Response to the Interim Draft of the NTP-CERHR report on the Reproductive and Developmental Toxicity of Bisphenol A

Prepared by
Laura N. Vandenberg, Maricel V. Maffini, Beverly S. Rubin, and Ana M. Soto

Tufts University School of Medicine
Department of Anatomy and Cellular Biology
Boston, MA 02111

June 20, 2007

Table of Contents

Summary of Findings pages 2-15
Detailed comments on Section 1.0
Chemistry, Use and Human Exposure pages 16-19
Detailed comments on Section 2.0
General Toxicology and Biological Effects pages 20-28
Detailed comments on Section 3.0
Developmental Toxicity Data pages 29-47
Detailed comments on Section 4.0
Reproductive Toxicity Data pages 48-60

Summary of Findings

1. The criteria established by the panel are arbitrary.

Most of the criteria established by the panel for considerations of study adequacy are without explanation or merit.

1.1 The use of DMSO as a vehicle

> From page 130: “Dimethyl sulfoxide (DMSO) has significant biological activities of its own (Santos et al. Biochem. Pharm. 65(2003)1035-1041references), and its use as a vehicle for *in vivo* studies raises significant concern about the relevance of those results for the human oral exposure situation. Those studies which used DMSO to solubilize Bisphenol A and injected or implanted that mixture were deemed of little or no utility because of this double limitation (injection + DMSO).”

1.1.1 The dose of DMSO must be considered

The panel has decided to disregard all studies that utilize DMSO as a vehicle due to the potential effects reported for DMSO administration. This idea must be revisited and the level of exposure to DMSO must be considered. The alzet pumps used for subcutaneous administration of BPA release at the rate of 0.25 µl/hour. Because the vehicle is 50% DMSO, approximately 0.125 µl of DMSO is released per hour. These levels of BPA exposure are far below the levels routinely reported to have significant systemic effects when infused intravenously and they are below the dose of DMSO vehicle used in most studies present in the literature.

Clearly DMSO is not inert and can have effects on its own- some desirable and others not desirable. There are many studies in the literature that have documented the potential effects of DMSO administration and the effects appear to be related to the level of exposure, the i.v. route of administration, and the high flow delivery rate [1]. The problematic effects of DMSO reviewed in the commentary article cited by the panel [2], including systemic side effects such as nausea and vomiting, diarrhea, hemolysis, anaphylactic reactions, renal failure, etc. were observed when DMSO was used as a cryoprotectant for cells or transplants, including autologous bone marrow transplants, that were directly infused into the bloodstream of patients. A more recent study of autologous stem cell transplants reported evidence of acute DMSO induced systemic toxicity in individuals that received intravenous infusions of stem cells suspended in DMSO [3]. In one case the patient received a 200 ml infusion and in another, a patient received a 500 ml infusion. In both cases the cells were suspended in 10% DMSO and administered directly into the bloodstream of very compromised patients.

1.1.2 DMSO does not induce maternal or developmental toxicity

Perhaps more relevant to the concerns expressed by the panel may be the results of a recent study of the effects of a DMSO metabolite (MSM) administered to pregnant dams from GD6-GD20 at doses from 25mg to 1000mg/kg BW [4]. None of the endpoints examined in that study revealed effects on maternal or developmental toxicity.

DMSO has been a vehicle of choice for many studies of prenatal, perinatal and postnatal exposure to estrogenic, androgenic, antisestrogenic and antiandrogenic compounds. A recent study [5] that examined the role of early hormone environment on sex differences in rhesus macaque vocal production found evidence for a role of prenatal androgens in the development of sex differences in scream production. In these studies, pregnant females were injected with 250 ul flutamide with DMSO as the vehicle or 250 ul DMSO

vehicle alone twice per day from gestation days 35-70 (early gestation) or days 110-115 (late gestation). No adverse effects of the treatment were reported.

1.1.3 DMSO does not produce effects when compared to untreated animals

Studies that used DMSO as a vehicle have reported clear differences between BPA-exposed groups and groups that received DMSO vehicle alone. When the negative controls in an experiment are negative there is no reason to exclude these studies. Additionally, in our experience, results from control groups that received DMSO only and those that were untreated were comparable. In fact, when the study by Markey et al, [6] that initially contained data from both control groups was reviewed, the editor requested that data from the two control groups be combined into a single control group because there were no differences between them. This is stated in the paper.

1.1.4 The use of other vehicles are flawed

The panel has disregarded the problems inherent in the administration of other vehicles, such as corn oil which is often known to contain estrogenic contaminants. Interestingly, in Section 4 page 343, the use of olive oil was questioned with the statement: "The stability/reactivity of bisphenol A was not determined and it is possible that bisphenol A interacted with olive oil, resulting in the observed findings." Of the fourteen or more studies cited that used olive oil as a vehicle, this comment appears only once (on the 13th study).

1.2 The subcutaneous route of exposure

> The report states on page 130: "Because human exposure is overwhelmingly oral, and because oral exposure produces an internal metabolite profile which is overwhelmingly dominated by the (inactive) glucuronide, the Panel concludes that injection studies, unless they proved otherwise, would produce metabolite profiles which would be skewed heavily towards higher levels of the parent compound, and would tend to produce "false positive" effects, from the point of view of the human oral situation. Thus, the Panel viewed those studies which injected Bisphenol A as providing "supplemental" information, unless they also analyzed the levels of parent compound and metabolites after the injection. The intent of this approach is limit [sic] the impact of those studies which produced an unrealistic and irrelevant internal metabolite profile (i.e., one which is significantly different from that experienced by humans)."

1.2.1 There are many potential routes of human exposure

The oral route of exposure is considered to be the main route for humans, but this does not negate the potential for exposure through other routes. A study (cited in the report examining preschoolers found BPA present in detectable levels in indoor and outdoor air samples, floor dust, and play area soil [7]. A follow-up study by the same group [8] verified that BPA could be detected in greater than 50% of indoor air and hand-wipe samples; the researchers estimated that daily inhalation exposure ranged from 0.24 – 0.41ng/kg. Dermal absorption exposures were not calculated in this study.

In a survey of 118 homes, BPA was found to be present in 86% of dust samples in the range of 0.2 to 17.6 micrograms/gram of dust [9]. An additional study from the same group detected BPA in air and dust of homes, offices, and a plastics manufacturing site [10]. Another study measured BPA levels in urban ambient air samples with an average level of 0.51ng/cubic meter of air (Matsumoto 2005). These studies indicate that BPA exposure through air and dust is likely.

BPA has been detected in landfill leachates [11,12,13] and in treated leachates [13]. BPA has also been detected in sewage treatment works effluents, rivers, creeks and

drinking water [14,15,16,17]. These authors demonstrated that BPA could be detected in rivers in concentrations ranging from 500pg/L to over 100ng/L. BPA levels in drinking water ranged from 300pg/L to 2 ng/L. These studies indicate that BPA exposure through water used for both drinking and bathing is likely.

1.2.2 Because of the range of products containing BPA, actual exposures remain unknown

To date, studies have estimated BPA intake either from dietary sources alone (EU 2002, [18] or from environmental contaminations (water, air and soil) and dietary sources [19]. However, no study has yet measured actual human exposures from all sources. BPA is used in many consumer products including digital media (CDs & DVDs), electrical and electronic equipment, automobiles, eyeglass lenses, sports safety equipment, bicycle helmets, etc. Exposures from these products are unknown.

1.2.3 The issue of oral versus parenteral exposure is overplayed.

First, the metabolic profile of BPA metabolites found in urine and feces of animals treated with 25 micrograms/kilo of BPA is qualitatively similar at 24 hours regardless of the route of administration [20]. Second, there is no single study that monitors the levels of parent compound after the administration of a single low dose of BPA orally and subcutaneously. The high dose studies are difficult to interpret due to the low bioavailability of oral BPA in these conditions (16% for 10 mg/kg and 5.6% at 100 mg/kg- see Upmeier et al). In a study cited many times in this report (Upmeier et al [21], supported by the chemical industry) it is clear that the plasma concentrations of parent compound after oral administration are far from being negligible. For example, while iv administration of 10mg/kg resulted in very high levels declining very fast (700 ng/ml at 1h and 100 ng/ml at 2h), oral administration of the same dose resulted in a maintained level of 20-30 ng/ml for 8 hours, with a second peak at 6 hours due to enterohepatic recirculation. Hence, the much-praised notion of rapid first pass and complete inactivation is not true. Third, conjugation reactions are reversible, and in the case of natural estrogens, the conjugates can be hydrolyzed in the target organs releasing the active parent compound [22,23].The fact that Zalko et al have shown a significant amount of the parent compound in feces (94%), bile (1%) and digestive tract (26%) 24 hours after subcutaneous administration of tritiated BPA suggests that the conjugates have been hydrolyzed by the intestinal bacteria, and have been secreted directly into the intestine. In both cases, it is likely that BPA would undergo enterohepatic recycling, as shown in several studies. Hence, to assume that all the conjugated metabolites have been irreversibly inactivated is not warranted. Due to all these facts, it is arbitrary to “downgrade” studies using the sc route of exposure, particularly when the overwhelming majority of truly low level exposure (ng/kg range) and a representative number of low dose (microgram/kg range) have used the sc route

1.2.4 Studies of metabolism following chronic, low-dose exposure are lacking

Because of the multiple sources of BPA and the consistent finding of BPA in human samples, human exposure is thought to be chronic and low-level. At this time, no pharmacokinetic study of chronic, low dose BPA has been performed.

1.2.5 The panel fails to make a distinction between adult and fetal exposure

A fetus does not eat; it is exposed to BPA through its mother’s blood, and studies of humans indicate that low levels of BPA are regularly detected in blood. Therefore, any route of exposure that allows BPA to circulate in the maternal blood is closely “replicating the human condition” as the panel desires.

1.2.6 The panel fails to identify problems inherent in other dosing paradigms
Any method of exposing animals is bound to have pros and cons, but the panel does not acknowledge or even consider the potential issues related to other exposure paradigms.

1.2.6.1 Oral gavage
Gavage allows for an oral exposure, but it is a highly stressful experience for the animal [24], particularly a pregnant animal. In addition to the stress of tube placement, this method requires daily handling of the pregnant females which adds an additional stress. A basic knowledge of endocrinology illustrates that stress responses during pregnancy can lead to unexpected phenotypes in offspring, and can confound the effects of low doses of hormones. An example of this is a study by vom Saal showing that prenatal stress can erase endocrine-mediated intrauterine positional effects [25].

1.2.6.2 Dosing in feed or water
This method of dosing is preferable because it allows for an oral exposure, but it is unreliable. First, the actual exposure level is only approximate: i) in group housing, the amount consumed by each animal is unknown; ii) water leaks from bottles, allowing for inaccurate calculations; iii) several studies cited by the panel indicate that animals exposed to higher doses of BPA actually decrease their intake by consuming less BPA-containing food or water.

1.2.6.3 Other methods of oral dosing
There are several additional methods of oral dosing including training animals to drink from a pipette or eat a BPA-dosed cookie. These exposure paradigms have less stress associated with them compared to gavage, but they still allow for a single bolus dose to be given once (or twice) per day, which is not directly relevant to human exposure.

