Estimating Burden and Disease Costs of Exposure to Endocrine Disrupting Chemicals in the European Union

Leonardo Trasande,1-4 R. Thomas Zoeller,5 Ulla Hass,6 Andreas Kortenkamp,7 Philippe Grandjean,8,9 John Peterson Myers,10 Joseph DiGangi,11 Martine Bellanger,12 Russ Hauser,7 Juliette Legler,13 Niels E. Skakkebaek,14 Jerrold J Heindel15

1New York University (NYU) School of Medicine, New York, New York, USA
2NYU Wagner School of Public Service, New York, New York, USA
3NYU Steinhardt School of Culture, Education and Human Development, Department of Nutrition, Food & Public Health, New York, New York, USA
4NYU Global Institute of Public Health, New York University, New York, New York, USA
5University of Massachusetts, Amherst, Massachusetts, USA
6National Food Institute, Technical University of Denmark, Søborg, Denmark
7Brunel University, Institute of Environment, Health and Societies, Uxbridge, Middlesex, United Kingdom
8Department of Environmental Health, Harvard School of Public Health, Boston, MA
9University of Southern Denmark, Odense, Denmark
10Environmental Health Sciences, Charlottesville, Virginia, USA
11IPEN, Gothenburg, Sweden
12EHESP School of Public Health, Paris, France
13Department of Chemistry and Biology, Institute for Environmental Studies, VU University, Amsterdam, Netherlands
Department of Growth and Reproduction, Rigshospitalet, EDMaRC and University of Copenhagen, Copenhagen, Denmark

National Institute of Environmental Health Sciences, Division of Extramural Research and Training, Research Triangle Park, NC, USA

Corresponding Author: Leonardo Trasande, MD, MPP, Associate Professor, Department of Pediatrics, New York University School of Medicine, 227 East 30th Street Rm 109, New York, NY 10016, leonardo.trasande@nyumc.org, phone 646-501-2520, fax 646-754-9688

Short running title: Methods to Quantify EDC Disease Burden and Costs

Keywords: endocrine disrupting chemicals, disease burden, economic costs, obesity, neurodevelopment, male reproductive health

DISCLOSURE STATEMENT: The authors have nothing to disclose.

Abbreviations: Attributable Fraction (AF); bisphenol A (BPA); endocrine disrupting chemicals (EDCs); diethylstilbestrol (DES); dichlorodiphenyldichloroethylene (DDE); Grading of Recommendations Assessment, Development and Evaluation (GRADE); polybrominated diphenyl ethers (PBDE); polychlorinated biphenyls (PCBs); Registration, Evaluation, Authorization and Restriction of Chemicals (REACH); Strategic Approach to International Chemicals Management (SAICM)

Acknowledgements: Research reported in this publication was supported by the Endocrine Society, the John Merck Fund, the Broad Reach Foundation, and the Oak Foundation. The funders and supporters had no role in the writing of the manuscript or the decision to submit it for publication. The content is solely the responsibility of the authors and does not necessarily
represent the official views of the NIEHS, the National Institutes of Health or the US government. We thank Charles Persoz, Robert Barouki and Marion Le Gal of the French National Alliance for Life Sciences and Health, and Barbara Demeneix, Lindsey Marshall, Bilal Mughal and Bolaji Seffou of the UMR 7221 Paris for providing technical and logistical support throughout the project. We also wish to thank Roberto Bertollini, Annette Pruss-Ustun and David Tordrup of the World Health Organization for their consultation and support in developing the guidelines for evaluating epidemiologic evidence. The contribution of JJH was supported by the NIEHS Division of Extramural Research and Training.

Word Count 5257
Abstract

Context: Rapidly increasing evidence has documented that endocrine disrupting chemicals (EDCs) contribute substantially to disease and disability.

Objective: To quantify a range of health and economic costs that can be reasonably attributed to EDC exposures in the European Union.

Design: A Steering Committee of scientists adapted the Intergovernmental Panel on Climate Change weight-of-evidence characterization for probability of causation based upon levels of available epidemiologic and toxicologic evidence for one or more chemicals contributing to disease by an endocrine disruptor mechanism. To evaluate the epidemiologic evidence, the Steering Committee adapted the WHO Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group criteria, while the Steering Committee adapted definitions recently promulgated by the Danish Environmental Protection Agency for evaluating laboratory and animal evidence of endocrine disruption. Expert panels used the Delphi method to make decisions on the strength of the data.

Results: Expert panels achieved consensus for probable (>20%) EDC causation for IQ loss and associated intellectual disability; autism; attention deficit hyperactivity disorder; childhood obesity; adult obesity; adult diabetes; cryptorchidism; male infertility and mortality associated with reduced testosterone. Accounting for probability of causation, and using the midpoint of each range for probability of causation, Monte Carlo simulations produced a median cost of €157 billion (1.23% of EU Gross Domestic Product) annually across 1000 simulations. Notably, using the lowest end of the probability range for each relationship in the Monte Carlo simulations produced a median range of €119 - 109 billion that differed modestly from base case probability inputs.
Conclusions: EDC exposures in the EU are likely to contribute substantially to disease and
dysfunction across the life course with costs in the hundreds of billions per year. These estimates
represent only those EDCs with the highest probability of causation; a broader analysis would
have produced greater estimates of burden of disease and costs.
Introduction

