Harvard Center for Risk Analysis

Using Decision Sciences to Promote Public Health

AUGUST 2004

Volume 12
Issue 3

Risk in Perspective

WEIGHT OF THE EVIDENCE EVALUATION OF LOW-DOSE REPRODUCTIVE AND DEVELOPMENTAL EFFECTS OF BISPHENOL A

GEORGE GRAY, PH.D.

JOSHUA COHEN, PH.D.

"...the evidence considered by the panel suggests that the weight of the evidence for low-dose effects is very weak."

Introduction

Bisphenol A (BPA) is a plastic monomer that is the starting material for the manufacture of polycarbonate plastics used in plastic bottles, jugs, baby bottles and many other products. It also lines the inside of food and beverage cans. As a result, exposure to BPA in the general population is widespread. Assessments based on evaluations by the U.S. Environmental Protection Agency (EPA) indicate that these exposures are well below the levels generally considered to be without harm. However, some scientists believe that BPA can act as an artificial estrogen, causing adverse reproductive and developmental toxicity at levels far lower than previously believed. Some laboratory animal studies support this hypothesis, while others do not.

A panel convened by the Harvard Center for Risk Analysis (HCRA) evaluated the weight of evidence in studies of BPA published as of April, 2002. Focusing on potential male reproductive impacts but also considering other end points associated with suspected hormone-like effects, the panel found no consistent affirmative evidence of low-dose BPA effects for any endpoint. It found that inconsistent responses by various rodent test species also raised doubts as to the generalizability of results to humans. And differences in the pattern of responses to BPA compared to known estrogenic compounds cast doubt on estrogenicity as a low-dose mechanism of action for BPA. Not only does this systematic evaluation provide regulatory decision makers with useful information in the case of BPA, but it also offers an example of how such issues can be addressed when the scientific data remain less than conclusive, a situation that is common.

The panel was chaired by Donald Mattison, former medical director of the March of Dimes and currently Senior Advisor to the Directors of the Center for Research for Mothers and Children at the National Institute of Child Health and Human Development. The work was funded by the American Plastics Council. This Risk in Perspective summarizes the full report, which has been peer-reviewed and appears in the journal *Human and Ecological Risk Assessment.*

For more information on HCRA visit our website at: www.hcra.harvard.edu

Human Exposure and the Suspected Hazard

Because BPA is used in the production of plastics, including food containers, there is widespread exposure to this substance. In addition, there may be local environmental exposure to BPA in the vicinity of polyvinylchloride manufacturing facilities or other industrial sites that use BPA in their production processes. BPA is also a component in some dental sealants and papers, although these products have been found to make little or no contribution to exposure.

European Union scientific bodies have estimated exposure from food and wine ranging from 0.0005 to 0.009 mg/kg-day, well below the reference dose (RfD) of 0.05 mg/kg-day identified by U.S. EPA as the average level of daily exposure expected to be free from harm. EPA identified their RfD from experiments using lab rats, where the lowest level of exposure associated with adverse effects was 50 mg/kg-day. The RfD was calculated by dividing this 50 mg/kg-day lowest observed adverse effect level (LOAEL) by a safety factor of 1,000.

Concerns have been raised that the identification of a LOAEL at 50 mg/kg-day might not be the end of the story and hence that EPA's RfD may not be sufficiently protective. In particular, some believe that BPA might be an endocrine modulating compound, specifically, a low-dose environmental estrogen. Experiments in animals and *in vitro* experiments using isolated tissues, cells, or cellular components from a number of species, including humans, have suggested that BPA exhibits estrogen-like activity, although it is considerably less potent than natural estrogen.

Recent investigations of live laboratory animals have suggested a broad range of effects from BPA exposure, hypothesized to be due to endocrine action. Studies from one group of investigators have received the most attention (see, for example, vom Saal *et al.,* 1998). These studies report changes in the weight of male reproductive organs (the prostate and epididymis) and changes in sperm production in offspring whose mothers had exposures as low as 0.002 to 0.02 mg/kg-day, levels that are comparable to estimated human exposures, far less than the LOAEL of 50 mg/kg-day in rats, and even less than EPA's RfD of 0.05 mg/kg-day.

