod 2011 May; 26(5):1012-6.

Bastomsky CH. Effects of a polychlorinated biphenyl mixture (arochlor 1254) and DDT on biliary thyroxine excretion in rats. Endocrinol. 1974; 95(4):1150-5.

Benbrahim-Tallaa L, Waalkes MP., 2008. Inorganic arsenic and human prostate cancer. Environ Health Perspect. 116(2),158-64.

Bianco AC, Kim BW. Deiodinases: implications of the local control of thyroid hormone action. J Clin Invest. 2006 Oct;116(10):2571-9.

Blaxill MF. What's going on? The question of time trends in autism. Public Health Rep 2004; 119(6):536-551.

Bonde JP, Ramlau-Hansen CH, Olsen J. Trends in sperm counts: the saga continues. Epidemiol. 2011 Sep;22(5):617-9.

Borgert CJ1, Baker SP, Matthews JC. Potency matters: thresholds govern endocrine activity. Regul Toxicol Pharmacol. 2013 Jul 6; pii: S0273-2300(13)00102-5.

Bradford Hill, A. The Environment and Disease: Association or Causation? Proc R Soc Med 1965 58:295-300.

Braun JM, Yolton K, Dietrich KN, Hornung R, Ye X, Calafat AM, Lanphear BP. Prenatal bisphenol A exposure and early childhood behavior. Environ Health Perspect. 2009 Dec;117(12):1945-52.

Bruner-Tran KL, Ding T, Osteen KG. Dioxin and endometrial progesterone resistance. Semin Reprod Med 2010 Jan; 28(1):59-68.

Bruner-Tran KL, Rier SE, Eisenberg E, Osteen KG. The potential role of environmental toxins in the pathophysiology of endometriosis. Gynecol Obstet Invest 1999; 48 Suppl 1:45-56.

Bryan G.W. The effect of oil-spill removers ('detergents') on the gastropod Nucella lapillus on a rocky shore and in the laboratory. J Mar Biol Ass UK 1968; 49: I 067 -I 092.

Cakmak H, Taylor HS. Molecular mechanisms of treatment resistance in endometriosis: the role of progesterone-hox gene interactions. Semin Reprod Med 2010 Jan; 28(1):69-74.

Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. BMJ 1992 Sep 12; 305(6854):609-13

Centers for Disease Control and Prevention (CDC). Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators. Prevalence of autism spectrum disorders--Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. MMWR Surveill Summ 2012; 61(3):1-19.

Cerná M, Krsková A, Cejchanová M, Spěváčková V. Human biomonitoring in the Czech Republic: an overview. Int J Hyg Environ Health. 2012 Feb;215(2):109-19.

Clark LS, Cowan DF, Pfeiffer DC. Morphological changes in the Atlantic bottlenose dolphin (Tursiops truncatus) adrenal gland associated with chronic stress. J Comparative Pathol. 2006 Nov; 135(4): 208-16.

Conrad JW Jr, Becker RA. Chemical safety studies: Conrad and Becker respond. Environ Health Perspect 2011b; 119:a508-a509.

Conrad JW Jr, Becker RA. Enhancing credibility of chemical safety studies: emerging consensus on key assessment criteria. Environ Health Perspect 2011a Jun; 119(6):757-64.

Costello MF, Sjoblom P, Haddad Y, Steigrad SJ, Bosch EG. No decline in semen quality among potential sperm donors in Sydney, Australia, between 1983 and 2001. J Assist Reprod Genet 2002 Jun; 19(6):284-90.

Danish Centre on Endocrine Disrupters (DTU Food), 2013. Input for the REACH-review in 2013 on endocrine disrupters (tærskelværdi-projekt, j.nr. MST-621-00050). Final report. Technical University of Denmark (DTU), National Institute of Food, Division of Toxicology and Risk Assessment

Dietrich DR, von Aulock S, Marquardt H, Blaauboer B, Dekant W, Kehrer J, Hengstler J, Collier A, Gori GB, Pelkonen O, Lang F, Nijkamp FP, Stemmer K, Li A, Savolainen K, Hayes AW, Gooderham N, Harvey A, Barile FA. Scientifically unfounded precaution drives European Commission's recommendations on EDC regulation, while defying common sense, well-established science and risk assessment principles. Toxicology in Vitro. 2013; pii:S0887-2333(13)00166-5.