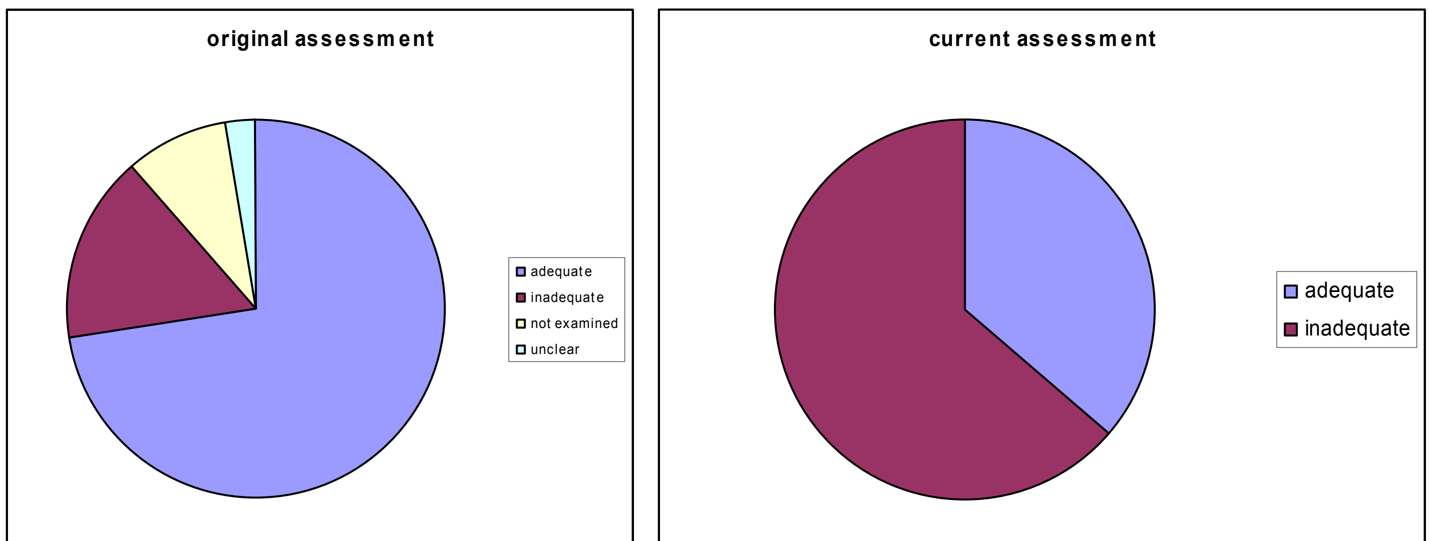
1.3 There is no explanation for the sample size requirement
The panel arbitrarily chose an n of 7 as an acceptable sample size for animal experiments without any justification. This capricious choice (“one size fits all”) is contrary to the understanding of statistical power and sample size analysis, which should be done by the experimenter a priori, i.e. before conducting the experiment.

1.4 Lack of understanding of non-monotonic dose response curves.
There were 86 studies cited in section 3 that examined more than one dose of BPA. Of these studies, 12 were noted by the panel to have “non-dose related” results, meaning that doses did not demonstrate a monotonic response. The expert panel members would benefit from understanding the importance and prevalence of non-monotonic dose response curves in many endocrine-related endpoints (from our lab, see [26,27,28,29,30,31]. Additionally, non-monotonic dose response curves are observed in many toxicological endpoints. A meta-analysis of 20,285 toxicology studies conducted between 1962 and 1998 found that only 1% of the published studies met the criteria set *a priori* to determine if a study was designed properly to detect a non-monotonic dose response [32] (i.e. most studies examine too few doses, an inadequate range of doses, insensitive endpoints, etc.). Of these studies, almost 40% satisfied the requirements for a non-monotonic dose response curve, supporting the idea that the occurrence of these responses is non-random and may even be more common than monotonic dose response curves [33,34]. Hence, the rejection of non-monotonic responses is unwarranted.

1.5 Lack of understanding of litter effects
Comments made by the expert panel in several sections indicate a lack of understanding about litter effects. If more than one animal is used per litter, a formula is available to compensate for the potential confounding factor of litter. However, if only 1 animal is used from a litter for each endpoint, no correction or differential statistics are required, because there is no confounding variable.

2. There is extensive evidence that the assessment criteria were not set a priori.

2.1 Evidence that assessment criteria were not established prior to examination of studies
In the first draft of this report, 73% of studies in section 3 were considered adequate and 16% were considered inadequate. In the current draft, 36% are considered adequate and 64% are considered inadequate. These data indicate that either no assessment criteria were used by the panel for the first draft, or the assessment criteria were drastically changed between the two drafts of this report. The reasoning behind this change is not presented by the panel and suggests a fundamental flaw, i.e. the criteria were tailored by the panel to remove or accept particular studies.



There are six examples from our own work where the initial review complimented our studies but the current review criticizes them. The statements of the Utility (Adequacy) for CERHR Evaluation Process are included in table 1.

2.2 There was a complete lack of assessment criteria in Sections 1 and 2
Criteria were not outlined for the studies cited in the first two sections of this report. There was a paucity of information on many of these studies, which were lacking limits of detection, sources of funding, and methods of detection. Additionally, the panel did not determine which studies were considered adequate or inadequate in these sections.

3. The assessment criteria are used inconsistently to weigh the evidence from each study to determine adequacy.

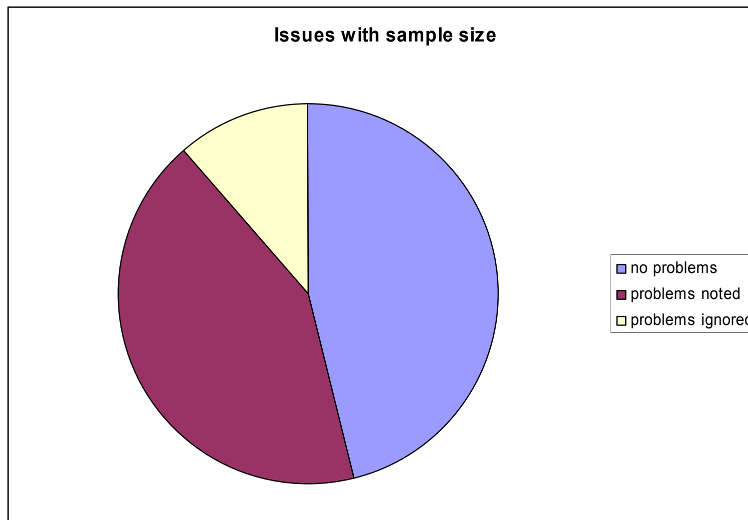
As stated in Point 1 above, many of the criteria used to determine adequacy are arbitrary or inappropriate. However, even considering this, the criteria were not applied consistently across all studies.

Table 1

Citation	1st draft comments	Current draft comments
Markey (page 231)	useful and shows tissue effect at extremely low dose levels	This paper is inadequate for the evaluation process
Markey (page 232)	This study is useful and showed significant effects, especially at the higher dose level.	This paper is inadequate for the evaluation process given exposure uncertainties.
Vandenberg (page 233)	This study is adequate and useful for the evaluation process and begins to catalogue the mammary compartmental changes induced by gestational bisphenol A exposure in mice.	This study is inadequate and not useful for the evaluation process
Markey (page 254)	This study is very useful	This paper is inadequate for the evaluation process given exposure uncertainties
Munoz-de-Toro (page 255)	This study is very useful	This paper is inadequate for the evaluation process given exposure uncertainties
Rubin (page 239)	this is a useful report for the evaluation process	This is inadequate for the evaluation process due to exposure and statistical concerns

3.1 The sample size requirement was not applied universally

Of the 124 studies examined in Section 3, 46% did not have an issue with adequate sample size. However, problems were noted in 43% of studies, where sample sizes were not considered large enough to be found acceptable. Importantly, 11% of studies described sample sizes of less than 7, yet this issue was not listed as a weakness by the expert panel.



Of the 53 studies where problems with sample size were noted, 34 (64%) were originally considered adequate, and 14 (26.4%) are still considered adequate. Additionally, of

the 14 studies with small sample sizes that were not noted as weaknesses, 11 (79%) were originally considered adequate, and 5 (36%) are currently considered adequate.

3.2 Several studies that used DMSO are considered adequate

Two of the 16 studies that used DMSO either for subcutaneous injections or in alzet pumps were considered adequate for inclusion. The first by Ramos in 2003, [35] was non-industry supported. The panel wrote that "This study is adequate for inclusion. However, the route of exposure is of concern and the results may be ancillary to the review process." In the second, Fukumori et al. 2003, the support was not indicated but the study examined was a translation provided by the American Plastics Council. Neither the route of exposure or the use of DMSO was mentioned in the strengths/weaknesses section by the panel.

3.3 Controlling for litter effects is not done universally

Of the 18 studies that did not use litter effect, 13 (72%) were originally considered adequate. In the current document, 1 of 18 studies (6%) is still considered adequate. There were an additional 9 studies that used litter effects for some endpoints and not others. In spite of this inappropriate statistical methodology, 6 (67%) were originally considered adequate and 2 (22%) are currently considered adequate. Finally, of the 30 studies with unclear statistical methods, 14 (47%) were originally considered adequate and 10 (33%) are currently considered adequate.

4. Is the panel purposefully misrepresenting data or grossly misunderstanding it?

The summaries provided for several studies are completely inaccurate and do not represent the experiments conducted and/or the results obtained.

4.1 Summaries of the results of many studies have been removed from this review

In Section 3, of the 124 studies examined, the results were removed from 38 studies without any explanation. Only one of these studies was industry funded. 32 of these studies were originally considered adequate. This action invites interpretations about the intentions of the panel, i.e. that the panel does not want the results from these studies to remain in the final document because they will be concerning to the public.

4.2 Metabolic studies of BPA

The report systematically misrepresents what is important about the papers reviewed. For example, the narrative on the paper by Zalko et al cites the percentage of the radioactive dose, rather than the concentration in each tissue (ng/g, ppb), a value that was also published in the paper and is much more representative of internal dose as it corrects for the weight of the organ (see Zalko et al [20], Table 1).

> For example, the reports states: "The highest percentages of radioactivity were detected in the digestive tract and contents (~45%) and feces (~21%). Less radioactivity was detected in the litter (~4%), liver (~2%), bile (~2%), urine (~6%), and carcass (~3%). Blood, ovaries, uterus, placenta, amniotic fluid, fat, and cage washes each contained <1% of the radioactive dose."

This gives the impression that there is precious little where it counts (blood, placenta, etc). When these data are expressed in ng/g or ng/ml of 3H-BPA equivalents at 24h, the data from Zalko et al illustrates values that are not negligible. Actually, for most tissues from the reproductive tract the detected concentrations are higher than in blood:

blood : 2.20 ppb
ovaries: 2.20 ppb

uterus:	3.45 ppb
placenta:	3.14 ppb
amniotic fluid:	4.85 ppb
fetus:	3.70 ppb

Moreover, the 4% detected in the litter should not be compared to the amount of BPA residues recovered in excreta, but to demonstrate that in this animal model, at a late stage of gestation, 1/25th of the administered low dose of [^{3}H]-BPA reaches the fetuses.

4.3 Summary of developmental toxicity studies

Perhaps some of the most blatant inadequacies in this report are found in the summary for Section 3.

4.3.1 A misrepresentation of results of anogenital distance studies

It is stated that "Although some sporadic effects were reported for anogenital distance in male and female rats, study authors concluded that the endpoint was not affected by prenatal, lactational, and/or post-weaning exposure to bisphenol A." The Rubin et al. 2001 rat study [36] is cited here, but this conclusion was never presented in this paper. This is a misrepresentation of the study and it is unacceptable.

4.3.2 A misrepresentation of results of LH serum levels

> It is stated that "There were some indications that bisphenol A exposure may affect serum LH levels in male rats after exposure to ≤1.2 mg/kg bw/day administered during gestational or postnatal periods, but the biological significance of the effect was uncertain because of questions regarding exposure characterization, lack of dose response relationships, and reproducibility of the effect (Rubin 2001, Akingbemi 2004)."

The Rubin et al. study [36]examined LH levels in female rats (exposed to BPA during fetal development until weaning) following long-term ovariectomy. A complaint about a lack of dose response relationships is unwarranted because significant effects were only observed for the highest administered dose. The second paper cited Akingbemi 2004 [37] may explain why the panel concluded that the effects observed were not reproducible, as this study examined LH levels in male rats (exposed to BPA during the pubertal period). Hence, there is no reason to compare these two studies, and it is inappropriate to expect similar responses when the timing of exposure and the sex of the animals are not comparable.

4.3.3 A misrepresentation of mammary gland studies

> "Hyperplastic or neoplastic changes (alone or in response to N-nitroso-N-methyl urea) were seen in mammary tissue of rats the dams of which were treated subcutaneously with bisphenol A 0.0025 or 0.025 mg/kg bw/day during gestation {Durando, 2007 #2450; Murray, 2007 #2451}. One single dose level study (using 50% DMSO vehicle in subcutaneous Silastic implants) reported that gestational exposure to 0.000250 mg/kg/d in mice caused quickened maturation of the mammary fat pad which was associated with increased ductal area and extension, decreased epithelial cell size, and altered collagen density in pubertal and adult mice."

