CrossMark

REVIEW ARTICLE

Human cost burden of exposure to endocrine disrupting chemicals. A critical review

Gregory G. Bond[1] · Daniel R. Dietrich[2]

Abstract Recently published papers have alleged that exposures to endocrine disrupting chemicals (EDCs) are causing substantial disease burdens in the EU and US and are consequently costing society hundreds of billions of dollars annually. To date, these cost estimates have not undergone adequate scientific scrutiny, but nevertheless are being used aggressively in advocacy campaigns in an attempt to fundamentally change how chemicals are tested, evaluated and regulated. Consequently, we critically evaluated the underlying methodology and assumptions employed by the chief architects of the disease burden cost estimates. Since the vast majority of their assigned disease burden costs are driven by the assumption that "loss of IQ" and "increased prevalence of intellectual disability" are caused by exposures to organophosphate pesticides (OPPs) and brominated flame retardants (PBDEs), we have taken special care in describing and evaluating the underlying toxicology and epidemiology evidence that was relied upon. Unfortunately, our review uncovered substantial flaws in the approach taken and the conclusions that were drawn. Indeed, the authors of these papers assumed causal relationships between putative exposures to EDCs and selected diseases, i.e., "loss of IQ" and "increased prevalence of intellectual disability", despite not having established them via a thorough evaluation of the strengths and weaknesses of the underlying animal toxicology and human epidemiology evidence. Consequently, the assigned disease burden costs are highly speculative and should not

be considered in the weight of evidence approach underlying any serious policy discussions serving to protect the public and regulate chemicals considered as EDCs.

Keywords Epidemiology · Endocrine · Chemicals · Intellectual disability · IQ · Neurobehavioral · Policy · Toxicology

Background

During the past 2 years, a series of seven inter-related papers published in three different journals (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) have provided estimates of the annual societal costs of exposures to endocrine disrupting chemicals (EDCs) in the hundreds of billions of Euros or US dollars. Although the validity of these estimates has been seriously questioned by others (Middelbeek and Veuger 2015; European Commission 2016a), they nevertheless continue to be widely cited in the popular press and in various policy forums in the US, EU and internationally. The proponents of the cost estimates are using them aggressively in advocacy campaigns to highlight a presumably missed health impact in the general population that is not controlled for by current safety evaluations and risk assessments, increasing public anxiety and fears, and thus finally attempting to fundamentally change how chemicals are tested, evaluated and regulated. As any statements regarding the potential impaired health of the public must be considered with the utmost care and urgency, it is equally important to understand the scientific basis that has led to the original statement(s). For this reason, a critical evaluation of these estimates is urgently needed.

✉ Gregory G. Bond
  gregory.g.bond@gmail.com

[1] Manitou View Consulting, LLC, Northport, MI, USA

[2] Human and Environmental Toxicology, University of Konstanz, Constance, Germany

Our review below explores how the cost estimates were made, restates the major findings as originally reported and highlights the multiple shortcomings of the methodology that was used. Since the vast majority of the estimated costs derive from alleged chemically mediated reductions in IQ and increased prevalence of intellectual disability, the evidence supporting these particular claims is more closely scrutinized.

How the cost estimates were derived

A series of four related papers estimating the societal costs to the EU from diseases purportedly caused by low-level exposures to EDCs was originally published in March 2015 in the *Journal of Clinical Endocrinology and Metabolism* (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015). They were followed by a fifth paper (Hunt et al. 2016) published in the same journal a year later which addressed estimated costs of female reproductive disorders and diseases attributable to EDC exposures. A sixth paper, which presented an update to the original EU cost estimates was published in the journal *Andrology* (Trasande et al. 2016). More recently, a seventh paper (Attina et al. 2016) which relies on the same evidence and methodology as applied in the original five papers has been published in the journal *The Lancet Diabetes and Endocrinology* and focuses on estimating the cost of exposures to EDCs in the US. An eighth paper which will focus on breast cancer, has yet to be published, but is anticipated at some point in the near future.

The first paper (Trasande et al. 2015) in the series describes the overall methodology that was used, the composite results and high-level conclusions. In brief, a self-appointed Steering Committee devised the overall methodology for assigning probability of causation between a focused list of alleged EDCs and selected disorders and diseases. The purported EDCs and disorders and diseases were chosen based on a 2012 UNEP/WHO sponsored review of the "state of the science" report (Bergman et al. 2013) which itself has been heavily criticized for failing to use systematic review methodology (Lamb et al. 2014; Vandenberg et al. 2016). The Steering Committee also developed the methodologies for assigning the fraction of disease burden in the population that is potentially attributable to exposure to EDCs as well as the methodology for assessing the potential costs to society from exposures to EDCs. Societal costs were ascribed using a human capital approach that measured the value of resources foregone and lost output due to illness, such as lost earnings or household contributions as a homemaker, as well as costs of medical treatment.

The Steering Committee selected scientists to participate on five panels and coached them on the methods to be used to assess the probability of causation between purported EDCs and disease outcomes. The panels, which ranged in size from four to eight scientists, focused on purported EDC exposures and (1) neurobehavioral deficits and diseases; (2) male reproductive disorders and diseases; (3) obesity and diabetes; (4) breast cancer; or (5) female reproductive disorders and diseases. Probability of causation between purported EDCs and selected medical conditions was assessed by each of the respective panels which evaluated the animal toxicology evidence separate from the human epidemiology evidence applying a modified Delphi technique in an effort to arrive at consensus. The fraction of disease attributable to EDC exposure was estimated, as were exposure–response relationships, often from the results of a single epidemiology study. Limited biomonitoring data were used to estimate levels of exposures to EDCs in the general population and to apply to the exposure–response relationships. Societal costs were then estimated using the total number of disease cases attributed to EDC exposure using estimated average cost per case estimates.

What were the reported results and conclusions?

Table 1 presents a summary of the findings reported to date from the seven studies. The expert panels reported consensus for "at least probable" (arbitrarily chosen as a greater than 19% probability) EDC causation for: IQ loss (ranging from 0.52 to 0.84 IQ points and 0.38–5.32 IQ points attributable to polybrominated diphenylether and organophosphates exposures, respectively) and associated intellectual disability, autism, attention-deficit disorder, childhood obesity, adult obesity, adult diabetes, cryptorchidism, male infertility, mortality associated with reduced Testosterone (T), uterine fibroids and endometriosis. Among the constellation of health outcomes that were studied, a "very low" (0–19%) probability of causation was determined only for PBDE exposure as a potential cause of testicular cancer.

The total annual cost of all medical conditions probably attributable to EDCs in the EU was then estimated to be 191B Euros, with sensitivity analyses suggesting costs ranging from 81.3 to 269B Euro annually (note, the female reproductive disorder costs, which were reported later (Hunt et al. 2016), raised these estimates by less than 1%).

Overwhelmingly, the total cost estimate was driven by the reported loss of IQ and associated intellectual disability purportedly attributable to prenatal organophosphate exposure which accounted for 146B Euros of the 191B Euro total (76.4%).

Table 1 Summary of findings from studies estimating societal costs attributable to EDC exposures (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Attina et al. 2016)

Exposure	Outcome	Strength of epidemiology evidence	Strength of toxicological evidence	Probability of causation, %	Base estimates of annual costs	
					EU (Billion Euros)	US (Billion USD)
PBDEs	IQ loss and intellectual disability	Moderate-to-high	Strong	70–100	9.6	266
Organophosphate pesticides	IQ loss and intellectual disability	Moderate-to-high	Strong	70–100	146	44.7
DDE	Childhood obesity	Moderate	Moderate	40–69	0.02	0.03
DDE	Adult diabetes	Low	Moderate	20–39	0.83	1.8
Di-2-ethylhexyl-phthalate	Adult obesity	Low	Strong	40–69	15.6	1.7
Di-2-ethylhexyl-phthalate	Adult diabetes	Low	Strong	40–69	0.61	0.09
BPA	Childhood obesity	Very low-to-low	Strong	20–69	1.5	2.4
PBDEs	Testicular cancer	Very low-to-low	Weak	0–19	0.85	0.08
PBDEs	Cryptorchidism	Low	Strong	40–69	0.13	0.04
Benzyl and butyl phthalates	Male infertility, resulting in increased assisted reproductive technology	Low	Strong	40–69	4.7	2.5
Phthalates	Low T, resulting in increased early mortality	Low	Strong	40–69	8.0	10.6
DDE	Fibroids	Low-moderate	Moderate	20–39	0.16	0.26
Phthalates	Endometriosis	Low-moderate	Moderate	20–39	1.3	47
Multiple exposures	ADHD	Low-to-moderate	Strong	20–69	1.7	0.70
Multiple exposures	Autism	Low	Moderate	20–39	0.20	2.0

ADHD attention-deficit hyperactivity disorder

Phthalate-attributable adult obesity was the second largest driver of costs in the EU at 15.6B Euro per year (8.2% of the total).

The authors (Trasande et al. 2015) concluded that EDC exposures in the EU are likely to contribute substantially to disease and dysfunction with costs in the hundreds of billions of Euros per year—more than 1% of annual GDP of the EU. Further, they asserted their belief that the actual costs are likely to be even higher, because their analysis was limited to what they regarded as a relatively small number of EDCs and disease conditions.

More recently (Attina et al. 2016), the same estimates of probability of EDC-health outcome links, attributable fractions of disease and exposure–response relationships that were generated for producing EU cost estimates, were applied to population and biomonitoring data specific to the US. From this information, the authors estimated the annual costs attributable to EDC exposures in the US alone are $340 billion. Again, the largest driver of costs was the purported loss of IQ and increased prevalence of intellectual disability, accounting for nearly 80% of the total costs. In this study, the estimated cost burden

was attributable to exposures to polybrominated diphenyl ethers (PBDEs). The next highest attributable cost, an estimated $43 billion annually (12.6% of total costs), was alleged to be attributed to endometriosis from phthalate exposure.