Duchan E, Patel DR. Epidemiology of autism spectrum disorders. Pediatr Clin North Am 2012; 59(1):27-43.

Du Preez L, Kunene N, Hanner R, Giesy J, Solomon K, Hosmer A, Van Der Kraak G. Population-specific incidence of testicular ovarian follicles in *Xenopus laevis* from South Africa: A potential issue in endocrine testing. Aquat Toxicol 2009 95: 10–16.

Elia J, Imbrogno N, Delfino M, Rossi T, Mazzilli R, Nofroni I, Toscano V, Mazzilli F. Comparative study of seminal parameters between samples collected in 1992 and samples collected in 2010. Arch Ital Urol Androl 2012 Mar; 84(1):26-31.

Elsabbagh M, Divan G, Koh YJ, Kim YS, Kauchali S, Marcin C, Montiel-Nava C, Patel V, Paula CS, Wang C, Yasamy MT, Fombonne E. Global prevalence of autism and other pervasive developmental disorders. Autism Res 2012; 5(3):160-179.

European Food Safety Authority (EFSA). Scientific Opinion on the hazard assessment of endocrine disruptors: Scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. EFSA Journal 2013a; 11(3): 3132.

European Food Safety Authority (EFSA). DRAFT Scientific Opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs – Part: exposure assessment. EFSA Journal 2013b.

European Food Safety Authority (EFSA). Scientific Report of the Endocrine Active Substances Task Force. EFSA Journal 2010; 8(11):1932.

European Food Safety Authority (EFSA). Opinion of the Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food on a request from the Commission related to 2,2-Bis(4-hydroxyphenyl)propane (Bisphenol A). The EFSA Journal 2006, 428:1-75.

Ferriby LL, Knutsen JS, Harris M, Unice KM, Scott P, Nony P, Haws LC, Paustenbach D. Evaluation of PCDD/F and dioxin-like PCB serum concentration data from the 2001-2002 National Health and Nutrition Examination Survey of the United States population. J Expo Sci Environ Epidemiol. 2007 Jul;17(4):358-71.

Fisch H, Braun SR. Trends in global semen parameter values. Asian J Androl 2013 Mar; 15(2):169-73.

Fisch H, Hyun G, Hensle TW. Rising hypospadias rates: disproving a myth. J Pediatr Urol. 2010 Feb;6(1):37-9.

Fombonne E, Quirke S, Hagen, A. Prevalence and interpretation of recent trends in rates of pervasive developmental disorders. McGill J Med 2009; 12(2):99-107.

Gaido, K.W., Leonard, L.S., Lovell, S., Gould, J.C., Babaï, D., Portier, C.J., McDonnell, D.P., 1997. Evaluation of chemicals with endocrine modulating activity in a yeast-based steroid hormone receptor gene transcription assay. Toxicol. Appl. Pharmacol. 143(1),205-212.

Gentile I, Zappulo E, Militerni R, Pascotto A, Borgia G, Bravaccio C. Etiopathogenesis of autism spectrum disorders: Fitting the pieces of the puzzle together. Med Hypotheses 2013; 81(1):26-35.

Gibbs PE, Green JC, Pascoe PC. A massive summer kill of the dog-whelk, Nucella Lapillus, on the north Cornwall coase in 1995: freak or forerunner? J Mar Biol Ass UK 1999; 79, 103-109.

Glass AG, Lacey JV Jr, Carreon JD, Hoover RN. Breast cancer incidence, 1980-2006: combined roles of menopausal hormone therapy, screening mammography, and estrogen receptor status. J Natl Cancer Inst 2007 Aug 1; 99(15):1152-61.

Goldey ES, Kehn LS, Lau C, Rehnberg GL, Crofton KM. Developmental exposure to polychlorinated-biphenyls (Arachlor-1254) reduced circulating thyroid-hormone concentrations and causes hearing deficits in rats. Toxicol Appl Pharmacol 1995; 135(1):77-88.

Goodman JE, Dodge DG, Bailey LA. A framework for assessing causality and adverse effects in humans with a case study of sulfur dioxide. Regul Toxicol Pharmacol 2010 Nov; 58(2):308-22.