The European Union defines an endocrine disrupting chemical (EDC) as an “exogenous substance that causes adverse health effects in an intact organism, or its progeny, secondary to changes in endocrine function” (1-3). EDCs are diverse in their chemical structure, but all known EDCs interfere with hormone action to cause adverse effects, resulting in increased incidence of disease/dysfunction (3). For example, the water contaminant perchlorate is an EDC because it directly inhibits thyroid hormone synthesis, restricting the availability of thyroid hormone in target tissues, thereby interfering with thyroid hormone action (e.g., (4)); while BPA is an EDC in part because it can act through the estrogen-related receptor-gamma to alter insulin production and release, thus contributing to the pathogenesis of insulin resistance and type 2 diabetes (5).

The past twenty years have produced a great deal of new information from experimental studies focused on molecular, cellular and animal experiments (6) as well as epidemiological studies demonstrating that a wide array of chemical structures — pharmaceuticals, personal care products, commercial chemicals and environmental pollutants – can interfere with hormone action. Among the chemicals known to be EDCs are diethylstilbestrol (DES) (7), polychlorinated biphenyls (PCBs), dioxins, perfluoroalkyl compounds, solvents, phthalates (8), bisphenol A (BPA) (9), dichlorodiphenyldichloroethylene (DDE) (10), organophosphate and organochlorine pesticides (11), and polybrominated diphenyl ethers (PBDE) (12, 13). These chemicals have been shown to interfere with a variety of endocrine pathways including estrogen (14), androgen (14), thyroid (15, 16), retinol (17), aryl hydrocarbon and peroxisome proliferator-activated receptor pathways (18). The chemicals are widely used in consumer products, electronics and agriculture and widespread human exposures occur. Many EDCs are food
contaminants (e.g., pesticides, BPA and phthalates), though inhalation and dermal absorption are known pathways for human exposure. Potential consequences of exposure to EDCs include infertility and male and female reproductive dysfunctions(19), prostate and breast cancer(20), birth defects(21), obesity(22, 23), diabetes, cardiopulmonary disease, neurobehavioral and learning dysfunctions and immune dysregulation (24). Laboratory data on these associations are supplemented by varying levels of epidemiologic evidence for each chemical-disease/dysfunction dyad. In part due to uncertainty of causation, no estimate of the health or economic burden of EDCs has been made. Systematic estimates of burden of disease attributable to EDC exposures could help inform decision-making that protects public health.

The European Union is taking the lead on regulating EDCs, through legislation such as REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) and regulations on pesticides and biocides(25). The outcome of these policy discussions will be crucial not only for consumer and public health protection in the EU, but will also set scientific and regulatory policy precedents for other national policies including those consistent with implementation of global agreements such as SAICM (the Strategic Approach to International Chemicals Management)(26). A critical element of the regulation of EDCs in EU policy will be the criteria by which test outcomes for EDCs are translated into regulatory action. These criteria will determine, based on the functional properties of each chemical and responses measured in appropriate test systems, whether it will be restricted, phased out, or allowed to enter or remain on the EU market. The EU Commission has requested that an impact assessment be conducted to assess the economic implications of the criteria under discussion(27). The impact assessment is focused on the economic impact to industry of regulating EDCs in Europe. Our goal here is to
estimate the health and economic benefit of regulating EDCs in Europe, as based on current

evidence.

We now describe the general methods used to attribute disease and disability to EDCs, to weigh
the probability of causation based upon the available evidence, and to translate attributable
disease burden into costs. During a two-day workshop in April 2014, five expert panels
identified conditions where the evidence is strongest for causation, and developed ranges for
fractions of disease burden that can be attributed for EDCs. While accompanying manuscripts
describe in greater detail the bases for their estimates of disease attribution and probability of
causation, we present here an overview of the methods they applied as well as approaches
applied to estimate disease burden and costs attributable to EDCs in the EU based upon those
data inputs.

Methods

General Approach

In 1981, the Institute of Medicine developed a general approach to assess the "fractional
contribution" of the environment to causation of illness in the U.S., which remains widely used
to this day and is depicted in Equations 1 & 2 (28).
Attributable disease burden = Disease rate \times AF \times Population size \quad \text{(Equation 1)}

Attributable Costs = \text{Disease rate} \times \text{AF} \times \text{Population size} \times \text{Cost per case} \quad \text{(Equation 2)}

where "Cost per case" refers to discounted lifetime expenditures attributable to a particular disease including direct costs of health care, costs of rehabilitation, and lost productivity; Disease rate and Population size refer, respectively, to either the incidence or prevalence of a disease and the size of the population at risk; and AF is the Attributable Fraction, which is defined by Smith et al. in the context of environmental health as "the percentage of a particular disease category that would be eliminated if environmental risk factors were reduced to their lowest feasible concentrations." The AF is a composite value and is the product of the prevalence of a risk factor multiplied by the relative risk of disease associated with that risk factor, and is estimated using the following equation:

\[
\text{AF} = \frac{\text{Prevalence}_{\text{exposure}} \times (\text{RR} - 1)}{1 + (\text{Prevalence}_{\text{exposure}} \times (\text{RR} - 1))},
\quad \text{(Equation 3)}
\]