A critical evaluation of the underlying methodology employed

Given the aggressive media campaign that accompanied the Trasande et al. (2015) publication that was directed at influencing the ongoing debate in the EU about criteria for identifying EDCs, one could legitimately question whether it may have biased the results toward exaggerating costs (Runkle and Frankel 2012). However, our review will be specifically restricted to the merits of the science as described in the various published papers (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016). In this section, we undertake a critical evaluation of the methodology that was used to develop cost

estimates of the disease burden attributable to EDC exposure. This includes an examination of:

- the role of the Trasande et al. (2015) Steering Committee,
- the approaches that were taken to search the literature and select the underlying scientific studies that were relied on as primary data sources,
- the methods used to assess the animal data and human epidemiology data, and the framework for assessing probability of causation,
- how Attributable Fraction and exposure–response relationships were estimated and applied,
- sources and use of biomonitoring data,
- sources of and uses for the cost data, and
- the cumulative effect of the numerous assumptions inherent in each step of the process.

The Trasande et al. (2015) Steering Committee and its role

The entire project was overseen by a Steering Committee that consisted of a self-appointed group of eight scientists who have published research and actively engaged in public policy advocacy on the topic of EDCs. They designed the methodology that was used throughout the process. They chose the five health outcome groupings for evaluation (i.e., neurobehavioral deficits and disorders; male reproductive disorders and diseases; obesity and diabetes; breast cancer; and female reproductive disorders and diseases). They also selected the scientist members of the five separate associated panels, who were identified as "epidemiological and toxicological experts, based upon their scholarly contribution in the diseases under consideration and endocrine disruptor toxicology" (Trasande et al. 2015).

Teleconferences of indeterminate length were held biweekly over a three-month period with all of the participants to "encourage familiarity with the literature being reviewed, to describe the Delphi method (including the definition of terminology and interaction structure), and to identify expert group leaders."

The Trasande et al. (2015) Steering Committee then convened a two-day meeting where the associated panels met to discuss their interpretation of the literature and to assign probability of causation using a modified Delphi technique to arrive at a consensus. Although not transparently communicated, it is apparent that deliberations among panel members continued on for some indeterminate time period after the two-day meeting concluded, at the discretion of the panel leaders, to accomplish their work.

Given the complex nature of the project, the use of a Steering Committee was a logical path to take; however,

there are several concerns that arise from the approach taken:

- The Steering Committee was self-appointed. No description was provided for the process that was used to recruit members, their qualifications to serve or how eligibility was determined.
- Similar criticisms can be leveled at the selection and recruitment of the associated panel members and their leaders.

Above criticisms take on added significance when, as the authors of Trasande et al. (2015) admit, the assignment of probability of causation is a highly subjective process. The casual selection approach taken to such a critical step in the process stands in stark contrast to the rigorous processes employed by regulatory agencies in the US and the EU when they form advisory panels and must, among other requirements, try to ensure a balance of perspectives on an issue (US EPA 2016a; European Commission 2016b). A careful examination of the curriculum vitae of the expert panelists convened by Trasande et al. (2015) demonstrates a clear predisposition toward those who have exhibited strong biases in the favor of ascribing causality to associations observed between exposure to alleged EDCs and adverse health outcomes. Indeed, none of the panels encompassed scientists who have expressed critical views about the hypotheses being evaluated.

There is evidence that the Trasande et al. (2015) Steering Committee exerted much influence over the deliberations of the expert panels providing specific guidance on topics about which there remains significant debate among scientists. Here are a few selected examples:

"Because of the EU decision-making context, panelists were advised to adhere to the EU definition [of an EDC], but to add a further requirement that the chemicals interfere with hormone action (as elaborated in the The Endocrine Society definition)."

Note: the *EU definition* of an EDC is consistent with the WHO/IPCC definition of an EDC and requires that a chemical both alters function(s) of the endocrine system and consequently causes adverse health effects. *The Endocrine Society definition* does not require an adverse health effect, and thus includes a larger universe of chemicals defined as EDCs.

"They [associated panels] were asked to consider the reality of mixtures and complexity of attribution in that context." Although we acknowledge that mixed exposures are a reality, such guidance provided in this particular context is tantamount to advising the panels to take a precautionary approach which goes beyond their scientific mandate.

"Throughout the Delphi process, the panels were strongly encouraged to produce ranges that represent low and high bounds for the dose–response relationship and to evaluate potential non-linearity and non-monotonicity as well as the presence or absence of threshold effects when appropriate." Regulatory agencies in the US (Gray 2013; US FDA 2014) and the EU (Beausoleil et al. 2016) still consider the likelihood of low dose effects to be unproven or even non-existent, and again, providing such vague guidance to the expert panels in this context seems inappropriate.

Exactly how much and which direction these directives influenced the panels is unknown, but the cumulative effect may have biased the results toward exaggerating the strength of evidence.

Search for and selection of studies to include in the review

Increasingly, systematic-review methods are being applied to address environmental health questions and specific guidance has been published for how to employ them (Rooney et al. 2014). Such methods can provide enhanced objectivity and transparency to the process of collecting and synthesizing scientific evidence in researching conclusions on specific research questions, although Ioannidis (2016) has identified potential abuses that are becoming more common. When correctly conceived and executed systematic reviews also allow others sufficient information to be able to independently replicate the work, an important part of the scientific process. Key initial steps include problem formulation and protocol development and the search for and selection of the studies for inclusion in the review. Unfortunately, the authors of the Trasande et al. (2015) series of cost estimate papers (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) did not employ systematic review methods and more critically they neither provided a description of the methods they used to search the literature nor did they describe the processes they used to select studies for evaluation. Absent adequate descriptions the authors leave themselves open to concerns of bias and "cherry picking" the literature.

Weight of evidence analysis

In addition to ensuring a transparent and unbiased selection of studies for review, i.e., not choosing only studies that demonstrate a positive association between exposure and disease outcome and ignoring those which fail to demonstrate an association, it is equally important to employ a system of weighing individual studies according to their quality, reliability and relevance (concentrations,

formulations and exposure routes chosen) as well as the reproducibility of reported effects, the pattern of effects across and within studies, number of species showing the same or similar effects, likelihood that the species showing the effect are physiologically similar to humans and not predisposed to higher susceptibility of effects due to species-specific physiology, time of onset of effects and life stage affected (European Commission 2013; US EPA 2011). Although the Trasande et al. (2015) associated panels (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) claimed to have employed a weight of the evidence approach to examining the body of evidence linking exposures to alleged EDCs and disease, critical analysis was severely lacking.

Evaluation of animal toxicological evidence

To evaluate the animal toxicological evidence, the Steering Committee "adapted" criteria recently proposed by the Danish Environmental Protection Agency (Danish-EPA) for evaluating laboratory and animal evidence of endocrine disruption (Danish Ministry of the Environment 2011). In brief, those criteria identify a substance as a Confirmed EDC (assigned a score of Strong Animal Evidence by the Steering Committee) when the substance is known to have caused endocrine mediated adverse effects in humans or animal species living in the environment or when there is evidence from animal studies, possibly supplemented with other information, to provide a strong presumption that the substance has the capacity to cause adverse endocrine-mediated effects in humans or animals living in the environment.

When there is mechanistic information that raises doubt about the relevance of the effect for humans or the environment, chemical classification of Suspected EDC (assigned a score of Moderate Animal Evidence) is more appropriate. Chemical classification as a Suspected EDC requires either: (a) the presence of an endocrine mode of action without clear corroboration of the mode of action producing the expected adverse effects in laboratory or animal studies; or (b) the presence of the adverse effect in laboratory or animal studies with a suspected endocrine mode of action. The animal evidence supported a chemical to be classified as Potential EDC (assigned a score of Weak Animal Evidence) when there was evidence of adverse effects in animal studies that could have either an endocrine mode of action or a non-endocrine mode of action or in vitro/in silico evidence indicating a potential for endocrine disruption in an intact organism.

It is notable that the Danish-EPA's criteria specifically do not include a category for designating a substance as not an EDC, regardless of the evidence available that it does not interfere or interact with any component of the endocrine

system. Thus, technically speaking by their criteria every substance is, at a minimum, a Potential EDC. This is not particularly useful for the purposes of regulating chemicals, or for introducing new chemicals intended to substitute for more hazardous ones.

Another criticism that has been leveled at the proposed Danish-EPA's criteria is that they fail to consider potency and therefore, do not distinguish between substances that cause only transitory endocrine effects from true EDCs that are candidates for regulatory intervention (Borgert et al. 2013; DE-UK 2011). The authors of the Danish-EPA criteria also described the type of animal evidence required to fulfill the criteria, which are heavily weighted toward the OECD conceptual framework for testing and assessment of endocrine disruptors (OECD 2012). The authors of the cost estimate papers did not explicitly state how they weighted evidence derived from validated OECD test guideline studies versus un-replicated non-guideline studies, or how they handled issues of inconsistent findings across multiple studies, studies with small numbers of animals, studies with very few dose groups, or other concerns about study design including the statistical approach taken and execution.

Evaluation of human epidemiology evidence

The Trasande et al. (2015) Steering Committee adapted the GRADE Working Group criteria as were recently applied in evaluating indoor air quality criteria by the World Health Organization (Bruce et al. 2014). This WHO revised approach, has been termed by its authors (Bruce et al. 2014) as "Grading of Evidence for Public Health Interventions' (GEPHI). The use of GEPHI by the Trasande et al. (2015) EDC cost Steering Committee biased their assessments as in contrast to the much more widely recognized and accepted GRADE methodology which weighs evidence from randomized controlled trials (RCTs) most heavily and typically treats observational epidemiology evidence, i,e., the only type of evidence which was available to the Trasande et al. (2015) associated panels, as of low or very low quality and consequently would have relegated their probability of causation estimates into the lowest tiers.