Goodman JE, Witorsch RJ, McConnell EE, Sipes IG, Slayton TM, Yu CJ, Franz AM, Rhomberg LR. Weight-of-evidence evaluation of reproductive and developmental effects of low doses of bisphenol A. Crit Rev Toxiocl 2009; 39:1-75.

Goodman JE, McConnell EE, Sipes IG, Witorsch RJ, Slayton TM, Yu CJ, Lewis AS, Rhomberg LR. An updated weight of the evidence evaluation of reproductive and developmental effects of low doses of bisphenol A. Crit Rev Toxicol 2006; 36:387-457.

GRADE, 2011. GRADE Working Group. Frequently asked questions. Available at: http://www.gradeworkinggroup.org/FAQ/index.htm. Last accessed November 8, 2013.

Guinchat V, Thorsen P, Laurent C, Cans C, Bodeau N, Cohen D. Pre-, peri- and neonatal risk factors for autism. Acta Obstet Gynecol Scand 2012; 91(3):287-300.

Guo SW. Epigenetics of endometriosis. Mol Hum Reprod 2009 Oct; 15(10):587-607.

Hammoud AO, Wilde N, Gibson M, Parks A, Carrell DT, Meikle AW. Male obesity and alteration in sperm parameters. Fertil Steril. 2008 Dec;90(6):2222-5.

Hardell L, Andersson SO, Carlberg M, Bohr L, van Bavel B, Lindström G, Björnfoth H, Ginman C. Adipose tissue concentrations of persistent organic pollutants and the risk of prostate cancer. J Occup Environ Med 2006; Jul:48(7):700-7.

Harvey PW, Sutcliffe C. Adrenocortical hypertrophy: establishing cause and toxicological significance. J Appl Toxicol 2010; Oct:30(7):617-26.

Hayes T, Haston K, Tsui M, Hoang A, Haeffele C, Vonk A. Atrazine-induced hermaphroditism at 0.1 ppb in American leopard frogs (Rana pipiens): laboratory and field evidence. Environ Health Perspect 2003 Apr; 111(4):568-75

Helsinki Commision (HELCOM). Baltic Sea Environment Fact Sheets for 2012, Helsinki Commission, Baltic Marine Environment Protection Commission, 2012.

Hengstler JG, Foth H, Gebel T, Kramer P-J, Lilienblum W, Schweinfurth H, Völkel W, Wollin K-M, Gundeert-Remy U. Critical evaluation of key evidence on the human health hazards of exposure to bisphenol A. Crit Rev Toxicol 2011; 41(4):263-291.

Henny CJ, Grove RA, Kaiser JL, Johnson BL. North American osprey populations and contaminants: Historic and contemporary perspectives. J Toxicol Environ Health B Crit Rev 2010; 13(7-8):579-603.

Hill AB. The environment and disease: association or causation? Proc R Soc Med. 1965 May;58:295-300.

Howdeshell KL. A model of the development of the brain as a construct of the thyroid system. Environ Health Persp 2002; 100(suppl 3):337-348.

Institute of Medicine (IOM). Breast cancer and the environment: A life course approach. Committee on Breast Cancer and the Environment: The Scientific Evidence, Research Methodology, and Future Directions. 2012. The National Academies Press; Washington, D.C.

International Agency for Research on Cancer (IARC). Arsenic in drinking water. 2004; Vol 84.

International Programme on Chemical Safety (IPCS). IPCS risk assessment terminology. Harmonization Project Document No. 1. 2004. World Health Organization, Geneva.

Itoh H, Iwasaki M, Hanaoka T, Sasaki H, Tanaka T, Tsugane S. Urinary phthalate monoesters and endometriosis in infertile Japanese women. Sci Total Environ 2009; Dec 15; 408(1):37-42.

Jørgensen N, Andersen AG, Eustache F, Irvine DS, Suominen J, Petersen JH, Andersen AN, Auger J, Cawood EH, Horte A, Jensen TK, Jouannet P, Keiding N, Vierula M, Toppari J, Skakkebaek NE. Regional differences in semen quality in Europe. Hum Reprod 2001; May; 16(5):1012-9.