Where RR is the relative risk of morbidity associated with the exposure. An alternative formulation of Equation 1 would presuppose an exposure-outcome relationship that would result in discrete calculations of the increment in disease or disability over and above a comparison, unexposed group, and is presented in Equation 4:
Disease burden = Incremental prevalence or incidence x Population (Equation 4) x Population size

Accounting for Uncertainty and Probability of Causation

In the past, certainty of causation, however defined, has been presumed a requirement prior to pursuing estimates of attributable disease burden or costs, when in reality causation is not simply binary. In his widely cited work about the criteria for causation, Sir Austin Bradford Hill acknowledged the reality that “[a]ll scientific work is incomplete – where it be observational or experimental,” noting that uncertainty “does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time.” (31) The Intergovernmental Panel on Climate Change (IPCC) has managed uncertainty by applying a weight-of-evidence characterization for probability of causation. (32) A steering committee of scientists overseeing the project (MB, JD, PG, JH, AK, PM, LT, RZ) adapted the IPCC approach to assessing probability of causation based upon the available epidemiologic and toxicologic evidence for one or a group of chemicals contributing to disease by an endocrine disruptor mechanism. The schema is presented in Table 1, and subsequent paragraphs delineate the approach to evaluating epidemiologic and toxicologic evidence.

To evaluate the epidemiologic evidence, the GRADE Working Group criteria (33, 34) were adapted as they were recently applied in evaluating indoor air quality criteria by the World Health Organization (35). As described in Table 2, the criteria utilize study designs as a primary basis for distinguishing strength of evidence, with factors specific to the studies (both...
individually and in the aggregate) such as potential bias, limitations, strength of dose-response relationships, residual confounding, consistency and analogy permitting upward and downward grading of the quality of evidence.

To evaluate the toxicologic evidence, the steering committee adapted criteria recently promulgated by the Danish Environmental Protection Agency for evaluating laboratory and animal evidence of endocrine disruption (36). The schema is presented in Table 3. Identification of an endocrine mechanism/mode of action and corroboration of toxicity in laboratory model studies was required to assess the toxicological evidence for the exposure-outcome association as Group 1 (Endocrine disruptor). Group 2A (Suspected endocrine disruptor) required either (1) the presence of endocrine disruptor mode of action without clear corroboration of the mode of action producing the expected adverse effects in laboratory or animal studies, or (2) the presence of the adverse effects in laboratory animal studies with a suspected endocrine mode of action. Exposure-outcome associations were evaluated to have group 2B (Potential endocrine disruptor) toxicological evidence when there was evidence of adverse effects in animal studies that could have either endocrine mode of action or a non-endocrine mode of action or in vitro/in silico evidence indicating a potential for endocrine disruption in intact organisms.

Quantifying Attributable Burden

The steering committee noted three general approaches on which to base attribution to EDCs: (1) trends in incidence/prevalence over and above a baseline that would be difficult to attribute to genetics accompanied by information on likely causal mechanisms by EDCs and/or increasing
exposure, (2) data from genetic studies that permit quantification of the remaining environmental
correlation (within which one might posit EDC to contribute a portion), and (3) dose-response
relationships from the epidemiologic literature. In general, the steering committee prioritized the
third approach. In the absence of epidemiologic evidence for a dose-response relationship, the
presence of toxicologic data documenting effect and mechanism and/or other data might suggest
a strong basis from which to reason an incremental effect in humans. In this scenario, the first
two lines of evidence would add support to an estimate of the degree that one or more EDCs
might contribute to the condition under consideration.

While chemicals banned by Europe (e.g., under the Stockholm Convention) have been
documented to be endocrine disruptors and contribute to disease and disability, panels were
advised not to examine effects of these exposures unless there was a compelling case that
interventions outside Europe could influence disease and disability in Europe. For example, the
obesity panel did not quantify the obesogenic and diabetogenic effects of other EDCs that
continue to contaminate the EU general population (e.g., polychlorinated biphenyls and
hexachlorobenzene) because they are banned under the Stockholm Convention (37, 38). In
contrast, DDE-attributable obesity and diabetes could be prevented through further reductions in
DDT use globally, which is substantially relevant due to the current use of this chemical for
malaria control and its long-range transport and persistence in the environment (39).

Panels were advised to consider all possible developmental windows of vulnerability, but to
focus on exposure timing and duration with the strongest evidence for causation from
toxicological and epidemiologic data. When a dose-response relationship was identified for a
particular exposure period, this relationship was applied to the EU population based upon biomarker data available from large surveys or pooled data from multiple studies in individual countries. Biomarkers were then estimated for quantiles (usually 0-9th, 10-24th, 25-49th, 50-74th, 75-89th, 90-99th) in recognition that narrower quantiles might reduce precision of estimates. In the rare circumstance that there were no epidemiologic studies on which to assess a dose-response relationship, but there existed enough evidence to suggest an effect in a portion of the appropriate population, a relative risk was estimated, and a prevalence of exposure was identified in order to estimate an attributable fraction, using Equation 3. Whenever possible, the most population-representative data were used for appropriate exposure and/or biomarker inputs, as convenience samples may have unmeasurable biases resulting in misestimation of exposure, and these inputs were applied consistently across all the exposure-outcome associations studied.