There are several significant problems with the Trasande et al. (2015) Steering Committee's use of the GEPHI methodology. First, GEPHI was developed by its original authors (Bruce et al. 2014) to recognize the distinction between traditional observational epidemiology studies and a "before and after, quasi-experimental design" that has been used in some studies evaluating different types of cooking stoves or fuels on indoor air quality. The GEPHI approach (Bruce et al. 2014) was influenced by the recognition that, while not of the same quality as randomized

controlled trials (RCTs), these "before and after" designs do provide evidence which is of a superior nature and quality than could be gleaned from simple observational comparisons. Even so, the group that developed GEPHI (Bruce et al. 2014) for the purposes of evaluating the epidemiology evidence on indoor air quality never rated the quality of their evidence above moderate. By contrast none of the epidemiology evidence available to the Trasande et al. (2015) associated endocrine panels came from a "before and after design", and thus using the GEPHI methodology, the associated panels were not justified in rating any of the epidemiology evidence they considered above low or very low. Counter to the latter and as shown in Table 1, however, the associated panels (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) did exactly this for six of the alleged EDC-health outcome relationships. They rated the epidemiology evidence linking PBDEs and OPs to IQ loss and intellectual disability as moderate to high, the DDE-obesity evidence as moderate, and the multiple EDC exposure-ADHD evidence, DDE/Fibroids and the phthalate/endometriosis evidence as low to moderate.

Moreover, the GEPHI approach was developed for and applied to a very large, extensive and diverse body of scientific evidence available on indoor air pollutants, whereas the associated endocrine panels admittedly had a much smaller scientific evidence base to work with. For example, Bellanger et al. (2015) stated, "We endeavored to make our estimates as precise as possible, but were limited due to sizable uncertainties in the evidence". In light of these uncertainties, there was insufficient scientific support to either utilize the GEPHI approach or rate the assessments of the strength of the epidemiology evidence based on consistency or analogy.

The authors of GEPHI (Bruce et al. 2014) even issued caution about generalizing the use of their approach, the serious consideration of which should have disqualified its application by the Trasande et al. (2015) Steering Committee right from the start. Indeed they clearly state, "Even with the modifications to GRADE described above, not all of this evidence is amenable to this method of evidence review. The approaches adopted are specific to the topics and evidence reviewed for these guidelines, and a generic description is neither possible nor useful."

Consequently, we find the Trasande et al. (2015) Steering Committee's claim that they applied GEPHI to rate the strength of the epidemiology evidence to be disingenuous and highly misleading. Their approach was not as firmly grounded in work done by others as was implied in their respective publications, and thus severely limits, if not prohibits, any conclusions drawn from the approach chosen.

Framework for evaluating probability of causation

Trasande et al. (2015) admitted that the approach they took in this series of published reports is atypical: "Although analyses like these are highly valuable, they have typically been limited to associations where causation is certain. Decades of epidemiological data typically are required before causation has been acknowledged and attributable disease burden calculated." Regrettably, the atypical and speculative nature of the approach was not mentioned in the more widely read report abstracts or in the press releases that were issued which accompanied the publication of the articles, nor was there any evidence brought forth that would support the "certainty of causation". Furthermore, a logical question critical to be answered is: at what point should one conclude that the available data are so sparse, weak and conflicting that undertaking any effort to assign probability of causation would be an exercise in speculation?

The Trasande et al. (2015) Steering Committee repeatedly asserted that they used the Intergovernmental Panel on Climate Change (IPCC) guidance for authors document (Intergovernmental Panel on Climate Change 2005) to justify their approach to assigning probability of causation in the face of uncertainty. However, a careful review of that document does not offer support for the approach that Trasande et al (2015) have taken. In fact, the IPCC document cautions authors to "communicate carefully, using calibrated language" in stark contrast to the often sweeping generalizations made by the authors of this series of papers (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016).

Table 2 is reproduced from the IPCC document (2005), but it is not clear how it influenced the Trasande et al. (2015) Steering Committee and associated panels (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) in their approach. One could argue that based on this guidance, the panels should not have calculated any cost estimates for disorders and diseases for which they judged the probability of EDC causation less than 66% (and greater than or equal to 33%), because the probability of causation is described as "about likely as not." Most certainly, the should not have calculated costs for diseases and disorders where probability of causation was judged to be less than 33%.

Instead, it is apparent that the Trasande et al. (2015) Steering Committee developed its own unique approach to weighing the strength of epidemiology and toxicology evidence to assign probability of causation. It is, therefore, inaccurate for the Committee members to claim that they adopted the IPCC approach. Table 3 presents the actual framework that the Trasande et al. (2015) Steering Committee used as the basis for assigning probability of causation and also demonstrates that this basis is strikingly dissimilar to the IPCC Guidance to Authors document.

Moreover, in practice the Trasande et al. (2015) associated panels actually deviated from the framework presented in Table 3 when they could not achieve consensus in interpreting the strength of the epidemiology evidence. This occurred when assessing seven of the purported EDC-health outcome links. When faced with lack of unanimity, the panels were coached by the Trasande et al. (2015) Steering Committee to develop a hybrid probability range combining two categories (i.e., moderate-high, low-moderate, and very low-low).

Table 2 Likelihood scale (Intergovernmental Panel on Climate Change 2005)

Terminology	Likelihood of the occurrence/outcome
Virtually certain	>99% probability of occurrence
Very likely	>90% probability
Likely	>66% probability
About likely as not	33–66% probability
Unlikely	<33% probability
Very unlikely	<10% probability
Exceptionally unlikely	<1% probability

Table 3 Framework for evaluating probability of causation: comparison of the IPCC likelihood of occurrence/outcome terms (25) with those employed by Trasande et al. (2015) and derived publications (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016)

IPCC terminology (Table 2)		Trasande et al. (2015) terminology			
		Epidemiological evaluation	Toxicological evaluation		
Probability term used	Likelihood of occurrence/outcome	Probability term used	Strong (group 1)	Moderate (group 2A)	Weak (group 2B)
Very likely	>90% probability	High	Very high (90–100%)	High (70–89%)	Medium (40–69%)
Likely	>66% probability	Moderate	High (70–89%)	Medium (40–69%)	Low (20–39%)
About likely as not	33–66% probability	Low	Medium (40–69%)	Low (20–39%)	Very low (0–19%)
Unlikely	<33% probability	Very low	Low (20–39%)	Very low (0–19%)	Very low (0–19%)

In providing their range of cost estimates, the authors failed to acknowledge that the level of uncertainty in their methodology was so high that they cannot exclude the possibility that the costs of exposure to endocrine disrupting chemicals may be as low as zero. The latter is due to the fact that they have not convincingly made the case that any of the current disease burden in the population is indeed caused by exposure to chemicals they allege are EDCs. The Trasande et al. (2015) Steering Committee directed the associated panels to proceed to develop estimates of the attributable fraction and societal costs even when their collective opinion was that the probability of causation estimate was as low as 0–19%. We contend that no effort should have been made to calculate cost estimates for any exposure-outcome combination when the probability of causation was less than or equal to 50%, i.e., the same likelihood as flipping a coin.

Trasande et al. (2015) also included a calculation of the probability that at least one or more of the exposure-outcome relationships they evaluated were indeed causal and concluded that it exceeded 99%. We have three significant concerns with the latter statement:

- first, to make such a calculation they chose to use single point estimates to represent what were actually broad ranges of estimates of probability, and thus the final result implies a level of precision that simply does not exist;
- second, they arbitrarily decided to exclude from their calculations several of the exposure-outcome relationships which were judged to be of low probability and whose inclusion would have resulted in a lower overall estimate of probability and;
- third, the entire calculation and their conclusion from it make no substantive contribution to enhancing understanding about EDC exposures and adverse health outcomes.

Even assuming that their calculation and conclusion are true and at least one of the associations they evaluated is indeed causal, without knowing which one(s) is/are true, the conclusion would not be actionable from a public health standpoint. Moreover, the costs burden estimates vary by four or more orders of magnitude among the 15 exposure-health outcome relationships studied. Our contention is that the calculation and the consequent speculative conclusion from it should have been omitted from the papers.

The modified Delphi technique

The Trasande et al. (2015) Steering Committee coached the associated panels to use a modified Delphi technique in an effort to reach consensus on the probability of causation.

The Delphi technique is a widely used and accepted method for achieving convergence of opinion concerning real-world knowledge solicited from experts within certain topic areas. However, it has been noted by Hsu and Sanford (2007), that there are some very important considerations and limitations in its practical application. Subject selection, time frames for conducting and completing a study, the possibility of low response rates, and unintentionally guiding feedback from the respondent group are areas which should be considered when designing and implementing a Delphi study Most of these considerations did not inform the authors of this series of cost estimates (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) in their application of the technique.

Indeed, the selection of panelists is the most important step in the entire process because it directly relates to the quality of the results generated. They must truly be experts in the field. They need to have diverse perspectives on the topic. And there must be a sufficient number of them (certainly >10 according to Hsu and Sanford 2007) to ensure a representative pooling of judgments. The Trasande et al. (2015) Steering Committee violated all three of those criteria in their selection of the associated panels that ranged in size from 4 to 8 members.

Another major drawback of the Delphi method is the "subtle pressure to conform with group ratings." Delphi investigators need to be cognizant, exercise caution, and implement the proper safeguards in dealing with this issue. The IPCC Guidance for Authors document (Intergovernmental Panel on Climate Change 2005) also explicitly warns of the dangers of this phenomenon: "Be aware of a tendency for a group to converge on an expressed view and become overconfident in it." It is clear that either the associated panel group leaders or that members of the Trasande et al. (2015) Steering Committee failed to exercise such caution. A final practical concern with application of the Delphi method is an assumption that the panelists are equivalent in knowledge and experience. The biographies of the members of the "associated panels" indicates that they were clearly not "equivalent in knowledge and experience" regarding the topics about which they were asked to opine. Hsu and Sanford (2007) warned "Some panelists may have much more in-depth knowledge of certain topics, whereas other panelists are more knowledgeable about different topics". Therefore, panelists who have less in-depth knowledge of certain topics may be less inclined to challenge the statements of more knowledgeable panelists.