Jørgensen N, Joensen UN, Jensen TK, Jensen MB, Almstrup K, Olesen IA, Juul A, Andersson AM, Carlsen E, Petersen JH, Toppari J, Skakkebæk NE. Human semen quality in the new millennium: a prospective cross-sectional population-based study of 4867 men. BMJ Open 2012; Jul 2; 2(4).

Jørgensen N, Vierula M, Jacobsen R, Pukkala E, Perheentupa A, Virtanen HE, Skakkebaek NE, Toppari J. Recent adverse trends in semen quality and testis cancer incidence among Finnish men. Int J Androl 2011 Aug; 34(4 Pt 2):e37-48.

Jörundsdóttir K, Svavarsson K, Leung KMY. Imposex levels in the dogwhelk Nucella lapillus (L.) - continuing improvement at high latitudes. Marine Pollution Bulletin 2005; 51(8-12):744-749.

Kabbarah O, Sotelo AK, Mallon MA, Winkeler EL, Fan MY, Pfeifer JD, Shibata D, Gutmann DH, Goodfellow PJ. Diethylstilbestrol effects and lymphomagenesis in Mlh1-deficient mice. International Journal of Cancer 2005; 115(6):666-669.

Kato Y, Haraguchi K, Ito Y, Fujii A, Yamazaki T, Endo T, Koga N, Yamada S, Degawa M. Polychlorinated biphenyl-mediated decrease in serum thyroxine level in rodents. Drug Metab Dispos 2010 Apr; 38(4):697-704.

Kato Y, Ikushiro S, Haraguchi K, Yamazaki T, Ito Y, Suzuki H, Kimura R, Yamada S, Inoue T, Degawa M. A possible mechanism for decrease in serum thyroxine level by polychlorinated biphenyls in Wistar and Gunn rats. Toxicol Sci 2004 Oct; 81(2):309-15.

Keller DA, Juberg DR, Catlin N, Farland WH, Hess FG, Wolf DC, Doerrer NG. Identification and characterization of adverse effects in 21st century toxicology. Toxicological Sciences 2012; 126(2):291-297.

Khan KS, Kunz R, Kleijnen J, Antes G. Five steps to conducting a systematic review. J R Soc Med, 2003 Mar; 96(3):118-21.

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

Kloas W, Lutz I, Springer T, Krueger H, Wolf J, Holden L, Hosmer A. Does atrazine influence larval development and sexual differentiation in *Xenopus laevis*? Toxicological Sciences 2009 107:376-384.

Kinast K, Peeters D, Kolk SM, Schubert D, Homberg JR. Genetic and pharmacological manipulations of the serotonergic system in early life: Neurodevelopmental underpinnings of autism-related behavior. Front Cell Neurosci 2013; 7:72.

Kollara A, Brown TJ. Functional interaction of nuclear receptor coactivator 4 with aryl hydrocarbon receptor. Biochem Biophys Res Commun. 2006 Jul 28; 346(2):526-34

Koutros S, Alavanja MC, Lubin JH, Sandler DP, Hoppin JA, Lynch CF, Knott C, Blair A, Freeman LE. An update of cancer incidence in the Agricultural Health Study. J Occup Environ Med 2010 Nov; 52(11):1098-105.

Kripke C. Lower- vs higher-dose estrogen for contraception. Am Fam Physician. 2005 Oct 1;72(7):1224-5

Kuiken T, Hofle U, Bennett PMT, Allchint CR, Kirkwood JK, Baker JR, Appleby EC, Lockyer CH, Walton MJ, Sheldrick MC. Adrenocortical Hyperplasia, Disease and Chlorinated Hydrocarbons in the harbour porpoise (*Phocoena phocoena*). Marine Pollution Bulletin 1993 Volume 26, No. 8.

LaFiandra EM, Babbitt KJ, Sower SA. Effects of atrazine on anuran development are altered by the presence of a nonlethal predator. J Toxicol Environ Health A 2008; 71(8):505-11.

Lair S1, Béland P, De Guise S, Martineau D. Adrenal hyperplastic and degenerative changes in beluga whales. J Wildl Dis. 1997 Jul;33(3):430-7.