First, it should be stated that the Murray 2007 [38] paper examined 0.0025, 0.025, 0.250 and 1 mg BPA/kg/day. Because no study was cited for this second statement, it is difficult to determine which experiment is being referenced. More importantly, these findings do not relate to any study we have conducted. First, our lab has not used subcutaneous silastic implants for BPA dosing purposes. Second, the results described have not been observed in pubertal or adult mice; instead, they seem to refer to our

study on gestational day 18 mice [39], while the Durando and Murray studies were performed in rats.

5. In spite of the reasonable statement in page 130 about the usefulness of positive controls, particularly when the test compound produced no effects, the weighing of reports according to this criterion is not consistent. In addition there is a lack of understanding or acknowledgement about the importance of negative controls.

Controls are included in studies to verify that all experimental procedures were conducted properly. When positive controls fail, the experiment should be outright rejected; it should not be included but given less weight. Additionally, if negative controls are not conducted, entire studies may be considered suspect, and more so if the test compound fails to produce any effect.

5.1 Malfunctioning positive controls

Of the 124 studies examined in Section 3, only 49 (39.5%) used a positive control. Of these studies, 61% had a properly functioning positive control, in 27% it is indeterminable if the positive control functioned properly (because it was not mentioned or the results section was deleted), and in 12%, the positive control did not work. In the 6 studies where the positive control did not work properly, 5 (83%) were originally considered adequate and 1 (17%) had an unclear assessment. In the current document, 4 (67%) are considered adequate and 2 (33%) are considered inadequate.

5.2 A misunderstanding of negative controls

The panel has systemically ignored issues of contamination, i.e. of phytoestrogens in feed, possible estrogenic activity of oil vehicles, plastic usage in animal environments, and a lack of understanding of the controls necessary in tissue culture experiments.

5.2.1 Many studies do not report the type of feed used in animal studies

Only 61% of the studies summarized in Section 3 had some level of information about the food used in animal experiments. In many of these cases, actual types of chow used were not described. Of those studies that did not provide information about the feed, the expert panel considered this a weakness only 14% of the time. Importantly, of the 75 studies that gave information about the type of feed used, 8% were criticized for their choice of chow.

5.2.2 Oil vehicles, especially corn oil, are at risk for contamination

43% of the studies summarized in Section 3 used unstripped oil of some type (arachis, olive, corn, peanut, sesame, etc.) and an additional 8% used oil that had been chemically stripped to remove estrogenic contaminants. Very few studies that used oil vehicles also examined completely unexposed (control) animals, and in those studies that did use unexposed animals, the results obtained from those animals (and any differences from the oil-exposed animals) were not presented by the panel.

5.2.3 Plastic cages and water bottles are often used in animal experiments

Many of these studies use polycarbonate cages and bottles, which are certain to contain (and likely leach) BPA. Studies by Pat Hunt's group [40] (and others) have demonstrated that leaching from cages can have significant impacts on reproductive endpoints.

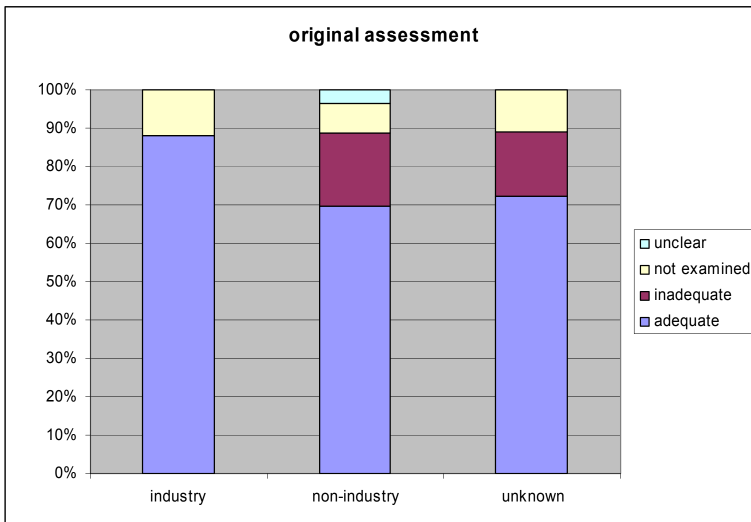
5.2.4 Controls for estrogenicity in cell culture experiments are incomplete

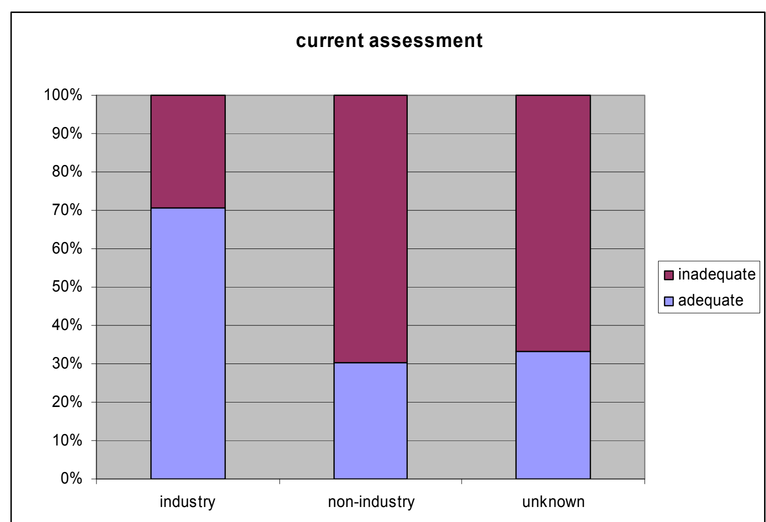
The panel states that appropriate controls for estrogenic contamination in cell culture experiments are the use of charcoal-dextran stripped serum and phenol red-free media. These criteria ignore the potential contamination of cell cultures from plasticware, which often contain BPA or other estrogenic contaminants.

6. There is evidence of bias in the assessment of studies based on source of funding.

Of the 124 studies examined, 17 (14%) were funded by industry, 2 (2%) were funded by US non-profits, 26 (21%) were funded by US government agencies, 61 (49%) were funded by non-US governments, and 18 (15%) had an unknown or not-indicated funding source. To simplify, 14% of studies were industry-funded, 72% were non-industry funded, and 15% were funded by unknown sources. The percent of studies that were originally deemed adequate separated by funding source has changed drastically between the two drafts of the evaluation.

In the original report, a similar percentage of studies were considered adequate in these three categories. In the current document, 12 of 17 (71%) industry-funded studies are considered adequate and 27 of 89 (30%) non-industry funded studies are considered adequate.





7. Many sections in this report illustrate a disregard for the nature of science.

This document addresses the research database as if all that is needed to know about basic biology is already known. Science is an ongoing project, rather than a completed one. For example, we are just beginning to learn how fetal malnutrition increases the propensity to diabetes and other diseases manifested during adulthood. These propensities are not “spiritual” entities, but material ones, hence, they have to manifest by producing changes during the period of exposure. Sooner or later, we will learn which are the subtle changes occurring during fetal life that mediate these adult-life outcomes. Hence, a change should be considered potentially pathological until further research clarifies whether or not this is the case. There is a fundamental understanding among developmental biologists that alterations in development are always significant. Changes in organ weight, cell number, etc. are not necessarily pathological, but are not devoid of developmental consequences including overt pathology at a later age.

References

1 Chaloupka JC, Huddle DC, Alderman J, Fink S, Hammond R, Vinters HV. A reexamination of the angiotoxicity of superselective injection of DMSO in the swine rete embolization model. Am J Neuroradiol 1999;**20**:401-10.

2 Santos NC, Figueira-Coelho J, Martins-Silva J, Saldanha C. Multidisciplinary utilizatoin of dimethyl sulfoxide: pharmacological, cellular and molecular aspects. Biochem Pharmacol 2003;**65**:1035-41.

3 Chen-Plotkin AS, Vossel KA, Samuels MA, Chen MH. Encephalopathy, stroke, and myocardial infarction with DMSO use in stem cell transplantation. Neurology 2007;**68**:859-61.

4 Magnuson BA, Appleton J, Ryan B, Matulka RA. Oral developmental toxicity study of methylsulfonylmethane in rats. Food Chem Toxicol 2007;**45**:977-84.

5 Tomaszycki ML, Gouzoules H, Wallen K. Sex differences in juvenile rhesus macaque (Macaca mulatta) agnostic screams: life history differences and effects of prenatal androgens. Developmental Psychobiology 2007;**47**:318-27.

6 Markey CM, Michaelson CL, Veson EC, Sonnenschein C, Soto AM. The mouse uterotropic assay: a re-evaluation of its validity in assessing the estrogenicity of bisphenol A. Environ Health Perspect 2001;**109**:55-60.

7 Wilson NK, Chuang JC, Lyu C, Menton R, Morgan MK. Aggregate exposures of nine preschool children to persistent organic pollutants at day care and at home. Journal of Exposure Analysis and Environmental Epidemiology 2003;**13**:187-202.

8 Wilson NK, Chuang JC, Morgan MK, Lordo RA, Sheldon LS. An observational study of the potential exposures of preschool children to pentachlorophenol, bisphenol-A, and nonylphenol at home and daycare. Environmental Research 2007;**103**:9-20.

9 Rudel RA, Camann DE, Spengler JD, Korn LR, Brody JG. Phthalates, alkylphenols, pesticides, polybrominated diphenyl ethers, and other endocrine-disrupting compounds in indoor air and dust. Environ Sci Technol 2003;**37**:4543-55.

10 Rudel RA, Brody JG, Spengler JD, Vallarino J, Geno PW, Sun G et al. Identification of selected hormonally active agents and animal mammary carcinogenesis in commercial and residential air and dust samples. Journal of the Air and Waste Management Association 2001;**51**:499-513.

11 Kawagoshi Y, Fujita Y, Kishi I, Fukunaga I. Estrogenic chemicals and estrogenic activity in leachate from municipal waste landfill determined by yeast two-hybrid assay. J Environ Monit 2003;**5**:269-74.

12 Yamamoto T, Yasuhara A, Shiraishi H, Nakasugi O. Bisphenol A in hazardous waste landfill leachates. Chemosphere 2001;**42**:415-8.

13 Coors A, Jones PD, Giesy JP, Ratte HT. Removal of estrogenic activity from municipal waste landfill leachate assessed with a bioassay based on reporter gene expression. Environ Sci Technol 2003;**37**:3430-4.

14 Rudel RA, Melly SJ, Geno PW, Sun G, Brody JG. Identification of alkylphenols and other estrogenic phenolic compounds in wastewater, septage, and groundwater on Cape Cod, Massachusetts. Environ Sci Technol 1998;**32**:861-9.

15 Kuch HM, Ballschmiter K. Determination of endocrine-disrupting phenolic compounds and estrogens in surface and drinking water by HRGC-(NCI)-MS in the picogram per liter range. Environ Sci Technol 2001;**35**:3201-6.

16 Belfroid A, van Velzen M, van der Horst B, Vethaak D. Occurrence of bisphenol A in surface water and uptake in fish: evaluation of field measurements. Chemosphere 2002;**49**:97-103.

17 Kolpin DW, Furlong ET, Meyer MT, Thurman EM, Zaugg SD, Barber LB et al. Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999-2000: A national reconnaissance. Environ Sci Technol 2002;**36**:1202-11.

18 Thompson BM, Cressey PJ, Shaw IC. Dietary exposure to xenoestrogens in New Zealand. J Environ Monit 2003;**5**:229-35.

19 Kang J-H, Kondo F, Katayama Y. Human exposure to bisphenol A. Toxicology 2006;**226**:79-89.

20 Zalko D, Soto AM, Dolo L, Dorio C, Ratahao E, Debrauwer L et al. Biotransformations of bisphenol A in a mammalian model: answers and new questions raised by low-dose metabolic fate studies in pregnant CD1 mice. Environ Health Perspect 2003;**111**:309-19.