Attributable fraction estimates

To calculate their estimates of the fraction of disease attributable to EDC exposures the authors of the Trasande et al. (2015) derived publications (Bellanger et al. 2015; Hauser

et al. 2015; Legler et al. 2015; Hunt et al. 2016) relied on point estimates of relative risk derived from a single epidemiology study. In several important instances (e.g., the link between organophosphates and IQ loss and intellectual disability) there were only two or three relevant epidemiology studies available to choose from. In other instances (e.g., modeling phthalate-attributable decreases in T), they relied on estimates from a single cross-sectional epidemiology study, among the weakest of designs because it cannot establish that exposures actually preceded in time the initiation of disease, a requisite for establishing causation. Point estimates of risk from a single epidemiology study, or even a few taken together, are notoriously imprecise (Ioannidis 2005). The authors of the Trasande et al. (2015) derived publications (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016;) failed to adequately acknowledge such imprecision and the impacts it had on their cost estimates.

Exposure response relationship estimates

In a few instances, the authors of the Trasande et al. (2015) derived publications (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) extracted dose–response relationships from epidemiology studies (often a single study) and then applied them to population biomonitoring data in an effort to further delineate an estimated number of cases of disease that could be attributed to exposure. Once again, this represents an over-reliance on sparse data taken from one or a few studies that is perilous and implies a level of precision in the data that is simply cannot be justified.

Sources and uses of biomonitoring data

Attributable fraction estimates and exposure–response relationships, if available, were applied to the EU or US population based upon biomonitoring data available from surveys or pooled data from multiple studies in individual countries. Exposure levels were then estimated for quantiles (usually 0–9th, 10–24th, 25–49th, 50–74th, 75–89th, 90–99th) of the at-risk population.

Biomonitoring data for the paper that focused on generating cost estimates for the US (Attina et al. 2016) were derived from the National Health and Nutrition Examination Survey (NHANES) which is designed to be nationally representative (Department of Health and Human Services Centers for Disease Control and Prevention 2009). The authors (Attina et al. 2016) used biomonitoring data on PBDEs (Total PBDE 47 in serum when there were measurement data available for 11 types of BDEs), Dichlorodiphenyldichloroethylene (DDE) and OPPs (total urinary dialkylphosphate) extracted from the 2007–2008 NHANES and on Bisphenol A (BPA) and phthalates extracted from

the 2009–2010 NHANES. The authors (Attina et al. 2016) acknowledged that their models extrapolate from a lesser-brominated form of PBDEs, which was banned in the US and the EU much earlier than the more highly brominated PBDEs, and which is also reflected by a much lower human body burden of lesser-brominated form of PBDEs in Europe than in the US (WHO Regional Office for Europe 2015), but they rationalize this by pointing out that deca-PBDEs which have been banned or phased out more recently, are "…de-brominated by UV light and microbial and vertebrate organisms, and commercial mixtures that contain only lower-brominated congeners might represent relevant sources of exposure." The authors do not, however, discuss how this may have affected their results.

NHANES reports biomonitoring levels by age (6–11 years, 12–19 years and 20 years or older), gender, and race (Department of Health and Human Services Centers for Disease Control and Prevention 2009). The manuscript in question neither provides any level of detail as to whether the authors of this Trasande et al. (2015) derived publication (Attina et al. 2016) adjusted for these factors in applying the biomonitoring data, nor does it explain how biomonitoring data were used to estimate perinatal exposures which is critical to the estimation of neurodevelopmental outcomes.

Biomonitoring data for the papers that focused on cost estimates for the EU (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016) were derived from individual studies reported in the literature. A single study of 731 individuals from a general adult population (18–74 years) collected in 2002 in Catalonia (north-eastern Spain) was used to estimate PBDE exposures across the EU (Mercè Garí and Grimalt 2013). A single study of urine specimens collected in 2003–2006 from 600 children and adolescents aged 7–14 from Germany was used to estimate urinary dialkylphosphate as a surrogate for organophosphate exposure across the EU (Becker et al. 2007). Estimates of DDE exposures were derived from maternal and cord blood and breast milk samples of 7990 women enrolled in 15 study populations from 12 European birth cohorts from 1990 through 2008 (Govarts et al. 2012). Phthalate exposures were estimated from a demonstration project conducted in 17 European countries that measured phthalate metabolites in urine of 1844 children (5–11 years of age) and their mothers from specimens collected over a 5-month period in 2011–2012 (Den Hond 2015). BPA exposures were estimated from analysis of urine collected from 674 child (5–12 years of age)—mother pairs recruited through schools or population registers from six European member states (Belgium, Denmark, Luxembourg, Slovenia, Spain and Sweden) (Covaci et al. 2014). Notably, the authors of the studies measuring phthalates (Den Hond et al. 2015)

and BPA (Covaci et al. 2014) reported that exposure levels in nearly every case were found to be far below levels of health concern, whereas the authors of the studies measuring PBDEs (Mercè Garí and Grimalt 2013), organophosphates (Becker et al. 2007) and DDE (Govarts et al. 2012) did not make any comparisons to health risk values.

None of the Trasande et al. derived publications (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) provided a critical analysis of the strengths and limitations of the biomonitoring data which they relied on with the exception of this terse summary reproduced verbatim in each of the individual papers: "Likewise, biomarker data were not available for all EU countries, and judgment was used in extrapolating to the EU as a whole. By this approach, the authors attempted to avoid underestimating the burden of disease simply because of insufficient or missing data. On the other hand, the calculations could not take into account potential differences between exposure levels in the member states."

The onus should rightfully fall on the authors of the Trasande et al. (2015) derived publications (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) to investigate whether differences in exposure levels might be expected across the EU and to then to discuss the potential impacts on their results. This was not done. Yet, in the demonstration biomonitoring study conducted in 17 countries, its authors (Den Hond et al. 2015) concluded that levels of phthalate exposures varied markedly across the EU with clustering following geographical groupings. The Southern European countries (Spain, Portugal) clustered separately from the other countries; Eastern European countries (Romania, Hungary, Poland, the Czech Republic, and the Slovak Republic) formed a further cluster; Western European countries (Germany, Belgium, Luxembourg, and Denmark) also showed fairly good resemblance. None of the relevant papers (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) discussed this phenomena and how it might have impacted the reported results. Nor did the authors discuss how population exposures may have subsequently changed in the intervening years since the biomonitoring data were all collected between 2008 and 2010. Presumably, exposures will continue to decrease over time for several of these alleged EDCs as they have been banned or voluntarily phased out, as was repeatedly demonstrated for PBCs and DDE in the WHO/UNEP survey 2000–2012 (WHO regional office for Europe 2015).

Sources and uses of cost data

For the purposes of this review we will not assess the validity of the sources of data or methods used to assign societal costs. However, we encourage others with specialized expertise and familiarity with the literature on cost analysis to do so. We would point out that the costs of alleged lost IQ points and cases of intellectual disability, which constitute the majority of all societal costs estimated to be attributable to EDCs, are largely due to what was considered lost productivity and owe their roots in methodology developed by several of the same authors (Attina and Trasande 2013; Trasande and Liu 2011) responsible for the Trasande et al derived publications (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) rather than from any independent scientists. Perhaps for this reason alone, they deserve further scrutiny.

Cumulative effect of numerous assumptions and uncertainties

As has been detailed in the preceding sections, the sheer number and variety of assumptions used by the authors of the Trasande et al derived publications (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) to develop their cost estimates are staggering. Probability of causation is highly subjective and is dependent upon a number of tenuous assumptions the Trasande et al. (2015) Steering Committee and the associated panels simply did not clearly or adequately explain. The same is true for the processes of assigning Attributable Fraction, selection and application of exposure–response relationships, use of biomonitoring data and estimation of costs per health outcomes.

Recognizing that attributable cost estimates were accompanied by a probability, the Trasande et al. (2015) associated panels (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016) did perform a series of Monte Carlo simulations to produce ranges of probable costs across all the exposure-outcome relationships assuming independence of each probabilistic event. However, those simulations were confined to the ranges of probabilities assigned by the panels that, as has already been demonstrated, were based on incomplete and flawed assessments of the strength and weight of the available evidence. Such simulations also ignored substantial uncertainties that exist in the Attributable Fraction estimates, dose–response relationships that were chosen for use, the biomonitoring data, and estimates of costs associated with each health outcome evaluated.

Critical review of evidence for loss of IQ and intellectual disability

Because the neurobehavioral deficits and diseases cost estimate overwhelmingly drives the total estimated cost attributable to EDCs (>75% of the total in the EU (Trasande et al. 2015) and >79% of the total in the US (Attina et al. 2016), the corresponding paper by Bellanger et al. (2015) was examined in greater detail. Special attention was paid to the purported links between exposure to organophosphates (OPs) and separately exposures to PBDEs and lowered IQ and increased prevalence of intellectual disability.

Organophosphate exposures and loss of IQ and intellectual disability

Bellanger et al. (2015) are not the first group of scientists to review the available evidence on this topic. Li et al. (2012) conducted an exhaustive critical review of the epidemiology and animal data available on chlorpyrifos (CPF), amongst the most widely used organophosphate pesticides (and the one Bellanger et al. (2015) cite as the most widely used in the EU) and made several conclusions. First, there was insufficient evidence that human developmental exposures to CPF produced adverse neurobehavioral effects in infants and children across different cohort studies that may be relevant to CPF exposure. In animals, neurodevelopmental behavioral, pharmacological, and morphologic effects occurred only at doses that produced significant brain or red blood cell acetylcholinesterase inhibition in dams or offspring. Chlorpyrifos is regulated based on protecting against red blood cell cholinesterase inhibition and this occurs at a level below which these other effects were reported, hence protecting the public against all other toxicities.