Approach to Evaluating Evidence

Following the WHO/UNEP State of the Science of Endocrine Disrupting Chemicals, which identified three distinct sets of health endpoints with the most substantial evidence for EDC attribution (obesity/diabetes, male reproductive health and neurodevelopmental disability) (24), the steering committee convened expert panels for each of the domains composed of four to eight scientific experts. Two expert panels were also convened for breast cancer and female reproductive conditions; their deliberations followed an identical process to that described below, are nearing completion, and will be the basis for future reports. The steering committee identified epidemiologic and toxicologic experts based upon their scholarly contribution in the diseases under consideration and endocrine disruptor toxicology, and invited them to attend a
two-day scientific meeting in Paris, which was held at the French National Alliance for Life Sciences and Health from April 28-29, 2014.

During this meeting, the steering committee applied a modified Delphi approach (40) to evaluating the strength of the epidemiologic and toxicological evidence, and the nature of the association between exposures and outcomes. The Delphi method was developed on the premise that group judgments are more valid than those of individuals. While named after the oracle at the sanctuary dedicated to Apollo in the 5th century BC, the method is not mystical and was first developed at the beginning of the Cold War to forecast technological impacts on warfare (41). Helmer, Dalkey and Rescher at the RAND Corporation formalized the method in the 1950s for science and technology forecasting (42). It has been applied successfully and with high consistency and rigor across many disciplines including health and education (43-46).

Teleconferences were held biweekly over a three-month period with participants to encourage familiarity with literature being reviewed, to describe the Delphi method (including definition of terminology and interaction structure), (41) and to identify group leaders (PG, RH, JL). An initial presentation at the beginning of the two-day meeting provided an overview of the process, and further clarified the definition of EDC to be used. The Endocrine Society defines EDCs somewhat differently than the European Union, as an exogenous chemical, or mixture of chemicals, that can interfere with any aspect of hormone action (3). Because of the EU decision making context, panelists were advised to adhere to the EU definition, but to add a further requirement that the chemicals interfere with hormone action (as elaborated in the Endocrine Society definition).
Panels began by selecting the association for which the evidence was judged to be the strongest to promote familiarity in subsequent iterations. For each exposure-outcome association, the process in each group began with presentation of epidemiologic and toxicologic reviews of the literature, and discussing the approach to identifying the overarching issues in attributing individual EDC exposures to the subject outcome. Expert panelists were then asked to provide their opinion about strength of the epidemiologic and toxicologic evidence for the exposure-outcome relationship, and the nature of that relationship. Responses were submitted to the leader anonymously.

Each leader then provided a summary of the findings from initial questionnaires and reasons for the judgments. Panelists were encouraged to refine their answers anonymously in light of replies of other experts on the panel, with a goal of convergence towards a consensus in subsequent rounds of questionnaires. Panelists were advised to consider the Smith et al definition of AF, i.e., "the percentage of a particular disease category that would be eliminated if environmental risk factors were reduced to their lowest feasible concentrations."(29) Recognizing that naturally occurring EDCs in the environment such as phytoestrogens do exist, the steering committee encouraged estimation of AFs attributable to anthropogenic activities, recognizing that naturally occurring exposures (e.g., phytoestrogen exposure from soy milk) may also contribute.(47) Panelists were asked to focus on EU populations, identifying the population affected (including age and demographic subgroups) as part of their iterative process, in addition to the population in which the outcome is being assessed. They were asked to consider the reality of mixtures and complexity of attribution in that context.
Management of ongoing discussions and trigger of subsequent rounds of questionnaires were determined by the expert panel leads. Consistent with application of the Delphi method to aspects of medical care, (45, 46, 48) pre-defined stop criteria included: a minimum of three questionnaire rounds, achievement of majority consensus, and stability of results across rounds. Converging answers for each EDC-outcome relationship formed the basis for manuscripts accompanying this overview, which describe each expert panel process and were prepared by the expert panel leads in collaboration with the other members after the meeting. 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 presence or absence of threshold effects when appropriate. Non-monotonicity did not influence strength of evidence when supported in its biological plausibility, though it could yield differences in the estimated disease burden. While unanimity was encouraged, in the event of non-unanimity, the range of strength of evidence evaluations from all participants was input to develop a range of results for probability of causation.

General Approach to Economic Estimation

We applied a human capital approach (49, 50), which is currently the most widely used method to calculate the costs of illness (51, 52). This approach measures the value of resources foregone and output lost due to illness, such as lost earnings or household contributions as a homemaker, and costs of medical treatment. With this method, costs were divided into direct and indirect costs. In
calculating these costs, we followed the widely cited costing guidelines recommended by the Panel on Cost Effectiveness and Medicine (53). Direct costs are the value of resources that could be allocated to other uses in the absence of disease. These include expenditures for hospitalization, physician services, nursing home care, medical appliances, and related items. Indirect costs are the value of the lost output of workers and retirees suffering premature death or disability. We assumed the societal perspective, as opposed to the perspective of the health care payer, (54) and our measures of costs adhered as closely as possible to the economic definition of costs, where cost is represented by foregone opportunities.