21 Upmeier A, Degan GH, Diel P, Michna H, Bolt HM. Toxicokinetics of bisphenol A in female DA/Han rats after a single i.v. and oral administration. Toxiokinetics and Metabolism 2000;**74**:431-6.

22 Raftogianis R, Creveling C, Weinshilboum R, Weisz J. Estrogen Metabolism by Conjugation. Journal of the National Cancer Institute Monographs 2000;**27**:113-24.

23 Coughtrie MW, Sharp S, Maxwell K, Innes NP. Biology and function of the reversible sulfation pathway catalysed by human sulfotransferases and sulfatases. Chem Biol Interactions 1998;**109**:3-27.

24 Balcombe JP, Barnard ND, Sandusky C. Laboratory routines cause animal stress. Contemporary Topics in Lab Animal Science 2004;**43**:42-51.

25 vom Saal FS, Quadagno DM, Even MD, Keisler LW, Keisler DH, Khan S. Paradoxical effects of maternal stress on fetal steroids and postnatal reproductive traits in female mice from different intrauterine positions. Biol Reprod 1990;**43**:751-61.

26 Sonnenschein C, Olea N, Pasanen ME, Soto AM. Negative controls of cell proliferation: human prostate cancer cells and androgens. Cancer Res 1989;**49**:3474-81.

27 Soto AM, Lin TM, Sakabe K, Olea N, Damassa DA, Sonnenschein C. Variants of the human prostate LNCaP cell line as a tool to study discrete components of the androgen-mediated proliferative response. Oncology Res 1995;**7**:545-58.

28 Geck P, Szelei J, Jimenez J, Lin TM, Sonnenschein C, Soto AM. Expression of novel genes linked to the androgen-induced, proliferative shutoff in prostate cancer cells. J Steroid Biochem Molec Biol 1997;**63**:211-8.

29 Geck P, Maffini MV, Szelei J, Sonnenschein C, Soto AM. Androgen-induced proliferative quiescence in prostate cancer: the role of AS3 as its mediator. Proc Nat Acad Sci USA 2000;**97**:10185-90.

30 Vandenberg LN, Wadia PR, Schaeberle CM, Rubin BS, Sonnenschein C, Soto AM. The mammary gland response to estradiol: monotonic at the cellular level, non-monotonic at the tissue-level of organization? J Steroid Biochem Molec Biol 2006;**101**:263-74.

31 Wadia PR, Vandenberg LN, Schaeberle CM, Rubin BS, Sonnenschein C, Soto AM. Perinatal Bisphenol-A exposure increases estrogen sensitivity of the mammary gland in diverse mouse strains. Environ Health Perspect 2007;**115**:592-8.

32 Calabrese EJ, Baldwin LA. The frequency of U-shaped dose responses in the toxicological literature. Toxicol Sci 2001;**330**:338.

33 Calabrese EJ, Baldwin LA. The hormetic dose-response model is more common than the threshold model in toxicology. Toxicol Sci 2003;**71**:246-50.

34 Calabrese EJ, Baldwin LA. Hormesis: the dose-response revolution. Annu Rev Pharmocol Toxicol 2003;**43**:175-97.

35 Ramos JG, Varayoud J, Kass L, Rodriguez H, Costabel L, Munoz de Toro MM et al. Bisphenol a induces both transient and permanent histofunctional alterations of the hypothalamic-pituitary-gonadal axis in prenatally exposed male rats. Endocrinology 2003;**144**:3206-15.

36 Rubin BS, Murray MK, Damassa DA, King JC, Soto AM. Perinatal exposure to low doses of bisphenol-A affects body weight, patterns of estrous cyclicity and plasma LH levels. Environ Health Perspect 2001;**109**:675-80.

37 Akingbemi BT, Sottas CM, Koulova AI, Klinefelter GR, Hardy MP. Inhibition of testicular steroidogenesis by the xenoestrogen bisphenol A is associated with reduced pituitary luteinizing hormone secretion and decreased steroidogenic enzyme gene expression in rat Leydig cells. Endocrinology 2004;**145**:592-603.

38 Murray TJ, Maffini MV, Ucci AA, Sonnenschein C, Soto AM. Induction of mammary gland ductal hyperplasias and carcinoma in situ following fetal Bisphenol A exposure. Reproductive Toxicology 2007;**23**:383-90.

39 Vandenberg LN, Maffini MV, Wadia PR, Sonnenschein C, Rubin BS, Soto AM. Exposure to the xenoestrogen bisphenol-A alters development of the fetal mammary gland. Endocrinology 2007;**148**:116-27.

40 Hunt PA, Koehler KE, Susiarjo M, Hodges CA, Ilagan A, Voigt RC et al. Bisphenol A exposure causes meiotic aneuploidy in the female mouse. Current Biology 2003;**13**:546-53.

Comments on Section 1.0 Chemistry, Use, and Human Exposure

Page 2: "Purity of bisphenol A was reported at 99-99.8%" **No actual reference given.**

Page 2: "Use of LC-tandem mass spectrometry (MS/MS) with and without hydrolysis of bisphenol A glucuronide permits determination of free and total bisphenol A with a limit of quantification of 1 µg/L (Volkel 2005)." **Volkel's method uses one of the least sensitive methods since 1999. There are many studies that have examined BPA in serum with limits of detection of 0.5ng/ml and below (reviewed in Welshons 2006). Volkel's limit of detection was 1.14ng/ml, more than 2x less sensitive.**

Page 2: **"**Bisphenol A glucuronidate can be an unstable product that is hydrolyzed to free bisphenol A at neutral pH and room temperature in diluted urine of rats and in rat placental and fetal tissue homogenates at room temperature. Bisphenol A glucuronide can also be hydrolyzed and in some cases degraded to unknown components either in acidic or basic pH solutions of diluted urine, adding another potential source of error in the measurement of sample levels of bisphenol A and its conjugates (Waechter, 2007)." This reference is not available **through PubMed. The lack of a reference list with this draft is both frustrating and impractical. However, the implication that BPA-gluc is degraded to *unknown* components seems unlikely.**

Page 7-9: Table 2: Examination of Bisphenol A in Polycarbonate Food Contact Surfaces. **The sources of funding for these studies are not given, so it is unclear if there are conflicts of interest. Also, the LIMITS OF DETECTION and METHODS OF DETECTION were not given for these studies, so it is unclear if no BPA was detected in some studies because of insensitive methods. The units reported for different studies are not the same, making comparisons very difficult. Also, the legend at the end of the table refers to a publication mistake in the Biles 1997 reference. Has this mistake been verified by the panel and, if so, why hasn't the table been corrected?**

Page 10: "High molecular weight, heat-cured bisphenol A-based epoxy resins are used as protective linings in cans for food and beverages and occasionally in wine storage vats (European-Union 2003)." **The report does not specify who conducted these studies, and it is unclear what "occasionally" means.**

Page 10: Table 3: Surveys [sic] of Bisphenol A Concentrations in Canned Infant Formulas of Foods. **Again, the sources of funding, detection methods and limits of detection are missing.**

Page 11-13: Table 4: Surveys [sic] of Bisphenol A Concentrations in Canned or Bottled Foods or Food Stimulants. **Again, the sources of funding, detection methods and limits of detection are missing.**

Page 13: [It is unexpected that many bisphenol A concentrations exceeded concentrations detected in canned foods (Table 4). Study authors did not compare their findings with those of other studies examining bisphenol A concentrations in foods.] **Just because a finding is unexpected does not make it false. If this study is flawed because it did not test canned foods, are the studies of canned foods flawed because they did not test fresh products?**

Page 14: “Sealants are comprised of an organic matrix, while composites contain inorganic filler in addition to the organic matrix. The British Dental Association therefore concluded that because composites contain less resin than sealers, there is likely to be lower exposure to bisphenol A from composites than sealants.” **This statement is not based on an actual study, so it has no factual basis. While this is an interesting hypothesis, the panel should not use it as fact.**

Page 15: (reference to Fung 2000). “No bisphenol A was detected in saliva samples at 24 hours after treatment or in serum samples at any point.” **Many studies have detected baseline BPA levels in serum, regardless of dental sealant treatment. Fung et al used an HPLC method with a limit of detection of 5ng/ml, which is very insensitive compared to most methods used in serum. See Joskow 2006 which shows a significant increase in the level of BPA in urine after sealant application, a level which remained elevated after 24h, demonstrating that, internal exposure had indeed occurred..**

Page 16: “As noted in greater detail in Section 2, the majority of bisphenol A is systemically absorbed and circulated as a glucuronide compound.” **At this time, there is no study to verify that this is true following chronic BPA exposure. There have been a few metabolic studies performed in humans, and they used single (acute) large doses. In any case, that the majority of BPA is conjugated does not mean that the parent compound is not present. See studies by Upmeier, for example.**

Page 17: “Two studies (Mao 2004; Yang 2003) reported urinary bisphenol A concentrations that were orders of magnitude higher than commonly observed concentrations, despite the use of apparently reliable analytical techniques. Goodman et al. (2006) stated that reported hormone concentrations for the study volunteers were also higher than expected, indicating the possibility of laboratory or reporting error.” **The other possibility is that these individuals were exposed to high levels of BPA and also had high hormone levels. Unless another study from an independent group has examined the same samples and obtained another result, Goodman’s suggestions are mere conjecture.**

Page 17-18: Table 5: Blood Concentrations of Bisphenol A in Adults. **Again, the sources of funding and limits of detection are missing. There are several references missing: Inoue 2001, Schonfelder 2002, Yamada 2002 and Takeuchi 2004. Inoue and Schonfelder used LC-MS and GC-MS, respectively.**

Page 19-20: Table 6: Urinary Concentrations of Bisphenol A and Metabolites in Adults of Children. **Again, the sources of funding, detection methods and limits of detection (for all but one study) are missing. There are also references missing: Matsumoto 2003, Tsukioka 2003, Volkel 2005 and Joskow 2006.**

Page 24: The concentration of BPA in human milk is listed as negligible. **The following studies have detected BPA in breast milk: (Otaka 2003- GC-MS; Sun 2004- HPLC; Kuruto-Niwa 2006- ELISA, Ye 2006- HPLC & MS/MS, Ye, 2006, from the CDC). It appears that these studies were not reviewed by the CERHR. What was the limit of detection for the cited study? In particular, Ye et al is relevant, as they find that free BPA is prevalent in human milk.**

Page 26-28: Table 11. **The reference Kang 2006 is not present here. Kang used literature on environmental contamination levels (air, water, soil, can surfaces, plastic containers) to estimate the daily human intake of BPA to be ~ 1 ug/kg/day.**

Page 28: “Goodman et al. {Goodman, 2006 #2234} noted that total urinary bisphenol A concentrations were useful for estimating bisphenol A intake. Because nearly 100% of bisphenol A is excreted in urine within 24 hours {Tsukioka, 2003 #582}, bisphenol A intake can be estimated by measuring bisphenol A in urine over a specified time interval.” **Metabolic studies of BPA excretion in humans have only used acute, large doses of BPA. There is no study to confirm that these excretion rates are the same following chronic exposure. Additionally, the Tsukioka 2003 paper is a methods paper for determining trace amount of BPA in urine. This paper does not show that 100% of BPA is excreted in urine within 24h. Also, mouse studies revealed that significant amounts of BPA remained in the carcass 24 h after an acute exposure (Zalko, 2003)**

Page 29: “Goodman et al. {Goodman, 2006 #2234} concluded typical daily intakes of bisphenol A are ~0.02–0.06 g /kg bw/day in adults, based on urinary concentrations. [Consistent with adult findings, estimated mean exposure based on urinary bisphenol A concentrations in 6–8-year-old girls was 0.059 g /kg bw/day].” **There appears to be a missing symbol or some other error here, otherwise the panel is suggesting that Goodman found intakes of 0.02 g/kg/day (equal to 20,000 micrograms/kg/day).**

Page 32: “Possible exposure to bisphenol A during PVC manufacture has been considered, but the European Union {European-Union, 2003 #2146} stated that the application was being phased out.” **Is there a citation to indicate that the use of BPA in PVC manufacture will be phased out in the US also?**