Eaton et al. (2008) also undertook an exhaustive critical review of toxicological and epidemiological information on human exposures to CPF, with an emphasis on the controversial potential for CPF to induce neurodevelopmental effects at low doses. They concluded that, based on the weight of the scientific evidence, it is highly unlikely that current levels of CPF exposure in the United States would have any adverse neurodevelopmental effects in infants exposed in utero to chlorpyrifos through the diet.

Independent reviews by Ntzani et al. (2013), Prueitt et al. (2011) and Burns et al. (2013) arrived at similar conclusions to those cited above. Inexplicably, Bellanger et al. (2015) neither make reference to these other reviews nor do they attempt to explain why the other authors

came to conclusions that differ from their own and more importantly, do not explain why these reviews were not included in their assessment. More recently, a US EPA FIFRA Science Advisory Panel (2016) examined the evidence and concluded, "The assumption that the impaired working memory and lower IQ measures observed are caused primarily by a single insecticide (chlorpyrifos) and predicted by the blood levels at time of delivery is not supported by the scientific weight of evidence".

Furthermore, the US EPA recently released the results of its Endocrine Disruption Screening Program (EDSP) Tier 1 weight-of-the-evidence evaluation of the potential interaction of chlorpyrifos with the estrogen, androgen or thyroid hormone signaling pathways (US EPA 2015a). The Agency concluded that further testing of chlorpyrifos under the EPA's EDSP Tier 2 program was not recommended since they found no evidence of potential interaction with any of the three pathways. Thus, EPA has concluded that chlorpyrifos is unlikely to be an EDC in contrast to what Bellanger et al. (2015) have alleged. Moreover, EPA also released weight of the evidence conclusions under EDSP Tier 1 for five other organophosphate pesticides (acephate, dimethoate, ethopropos, phosmet, and tetrachlorvinphos) for which it was similarly concluded that further testing under Tier 2 was not warranted (US EPA 2015b).

In direct contradiction to these other assessments, Bellanger et al. (2015) concluded that the epidemiology evidence linking OPs to lowered IQ and intellectual disability was moderate to high and that the toxicological evidence was strong, assigning the probability of causation to be 70–100%. This is remarkable since the epidemiological evidence derived from three epidemiology studies has been heavily scrutinized by others (Li et al. 2012; Eaton et al. 2008, Ntzani et al. 2013; Prueitt et al. 2011; Burns et al. 2013, US EPA FIFRA Science Advisory Panel 2016) and criticized for problems related to their design, exposure and outcome measurements, and inadequate control of potential confounding variables (i.e., other risk factors). The paper by Bellanger et al. (2015) offers no critical analysis of the studies upon which it relied. Moreover, the original studies actually report relatively few statistically significant findings suggestive of adverse effects, and no consistent patterns can be found across them. They also reported small differences in test scores for the children in the three groups of women with relatively higher OP exposures than those typically measured in the general population.

Childhood IQ and neurobehavioral development are strongly influenced by the complex interaction of many other well-established risk factors (genetic conditions, alcohol and/or drug use in pregnancy, prematurity and low birth weight, childhood diseases, poverty and cultural deprivation, iodine deficiency, and others), the effects of which

the researchers could not sufficiently control in their studies. The reported differences in test scores between children with presumed high vs. low OP exposures were small enough to be explained by normal variation or by the interactions of other factors.

None of the epidemiology studies Bellanger et al. (2015) relied upon reported links between OP exposure and intellectual disability. Rather than being constrained by this limitation, however, Bellanger et al. (2015) then proceeded to define intellectual disability as an IQ <70 and used a questionable approach of modeling the reported increases in intellectual disability assuming a normal distribution with a mean IQ of 100 and standard deviation of 15. Then, within each exposure grouping, they used a statistical function to identify increases in associated intellectual disability. Note that the definition of intellectual disability (IQ <70) used by Bellanger et al. (2015) is to some extent in conflict with established definitions (Schalock et al. 2010) which define intellectual disability as a below-average cognitive ability with three characteristics: (1) IQ is between 70 and 75 or below; (2) Significant limitations in adaptive behaviors (the ability to adapt and carry on everyday life activities such as self-care, socializing, communicating, etc.); and (3) The onset of the disability occurs before age 18.

PBDE exposures and loss of IQ and intellectual disability

Although the costs attributed to IQ loss and intellectual disability due to PBDE exposures in the EU accounted for less than 6% of the alleged total EDC induced cost burden there (Bellanger et al. 2015), it accounted for 78% of the alleged total costs in the US (Attina et al. 2016). For this reason, the purported causal link between PBDEs and neurobehavioral outcomes also deserves further scrutiny. Bellanger et al. (2015) judged the animal toxicology evidence to be strong because of their stated belief that PBDEs interfere with thyroid action and this consequently causes IQ loss.

One of the main issues when considering the potential of compounds to interfere with thyroid hormones is to look at the differences between humans and rodents with regard to thyroid function (Swenberg et al. 1992; Robbins and Rall 1979; Dohler et al. 1979; Dingemans et al. 2011; Huwe and Smith 2007). Indeed, rodents do not express thyroid-binding globulin (Swenberg et al 1992), i.e., the predominant plasma protein that binds and transports thyroid hormone in the blood. Thyroxine (T4) binds to three plasma proteins, thyroid binding globulin, pre-albumin and albumin, with binding constants of 10-1, 10-7 and 10-5, respectively (Robbins and Rall 1979).

The lack of thyroid-binding globulin in the rodent, in which albumin and pre-albumin have three and five orders of magnitude less binding affinity for thyroxin, may be one of the more important factors that could render rodents more susceptible to compounds interfering with thyroid action. Indeed, the half-life of T4 is 12 h in rats and five to nine days in humans, and the serum TSH level is 25 times higher in rodents than in man (Dohler et al. 1979). The rodent thyroid gland thus has much higher activity than that of the primate, a conclusion that is also supported by the histological appearance of the thyroid gland (Swenberg et al. 1992). Both the physiological parameters and the histological appearance indicate that the rodent thyroid gland is markedly more active and operates at a considerably higher level with respect to thyroid hormone turnover as compared to the primate. There are even greater differences in thyroid homeostasis and the importance of the thyroid hormone system between rats and humans with regard to reproductive tract and neurological development, especially with regard to the timing during development and downstream effects resulting from correctly regulated interactions of thyroid hormones (Choksi et al. 2003). Moreover, it is important to note that many quantitative parameters of the rodent differ from those of the primate by orders of magnitude, and thus making any extrapolations from rat studies to humans, especially with regard to the development of the human nervous system, a highly complex if not impossible exercise (Crofton et al. 2008). As a consequence of the latter, the ototoxicity, and thus hearing loss observed in early postnatal exposed rats to polychlorinated biphenyls (PCBs) with subsequent upregulation of hepatic uridine diphosphoglucuronyltransferases and subsequent hypothyroxinemia during the critical period of rat cochlear development, is not readily extrapolated to the human exposure scenario. Indeed it is the interspecies differences in toxicodynamic and kinetic factors and the order of magnitude lower concentrations of PCBs that humans are exposed to under normal ambient conditions that greatly increase confidence that this type of neurotoxicity would not occur in humans (Crofton and Zoller 2005).

Based on the latter and due to its similarity with PCBs, it is difficult to extrapolate the findings on PBDE in rodents on the thyroid i.e., its thyroidogenic activity, and the indirect effects reported on behavioral effects and neurotoxicity (Dingemans et al. 2011), usually obtained after exposure to very high doses, with the very low concentrations observed in humans (WHO Regional Office for Europe 2015).

One of the main troubling issues when using rodent in vivo studies to extrapolate to the humans is the fact that very few if any of the studies reporting behavioral and neurotoxic effects secondary to thyroidogenic activity looked at PBDE distribution and kinetics within the same exposure experiment. Indeed PBDE levels in the plasma, pituitary, thyroid etc., are critical for comparison with human levels found following long-term exposure before any adverse effect noted can be interpreted and hazard or risk can be extrapolated (WHO Regional Office for Europe 2015).

Moreover, PBDEs appear to undergo extensive metabolism and excretion in rodents (Huwe and Smith 2007; Staskal et al. 2006), whereby upon hydroxylation most PBDE metabolites are excreted via the urine as conjugates.

In humans, similar hydroxylated PBDE-metabolites are formed (Stapleton et al. 2009), albeit the excretion kinetics of individual PBDE metabolites and their parent compounds in humans are yet unknown. Thus, to compare rodent in vivo findings and adverse effects reported with the human exposure situation one would have to determine the biologically available concentrations of individual PBDE parent compounds and their active metabolites prior to inferring any risk or even consider a causal link to human developmental neurotoxicity.

Recently, efforts were made to establish better hazard identification using human neuronal systems for the in vitro study of the potential (developmental) neurotoxic effects of PBDEs (and resulted in the identification of PBDE-47 and PBDE-99 as potential developmental neurotoxicants (DNT) with relevance for humans (Aschner et al. 2016). The issue with the latter assessments is, however, the fact that all in vitro systems employed to determine the DNT lacked lipids in the cell culture media used, thus not representing a plausible system where blood lipids would compete with cell membranes and enterocytes for the highly lipophilic PBDEs and their metabolites. The latter, as well as due to the fact that high PBDE concentrations (10 μM) were used in conjunction with multiple replacement of cell culture media, resulted in an over-proportional distribution of PBDE from the cell culture media into the cells (most likely merely the plasma membrane) (Schreiber et al. 2010; Fritsche et al. 2005; Polloca et al. 2016), thus severely questioning the relevance of the in vitro findings for human hazard and risk assessment. Whether the most recent report of PBDE-44 and -47 inhibition of axonal growth in primary rat hippocampal neuron-glia co-cultures via ryanodine receptor (RyR)-dependent mechanisms (Chen et al. 2016) has any bearing for human risk assessment is yet to be established with human equivalent in vitro systems, while bearing in mind the distribution kinetics of PBDEs, especially as the concentration-effect relationships of PBDE-47 and PBDE-49 on axonal growth were reported to be comparable despite reported differences in their potency at the RyR.