Whenever possible, we utilized European data sources for cost-of-illness inputs, and relied upon already published estimates when available. Our preference was to identify incremental costs associated with a condition, rather than average costs, as these tend to produce overestimates (55). When European data were not available, we extrapolated from available US estimates, applying a correction factor representing the ratio of the per capita gross domestic product purchasing power parity of each European country compared to the United States. All results are presented in 2010 Euros, and represent costs as estimated to occur in 2010, the most recent year for which prevalence/incidence data could permit robust estimation.

Finally, recognizing that attributable cost estimates were accompanied by a probability, we performed a series of Monte Carlo simulations to produce ranges of probable costs across all the exposure-outcome relationships, assuming independence of each probabilistic event. Separate random number generation events were used to assign (1) causation or not causation, and (2) cost
given causation, using the base case estimate as well as the range of sensitivity analytic inputs produced by the expert panel. To illustrate with an example, for an exposure-outcome relationship with an 80% probability of causation, random values between zero and one in each simulation led to the first step, which either assigned no costs (random value ≤ 0.2) and costs (random value > 0.2). For relationships in which lower and/or higher bound estimates of costs were identified in addition to base case costs, a second random number generation was used to assign costs in the scenario of causation. For those relationships with a lower and outer bound estimate, equal probabilities were assigned to values below and above the base case estimate, with costs linearly interpolated across the remaining probability range. For relationships for which only a higher or lower bound estimate was available, a 50% probability was assigned for the base case estimate, while the remaining 50% probability was applied over the range of the higher/lower bound.

Recognizing that probability of causation could be highly influential on cost estimates, we performed three sets of 1000 simulations, using the midpoints of the ranges for probability of causation for each exposure-outcome relationship as a base case scenario, and low and high bounds of the probability range as alternate scenarios, to assess the sensitivity of Monte Carlo simulations to this input. For each of the three sets of simulations, we produced ranges of burden and disease costs associated with EDCs. We developed a 95% confidence interval as well as the interquartile range and first and ninth deciles to convey the spread of possible scenarios.

**Results**
Expert panels achieved consensus for probable (>20%) EDC causation for IQ loss and associated intellectual disability; autism; attention deficit hyperactivity disorder; childhood obesity; adult obesity; adult diabetes; cryptorchidism; male infertility and mortality associated with reduced testosterone (Table 4). Only for testicular cancer was 0-19% probability of causation identified. We refer the reader to accompanying manuscripts which describe specific results from each of the expert panels (56-58), but to illustrate we present burden of disease results from a few examples here.

The neurodevelopment panel estimated a strong probability (70-100%) that, each year in Europe, 13.0 million IQ points are lost (sensitivity analysis: 4.2-17.1 million) due to prenatal organophosphate exposure, and 59,300 additional cases of intellectual disability (sensitivity analysis: 16,500-84,400). With more modest probabilities, 316 cases of autism and 19,400-31,200 new cases of ADHD annually are attributable to EDCs (sensitivity analysis: 126-631).

The male reproductive panel identified male infertility attributable to phthalate exposure to have a 40-69% probability of causing 618,000 additional assisted reproductive technology procedures annually in Europe. A 40-69% probability of lower testosterone concentrations in 55-64 year old men due to phthalate exposure was identified, with 24,800 associated deaths annually. The obesity/diabetes panel identified a 40-69% probability of phthalate exposure causing 53,900 cases of obesity and 20,500 new-onset cases of diabetes in older women annually. Prenatal BPA exposure was identified to have a 20-69% probability of causing 42,400 new cases of childhood obesity annually, with associated lifetime costs of €1.54 billion.
The most substantial costs were related to loss of IQ and intellectual disability attributable to prenatal organophosphate exposure; base case estimates identified €146 billion in attributable costs, while sensitivity analyses suggested that costs might actually range from €46.8-195 billion annually. Phthalate attributable adult obesity was the second largest driver of costs, at €15.6 billion per year. The total costs of all conditions probably attributable to EDCs were €191 billion, with sensitivity analyses suggesting costs ranging from €81.8-269 billion annually.

Accounting for probability of causation, the base case Monte Carlo simulation using the midpoint of each range for probability of causation produced costs between €3.3-2.5-244 billion annually across the 1000 simulations (median, €157 billion; Figure 1). Using the 2010 EU purchasing-power-parity corrected Gross Domestic Product (GDP) estimate of €127.9 billion(59), the estimated costs comprise 1.23% of GDP. There is a 5% probability that costs of EDC exposures are less than €20-21 billion annually, a 90% probability that costs are at least €32.4-0 billion, a 75% probability that costs are at least €9665.4-0 billion/year, a 25% probability of costs at least €194 billion/year, and a 10% probability of costs over €244-212 billion/year.