The reporting of positive findings (or negative findings) from such studies, however, represents only the beginning of the scientific debate. The HCRA panel was convened because there are a multitude of factors that can make it unclear whether such a relationship exists. Subtle problems with study design can make it appear that BPA is causing an effect even if it is not. Other factors can disguise the existence of such a relationship. For example, errors in making fine measurements of organ weight can obscure an impact on this quantity. Finally, even if effects are verified in laboratory animals, it is not clear if they can be extrapolated from animals to humans or, when necessary, from the high doses sometimes used in studies to the relatively low levels of exposure typically experienced by people. The panel's task was to look at the scientific literature as a whole in an effort to sort out the puzzle in the case of BPA.

Methods of Assessing the Evidence

To do so, the Harvard panel reviewed the 19 laboratory animal studies published through April, 2002 that investigated any health effects in laboratory animals following low dose exposure to BPA (i.e., no more than 5 mg/kg-day, which is a factor of 10 below the LOAEL). Although some of the studies reported that BPA exposure was associated with health effects in the laboratory animals, many others did not. Comparing different studies was complicated by differences in the type of animals used, differences in the health effects investigated, as well as other differences in methodology. Even the type of feed administered to the animals in these experiments could have affected the results.

The panel chose to use a framework developed by an expert panel, and described in Gray et al. (2001), to guide their assessment of the weight of evidence for low-dose effects of BPA and the potential relevance given common exposures.

In order to evaluate whether the effects observed were really occurring in laboratory animals, the panel placed weight on findings based on the following criteria:

- Corroboration – Replication of findings among similar studies and the observation of similar effects under relevant conditions increases the confidence that the findings represent a real effect in experimental animals. Lack of corroboration is grounds to doubt the validity of a single finding. In multi-generational studies, corroboration is supported if the effect appears across generations, and is challenged if it does not.
- Rigor – Studies must be evaluated for their conduct and analysis. Greater weight is given to better-conducted studies and those that follow the codes of good laboratory practice (GLP).

- Power – The statistical power of an experimental design should be examined for its ability to detect effects of a certain magnitude, especially for "negative" studies, where a low level of response could be mistaken for lack of response.

As to the question of whether positive findings in test animals can be generalized across species to humans, and from high experimental doses to relatively low environmentally relevant exposures for humans, the criteria are:

- Universality – The degree to which an effect is consistently reproduced in multiple valid test species increases confidence that the result might apply to humans. If, however, the result shows up in only one species or one strain, or one route of administration (oral, subcutaneous injection, respiratory, etc.), and not in others, generalizability is challenged.
- Proximity – Effects in species closer to humans, at doses and by routes of administration similar to human exposures, weigh more heavily in favor of generalizability.

Finally, a key component to the evaluation of any scientific hypothesis is its plausibility. In the case of BPA, the question is whether this substance shares characteristics with other estrogenic substances, and whether its molecular mechanisms would operate in humans. The formal criteria for evaluating biological plausibility were:

- Cohesion – The extent to which all the data are consistent and are subject to a single, biologically plausible explanation increases weight compared to a situation where inconsistencies require *ad hoc* explanations and exceptions to general patterns.
- Relevance – From what is known about the underlying biological basis for a toxic response in test animals and the biology of humans, it may be possible to judge whether similar metabolism, mechanisms of damage and their repair, and molecular targets of toxic action should be expected to operate in humans as they do in the test animals.

Findings

In general, the findings in the literature the panel examined failed to meet one or more of the above criteria, which is why the panel found that there is no consistent affirmative evidence of low-dose BPA effects for any endpoint.

As noted earlier, the potential association between BPA and prostate weight has been regarded as particularly important. Some of the studies of this effect have not found any association with BPA exposure. The panel found a number of factors that may have masked this association, including:

- Differences in the way the studies weighed the prostates.
- Inadequate statistical power, although the studies that did not find an association were larger and hence more statistically powerful than the studies that did find an association.
- Differences in the animal strains used, with the possibility that some studies used animals that are naturally more sensitive to BPA's effects than other strains. Whether humans share the characteristics of putatively sensitive strains is not known.
- Other study design factors, such as the type of feed used and the age at which the animals were sacrificed and tested.