Leclère B, Molinié F, Trétarre B, Stracci F, Daubisse-Marliac L, Colonna M. Trends in incidence of breast cancer among women under 40 in seven European countries: A GRELL cooperative study. Cancer Epidemiol 2013 Jun 11.

Lidsky TI, Schneider JS. Lead neurotoxicity in children: basic mechanisms and clinical correlates. Brain 2003 Jan; 126(Pt 1):5-19.

Lind PM, Bergman A, Olsson M, Orberg J. Bone mineral density in male Baltic grey seal (Halichoerus grypus). Ambio 2003 Sep; 32(6):385-8.

Marimuthu P, Kapilashrami MC, Misro MM, Singh G. Evaluation of trend in semen analysis for 11 years in subjects attending a fertility clinic in India. Asian J Androl 2003 Sep; 5(3):221-5.

Martin L, Klaassen CD. Differential effects of polychlorinated biphenyl congeners on serum thyroid hormone levels in rats. Toxicol Sci 2010 Sep; 117(1):36-44.

Masuda Y. The Yusho rice oil poisoning incident. Dioxins and Health. A. Schecter. New York, Plenum Press 1994; 633-659.

Mayes BA, McConnell EE, Neal BH, Brunner MJ, Hamilton SB, Sullivan TM, Peters AC, Ryan MJ, Toft JD, Singer AW, Brown JF Jr, Menton RG, Moore JA. Comparative carcinogenicity in Sprague-Dawley rats of the polychlorinated biphenyl mixtures Aroclors 1016, 1242, 1254, and 1260. Toxicol Sci 1998 Jan; 41(1):62-76.

McDaniel TV, Martin PA, Struger J, Sherry J, Marvin CH, McMaster ME, Clarence S, Tetreault G. Potential endocrine disruption of sexual development in free ranging male northern leopard frogs (Rana pipiens) and green frogs (Rana clamitans) from areas of intensive row crop agriculture. Aquat Toxicol 2008 Jul 30; 88(4):230-42.

Morton B. Recovery from imposex by a population of the dogwhelk, Nucella lapillus (Gastropoda: Caenogastropoda), on the southeastern coast of England since May 2004: a 52-month study. Marine Pollution Bulletin 2009; 58(10):1530-1538.

Murphy MB, Hecker M, Coady KK, Tompsett AR, Du Preez LH, Everson GJ, Solomon KR, Carr JA, Smith EE, Kendall RJ, Van Der Kraak GJ, Giesy JP. Atrazine concentrations, gonadal gross morphology and histology in *ranid* frogs collected in Michigan agricultural areas. Aquatic Toxicology 2006 76:230-245.

Nakamura Y, Tsuji S, Tonogai Y. Determination of the levels of isoflavonoids in soybeans and soy-derived foods and estimation of isoflavonoids in the Japanese daily intake. Journal of AOAC International 2000; 83(3):635-650.

Nakatani T, Yamamoto A, Ogaki S. A survey of dietary intake of polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans, and doxin-like coplanar polychlorinated biphenyls from food during 2000-2002 in Osaka City, Japan. Archives of Environmental Contamination and Toxicology 2011; 60(3):543-555.

National Research Council (NRC). Health Implications of Perchlorate Ingestion. Committee to Assess the Health Implications of Perchlorate Ingestion. 2005. The National Academies Press; Washington D.C.

National Toxicology Program (NTP). Toxicology and carcinogenesis studies of 3,3',4,4',5-pentachlorobiphenyl (PCB 126) (CAS No. 57465-28-8) in female Harlan Sprague-Dawley rats (gavage studies). 2006.

National Toxicology Program (NTP). Toxicology and carcinogenesis studies of 2,3',4,4'5-pentachlorobiphenyl (PCB 118) (CAS No. 31508-00-6) in female Harlan Sprague-Dawley rats (gavage studies). 2009.

Nelson CM, Bunge RG. Semen analysis: evidence for changing parameters of male fertility potential. Fertil Steril 1974 Jun; 25(6):503-7.

Nilsson R. Endocrine modulators in the food chain and environment. Toxicol. Pathol. 2000; 28(3):420-431.

Nohynek GJ1, Borgert CJ, Dietrich D, Rozman KK. Endocrine disruption: fact or urban legend? Toxicol Lett. 2013 Dec 16;223(3):295-305.