Notably, using the lowest end of the probability range for each relationship in the Monte Carlo simulations produced a range of €24.6-11.0 million-236-235 billion (median, €149-109 billion) that differed modestly from the base case probability inputs. There is a 5% probability that costs of EDC exposures are less than €810.8-0 billion annually, a 90% probability that costs are at least €15.6-8 billion, a 75% probability that costs are at least €3430.9-8 billion/year, a 25% probability of costs at least €129-181 billion/year, and a 10% probability of costs over €202-204 billion/year. Applying the lowest end of the probability range and assuming all the relationships
are independent, multiplying each of the probabilities for the exposure-outcome relationships
suggests a very high (>99% =1-0.3x0.3x0.6x0.8x0.6x0.8x0.6x0.6x0.6x0.6x0.6x0.8x0.8)
probability that EDCs contribute to disease in Europe. Leaving aside the highly probable costs of
developmental neurotoxicity from organophosphate pesticide and brominated flame retardants,
there is still a substantial probability (>98%) that one or more of the other exposure-outcome
relationships are causal. Using the highest end of the probability ranges narrowed the range of
costs more substantially (€621.7-246 billion; median €176-180 billion). There was a
421.80% probability of costs under €100 billion, and a 2831.45% probability of costs over €200
billion.

Discussion

The primary finding of this manuscript is that there is a substantial probability of very high
disease costs across the lifespan associated with EDC exposure in the European Union. For
some perspective, the median €157 billion cost/year we identified is approximately one sixth the
€798 billion European cost of brain disorders in 2010 (60), and 1.23% of GDP. These costs will
accrue annually insofar as exposures that are harmful continue unabated. Thus, regulatory action
to limit exposure to the most widely prevalent and potentially hazardous EDCs is likely to
produce substantial economic benefits. These economic benefits should inform decision-making
on measures to protect public health.

Calculations of the health and economic benefits associated with reducing exposure to
environmental chemicals have proven extremely informative to regulatory decision-making.
Estimates of the benefits associated with removal of lead-based paint hazards informed funding of federal lead hazard control grants in the early 2000s (61), and measurement of the benefits associated with reduced prenatal methylmercury toxicity (62) informed formulation of the global mercury treaty. Though analyses like these are highly valuable, they have been typically limited to associations where causation is certain. Decades of epidemiologic data typically are required before possible causation has been acknowledged and attributable disease burden calculated (63, 64). Failure of the current approach in assessing the economic costs of environmental health hazards is especially acute for EDCs, for which longitudinal studies of early life exposures are only beginning to be completed. The approach we have taken will potentially transform decision making in environmental health, by providing a new model for evaluating environmental health risks and permitting a complete assessment of the potential costs of failing to prevent chronic disease through use of safer alternatives to EDCs. It produces substantial insights regarding strength of the epidemiologic and toxicologic data, placing them alongside the cost of the disease as never done before. This approach also documents data gaps in both the epidemiology and toxicology of EDCs, which has only been documented through systematic reviews.

We used an expert elicitation approach to estimating the probability that EDCs contribute to disease and disability. While the Global Burden of Disease project does rely on expert opinion, it has focused on a small subset of exposure-outcome relationships with the strongest causation. In preparation for this work, we considered the International Agency for Research and Cancer (65) and World Cancer Research Fund grading systems (66), but these approaches could not be readily adapted to account for contribution by an endocrine disruptor mechanism for this project.
Expert opinion is of course not a substitute for solid epidemiologic evidence regarding the relationships between EDCs and disease, or for systematic toxicological documentation regarding endocrine disruption as the mechanism by which EDCs act to promote disease. Yet, uncertainty is a reality across aspects of decision-making in science and public policy, and we relied upon widely accepted and used methods for accounting for uncertainty. In the course of a two-day workshop and associated conference calls, the panels could not be comprehensive in their examination of the panorama of EDCs and potential effects. While each accompanying manuscript endeavors to call attention to the limited scope of the chemicals and outcomes assessed, it bears emphasis that the present work focused only on the conditions and exposures with the strongest evidence for causation, within the three disease areas for which the steering committee judged the investment in assembling an expert panel to be appropriate.

In addition to producing ranges of probability of causation based upon strength of evidence, we also endeavored to incorporate the substantial uncertainty in EDC-disease relationships using sensitivity analyses to model impacts of key uncertainties on estimates of burden of disease and costs that produced a wide range of potential costs associated with EDCs. The estimates presented in this report are uncertain, and the range of likely costs has been expressed as allowed by the evidence available. Clearly, more research would allow calculation of better estimates, but would take time and substantial investment. Given the current concerns about regulation of endocrine disruptors, the present report aims at providing the best possible documentation for possible decision-making at this time. Though the analysis was limited to the EU, if similar exposures and effects are identified in the US and other areas of the world, then the burden of disease and costs attributable to EDCs elsewhere is likely to be on the same order of magnitude.
Additional investment across the world in research to identify how and which EDCs are harmful is also indicated.