On the other hand, the panel identified factors that may have given rise to the appearance of an effect associated with BPA exposure even if no such effect actually exists. These factors include:

- Inadequate control for confounders like body weight of individual animals, and individual vs. group housing of test animals.
- Failure to account for potential effects of intrauterine position (variations in natural hormonal exposures due to proximity to males or females during prenatal development)

Given the contradictory nature of the studies and the potential problems clouding their findings, the biological plausibility of the low-dose estrogen hypothesis was particularly important to the panel. The panel's review of the literature indicated that BPA does not exhibit several key characteristics that are typical of estrogenic agents. For example, natural estrogen causes cancer at high doses, whereas BPA does not. The panel concluded that because BPA does not share these other well-established estrogenic characteristics, it is unlikely to be exhibiting estrogenic characteristics in the case of the disputed impacts on the male reproductive tract.

In any case, the panel noted that the inconsistency of effects across species casts doubt on whether even a real effect observed in mice could be extrapolated to humans. Rat studies, including two large, well-conducted multiple generation reproductive and developmental studies in rats, have generally reported an absence of effects at low levels of BPA exposure.

Harvard Center for Risk Analysis
Harvard School of Public Health
718 Huntington Avenue
Boston, Massachusetts
02115-5924
617 432-4497
www.hcra.harvard.edu

Conclusions

Many important public health questions confront conflicting evidence and interpretation. Decisions often must be made before the answers are clear. We believe that the organized and systematic approach for assessing the weight of evidence described here can provide a more reliable characterization of the current 'best estimate' of the scientific evidence, and do more, sooner, to help inform risk management decision making in the face of uncertainty. The characterization of the state of the science is an important step in the risk assessment process. It should be recognized, however, that the risk management process may evaluate other factors, such as the relative significance of false positive or false negative results, in coming to a decision.

In the case of BPA, the evidence considered by the panel suggests that the weight of the evidence for low-dose effects is very weak. Studies are conflicting, the effects are subtle with questionable functional importance even if real, and there are conflicting data as to the proposed mode of action (*i.e.,* whether BPA acts as an estrogen).

The panel suggested follow-up research on the low-dose effects of BPA that:

- Uses similar species and strains of test animals.
- Homogenizes methodological approaches such as dose range, route of exposure, housing, feed, and many other design issues.
- Develops models to further determine applicability of rodent results to humans. A widely available, peer-reviewed physiologically-based pharmacokinetic (PBPK) model would be helpful in understanding potential similarities and differences in the responses of different species.

FURTHER READING/REFERENCES:

Gray, GM, Baskin, SI, Charnley, G, et al. (2001) The Annapolis accords on the use of toxicology in risk assessment and decision-making: An Annapolis Center workshop report. Toxicol Methods 11(3): 225-231

Gray, GM, Cohen, JT, Cunha, G, Hughes, C, McConnell, EE, Rhomberg, L, Sipes, IG, and Mattison, D (2004) Weight of the evidence evaluation of low-dose reproductive and developmental effects of bisphenol A. Human and Ecological Risk Assessment 10: XX-XX

Vom Saal, FS, Cooke, PS, Buchanan, DL, *et al.* (1998) A physiologically based approach to the study of bisphenol A and other estrogenic chemicals on the size of reproductive organs, daily sperm production, and behavior. Toxicol Ind Health 14(1-2): 239-60

PEER REVIEWER:
James K. Hammitt, PH.D

Other authors of the full panel report include Gerald Cunha, Department of Anatomy, University of California, San Francisco; Claude Hughes, Quintiles Inc.; Ernest E. McConnell, ToxPath Inc.; Lorenz Rhomberg, Gradient Corp.; I. Glenn Sipes, Dept of Pharmacology and Toxicology, University of Arizona; and Donald Mattison, Center for Research for Mothers and Children, National Institute of Child Health and Human Development. Other members of the panel who chose to be acknowledged in the paper for their contributions but not to be listed as authors are Paul Foster, NIEHS, Marvin Meistrich, University of Texas at Houston, Heinz Nau, Veterinary Medical University of Hannover, and Richard Sherins, Genetics and IVF Institute.