Nøst TH1, Breivik K, Fuskevåg OM, Nieboer E, Odland JØ, Sandanger TM. Persistent organic pollutants in Norwegian men from 1979 to 2007: intraindividual changes, age-period-cohort effects, and model predictions. Environ Health Perspect. 2013 Nov-Dec;121(11-12):1292-8.

Oka T, Tooi O, Mitsui N, Miyahara M, Ohnishi Y, Takase M, Kashiwagi A, Shinkai T, Santo N, Iguchi T. Effect of atrazine on metamorphosis and sexual differentiation in Xenopus laevis. Aquat Toxicol 2008 May; 30; 87(4):215-26.

Organization for Economic Cooperation and Development (OECD), 2005. Manual for investigation of HPV chemicals. Chapter 3: data evaluation. OECD Cooperative Chemicals Assessment Programme (CoCAP).

Pal PC, Rajalakshmi M, Manocha M, Sharma RS, Mittal S, Rao DN. Semen quality and sperm functional parameters in fertile Indian men. Andrologia 2006 Feb; 38(1):20-5.

Rhomberg LR, Goodman JE. Low-dose effects and nonmonotonic dose-responses of endocrine disrupting chemicals: Has the case been made? Regul Toxicol Pharmacol 2012; 64:130-133.

Rhomberg LR, Goodman JE, Foster WG, Borgert CJ, Van Der Kraak G. A critique of the European Commission document, “State of the Art Assessment of Endocrine Disruptors”. Critical Reviews in Toxicology 2012; 42(6):465-473.

Rice CD. Prevalence of autism spectrum disorders--autism and developmental disabilities monitoring network, six sites, United States, 2000. MMWR Surveill Summ. 2007; 56(1):1-11.

Rice CE, Rosanoff M, Dawson G, Durkin MS, Croen LA, Singer A, Yeargin-Allsopp M. Evaluating changes in the prevalence of the autism spectrum disorders (ASDs). Public Health Rev 2012; 34(2): 1-22.

Ritchie JM, Vial SL, Fuortes LJ, Guo H, Reedy VE, Smith EM. Organochlorines and risk of prostate cancer. J Occup Environ Med 2003 Jul; 45(7):692-702.

Robertson A. Effects of a toxic bloom of Chrysochromulina polylepis on the common dog-whelk, Nucella lapillus on the Swedish west coast. J Mar Biol Assoc UK 1991; 71, pp 569-578.

Sawada N, Iwasaki M, Inoue M, Itoh H, Sasazuki S, Yamaji T, Shimazu T, Tsugane S. Plasma organochlorines and subsequent risk of prostate cancer in Japanese men: a nested case-control study. Environ Health Perspect 2010 May; 118(5):659-65.

Schmidt RJ, Hansen RL, Hartiala J, Allayee H, Schmidt LC, Tancredi DJ, Tassone F, Hertz-Picciotto I. Prenatal vitamins, one-carbon metabolism gene variants, and risk for autism. Epidemiology 2011; 22(4):476-485.

Séradour B, Allemand H, Weill A, Ricordeau P. Changes by age in breast cancer incidence, mammography screening and hormone therapy use in France from 2000 to 2006. Bull Cancer 2009 Apr; 96(4):E1-6.

Sharpe RM. Environmental/lifestyle effects on spermatogenesis. Philos Trans R Soc Lond Biol Sci 2010 May 27; 365(1546):1697-712.

Sharpe, R.M., 2003. The 'oestrogen hypothesis'- where do we stand now? Int. J. Androl. 26,2–15.

Sirot V, Tard A, Venisseau A, Brosseaud A, Marchand P, Le Bizec B, Leblanc JC. Dietary exposure to polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans and polychlorinated biphenyls of the French population: Results of the second French Total Diet Study. Chemosphere 2012; 88(4):492-500.

Skelly, DK, Bolden, SR, and Dion, KB. 2010. Intersex frogs concentrated in suburban and urban landscapes. *EcoHealth* 7, 374-379.