Three additional issues should be considered when evaluating our findings. First, the approach we took to quantifying the probability of costs fails to account for risk aversion. Generally, societies value small probabilities of costs (e.g., 10% of $1,000) more than the weighted average ($100 = (10% x $1,000) + (90% x $0)). A major driver for health insurance is that people may value investment on behalf of preventing even a rare but uncertain outcome more than the weighted-average likelihood of the consequences of the outcome. Because people generally prefer to pay more in such a scenario, societies are described as risk-averse (67). We did not account for risk aversion in the present work. Indeed, the societal value of the uncertain health effects analyzed here may be much higher than our calculations. Second, cost-of-illness approaches fail to capture the complete scope of economic costs associated with illness (especially psychological and other indirect or intangible costs that are difficult to assess), thus our cost-of-illness estimate of EDCs must be considered an underestimate (68-71). Finally, when considering the costs of safer alternatives, it is important to keep in mind that estimates of the cost of safer alternatives produced by those who create environmental toxicants may overestimate costs of prevention because they do not account fully for ongoing technological innovation that may reduce future costs of safer alternatives (72). The costs of such innovations are often one-time costs, whereas the benefits of prevention accumulate over time, as has been documented with the annual economic benefit of eradicating lead from gasoline (73).
Finally, the findings described here suggest potentially large burdens of disease and associated costs in the developed world, insofar as exposures are similar. Future studies could extend and apply this approach to the United States, where the National Health and Nutrition Examination Survey among other studies offer arguably more comprehensive and national reference points for extrapolation. In the industrializing world, the attributable disease burden and costs could well be higher in a much weaker regulatory framework (74). A major challenge to documenting the scope of EDC-attributable disease in these more vulnerable populations is the absence of biomarker or other exposure data to support similar estimates. The World Health Organization and United Nations Environment Programme can catalyze and coordinate such efforts, which will require substantial resources for its proper execution.
Table 1. Framework for Evaluating Probability of Causation.

<table>
<thead>
<tr>
<th>Epidemiologic Evaluation</th>
<th>Toxicologic Evaluation</th>
<th>Strong (Group 1)</th>
<th>Moderate (Group 2A)</th>
<th>Weak (Group 2B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Very High (90-100%)</td>
<td>High (70-89%)</td>
<td>Medium (40-69%)</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>High (70-89%)</td>
<td>Medium (40-69%)</td>
<td>Low (20-39%)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>Medium (40-69%)</td>
<td>Low (20-39%)</td>
<td>Very Low (0-19%)</td>
<td></td>
</tr>
<tr>
<td>Very Low</td>
<td>Low (20-39%)</td>
<td>Very Low (0-19%)</td>
<td>Very Low (0-19%)</td>
<td></td>
</tr>
</tbody>
</table>

Adapted from (32).
Table 2. Criteria for Evaluating Epidemiologic Evidence.

<table>
<thead>
<tr>
<th>Quality of evidence</th>
<th>Interpretation</th>
<th>Study design</th>
<th>Lower the quality in presence of</th>
<th>Raise the quality in presence of</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>We are very confident that the true effect lies close to that of the estimate of the effect.</td>
<td>Randomized trial</td>
<td>Study limitations:</td>
<td>Strong association:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1 Serious limitations</td>
<td>+1 Strong, no plausible confounders, consistent and direct evidence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-2 Very serious limitations</td>
<td>+2 Very strong, no major threats to validity and direct evidence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1 Important inconsistency</td>
<td>+1 Evidence of a dose-response gradient</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Directness:</td>
<td>+1 All plausible confounders would have reduced effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1 Some uncertainty</td>
<td>Additional criteria (applied across a body of evidence based on multiple study designs):</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-2 Major uncertainty</td>
<td>+1 Consistency across multiple studies in different settings</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1 Imprecise data</td>
<td>+1 Analogy across other exposure sources</td>
</tr>
<tr>
<td>Moderate</td>
<td>We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.</td>
<td>Quasi-experimental (with controls) and before and after (uncontrolled) studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect</td>
<td>Observational study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very low</td>
<td>We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect</td>
<td>Any other evidence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adapted from (33, 75).
### Table 3. Criteria for Evaluating Toxicologic Evidence.

<table>
<thead>
<tr>
<th>Quality of evidence</th>
<th>Interpretation</th>
<th>Study design</th>
</tr>
</thead>
</table>
| Strong, Group 1 (Endocrine disruptor) | There is a strong presumption that the chemical has the capacity to cause the health effect through an endocrine disruptor mechanism. | The animal studies provide clear evidence of the ED effect in the absence of other toxic effects, or if occurring together with other toxic effects, the ED effects should not be a secondary non-specific consequence of other toxic effects. However, when there is e.g. mechanistic information that raises doubt about the relevance of the effect for humans or the environment, Group 2 may be more appropriate. Substances can be allocated to this group based on:  
• Adverse \textit{in vivo} effects where an ED mode of action is plausible  
• ED mode of action \textit{in vivo} that is clearly linked to adverse \textit{in vivo} effects (by e.g. read-across) |
| Moderate, Group 2a (Suspected endocrine disruptor) | There is some evidence from experimental animals, yet the evidence is not sufficiently convincing to place the substance in Group 1. | The health effects are observed in the absence of other toxic effects, or if occurring together with other toxic effects, the ED effect should be considered not to be a secondary non-specific consequence of other toxic effects. Substances can be allocated to this group based on:  
• Adverse effects \textit{in vivo} where an ED mode of action is suspected  
• ED mode of action \textit{in vivo} that is suspected to be linked to adverse effects in vivo  
• ED mode of action \textit{in vitro} combined with toxicokinetic in vivo data (and relevant non test information such as read across, chemical categorisation and QSAR predictions) |
| Weak, Group 2b (Potential endocrine disruptor) | There is some evidence indicating potential for endocrine disruption in intact organisms. | There is some \textit{in vitro}/\textit{in silico} evidence indicating a potential for endocrine disruption in intact organisms or effects in vivo that may, or may not, be ED-mediated. |