Obviously the currently available evidence suggesting DNT due to thyroidogenic interaction of PBDE cannot be summarily discounted, especially as slightly higher body burdens were observed in infants than in older children (Linares et al. 2015). However, the current evidence must be placed in proper context with the high concentrations needed to achieve these adverse effects in vivo and in vitro, the low concentrations of PBDEs and their metabolites reported in humans (WHO Regional Office for Europe 2015), the uptake and elimination kinetics prevailing in humans, the different physiological predisposition and thyroid function of rodents and humans, and thus different bioavailabilities of the compounds and metabolites, and finally the fact that humans are exposed daily to potentially thyroidogenic compounds of natural origin, e.g., Bisphenol-F from sweet mustard (Dietrich and Hengstler 2016; Rochester and Bolden 2015), including PBDEs from marine biota that are readily transferred to higher trophic levels including fish for human consumption (Agarwal et al. 2014; Teuten et al. 2005). Moreover, as already stated above, most body burdens determined were related to PBDE-47 and PBDE-99 that have been banned from markets in the US and Europe and would be demonstrating a decreasing body burden over the past and next decades.

The epidemiology evidence available to the Trasande et al. (2015) associated panel (Bellanger et al. 2015) derived from three studies (Herbstman et al. 2010; Eskenazi et al. 2013; Chen et al. 2014) and was judged to be as moderate to high strength. Once again, Bellanger et al. (2015) neither provided a critical review of these three studies, nor did they cite existing critical reviews already published in the literature (Kim et al. 2014; Roth and Wilks 2014) that arrived at conclusions different than their own.

Kim et al. (2014) conducted a systematic review of the epidemiology literature linking exposure to brominated flame retardants (BFRs) to various health effects, including alteration in thyroid function, diabetes, cancer, reproductive disorders and diseases and neurobehavioral and developmental disorders. Evidence for a causal relationship between BFRs and health outcomes was evaluated within the Bradford Hill's framework. Although they found suggestive evidence that exposure to BFRs is harmful to health, they concluded that "…more well-designed research is needed to support these tentative but biologically plausible associations". They noted possible publication bias in the literature available to them to review; small population sample sizes; incomplete control of confounding variables; weak and inconsistent findings across studies; and the likelihood of measurement errors in assessing both exposure and putative health effects as reasons for why the evidence could not be judged to be stronger.

Similarly, Roth and Wilks (2014) conducted a systematic review of the epidemiology literature on PBDEs and perfluorinated chemicals (PFCs). They concluded that "Collectively, the epidemiological evidence does currently not support a strong causal association between PBDEs and PFCs and adverse neurodevelopmental and neurobehavioural outcomes in infants and children." They noted the following shortcomings among the studies: the lack of consideration of confounding factors; uncertainties regarding exposure characterization; inadequate sample size; the lack of a clear dose–response; and the representativeness/generalizability of the results.

Discussion

Our critical review of a series of seven inter-related papers (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) alleging hundreds of billions of euros and dollars in annual societal costs purportedly attributable to exposures to EDCs in the EU and the US, respectively, has demonstrated substantial shortcomings with the underlying methodology employed that render the cost estimates highly uncertain. In our opinion, the cost estimates are so speculative and flawed that they should not be accorded any weight in serious policy discussions. To its credit, the European Commission recently concluded, "Estimates on costs of diseases related to exposure to EDs [EDCs] which were recently published should be taken with caution. There are concerns over the validity of these estimates and the methods used to calculate them, which are linked to the scattered evidence" (European Commission 2016a).

The fundamental problem with the cost estimates is that they derive from the assumptions of causal relationships between putative exposures to EDCs that have not been established through any serious consideration of the strengths and weaknesses of the underlying animal toxicology and human epidemiology evidence. The authors readily acknowledge that the approach they have taken to assigning probability of causation is atypical and involves a high degree of subjective judgment. Such cautionary language was neither included in the abstracts of the individual papers nor was it mentioned in the aggressive press releases that accompanied the publications Unfortunately, it was and continues to be largely hidden from the public and policymakers.

Regrettably, this lack of transparency and failure to clearly communicate findings is becoming all too common in science. For example, Sumner et al. (2014) compared 462 biomedical/health-related science press releases issued by 20 leading UK-based universities to the associated peer-reviewed research papers and subsequent news stories and found that 40% of the press releases contained exaggerated advice, 33% contained exaggerated causal claims and that 36% contained exaggerated inference to humans from animal studies. When the press releases contained such exaggeration, 58% of the subsequent news stories did as well.

Further aggravating the problem, in an apparent attempt to bolster the credibility of their chosen approach the authors of the Trasande et al. derived publications (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) claimed they had adapted methods employed by the WHO and by the IPCC; however, those claims fail to pass closer scrutiny, and it is

quite apparent that the authors instead devised their own unique approach without a firm grounding in science or precedence. The papers (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) are also fraught with numerous deficiencies including: a failure to use state of the art systematic review methodology, a lack of transparency in reporting how the literature was searched and which studies were selected for review, a failure to achieve a balance of perspectives in selecting membership for the panels, deviation from best practices for applying the Delphi technique in an effort to gain consensus among panel members, a lack of serious critical discussion of the strengths and weaknesses of the individual studies relied upon, over-reliance on sparse and unrepresentative data for assigning Attributable Fraction and choosing dose–response relationships, and on unrepresentative biomonitoring data for modeling the number of cases of disease attributable to EDC exposures.

The largest driver of estimated costs was IQ loss and intellectual disability purportedly caused by exposures to OPP insecticides in the EU and PBDE exposures in the US (Bellanger et al. 2015). As we have shown this was based on evidence from just a few observational epidemiology studies of relatively small, and unrepresentative population samples, which reported weak and inconsistent findings. Other independent scientists have reviewed the same literature, have enumerated numerous limitations of those studies, and have arrived at very different conclusions than those of the expert panel assembled to review the neurobehavioral deficits and diseases evidence. In the case of OPPs, the independent reviews concluded that, at the exposure levels seen in the general population, it is unlikely that they cause lost IQ and intellectual disability. For the purported PBDE link, they concluded that the evidence was inadequate to establish a causal relationship. Remarkably, none of these independent reviews was cited or discussed by the authors of the Trasande et al. associated panel (Bellanger et al. 2015) an obvious major omission. When communicating with the public, scientists should be clear about uncertainties, presenting competing views or interpretations of data, and stating the limitations of the data presented (Runkle and Frankel 2012). The authors of the Trasande et al. (2015) derived publications (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) fall well short of that standard.

We did not undertake a detailed review of the underlying toxicology and epidemiology evidence considered by the three other Trasande et al. associated panels that explored links between EDC exposures and male and female reproductive diseases and disorders, and obesity and diabetes (Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016).

We cannot comment specifically upon them. However, we would highlight that those panels used the same flawed methodology highlighted herein, and a cursory review of the evidence suggests a similar over-reliance on a few human epidemiology studies that have some of the same limitations noted above.

A recent systematic review and meta-analysis of the link between EDCs and male reproductive diseases and disorders (Bonde et al. 2016) concluded. "The widely stated view that ubiquitous endocrine disrupting chemicals in our environment play a substantial role in the development of male reproductive disorders through prenatal and perinatal mechanisms is to some extent challenged by this review. Although the current epidemiological evidence is compatible with a small increased risk of male reproductive disorders following prenatal and postnatal exposure to some persistent environmental chemicals classified as endocrine disruptors, the evidence is limited. In this light, estimates of the burden of disease and costs of exposure to endocrine chemicals (Trasande et al. 2016) seem highly speculative, at least with respect to male reproductive disorders." (Note the conclusion that evidence was compatible with a small increased risk of make reproductive disorders (Bonde et al. 2016) was based on a meta-analysis generated overall odds ratio reported to be 1.11 with 95% confidence Intervals ranging from 0.91 to 1.35, which most epidemiologists would regard as very weak.) We encourage others to scrutinize this review and other epidemiology evidence purportedly linking EDCs and adverse health effects in greater detail.

Pertinent to this discussion, LaKind et al. (2015) have reviewed some general problems with the environmental epidemiology literature that often precludes an ability to draw robust conclusions regarding the presence or absence of causal links between specific exposures and human health effects. They note that, to develop policies that are protective of public health and that can withstand scrutiny, the investigations need to be of sufficiently high quality in terms of exposure assessment, health outcome ascertainment, data analysis and reporting of results. They propose a three-part approach addressing methods for improving the quality and accessibility of systematic reviews, access to information on ongoing and completed studies, and principles for reporting study results. This approach could certainly improve the impact that environmental epidemiology research has on future chemical hazard and risk assessments.

The proponents of the cost estimates reviewed herein have also asserted that they have in all likelihood underestimated the total societal costs attributed to EDC exposures, noting they only focused on the 15 exposure-outcome relationships that had the highest probability of causation, and that a broader analysis would have produced even greater estimates of burden of disease and of costs. As discussed, however, the cost estimates presented are highly uncertain because the authors have not even established a compelling case for causation for the 15 exposure-outcome relationships they did include for study.

The timing of the release of the six papers (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016) which focused on estimated disease burden and costs in the EU appears to coincide with other efforts being made to influence the EU Commission's proposed criteria to identify EDCs for regulatory purposes (European Commission 2016a). By statute, the EU has committed to taking a hazard-only rather than a risk-based approach to regulating pesticides and biocides that are identified as EDCs. Broadly defined criteria could substantially increase the number of cost effective and efficacious chemicals that are misidentified as EDCs and subjected to bans and/or restrictions, thereby limiting consumer choice and increasing costs. However, such increased costs might look much less onerous when balanced against the massive societal costs attributed to EDC exposures as have been alleged by this series of papers. Consequently, for the purposes of policy decisions it is important to place the alleged cost of inaction estimates into proper perspective.