Smith EE, du Preez LH, Gentles BA, Solomon KR, Tandler B, Carr JA, Van Der Kraak GJ, Kendall RJ, Giesy JP, Gross TS. 2005. Assessment of laryngeal muscle and testicular cell types in *Xenopus laevis* (AnuraPipidae) inhabiting maize and non-maize growing areas of South Africa. *African Journal of Herpetology* 54:69-76.

Soto AM, Lin TM, Sakabe K, Olea N, Damassa DA, Sonnenschein C. Variants of the human prostate LNCaP cell line as tools to study discrete components of the androgen-mediated proliferative response. Oncol Res. 1995;7(10-11):545-58.

Spolyarich N, Hyne RV, Wilson SP, Palmer CG, Byrne M. 2011. Morphological abnormalities in frogs from a rice-growing region in NSW, Australia, with investigations into pesticide exposure. *Environmental Monitoring and Assessment* 173:397-407.

Storrs SI, Semlitsch RD. Variation in somatic and ovarian development: predicting susceptibility of amphibians to estrogenic contaminants. Gen Comp Endocrinol 2008 May 1;156(3):524-30.

Stump DG, Beck MJ, Radovsky A, Garman RH, Freshwater L, Sheets LP, Marty MS, Waechter JM, Dimond SS, Van Miller JP, Shiotsuka RN, Beyer D, Chappelle AH and Hentges SG. Developmental neurotoxicity study of dietary bisphenol A in Sprague-Dawley rats. 2010. Toxicological Sciences 115, 167-182.

Swan SH, Elkin EP, Fenster L. Have sperm densities declined? A reanalysis of global trend data. Environ Health Perspect. 1997 Nov;105(11):1228-32.

Teeguarden JG, Hanson-Drury S. A systematic review of Bisphenol A "low dose" studies in the context of human exposure: a case for establishing standards for reporting "low-dose" effects of chemicals. Food Chem Toxicol. 2013 Dec;62:935-48.

Tinwell, H., Colombel, S., Blanck, O., Bars, R. The screening of everyday life chemicals in validated assays targeting the pituitary-gonadal axis. Regul. Toxicol. Pharmacol. 2013. 66(2),184-196.

Titley-O'Neal CP, Munkittrick KR, Macdonald BA. The effects of organotin on female gastropods. J Environ Monit 2011; 13(9):2360-2388.

Tyl, R.W., Myers, C.B., Marr, M.C., Sloan, C.S., Castillo, N.P., Veselica, M.M., Seely, J.C., Dimond, S.S., Van Miller, J.P., Shiotsuka, R.N., Beyer, D., Hentges, S.G., Waechter, J.M. Jr. Two-generation reproductive toxicity study of dietary bisphenol A in CD-1 (Swiss) mice. Toxicol. Sci. 2008 104(2),362-384.

Tyl, R.W., Myers, C.B., Marr, M.C., Thomas, B.F., Keimowitz, A.R., Brine, D.R., Veselica, M.M., Fail, P.A., Chang, T.Y., Seely, J.C., Joiner, R.L., Butala, J.H., Dimond, S.S., Cagen, S.Z., Shiotsuka, R.N., Stropp, G.D., Waechter, J.M. Three-generation reproductive toxicity study of dietary bisphenol A in CD Sprague-Dawley rats. Toxicol. Sci. 2002 68(1),121-146.

US Environmental Protection Agency (USEPA). Vocabulary Catalog List Detail - Integrated Risk Information System (IRIS) Glossary. 2014. Available at http://ofmpub.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=IRIS%20Glossary Last accessed January 24, 2014.

US Environmental Protection Agency (USEPA). State of the Science Evaluation: Nonmonotonic Dose Responses as They Apply to Estrogen, Androgen, and Thyroid Pathways and EPA Testing and Assessment Procedures. Draft. June 2013. Available at http://epa.gov/ncct/download_files/edr/NMDR.pdf.

US Food and Drug Administration (FDA). Memorandum to M. Hepp (Division of Food Contact Notifications, Regulatory Group 1) Re: Exposure to bisphenol A (BPA) for infants, toddlers and adults form the consumption of infant formula, toddler food and adult (canned) food. FDA-2010-N-0100-0001. 2009. Available at http://www.regulations.gov/#!documentDetail;D=FDA-2010-N-0100-0009.

van Maele-Fabry G, Libotte V, Willems J, Lison D. Review and meta-analysis of risk estimates for prostate cancer in pesticide manufacturing workers. Cancer Causes Control 2006 May; 17(4):353-73. Review.

van Maele-Fabry G, Willems JL. Prostate cancer among pesticide applicators: a meta-analysis. Int Arch Occup Environ Health. 2004 Nov; 77(8):559-70.