Adapted from (36).
<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Strength of Human Evidence</th>
<th>Strength of Toxicologic Evidence</th>
<th>Probability of Causation</th>
<th>Base estimate</th>
<th>Low estimate</th>
<th>High estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polybrominated diphenyl ethers (PBDE)</td>
<td>IQ Loss and Intellectual Disability</td>
<td>Moderate-to-high</td>
<td>Strong</td>
<td>70-100%</td>
<td>€ 9,587,571,420</td>
<td>€ 1,577,449,522</td>
<td>€ 22,356,864,892</td>
</tr>
<tr>
<td>Organophosphate pesticides</td>
<td>IQ Loss and Intellectual Disability</td>
<td>Moderate-to-high</td>
<td>Strong</td>
<td>70-100%</td>
<td>€ 146,178,556,566</td>
<td>€ 46,760,988,423</td>
<td>€ 194,850,545,761</td>
</tr>
<tr>
<td>Dichlorodiphenytrichloroethylene (DDE)</td>
<td>Childhood obesity</td>
<td>Moderate</td>
<td>Moderate</td>
<td>40-69%</td>
<td>€ 24,610,041</td>
<td>€ 24,610,041</td>
<td>€ 86,448,264</td>
</tr>
<tr>
<td>Dichlorodiphenytrichloroethylene (DDE)</td>
<td>Adult diabetes</td>
<td>Low</td>
<td>Moderate</td>
<td>20-39%</td>
<td>€ 834,741,170</td>
<td>€ 834,741,170</td>
<td>€ 16,694,823,393</td>
</tr>
<tr>
<td>Di-2-ethylhexylphthalate (DEHP)</td>
<td>Adult obesity</td>
<td>Low</td>
<td>Strong</td>
<td>40-69%</td>
<td>€ 15,610,612,091</td>
<td>€ 15,610,612,091</td>
<td>€ 15,610,612,091</td>
</tr>
<tr>
<td>Bisphenol A</td>
<td>Childhood obesity</td>
<td>Very low-to-low</td>
<td>Strong</td>
<td>20-69%</td>
<td>€ 1,537,177,463</td>
<td>€ 1,537,177,463</td>
<td>€ 1,537,177,463</td>
</tr>
<tr>
<td>Polybrominated diphenyl ethers (PBDE)</td>
<td>Testicular cancer</td>
<td>Very low-to-low</td>
<td>Weak</td>
<td>0-19%</td>
<td>€ 847,975,932</td>
<td>€ 313,179,835</td>
<td>€ 847,975,932</td>
</tr>
<tr>
<td>Polybrominated diphenyl ethers (PBDE)</td>
<td>Cryptorchidism</td>
<td>Low</td>
<td>Strong</td>
<td>40-69%</td>
<td>€ 129,807,327</td>
<td>€ 116,841,584</td>
<td>€ 129,807,327</td>
</tr>
<tr>
<td>Benzyl and butylphthalates</td>
<td>Male Infertility, Resulting in Increased Assisted Reproductive Technology</td>
<td>Low</td>
<td>Strong</td>
<td>40-69%</td>
<td>€ 4,714,114,146</td>
<td>€ 4,714,114,146</td>
<td>€ 4,714,114,146</td>
</tr>
<tr>
<td>Phthalates</td>
<td>Low testosterone, Resulting in Increased Early Mortality</td>
<td>Low</td>
<td>Strong</td>
<td>40-69%</td>
<td>€ 7,958,358,238</td>
<td>€ 7,958,358,238</td>
<td>€ 7,958,358,238</td>
</tr>
<tr>
<td>Multiple exposures</td>
<td>ADHD</td>
<td>Low-to-moderate</td>
<td>Strong</td>
<td>20-69%</td>
<td>€ 1,743,332,686</td>
<td>€ 1,212,298,027</td>
<td>€ 2,861,405,410</td>
</tr>
<tr>
<td>Multiple exposures</td>
<td>Autism</td>
<td>Low</td>
<td>Moderate</td>
<td>20-39%</td>
<td>€ 199,339,876</td>
<td>€ 79,735,951</td>
<td>€ 398,679,753</td>
</tr>
</tbody>
</table>
The numbers on the X-axis denote cumulative probability across the 1000 simulations for base case probability of causation, as well as low and high bounds for probability of causation.
References

17. Colborn T 2003 Neurodevelopment and Endocrine Disruption. Environ Health Perspect 112
32. Intergovernmental Panel on Climate Change 2005 Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties. Available at


41. Juri P 1971 The Delphi method: Substance, context, a critique and an annotated bibliography. Socio-Economic Planning Sciences 5:57-71


46. Rao J, Anderson L, Sukumar B, Beauchesne D, Stein T, Frankel R 2010 Engaging communication experts in a Delphi process to identify patient behaviors that could enhance communication in medical encounters. BMC Health Services Research 10:97
48. Davies S, McDonald KM, Schmidt E, Schultz E, Geppert J, Romano PS 2011 Expanding the Uses of AHRQ's Prevention Quality Indicators: Validity From the Clinician Perspective. Medical Care 49:679-685 610.1097/MLR.1090b1013e3182159e3182165


70. Ried W 1996 Willingness to pay and cost of illness for changes in health capital depreciation. Health Economics 5:447-468