The timing of the release of the seventh paper (Attina et al. 2016) that focused on cost estimates in the US is somewhat curious as it lagged by several months the adoption of the Frank R. Lautenberg Chemical Safety for the 21st Century Act (US EPA 2016b). The authors of that paper criticize the Act because they claim it does not specifically mention endocrine disruption. They also allege that the law does not make a provision for endocrine testing programs and does not address the urgent threat posed by EDCs. Yet, they fail to mention that the Act provides the U.S. EPA with new and broad authority to require chemical testing for new and existing chemicals and requires EPA to "prohibit or impose restrictions to protect against unreasonable risk". There exist many opportunities for interested parties, including the authors of the Trasande et al. (2015) derived publications (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) to comment on how EPA implements the Act, especially with respect to prioritizing specific chemicals for evaluation and action. So, despite their protests, the authors of the Trasande et al. (2015) derived publications (Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) and indeed all stakeholders have an ability to exercise their right to try to influence the EPA without having to exaggerate what they perceive to be the costs of inaction.

US EPA also has ultimate responsibility for implementing the EDSP (US EPA 2017) for screening and testing chemicals for endocrine disrupting properties. The Agency has recently announced a pivot in the program to greatly accelerate the pace at which chemicals are screened and to reduce the use of laboratory animals. Once again, the program operates transparently and stakeholders are invited to engage constructively to improve upon it.

Conclusion

Some readers may still ask the questions: "What is the harm if the costs of EDCs are grossly exaggerated? Isn't it better to be safe than sorry?" We would argue, as have others (Runkle and Frankel 2012; Allen et al. 2016; Vergano 2015), that everyone loses when the scientific method, which demands a level of impartial objectivity, is not met as demonstrated by this series of papers. Not only can there be significant and lasting damage to the credibility of individual scientists and organizations, but public acceptance and support for science, which is already on shaky ground (Vergano 2015), becomes further eroded. Public policy choices based on poor science or isolated findings from unreplicated studies, even while well-intentioned, will have significant negative consequences for individuals and society. As Dietrich et al. (2016) noted, "few without scientific training realize that science progresses by the detection of, and subsequent elimination of, errors. This is why acting on findings in isolation, all too common an occurrence today, is an unsound strategy. Perhaps equally important, failure of decision makers to recognise this, leads to unnecessarily restrictive and potentially damaging regulation."

Our analysis has consistently demonstrated, that the cost estimates made in this series of papers (Trasande et al. 2015; Bellanger et al. 2015; Hauser et al. 2015; Legler et al. 2015; Hunt et al. 2016; Trasande et al. 2016; Attina et al. 2016) are based on flawed methodology, and a host of tenuous, unsupported assumptions and, as a consequence, they lack scientific validity. The estimates are so highly speculative that no weight should be ascribed to them in any serious policy discussions of EDCs.

References

Agarwal V et al (2014) Biosynthesis of polybrominated aromatic organic compounds by marine bacteria. Nat Chem Biol 10:640–647. doi:10.1038/nchembio.1564

Allen S et al (2016) An appeal for the integrity of science and public policy. Toxicology 371:1–11. doi:10.1016/j.tox2016.08.015

Aschner M et al (2016) Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: example lists and criteria for their selection and use. Altex 34:1. doi:10.14573/altex.1604201

Attina T et al (2016) Exposure to endocrine-disrupting chemicals in the USA: a population-based disease burden and cost analysis. Lancet Diabetes Endocrinol 4:996–1003. doi:10.1016/S2213-8587(16)30275-3

Attina TM, Trasande L (2013) Economic costs of childhood lead exposure in low- and middle-income countries. Environ Health Perspect 121:1097–1102

Beausoleil C et al (2016) Review of non-monotonic dose-responses of substances for human risk assessment. Support Publ 13(5). doi:10.2903/sp.efsa.2016.EN-1027

Becker K et al (2007) German environmental survey for children, 2003/06. Available at http://www.umweltbundesamt.de/sites/default/files/medien/publikation/long/3355.pdf. Accessed 4 Jan 2017

Bellanger M et al (2015) Neurobehavioral deficits, diseases, and associated costs of exposure to endocrine-disrupting chemicals in the European Union. J Clin Endocrinol Metab 100(4):1256–1266. doi:10.1210/jc.2014-432

Bergman A et al (2013) State of the science of endocrine disrupting chemicals, 2012. United Nations Environment Programme and the World Health Organization. http://apps.who.int/iris/bitstream/10665/78101/1/9789241505031_eng.pdf?ua=1

Bonde JP et al (2016) The epidemiologic evidence linking prenatal and postnatal exposure to endocrine disrupting chemicals with male reproductive disorders: a systematic review and meta analysis. Hum Reprod Update. doi:10.1093/humupd/dmw036

Borgert C et al (2013) Potency matters: thresholds govern endocrine activity. Regul Toxicol Pharmacol 67(1):83–88. doi:10.1016/j.yrtph.2013.06.007

Bruce N et al (2014) WHO Indoor air quality guidelines: household fuel combustion. Methods used for assessment. Available at http://www.who.int/indoor/guidelines/hhfc/Evidence_review_methods.pdf Accessed 1 Mar 2015

Burns C et al (2013) Pesticide exposure and neurodevelopmental outcomes: review of the epidemiologic and animal Studies. J Toxicol Environ Health 16(3–4):127–283. doi:10.1080/10937404.2013.783383

Chen A et al (2014) Prenatal polybrominated diphenyl ether exposures and neurodevelopment in US children through 5 years of age: the HOME Study. Environ Health Perspect 122(8):856–862

Chen H et al (2016) BDE-47 and BDE-49 inhibit axonal growth in primary rat hippocampal neuron-glia co-cultures via ryanodine receptor-dependent mechanisms. Toxicol Sci. doi:10.1093/toxsci/kfw259

Choksi NY et al (2003) Role of thyroid hormones in human and laboratory animal reproductive health. Birth Defects Res (Part B) 68:479–491. doi:10.1002/bdrb.10045

Covaci A et al (2014) Urinary BPA measurements in children and mothers from six European member states: overall results and determinants of exposure. Environ Res. doi:10.1016/j.envres.2014.08.008i

Crofton KM (2008) Thyroid disrupting chemicals: mechanisms and mixtures. Int J Androl 31:209–223. doi:10.1111/j.1365-2605.2007.00857.x

Crofton KM, Zoller RT (2005) Mode of action: neurotoxicity induced by thyroid hormone disruption during development—hearing loss resulting from exposure to PHAHs. Crit Rev Toxicol 35:757–769. doi:10.1080/10408440591007304

Danish EPA (2011) Establishment of criteria for endocrine disruptors and options for regulation. http://eng.mst.dk/media/mst/Attachments/DKEDcriteria110517_finalcorr1.pdf. Accessed on 2 Nov 2016

DE-UK (2011) Joint DE-UK Position Paper. Regulatory definition of an endocrine disrupter in relation to potential threat to human health. Proposal applicable in the regulatory context of plant protection products, biocidial products, and chemicals targeted within REACH. www.bfr.bund.de

Den Hond E et al (2015) First steps toward harmonized human biomonitoring in Europe: demonstration project to perform human biomonitoring on a European scale. Environ Health Perspect. doi:10.1289/ehp.1408616

Department of Health and Human Services Centers for Disease Control and Prevention (2009) Fourth national report on human exposure to environmental chemicals 2009. http://www.cdc.gov/exposurereport/pdf/fourthreport.pdf. Accessed on 2 Nov 2016

Dietrich D et al (2016) Allowing pseudoscience into EU risk assessment processes is eroding public trust in science experts and in science as a whole: the bigger picture. Chemico-Biol Interact 257(2016):1–3

Dietrich DR, Hengstler JG (2016) Highlight report: from bisphenol A to bisphenol F and a ban of mustard due to chronic low dose exposures? Arch Toxicol 90(2):489–491. doi:10.1007/s00204-016-1671-5

Dingemans MML et al (2011) Review: neurotoxicity of brominated flame retardants: (In)direct effects of parent and hydroxylated polybrominated diphenyl ethers on the (developing) nervous system. Environ Health Perspect. doi:10.1289/ehp.1003035

Dohler KD et al (1979) The rat as a model for the study of drug effects on thyroid function: consideration of methodological problems. Pharmacol Ther 5:305–318

Eaton D et al (2008) Review of the toxicology of chlorpyrifos with an emphasis on human exposure and neurodevelopment. Crit Rev Toxicol 38(Sup 2):1–125. doi:10.1080/10408440802272158

Eskenazi B et al (2013) In utero and childhood polybrominated diphenyl ether (PBDE) exposures and neurodevelopment in the CHAMACOS study. Environ Health Perspect 121:257–262

European Commission (2013) JRC scientific and policy reports: key scientific issues relevant to the identification and characterization of endocrine disrupting substances. Rep Endocr Disrupt Expert Advisory Group. doi:10.2788/8659 (online)

European Commission (2016a) Staff working document: defining criteria for identifying endocrine disruptors in the context of the implementation of the plant protection products regulation and biocidal products regulation http://ec.europa.eu/health/endocrine_disruptors/docs/2016_impact_assessment_en.pdf

European Commission (2016b) Rules of procedure for the scientific committees on consumer safety (SCCS) and health, environmental and emerging risks (SCHEER). http://ec.europa.eu/health/scientific_committees/docs/rules_procedure_2016_en.pdf. Accessed on 2 Nov 2016

Fritsche E et al (2005) Polychlorinated biphenyls disturb differentiation of normal human neural progenitor cells: clue for involvement of thyroid hormone receptors. Environ Health Perspect 113:871–876. doi:10.1289/ehp.7793