Page 32: “The European Union {European-Union, 2003 #2146} summarized occupational exposure data for bisphenol A in Europe and the US. In some cases, industry sources provided data on total inhalable or respirable particulates that were not specifically analyzed for bisphenol A. It was sometimes difficult to discern if values reported in the European Union report were specifically for bisphenol A or for particulates in general. Unless otherwise stated, this report contains only values the Expert Panel was reasonable certain were for bisphenol A.” **If there were difficulties examining the EU review, it seems that the panel should have examined the original documents used by the EU commission .**

Page 34: Several sections have been copied directly from page 2. **See earlier comments regarding use of HPLC, and the insensitive methodology used by Volkel et al, 2005.**

Page 35: “These considerations taken together, [sic] suggest that it is possible that free bisphenol A concentrations measured in biological samples may be overestimated.” **There is no evidence to indicate that this is the case, especially because the listed considerations are not entirely fact-based. It is also possible that the concentrations measured in samples are underestimates.**

Page 35: “Because of its low volatility and relatively short half-life in the atmosphere, bisphenol A is unlikely to be present in the atmosphere in large amounts (European Union, 2003).” **More than 100 tons are released into the atmosphere during manufacture annually. Also, BPA has been detected in air, dust, and particulates (Rudel 2001, Rudel 2003, Matsumoto 2005).**

Page 35: Table 15: Summary of Bisphenol A Concentrations in US Environmental Samples and Drinking Water. **The sources of funding, detection methods and limits of detection are missing.**

Page 36: Table 16. **It is unclear what is meant by levels are "generally" below a given level.**

Page 36: Table 16B. **This table is so incomplete and poorly constructed that it is not understandable. There are no given references, limits of detection, no units for the values given, etc.**

Page 36-37: Table 16C. **It is unclear what is meant by "probabilistic food and/or aggregate exposure estimates". How is this different from data published in Table 11? There are no references to showing where the exposure values were derived.**

Detailed comments on Section 2.0, General Toxicology and Biological Effects:

General Comments:

1. **This section is highly repetitive and cumbersome to the reader.**
2. **While several attempts were made by the Panel to point out studies indicating differences caused by mode of exposure, there were no consistent findings of differences between oral and s.c. dosing. This is mentioned many times. In studies where supposed differences by exposure route were found, graphical analysis of the data presented in tables often indicates no differences.**
3. **There were no standard utility measures used for any of the referenced studies. Some included BPA purity/radioactivity purity, while others did not. Some indicated the method of statistical analysis, while others did not- and some studies did not use statistics at all. Strengths and weaknesses of studies were not identified. Importantly, many studies did not indicate whether testing was done for extraneous BPA or other estrogenic exposures (i.e. cage type, water bottles, type of feed, etc.)**
4. **There were no summary statements made to bring together the conclusions from the many studies examined. Some researchers found sex differences in BPA metabolism, while others did not. Some researchers found differences in pregnant vs. non-pregnant animals while others did not. Some researchers found differences by modes of exposure while others did not. What conclusions can be drawn by the expert panel from these studies?**
5. **Nowhere does the expert panel acknowledge non-monotonic dose response curves. (The term does not appear in the entire document.) Additionally, the panel has therefore ignored the fact that low and high doses can have drastically different effects, a phenomenon often reported in endocrinological end points- this is despite a previous NTP panel finding for the validity of low dose effects.**
6. **For the most part, the metabolism studies examine high, acute dosing paradigms. There is no mention of the possible differences in metabolism following chronic exposure (the suspected case for humans).**
7. **There is no indication of the sources of funding for these studies. This information is crucial to rule out bias due to funding source.**

Specific Comments and Biological Effects

Page 38: "High performance liquid chromatography (HPLC) with ultraviolet, fluorecence [sic], or electrochemical detection is unable to make definitive identification of bisphenol A or bisphenol A glucuronides, because similar retention times may occur for the metabolites of other endogenous and exogenous compounds (Volkel 2005)." … "Use of LC-tandem mass spectrometry (MS/MS) with and without hydrolysis of bisphenol A glucuronide permits determination of free and total bisphenol A with a limit of quantification of 1 µg/L (Volkel 2005)." … **"**Bisphenol A glucuronidate can be an unstable product that is hydrolyzed to free bisphenol A at neutral pH and room temperature in diluted urine of rats and in rat placental and fetal tissue homogenates at room temperature. Bisphenol A glucuronide can also be hydrolyzed and in some cases degraded to unknown components either in acidic or basic pH solutions of diluted urine, adding another potential source of error in the measurement of sample levels of bisphenol A and its conjugates (Waechter, 2007)." **These statements appear on Page 38 for the third time. The fact that they are repeated this often**

appears to indicate that they are considered extremely strong arguments by the CERHR panel, but the opposite is true. Again, many studies have used HPLC with low detection limits. The use of the Volkel reference is inappropriate because Volkel's method uses one of the least sensitive methods since 1999 (1.37ng/ml). There are many studies that have examined BPA in serum with limits of detection of 0.5ng/ml and below (reviewed in Welshons 2006). Finally, the Waechter reference is not available through PubMed. The release of this document without a complete bibliography has made it impossible to verify that the CERHR panel has correctly cited papers.

Page 39: (A repeat of the reference to Fung 2000 from Section 1, page 15). "No bisphenol A was detected in saliva samples at 24 hours after treatment or in serum samples at any point." **Many studies have detected baseline BPA levels in serum, regardless of dental sealant treatment. Fung et al used an HPLC method with a limit of detection of 5ng/ml, which is very insensitive compared to most methods used in serum. See Calafat 2005 and Joskow 2006.**

Page 40: "Studies reporting bisphenol A concentrations in fetal and/or maternal compartments are summarized in ." **This sentence is incomplete.**

Page 41-42: Table 18: Concentrations of Bisphenol A in Maternal and Fetal Samples. **Source of funding and limits of detection are not given. A study of umbilical cords is missing (Tokada & Mori 2002).**

Page 43: "Concentrations of bisphenol A measured in maternal plasma and amniotic fluid are summarized in ." **and** "Results for pregnant women are summarized in ." **and** "As noted in , concentrations ranged from <0.10 to 4.05 µg/L" **These sentences are incomplete.**

Page 44: "Volunteers received 25 µg D16-bisphenol A in drinking water [0.00028-0.00063 mg/kg bw based on reported body weights], a dose reported to represent a worst-case human exposure." **A reference is missing. This same group of researchers, in aprevious study, gave subjects 5 mg (0.054-0.090 mg/kg bw).** "The dose was reported to be ~10-fold higher than the estimated human exposure level of 0.6 mg/day." **These two statements are contradictory.**

Page 44: With regard to Volkel et al 2005: "Bisphenol A concentrations exceeding the detection limit were detected in only 2 urine samples at concentrations of ~10 pmol [2ng]/mg creatine." **As stated earlier, the method used by Volkel was one of the least sensitive methods used. It was 10x less sensitive than several other methods (see Brock 2001, Calafat 2005, Joskow 2006 and others).**

Page 45: With regard to Volkel et al 2002: "D16-Bisphenol A glucuronide concentrations in plasma and urine fell below LOD at 24-34 hours post dosing." **Again, the LODs in this study were 1.37ng/ml and 2.28ng/ml for urine and serum, respectively. These are 10-20x higher than methods used by other researchers. If a more sensitive method of detection were used, BPA may have been detected in excretions from these human subjects for a longer period of time.**

Page 49: Table 22. **Was the limit of detection really 0.005-0.029 mg/L? This seems high.**

Page 51: With regard to Snyder 2000: 2-4 lactating dams/group were examined at 4 different time points. 8-16 pups/time point were also examined. **There was no mention of litter effect.**

Page 52: Table 24: **It is unclear why such large BPA concentrations were detected in the animals exposed to 0 mg/kg/day. There must be a contamination.**

Page 54 lines 34-39. **This document systematically misrepresents the points of emphasis of the papers reviewed. The narrative on the paper by Zalko et al cites the % of the radioactive dose, rather than the concentration in each tissue (ng/g, ppb), which was also published in the paper and is much more representative of internal dose as it corrects for the weight of the organ (both % and concentrations are published in Table 1). that the document states:**

"The highest percentages of radioactivity were detected in the digestive tract and contents (~45%) and feces (~21%). Less radioactivity was detected in the litter (~4%), liver (~2%), bile (~2%), urine (~6%), and carcass (~3%). Blood, ovaries, uterus, placenta, amniotic fluid, fat, and cage washes each contained <1% of the radioactive dose. "

This statement gives the impression that there is little radioactivity where it counts (blood, placenta, etc). When these data are expressed in ng/g or ng/ml of 3H-BPA equivalents at 24h, Zalko et al found values that are not negligible. Actually, for most tissues from the reproductive tract the detected concentrations are higher than in blood:

blood : 2.20 ppb
ovaries: 2.20 ppb
uterus: 3.45 ppb
placenta: 3.14 ppb
amniotic fluid: 4.85 ppb
fetus: 3.70 ppb

Moreover, the 4% of BPA-equivalentsdetected in the litter should not be compared to the amount of BPA residues recovered in excreta, but rather used to demonstrate that in this animal model, at a late stage of gestation, 1/25th of the administered low dose of [3H]-BPA reaches the fetuses.

Pages 54-70, Zalko and Jaeg studies: **No mention is made on pages 54-55 and 70 of the significant levels of the parent compound (3H-BPA) in maternal plasma, placenta, amniotic fluid and fetus. Similarly, no mention is made to the internal dose of parent compound for example, maternal plasma, 1.0; placenta, 16; fetus, 4.2, amniotic fluid, 0.9 ng/g or ng/ml at 0.5 hours) for the 0.5, 2 and 24 hour timepoints. Only the presence of conjugated metabolites is discussed, suggesting that these are non estrogenic entities of minor importance and that nearly no BPA reached the fetus. This is also suggested by the report at page 70 (lines 15-21). Not only does BPA reach the fetus (see Zalko et al., table 5, last column), but it should be stressed that a further deconjugation of conjugated metabolites (as biochemically possible for any conjugated metabolite) might lead to the liberation of additional amounts of free BPA.**

The two first paragraphs at page 70 comment on the studies of BPA metabolism by Zalko et al. and Jaeg et al., in vivo and in vitro, respectively. It seems that a special

care was taken in order not to mention (or not to discuss) any data from these studies showing that some metabolic pathways, namely oxidations, could be of major interest regarding the understanding of the toxicity and of the biological activity of BPA and its residues.

For Jaeg et al.'s in vitro study on mice S9 liver fractions, most of the metabolites identified are detailed (lines 23-26, page 70). However, no mention is made about the possible signification of the cited results, namely the fact that several metabolites are of toxicological interest, such as BPA catechol, as well as metabolites (glutathione conjugates, hydroxylated phenol, dimers) showing that reactive intermediates could be formed following a first-step oxidation of BPA. This is surprising since the article contains an extensive discussion about these issues. It should be stressed that Jaeg's et al work gives strong "metabolic" support to the understanding of previous reports showing that BPA is converted to DNA binding metabolites in vitro (Atkinson et Roy, 1995a) and generates reactive oxygen species in male rat liver (Bindhumol et al, 2003).