Vandenberg, L.N., Colborn, T., Hayes, T.B., Heindel, J.J., Jacobs, D.R. Jr., Lee, D.H., Shioda, T., Soto, A.M., vom Saal, F.S., Welshons, W.V., Zoeller, R.T., Myers, J.P. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. Endocr. Rev. 2012; 33(3),378-455.

Verhaegen Y, Monteyne E, Neudecker T, Tulp I, Smagghe G, Cooreman K, Roose P, Parmentier K. Organotins in North Sea brown shrimp (Crangon crangon L.) after implementation of the TBT ban. Chemosphere 2012; 86(10):979-984.

Wakai K, Egami I, Kato K, Kawamura T, Tamakoshi A, Lin Y, Nakamura T, Wada M, Ohno Y. Dietary intake and sources of isoflavones among Japanese. Nutrition and Cancer 1999; 33(2):139-145.

Weedon-Fekjær H, Bakken K, Vatten LJ, Tretli S. Understanding recent trends in incidence of invasive breast cancer in Norway: age-period-cohort analysis based on registry data on mammography screening and hormone treatment use. BMJ 2012 Jan 30; 344:e299.

Wilhelm M, Ewers U, Wittsiepe J, Fürst P, Hölzer J, Eberwein G, Angerer J, Marczynski B, Ranft U. Human biomonitoring studies in North Rhine-Westphalia, Germany. Int J Hyg Environ Health. 2007 May;210(3-4):307-18.

Wing L, Yeates SR, Brierley LM, Gould J. The prevalence of early childhood autism: Comparison of administrative and epidemiological studies. Psychol Med 1976; 6(1):89-100.

Wittassek M, Wiesmüller GA, Koch HM, Eckard R, Dobler L, Müller J, Angerer J, Schlüter C. Internal phthalate exposure over the last two decades—a retrospective human biomonitoring study. Int J Hyg Environ Health. 2007 May;210(3-4):319-33.

World Health Organization, International Programme on Chemical Safety (WHO-IPCS). Global Assessment of the State-of-the-Science of Endocrine Disrupters. Damstra T, Barlow S, Bergman A, Kavlock R, Van Der Kraak G, eds. 2002.

World Health Organization, United Nations Environment Programme (WHO-UNEP). State of the Science of Endocrine Disrupting Chemicals - 2012. Bergman A, Heindel JJ, Jobling S, Kidd KA, Zoeller RT, eds. 2012a; accessed at http://www.who.int/ceh/publications/endocrine/en/index.html, 289p.

World Health Organization, United Nations Environment Programme (WHO-UNEP). State of the Science of Endocrine Disrupting Chemicals 2012 Summary for Decision-Makers. Bergman, A., Heindel, J.J., Jobling, S., Kidd, K.A., Zoeller, R.T., eds. 2012b; accessed at http://www.who.int/ceh/publications/endocrine/en/index.html, 289p.

Xu X, Dailey AB, Talbott EO, Ilacqua VA, Kearney G, Asal NR. Associations of serum concentrations of organochlorine pesticides with breast cancer and prostate cancer in U.S. adults. Environ Health Perspect 2010 Jan; 118(1):60-6.

Yolton, K., Xu, Y., Strauss, D., Altaye, M., Calafat, A.M., Khoury, J., 2011. Prenatal exposure to bisphenol A and phthalates and infant neurobehavior. Neurotoxicol. Teratol. 33,558-566.

Younglai EV, Collins JA, Foster WG. Canadian semen quality: an analysis of sperm density among eleven academic fertility centers. Fertil Steril 1998 Jul; 70(1):76-80.

Zoeller RT, Dowling AL, Vas AA. Developmental exposure to polychlorinated biphenyls exerts thyroid hormone-like effects on the expression of RC3/neurogranin and myelin basic protein messenger ribonucleic acids in the developing rat brain. Endocrinology 2000; 141(1):181-189.