Govarts E et al (2012) Birth weight and prenatal exposure to polychlorinated biphenyls (PCBs) and dichlorodiphenyldichloroethylene (DDE): a meta-analysis within 12 European birth cohorts. Environ Health Perspect 120:162–170. doi:10.1289/ehp.1103767 (online 13 October)

Gray E (2013) Non-monotonic dose response curves (NMDRCs) are common after estrogen or androgen signaling pathway disruption. Fact or falderal? Presented at Risk Assessment Specialty Section, Society of Toxicology sponsored RASS Webinar, May 08, 2013. Available on-line at https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=257864

Hauser R et al (2015) Male reproductive disorders, diseases, and costs of exposure to endocrine-disrupting chemicals in the European Union. J Clin Endocrinol Metab 100(4):1267–1277. doi:10.1210/jc.2014-4325

Herbstman J et al (2010) Prenatal exposure to PBDEs and neurodevelopment. Environ Health Perspect 118:712–719

Hsu CC, Sanford B (2007) The Delphi technique: making sense of consensus. Pract Assess Res Eval 12(10):1–8

Hunt PA et al (2016) Female reproductive disorders, diseases, and costs of exposure to endocrine disrupting chemicals in the European Union. J Clin Endocrinol Metab. doi:10.1210/jc.2015-2873#sthash.uhE1pZY.dpuf

Huwe JK, Smith DJ (2007) Accumulation, whole-body depletion, and debromination of decabromodiphenyl ether in male Sprague−Dawley rats following dietary exposure. Environ Sci Technol 41(7):2371–2377. doi:10.1021/es061954d

Intergovernmental Panel on Climate Change (2005) Guidance notes for lead authors of the IPCC fourth assessment report on addressing uncertainties. http://www.ipcc.ch/meetings/ar4-workshops-express-meetings/uncertainty-guidance-note.pdf. Accessed on 2 Nov 2016

Ioannidis J (2005) Why most published research findings are false. PLoS Med 2(8):124

Ioannidis J (2016) We have an epidemic of deeply flawed meta-analyses, says John Ioannidis. Retraction Watch. http://retractionwatch.com/2016/09/13/we-have-an-epidemic-of-deeply-flawed-meta-analyses-says-john-ioannidis/

Kim YR et al (2014) Health consequences of exposure to brominated flame retardants: a systematic review. Chemosphere 106:1–19. doi:10.1016/j.chemosphere.2013.12.064

LaKind J et al (2015) Improving concordance in environmental epidemiology: a three-part proposal. J Toxicol Environ Health Part B 18(3):105–120. doi:10.1080/10937404.2015.1051612

Lamb J et al (2014) Critical comments on the WHO-UNEP state of the science of endocrine disrupting chemicals–2012. Regul Toxicol Pharmacol 69:22–40

Legler J et al (2015) Obesity, diabetes, and associated costs of exposure to endocrine-disrupting chemicals in the European Union. J Clin Endocrinol Metab 100(4):1278–1288. doi:10.1210/jc.2014-4326

Li A et al (2012) Evaluation of epidemiology and animal data for risk assessment: chlorpyrifos developmental neurobehavioral outcomes. J Toxicol Environ Health Part B 15(2):109–184. doi:10.1080/10937404.2012.645142Journal

Linares V et al (2015) Human exposure to PBDE and critical evaluation of health hazards. Arch Toxicol 89:335. doi:10.1007/s00204-015-1457-1

Mercè Garí M, Grimalt JO (2013) Inverse age-dependent accumulation of decabromodiphenyl ether and other PBDEs in serum from a general adult population. Environ Int 54:119–127. doi:10.1016/j.envint.2013.01.012

Middelbeek RJ, Veuger SA (2015) Letter to the Editor: Re: neurobehavioral deficits, diseases, and associated costs of exposure to endocrine-disrupting chemicals in the European Union. J Clin Endocrinol Metab. doi:10.1210/jc.2015-2010

Ntzani E et al (2013) Literature review of epidemiology studies linking exposure to pesticides and health effects. EFSA Supporting Publication EN:497, pages 159. http://onlinelibrary.wiley.com/store/10.2903/sp.efsa.2013.EN-497/asset/efs3497e.pdf?v=1&t=iv1eeoou&s=9b9545d83a1e7ed6c627ba935e169519e9058a8c

OECD (2012) Guidance document on standardised test guidelines for evaluating chemicals for endocrine disruption, series on testing and assessment, No. 150, ENV/JM/MONO 22 http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2012)22&doclanguage=en. Accessed 2 Nov 2016

Polloca G et al (2016) Identification of transcriptome signatures and biomarkers specific for potential developmental toxicants inhibiting human neural crest cell migration. Arch Toxicol 90(1):159–180. doi:10.1007/s00204-015-1658-7

Prueitt R et al (2011) Hypothesis-based weight of evidence evaluation of the neurodevelopmental effects of chlorpyrifos. Crit Rev Toxicol 41(10):822–903. doi:10.3109/10408444.2011.616877

Robbins J, Rall JE (1979) The iodine containing hormones. In: Gray CH, James YHT (eds) Hormones in blood, 3rd edn. Academic Press, London, p 576

Rochester JR, Bolden AL (2015) Bisphenol S and F: a systematic review and comparison of the hormonal activity of bisphenol A substitutes. Environ Health Perspect 123:643–650. doi:10.1289/ehp.1408989

Rooney A et al (2014) Systematic review and evidence integration for literature-based environmental health science assessments. Environ Health Perspect 122:711–718. doi:10.1289/ehp.1307972

Roth N, Wilks M (2014) Neurodevelopmental and neurobehavioural effects of polybrominated and perfluorinated chemicals: a systematic review of the epidemiological literature using a quality assessment scheme. Toxicol Lett 230(2):271–281. doi:10.1016/j.toxlet.2014.02.015

Runkle D, Frankel M (2012) Advocacy in science: summary of a workshop convened by the American Association for the Advancement of Science. Washington, DC. https://mcm-prodaaas.s3.amazonaws.com/s3fs-public/reports/Advocacy_Workshop_Report_FINAL.pdf

Schalock R et al (2010) Intellectual disability: Definition, classification, and systems of supports (11th Edition) association on intellectual and developmental disabilities. ISBN: 978-1-935304-04-3

Schreiber et al (2010) Polybrominated diphenyl ethers induce developmental neurotoxicity in a human in vitro model: evidence for endocrine disruption. Environ Health Perspect 118:572–578. doi:10.1289/ehp.0901435

Stapleton HM et al (2009) Detection of organophosphate flame retardants in furniture foam and US house dust. Environ Sci Technol 43(19):7490–7495

Staskal et al (2006) Toxicokinetics of polybrominated diphenyl ether congeners 47, 99, 100, and 153 in mice. Tox Sci 94(1):28–37. doi:10.1093/toxsci/kfl091

Sumner P et al (2014) The association between exaggeration in health related science news and academic press releases: retrospective observational study. Br Med J. doi:10.1136/bmj.g7015

Swenberg et al (1992) Species-specific mechanisms of carcinogenesis in: mechanisms of carcinogenesis in risk identification. Edited by H. Vainio. Lyon, Internat. Agency for Research on Cancer, pp 477–500

Teuten EL et al (2005) Two abundant bioaccumulated halogenated compounds are natural products. Science 307:917–920

Trasande L et al (2015) Estimating burden and disease costs of exposure to endocrine-disrupting chemicals in the European Union. J Clin Endocrinol Metab 100(4):1245–1255. doi:10.1210/jc.2014-4324

Trasande L et al (2016) Burden of disease and costs of exposure to endocrine disrupting chemicals in the European Union: an updated analysis. Andrology 4(4):565–572. doi:10.1111/andr.12178

Trasande L, Liu Y (2011) Reducing the staggering costs of environmental disease In children, estimated at $76.6 billion In 2008. Health Aff 30:863–870

US EPA FIFRA Scientific Advisory Panel (2016) Minutes No. 2016-01: A set of scientific issues being considered by the Environmental Protection Agency regarding chlorpyrifos: analysis of biomonitoring data April 19–21, 2016

US EPA (2011) Endocrine disruption screening program weight-of-evidence: evaluating results of EDSP tier 1 screening to identify the need for tier 2 testing. OPPT-2010-0877-0021

US FDA (2014) 2014 Updated safety assessment of Bisphenol A (BPA) for use in food contact applications. Memorandum dated June 17, 2014. https://www.fda.gov/downloads/NewsEvents/PublicHealthFocus/UCM424266.pdf

US EPA (2015a) Endocrine Disruption Screening Program (EDSP) Tier 1 weight of the evidence evaluation of the potential interaction of chlorpyrifos. https://www.epa.gov/sites/production/files/2015-06/documents/chlorpyrifos-059101_2015-06-29_txr0057162.pdf. Accessed on 2 Nov 2016

US EPA (2015b) Endocrine disruptor screening program tier 1 screening determinations and associated data evaluation records. https://www.epa.gov/endocrine-disruption/endocrine-disruptor-screening-program-tier-1-screening-determinations-and. Accessed on 2 Nov 2016

US EPA (2016a) Federal advisory committees at US EPA. https://www.epa.gov/faca/about-federal-advisory-committee-act-faca-epa, Accessed on 2 Nov 2016

US EPA (2016b) The Frank R. Lautenberg Chemical Safety for the 21st Century Act. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/frank-r-lautenberg-chemical-safety-21st-century-act

US EPA (2017) Endocrine Disruptor Screening Program (EDSP) Overview. https://www.epa.gov/endocrine-disruption/endocrine-disruptor-screening-program-edsp-overview

Vandenberg L et al (2016) A Proposed framework for the systematic review and integrated assessment (SYRINA) of endocrine disrupting chemicals. Environ Health 15:74–93. doi:10.1186/s12940-016-0156-6

Vergano D (2015) Poll reveals rift between scientists, regular folks. National Geographic. 29 Jan 2015

WHO Regional Office for Europe (2015). Human biomonitoring: facts and figures. Copenhagen