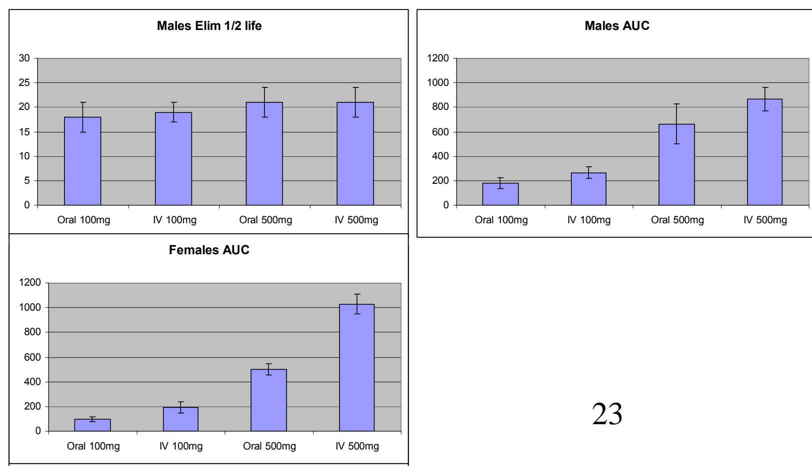
One could argue that in vivo, the major metabolic pathways of BPA are conjugation pathways, and that, consequently, conclusions from in vitro studies are of limited interest. This is more or less what is suggested by the way Zalko et al's results are presented (lines 9-14 pages 70) :

"major metabolites in urine were BPA glucuronide and hydroxylated BPA glucuronide", **and** "additional compounds detected in urine, feces, digestive tract or liver included a double glucuronide of BPA and sulfate conjugate".

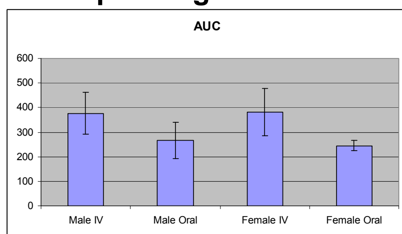
In fact, as extensively detailed in Zalko et al.'s article, BPA glucuronide and hydroxylated BPA glucuronide accounted for about 2/3rd of the urinary residues, but minor metabolites were also identified. The structure of several ones (metabolites 22a to 22d, table 2) indicates a highly probable in vivo formation of BPA catechol, in good accordance with in vitro results. As the formation of reactive species has also been postulated in vivo (Atkinson et Roy, 1995b; Hunt et al., 2003), there is no reason abridge this information from metabolic studies.

Page 56: "Halldin et al. (2001) examined distribution of bisphenol A in quail eggs or hens. After injection of fertilized quail egg yolk sacs with 67 ug/g 14C-bisphenol A egg [sic] on incubation day 3, <1% of radioactivity was detected in embryos at incubation day 6 or 9." **In which egg compartment was the radioactivity detected? If it was still in the yolk sac, was the embryo being continually exposed? The description of this study and its findings are lacking.**

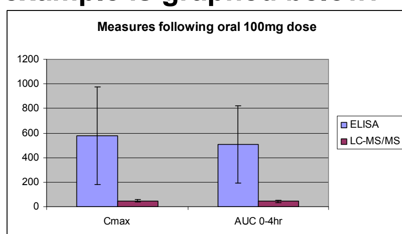
Page 63: Table 36: **There must be a mistake with either the units for this table, or the description of this study (page 62 line 13) because they do not agree. This table indicates that elimination ½ life and AUC for males is very similar between oral and iv dose. In females, elimination ½ life is similar but not AUC.**



Page 63-64: From Kurebayashi 2002: "Toxicokinetic endpoints are summarized in Table 37. Based on the toxicokinetic values, study authors concluded that absorption of bisphenol A following oral exposure was rapid and high, and terminal elimination half-lives of bisphenol A/metabolites were longer following iv than oral exposure." **This information is not apparent from Table 37. In fact, the endpoints given for iv exposure are not the same as endpoints given for oral exposure with the exception of AUC (graphed below).**



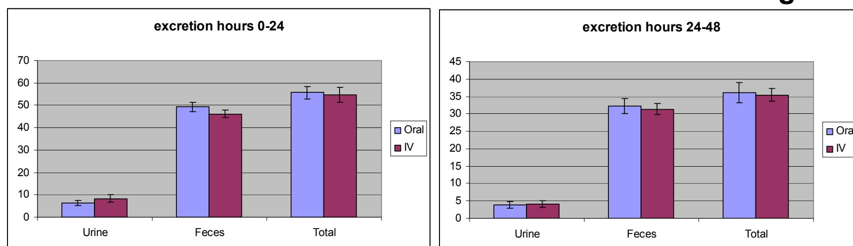
Page 64-65: Tables 38 and 39. **In Table 39, data is presented from only 1 or 2 rats for some endpoints. How can these low numbers be enough to make conclusions about the model and its reproducibility? The extreme differences in findings using the same samples but different methods was not discussed and is highly concerning. An example is graphed below:**



Page 66: "…hepatocytes were isolated from livers of adult female Wistar rats and incubated in dimethylsulfoxide (DMSO) vehicle or bisphenol A…" **What concentration or percentage was the DMSO?**

Page 67: "The European Union {European-Union, 2003 #2146} reviewed studies by Atkinson and Roy {Atkinson, 1995 #153;Atkinson, 1995 #26} that reported two major and several minor adducts in DNA obtained from the liver of CD-1 rats dosed orally or ip with 200 mg/kg bw bisphenol A. Chromatographic mobility of the two major adducts was the same as that observed when bisphenol A was incubated with purified DNA and a peroxidase or microsomal P450 activation system. The profile closely matched that of adducts formed with the interaction between bisphenol O-quinone and purified rat DNA deoxyguanosine 3´- monophosphate. Formation of the adduct appeared to be inhibited by known inhibiters of cytochrome P (CYP) 450. It was concluded that bisphenol A is possibly metabolized to bisphenol O-quinone by CYP450." **It is not obvious why this statement is included here. This is not really a metabolism study.**

Page 71: Table 41: **Data from this table (graphed below) appear to indicate similar modes and rates of excretion between oral and iv dosing.**



Page 75: "Teeguarden et al. {Teeguarden, 2005 #2114} developed a physiologically based pharmacokinetic (PBPK) model for bisphenol A. Rat toxicokinetic data for the model were obtained from the studies by Pottenger et al. {Pottenger, 2000 #1818} and Upmeier et al. {Upmeier, 2000 #1768}. Human toxicokinetic data were obtained from the study by Völkel et al. {Völkel, 2002 #589}." **This model should be disregarded for several reasons: 1) The Volkel study is flawed because of its use of insensitive detection methods. 2) The correlations used to determine the relationship between uterine weight and ER binding is not published or peer-reviewed. 3) The model disregards intestinal metabolism of BPA.**

Page 76: "This section includes information on general toxicity as well as information on estrogenicity and androgenicity; however, results of estrogenicity and androgenicity testing are not automatically interpreted as evidence of toxicity." **This sentence is confusing.**

Page 77: "Cecal enlargement occurring at doses ≥ 25 mg/kg bw/day was the most frequently observed effect in those studies but was not considered toxicologically significant by the European Union." **Does the CERHR agree with the European Union conclusions? Cecal enlargement is noted in several studies throughout this section. Cecal enlargement is often associated with digestive problems, iron deficiencies, bowel accumulations of osmotically active substances, diarrhea, etc.**

Page 77: "In a 90-day dietary study in dogs reviewed by the European Union, an increase in relative liver weight with no accompanying histopathological alterations was found to be the only effect at doses ≥270 mg/kg bw/day. This finding was considered by the European Union to be of doubtful toxicological significance." **Does the CERHR agree with the European Union conclusions?**

Page 78: "Yamasaki et al. 2002 examined the effects of nisphenool A exposure on male and female CD rats in a study conducted according to Good Laboratory Practices (GLP)." **Even with this GLP standard, the proper assurances for the controls are not described. This study used olive oil as a vehicle. Were any tests performed regarding the estrogenic content of the oil?**

Page 79-80: Table 48. **There is no indication of which differences were significantly different from controls.**

Page 80: With regard to the GE 1984 study: "No treatment-related clinical signs, ophthalmological changes, or death were observed during the study." **This sentence is unclear. Were the same clinical signs seen in the exposed and vehicle animals, or were the clinical signs not found in a monotonic dose-related manner?**

Page 81: With regard to the GE 1984 study: “The study authors did not consider any of the clinical chemistry changes to be biologically significant.” **In the opinion of this panel, is that an appropriate conclusion? (i.e. is the panel suggesting that differences in glucose levels, serum albumin levels, and alkaline phosphatase activity are not clinically relevant?)**

Page 82: “The first identification of bisphenol A as an estrogen has been attributed to Dodds and Lawson.” **This statement is misleading and ignores historical facts. Bisphenol A was, in fact, developed during a pharmacological search for potent estrogens and was developed for the purpose of pharmaceutical use. It was abandoned as a synthetic estrogen when diethylstilbestrol was synthesized and found to be more potent.** “Because the organ weight changes were not associated with microscopic changes in organs, the study authors concluded that the effects reflected decreases in body weight and were not toxicologically significant.” **This sentence is confusing. There were differences noted in absolute organ weights as well as relative organ weights. Both are indicative of a biological effect.**

Page 83: Figure 2. **It is unclear if these data includes potency relative to ER-alpha, ER-beta or both, but to combine the data for these two receptors would be inaccurate.**

Page 84-89: **There appears to be a lack of understanding about levels of biological complexity, and why some endpoints are more robust than others. (See Soto et al, Best Pract 2006)**

Page 98: “Increase in uterine weight 6 hours after treatment represents fluid imbibition and not true tissue growth.” **It is unclear if this statement is related to a finding from a single study, but it shows a lack of understanding of the uterotrophic assay. The classical uterotrophic assay measures an increase in the wet weight of the uterus; this response results from the integration of a large number of effects occurring at diverse levels of organization. Early events observable upon estradiol administration are water imbibition of the lamina propria and inhibition of cell death in the luminal epithelium. Following these events, cell number increases (hyperplasia) in the luminal and glandular epithelia of the endometrium, and the thickness of the lamina propria increases as well. The alterations observed in the lamina propria vary depending on the age of the animal; in ovariectomized adult rodents, alterations are due mostly to an accumulation of extracellular matrix and fluid, while in immature animals estradiol induces hyperplasia of cells in the stroma {Kang, Anderson, et al. 1975 4372 /id}.**

Page 98: “The greater activity of sc than oral bisphenol A is presumably due to glucuronidation of the orally administered compound with consequent loss of estrogenicity {Matthews, 2001 #459}… Matthews et al. {Matthews, 2001 #459} found a similar increase in uterine weight in rats given sc or oral bisphenol A at 800 mg/kg bw/day.” **These two statements appear contradictory.**

Page 98: “Nagel et al. {Nagel, 1997 #6;Nagel, 1999 #1208} noted that 17β-estradiol is extensively protein bound in vivo and bisphenol A is minimally protein-bound. They suggested that estrogenicity can be more accurately predicted by considering the free fraction of a chemical in serum. [The Expert Panel notes that does not suggest that bisphenol A is more potent than 17β-estradiol in vivo than in vitro. The Expert Panel also notes that Nagel et al. appeared to be referring primarily to prediction of developmental effects in the prostate rather than the estrogenic endpoints discussed in this section. The

developmental effects of bisphenol A in the prostate are discussed in Section 3.2.]" **Again, I think the Expert Panel is misunderstanding or ignoring the implications of these findings. BPA may have unexpected estrogenic activity because, unlike E2 which is bound by proteins in vivo, Nagel et al found that BPA is relatively unbound. The binding of the hormone makes it inactive until it is unbound; therefore, assays of E2 levels must determine the "free" fraction. The same is true for BPA. This is a basic fact in endocrinology.**

Page 99: With reference to the data presented in Table 51: "There was no apparent effect of strain on the sensitivity of the assay." **This is not true for immature rats, gavage treatment (effects seen at 1000mg/kg in Wistar but 375-600 in CD(SD)) or immature rats, sc treatment (effects seen as low as 10mg/kg in Sprague-Dawley but at 300mg/kg in Wistar.) The sample sizes for these studies are missing.**

Page 100: "[The assertions of some investigators notwithstanding, the Expert Panel notes that oral bisphenol A does not consistently produce estrogenic responses and, when seen, estrogenic effects after oral treatment occur at high dose levels.]" **What evidence is there that BPA is not estrogenic after oral exposure? This has not been presented here, at least not clearly. Maybe estrogenic activity has not always been demonstrated in the insensitive uterotrophic assay, however, it was observed in other end points such as mammary gland, vagina and pituitary at lower doses.**

Page 102-103: **Again, there is a clear lack of understanding of receptors within this summary. Membrane-bound ER is present in more cell types than the pancreas and pituitary as presented here.**

Page 103: "Estrogen-receptor demonstrates high constitutive activity that is maintained by bisphenol A in the presence of 4-hydroxytamoxifen, which otherwise decreases receptor activity." **This sentence is unclear.**

Page 104: With regard to the study by Kim 2002. **There is no mention of tests for background estrogenicity of the corn oil vehicle.**

Page 105: "Study authors concluded that bisphenol A did not exert androgenic effects; they characterized the effects of 3 and 50 mg/kg bw/day bisphenol A on epithelial cell height and luminal area of the prostate and seminal vesicles as "androgen-like" effects for which the mechanism was unclear." **This is a contradictory statement.**

Page 105: "[Although some unusual endpoints appear to have been affected, they are not validated measures of hormone action. Bisphenol A had no androgenic or anti-androgenic effects on organ weights.]" **How are changes in prostate epithelium considered "unusual endpoints"?**

Page 107: "DNA adduct formation was thought unlikely to be of concern for humans." **This is scientifically unclear. Is the panel saying that it's unlikely to occur because of the doses studied, or because DNA adduct formation will not cause problems to humans?**

Page 107: "Haighton et al. {Haighton, 2002 #391} concluded that results of standardized and validated genetic toxicity tests demonstrated the lack of mutagenic and genotoxic

activity of bisphenol A in vivo. Studies demonstrating distrupted microtubule formation or DNA adduct formation were noted, but because the studies used high doses, they were judged to be of limited relevance." **This statement is not relevant to this section, considering the high doses used in almost all studies cited in Section 2.0, unless of course, high doses are considered irrelevant in all contexts..**

Page 107: "Lastly it was concluded that bisphenol A had no structural features that suggested mutagenic activity." **This is a confusing statement, particularly when experts on the oxidative metabolism of estrogens and xenoestrogens claim the contrary.**

Page 116: **The same inaccurate statement about HPLC is included here, for the 4th (or 5th) time and is unnecessary.**

Page 118: Table 58. **It is unclear what is meant by dermal absorption, in vitro.**

Detailed Comments on Section 3.0 Developmental Toxicity

General Comments

1. **A major complaint is that huge sections of "results" have been removed for some studies, and it is unclear why this occurred. The reviews of many studies deemed "inadequate" were kept in entirety, including Tables.**
2. **A complete examination of literature on any topic requires the establishment of *a priori* criteria. The mere definition implies that criteria were selected before any review process began. This is obviously not the case, as illustrated by the many references that were originally deemed acceptable, but have now been deemed inadequate. Additionally, it is obvious that the same criteria are not being used for all reviews. These will be listed in the specific comments below.**
3. **Comments made by the expert panel regarding litter effect in several sections indicate a lack of understanding about litter effects. If more than one animal is used per litter, a formula is available to compensate for the potential confounding of litter. However, if only 1 animal is used from a litter for each endpoint, no correction or differential statistics are required, because there is no confounding variable. Additionally, the repeated requests/suggestions for authors to account for litter effect for postnatal dosing paradigms again shows a lack of understanding of this biological phenomenon. (For instance, when dams are ordered from an animal supplier, you cannot tell if these dams are sisters, but their litter effect is not essential.)**
4. **Including the purity of BPA is only helpful when studies actually examined it. Writing the purity level off the stock bottle is not helpful.**
5. **The Expert Panel has declared the use of DMSO to be an inadequate vehicle control; this is discussed in the summary statement. However, the panel fails to acknowledge the problems inherent in administering a potentially estrogenic chemical in plant oils, which often have contaminating estrogens. While some researchers used stripped oil, most do not. A good understanding of endocrinology indicates that even low levels of contaminating estrogens in the negative controls could confound results.**
6. **Most studies did not examine their housing conditions for extraneous estrogen exposure. Again, low levels of estrogens leaking from housing, food, water bottles, and bedding can lead to low levels of contamination of negative controls, which could potentially confound results, especially in experiments that found no effect.**
7. **Positive controls are included in studies to verify that all experimental procedures were conducted properly. When positive controls fail, the experiment should be outright rejected; it should not be included but given less weight.**

Specific Comments

Page 130: "Because human exposure is overwhelmingly oral, and because oral exposure produces an internal metabolite profile which is overwhelmingly dominated by the (inactive) glucuronide…" **Because BPA has been detected in water, air and dust, current studies cannot accurately predict all modes of BPA exposure. In addition, conjugation is not irreversible; studies in rodents have shown enterohepatic recycling, and the presence of parent compound in feces, which is reabsorbed by the intestine. The fact that Zalko et al have shown a significant amount of the parent compound, BPA in feces (94%),**

bile (1%) and digestive tract (26%) 24 hours after subcutaneous administration of tritiated BPA suggests that the conjugates have been hydrolyzed by the intestinal bacteria, and been secreted directly into the intestine. Moreover, conjugation reactions are reversible, and in the case of natural estrogens, the conjugates can be hydrolyzed in the target organs releasing the active parent compound (Chapter 6: Estrogen Metabolism by Conjugation Rebecca Raftogianis, Cyrus Creveling, Richard Weinshilboum, Judith WeiszJournal of the National Cancer Institute Monographs, No. 27, 113-124, 2000, © 2000 Oxford University Press), (Coughtrie MW, Sharp S, Maxwell K, Innes NP. Biology and function of the reversible sulfation pathway catalysed by human sulfotransferases and sulfatases. Chem Biol Interact 1998;109:3–27). Hence, to assume that all the conjugated metabolites have been irreversibly inactivated is not warranted.

Page 131: Morrissey et al. 1987 **This study used corn oil vehicle. It was stated that this study was performed according to GLP, however, there is no mention of the housing conditions, caging, food type, water source, bedding type or type of vehicle used. Evaluation of hard-tissue structures was listed as a strength when none were mentioned. A weakness was "**absence of information about the birth process**" but all dams were killed during pregnancy.**

Page 132: Kim et al. 2001**This study used corn oil vehicle. Both parametric and non-parametric statistics were used for n=20. A weakness was "**absence of information about the birth process**" but all dams were killed during pregnancy.**

Page 133: Kim et al. 2003 **This study used corn oil vehicle. The panel mentioned "**little evidence of a dose-response relationship**" and "**the absence of dose-related effects complicates interpretation at these low doses**" when perhaps the panel means that there was no evidence for a monotonic dose response relationship.**

Page 135: Talsness et al. 2000 **This study used corn starch vehicle. Also it was stated that** "the effects do not occur according to a classic dose-response curve, which is generally observed in toxicology studies." **The expert panel needs to reconsider the importance and prevalence of non-monotonic dose response curves in many endocrine-related endpoints (See Sonnenschein 1989, Soto 1995, Geck 1997, Geck 2000, Vandenberg 2006, and Wadia 2007). Additionally, non-monotonic dose response curves are observed in many toxicological endpoints. A meta-analysis of 20,285 toxicology studies conducted between 1962 and 1998 found that only 1% of the published studies met the criteria set *a priori* to determine if a study was designed properly to detect a non-monotonic dose response curve (Calabrese, 2001) (i.e. most studies examine too few doses, insensitive endpoints, etc.). Of these studies, almost 40% satisfied the requirements for a non-monotonic dose response curve, supporting the idea that the occurrence of non-monotonic responses is non-random and may even be more common than monotonic dose response curves (Calabrese, 2003).**

Page 138: Tinwell et al. 2002 **"**Support not indicated." **It is clear from their affiliations that these authors work in industry. This study used arachis oil vehicle. The rat chow used during gestation and lactation is likely to have a high phytoestrogen content** (18.5% soybean protein). **The sample size was listed as 7/group/strain, but the paper indicates groups of 6-7 (see Table 5 in the paper). The sample size for ethinyl estradiol treatment was only 3. It was stated that one strength of this study was "**the

utilization of 4 dose levels**", but only 3 were tested, plus vehicle and ethinyl estradiol. Another strength was "**the comparison between 2 strains of rat**" but no statistics were actually used to directly compare strains.**

Page 139: Schonfelder et al. 2002 **This study used corn starch vehicle without testing for background estrogenicity.**

Page 140: Wistuba et al. 2003 **This study used corn starch suspension vehicle without testing for background estrogenicity.**

Page 141: Thuillier et al. 2003 **This study used corn oil vehicle without testing for background estrogenicity.**

Page 142: Wang et al. 2004 "Male offspring from 3 independent litters were killed on GD21, PND3, or PND21. Western blot, RT-PCR, and immunohistochemistry techniques were used to measure expression of protein or mRNA for Hsp90, Hsp90a, p23, CYP40, Hsp70, and/or ER-beta… this study is inadequate by itself based on insufficient sample size." **In most scientific fields, n=3 is acceptable for RT-PCR analysis.**

Page 143: Funabashi et al. 2004 "The results suggest a disruption of the normal pattern of sexually dimorphic neurons, a result of critical importance to concerns about disruptions relevant to reproductive function and sexually dimorphic behaviors… This study is adequate for inclusion in the evaluation process, although utility is decreased due to the uncertain nature of the effect." **These two statements are contradictory.**

Page 145: Ramos et al. 2001 **The assertion that "**statistical approaches are inadequate**" contrasts with a previous comment which was deleted:** "statistical approaches appear 'conservative'**".**

Page 146: Ramos et al. 2003 **Under strengths/weaknesses:** "Altered values are given without addressing the normal range of variation or the likely functional significance of the changes." **First, this is a criticism that could be given to many of the studies examined here. Second, there is a fundamental understanding among developmental biologists that alterations in development are always significant. Changes in organ weight, cell number, etc. are not necessarily pathological, but are not without some developmental consequence. Hence, a change should be considered potentially pathological until further research clarifies whether or not this is the case.**

Page 147: Naciff et al. 2002 **How is this study both "**inadequate *and* ancillary**"?**

Page 147: Naciff et al. 2005 "expression of 15 genes was significantly altered in a dose-related manner" **and** "the dose response to bisphenol A was monotonic with no evidence of robust quantifiable responses at low doses." **Studies performed with estrogen indicate that i) gene expression endpoints are often less sensitive than tissue/organ based endpoints, and ii) gene expression endpoints can have a monotonic dose response curve while tissue/organ based endpoints from the same animals can demonstrate non-monotonic dose response curves (See Vandenberg 2006). Also:** "While endocrine disruption may certainly affect reproductive tissue development, of greater concern are potential disruptions in the neural control centers that are programmed in early development for performance at puberty and beyond." **This is an inappropriate statement. This study**

was designed to examine gene expression endpoints in the testes, epididymis, etc. not neural endpoints. For neurological effects of bisphenol A see Rubin et al. 2006.

Page 148: Saito et al. 2003 **This study used corn oil vehicle without testing for background estrogenicity.**

Page 149: Murray et al. 2007 **Why was the synopsis of the results removed? The inferences made by the panel are unjustified. As in all cases with papers from this group, the statistical unit is the dam. One animal per litter per time point is analyzed. The weakness mentioned,** "the uncertainty of the response rate of histopathology in the controls" **is not warranted. Intraductal hyperplasias are seen in young normal animals, hence a statistically significant increase upon prenatal exposured to BPA is a relevant finding. The use of 50% DMSO as vehicle is addressed in the first section of this document, summary of findings.**

Page 150: Durando et al. 2007 "The high single dose is a weakness." **In comparison to the studies earlier in this section that examined exposures of over 1000mg/kg/day, a dose of 25 micrograms/kg/day is quite low.**

Page 151: Hong et al. 2005 **This study used corn oil vehicle without testing for background estrogenicity. Also,** (top of page 152) "the level of calbindin-D9k expression in the uteri…" **This should say mRNA expression.**

Page 152: General Electric, 1976 **This study used 5 rats/sex/group, did not perform statistical analysis (or did not mention statistics), and has not been peer-reviewed. It is unclear why it is considered adequate.**

Page 153: General Electric, 1978 "Body weights of pups in the 750 ppm group were significantly higher [by ~10%] compared to controls on PND21, but the study authors did not consider the effect to be treatment related." **Is any other explanation given? Is this because the results did not follow a "typical" monotonic dose response curve? Also, as a strength:** "The use of multiple dose levels (going down to fairly low exposure levels) is a plus." **The lowest dose examined by this study is approximately 5mg/kg/day. It is not accurate to consider this a fairly low exposure level. Additionally, it was mentioned that a weakness was** "the use of conventional 'top-down' pathology evaluation, wherein the lower dose groups were examined only if effects were noted in the high dose." **This is a fundamental flaw, considering the CERHR's recognition of low dose effects of BPA. Together with this flaw (**"this study was not designed to find unusual effects or non-linear dose-response relationships"**), the large doses, and the fact the study has not been peer-reviewed, it is unclear why it is considered adequate.**

Page 155: Ema et al. 2001 **Housing, feed, etc. was not included in the panel's summary. Rats were gavaged, but it is not stated what vehicle was used. The summary does not state what ages were examined, what the sample size was for different endpoints, which statistical methods were used, etc. It is unclear why it is considered adequate when so much information is lacking.**

Page 155: Tyl et al. 2002 "Although some decreases in epididymal sperm concentration and daily sperm endpoints were each observed in 1 generation of males from the high-dose group, the study authors concluded there were no treatment-related effects on sperm endpoints or reproductive function." **This is confusing and seems contradictory.**

Page 156: Cagen et al. 1999 **What were the housing conditions (other than exposure to music at "**a low volume to provide background noise**")?** "Male sex ratio was increased in the 0.1 ppm bisphenol A group (56.7% males versus 48.4% in control) but the study authors did not consider the effect to be treatment related due to the lack of a dose response relationship." **See comments above about insufficient consideration of non-monotonic dose response curves. More importantly, this study showed no effects following DES (positive control) exposure, which should have made it obviously inadequate.**

Page 157: Elswick et al. 2000 **This paper is a review of unpublished, non peer-reviewed data and does not belong in this document. However, if the rest of the summary is kept, the summary of the results should definitely remain.**

Page 159: Takashima et al. 2001 "In 25-week-old rats that were not exposed to N-nitrosobis (2-hydroxypropyl)amine, prenatal bisphenol A exposure was associated with some thyroid-stimulating hormone elevations in males and females from the MF and soybean-free diet groups. According to a statement in the study abstract, the study authors did not consider the effects on thyroid-stimulating hormone to be related to bisphenol A exposure." **Was any other explanation given?**

Page 161: Kobayashi et al. 2002 **This study used corn oil vehicle without testing for background estrogenicity.**

Page 162: Ichihara et al. 2003 **This study used corn oil vehicle without testing for background estrogenicity. In strengths/weaknesses:** "Statistical accounting for litter effects is unclear for neonatal measures, body weight, and fertility endpoints, however this needs to be judged in the context of the absence of effects. Thus, failure to attend to litter effects may not be critical." **This statement is confusing.**

Page 163: Yoshida et al. 2004 **The use of HPLC to detect BPA in water, containers, etc. is commendable but does not indicate whether these animals were exposed to other extraneous estrogens.**

Page 164: Takagi et al (unknown date) **Polycarbonate cages are known to leach BPA. The results section should remain in this summary. The strengths/weaknesses section lists the use of 17β-estradiol as a control, but the text lists ethinyl estradiol as the positive control.**

Page 166: Akingbemi et al. 2004 **This study used polycarbonate cages and water bottles. This study used corn oil vehicle without testing for background estrogenicity. It was noted that** "weaknesses include an inadequate number of animals to obtain confidence about the hormonal changes (indeed, LH was decreased in the first experiment and increased in the third)…" **Experiment 1 dosed animals from PND21 to PND35, while experiment 3 dosed animals from PND21 through PND90.**

Page 167: Masutomi et al. 2004 **This study used polycarbonate cages. It was stated that** "statistical analyses were performed by Student or Welch t-test using values from the highest bisphenol A dose group and the control." **Why weren't statistical analyses required for the other doses? Also, sample size is given as 5-8, but this paper was nonetheless considered adequate.**

Page 168: Tan et al. 2003 **This study used Tween 80 in corn oil as a vehicle without testing for background estrogenicity. Also, a sample size of 6 per endpoint was used but this paper was still considered adequate.**

Page 169: Kobayashi et al. 2005 **This study used corn oil vehicle without testing for background estrogenicity.**

Page 169: Zoeller et al. 2005 **This study used plastic cages, but the type of plastic was not given. The expert panel mentioned there was no apparent dose response relationship for RC3/neurogranin mRNA- was this just a lack of a monotonic dose response curve? It was stated that** "weaknesses include the lack of litter-based analyses" **but it was stated that 1 animal/litter was used, so litter was not a confounding variable.**

Page 170: Kwon et al. 2000 **Animals were housed in polycarbonate cages. This study used corn oil vehicle without testing for background estrogenicity.**

Page 172: Kubo et al. 2003 **Results were removed with no explanation.**

Page 174: Facciolo et al. 2002 **This study used arachis oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**
Page 175: Facciollo et al. 2005 **This study used arachis oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**
Page 177: Aloisi et al. 2002 **This study used peanut oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**

Page 178: Negishi et al. 2003 **This study used olive oil vehicle without testing for background estrogenicity.** "Data analyzed at birth were presented and analyzed on a per litter basis. Postnatal data were apparently analyzed on a pup basis." **The measurements made on the day of birth were measured in the pups, so all of these endpoints are postnatal. Results were removed with no explanation. In the strengths/weaknesses, it says** "Doses were sufficiently high to produce gross body weight changes…" **This statement is misleading, since body weight changes have been observed following much lower doses than 4 mg/kg/day.**

Page 179: Negishi et al. 2004 **This study used corn oil vehicle without testing for background estrogenicity.**

Page 180: Farabollini et al. 1999 **This study used arachis oil vehicle without testing for background estrogenicity. Results were removed with no explanation. Also, this study was deemed inadequate because of absence of expected effects in the positive control (which was not otherwise mentioned) although other studies were not rejected for this reason. In the paper itself there was no positive control mentioned.**

Page 181: Farabollini et al. 2002 **This study used arachis oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**

Page 182: Dessi-Fulgheri et al. 2002 **This study used arachis oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**

Page 183: Porrini et al. 2005 **This study used peanut oil vehicle without testing for background estrogenicity. It was stated that** "Pups were weaned on day 21 [assumed to be PND21]." **Considering that pups cannot be weaned on embryonic day 21, this is an unnecessary comment. Results were removed with no explanation. It was also stated that** "The fostering of pups within treatment group prevents the evaluation of intrauterine effects." **It seems that the panel means litter effects. Litter effects and intrauterine effects are not the same.**

Page 184: Adriani et al. 2003 **This study used arachis oil vehicle without testing for background estrogenicity.**

Page 185: Carr et al. 2003 **This study used safflower oil vehicle without testing for background estrogenicity. With doses of 0.1 and 0.25 mg/kg/day, it was stated that** "though the doses were high, they were not damagingly so." **This statement is vague and its meaning is unclear.**

Page 186: Della Seta et al. 2006 **This study used safflower oil vehicle without testing for background estrogenicity. Some endpoints had a sample size of 12, or 7-8, and some had only 5-6.**

Page 187: Ceccarelli et al. 2007 **This study used peanut oil vehicle without testing for background estrogenicity.**

Page 188: Fisher et al. 1999 **This study used s.c. injection, corn oil vehicle without testing for background estrogenicity. The sample sizes were 3-7 in the treated group and 5-20 in the BPA group, depending on the endpoint. The animals were dosed with 37 mg BPA/kg, no info was provided on housing, and litter information was not given. It is unclear why this study was labeled** "a carefully performed study" **that was adequate for evaluation.**

Page 189: Nagao et al. 1999 **This study used corn oil vehicle without testing for background estrogenicity. Also, in the weaknesses, the report originally stated** "the use of only a single high dose level of bisphenol A" **but the word "high" has been removed, even though these authors were dosing animals with 300 mg BPA/kg.**

Page 190: Stoker et al. 1999 **This study used sesame oil vehicle without testing for background estrogenicity, and silastic implants for the estradiol positive control. For one endpoint sample size was 6, for the other it was 9-11. It was stated that** "the myeloperoxidase assay was reported to show a "trend" for lateral prostate inflammation in the bisphenol A group. [Trend was not defined; there was no statistical difference between the bisphenol A group and the control in the myeloperoxidase assay.]" **A "trend" generally refers to a p-value >0.05 and <0.1.**

Page 191: Atanassova et al. 2000 **This study used corn oil vehicle without testing for background estrogenicity. For the three different ages examined, sample size was 7-14, 4-12, and 6.**

Page 191: Williams et al. 2001 **This study used corn oil vehicle without testing for background estrogenicity. The study authors concluded that** "only high doses of potent estrogens induce gross abnormalities in the male reproductive system" **but this statement is not supported by a study that examined a single (high) dose of bisphenol A (range**

of 100mg/kg/day to 20mg/kg/day as the animals gained weight) nor a single dose of ethinyl estradiol (10ug/day). Additionally, the statement that a strength of this study is "the expertise of the group" **calls into question the biases of the CERHR panel. Because this panel was purposely not composed of people in the bisphenol A field, they are not qualified to comment on the expertise of each study's authors. Furthermore, this statement implies that each study is being examined based on information outside of the study itself, information that is not available to the general public nor the average journal reader. Opinions on article authors should have no bearing on the strength of an individual study. For this reason, these studies should have been examined with the expert panel blind to the identity of the authors.**

Page 192: Rivas et al. 2002 **This study used corn oil vehicle without testing for background estrogenicity. Only 3-10 animals/group were assigned to endpoints but it was still considered adequate.**

Page 193: Sharpe et al. 2003 **This study used corn oil vehicle without testing for background estrogenicity.**

Page 194: Khurana et al. 2000 **This study was rejected for** "lack of design clarity and small sample size." **The issues with litter effects are not relevant because these animals were dosed postnatally. The sample size was 8-10 pups/sex/group.**

Page 194: Fukumori et al. 2003 **This review is based on a translation provided by the American Plastics Council, which is concerning. The control animals were treated subcutaneously with DMSO; according to the panel criteria, it should be rejected, but it is instead considered adequate. It was stated that** "this is a translation of an apparently carefully performed study…" **yet also comments from the panel include** "[The number of rats treated and examined/group and the number of litters represented were not reported. No information was provided on purity of bisphenol A, type of feed, or composition of bedding and caging. The translated version of the report did not include figures from the original report.]" **Another unsupported statement from the panel was:** "It may be that some of the effects observed simply reflect mild retardation of development in the treated animals, which would be corrected with the passage of time."

Page 195: Kato et al. 2003 **This study used corn oil vehicle without testing for background estrogenicity. For all endpoints, the sample size was 5-8 rats. It was also stated that doses of ~26, 105 and 427 mg BPA/kg/day were** "relatively high." **Additionally, the expert panel made the statement that** "The changes in the female reproductive organs seen are well documented, but given the extremely high dose of agent used, broadly unsurprising." **If the panel considers the effects of exposure on the female reproductive tract to be well-documented, then why is this worth mentioning?**

Page 196: Toyama and Yuasa 2004 **This study used DMSO/olive oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**

Page 197: Kato et al. 2006 **This study used ethanol/corn oil vehicle without testing for background estrogenicity. This study was rejected because of** "route of administration and dosing on a per pup basis, this not adjusting for bodyweight." **Several other studies used this method but were not rejected: see Atanassova et al 2000, Williams et al. 2001 or Rivas et al. 2002 as examples.**

Page 198: Noda et al. 2005 **This study used olive oil vehicle without testing for background estrogenicity. Results were removed with no explanation. This study was also rejected because of s.c. dosing and dosing on a per-pup basis. See complaints for Kato et al. above.**

Page 199: Ho et al. 2006 **With regard to the increased incidence and severity of prostatic intraepithelial neoplasias in treated animals, the panel said** "in humans, this is a precursor lesion to prostate cancer, however in rodents it is a lesion of unknown significance." **The significance here is that PIN lesions are being observed in rodents. PINs are precancerous lesions in rodents as well as in humans (see: Shappell SB, Thomas GV, Roberts RL, Herbert R, Ittmann MM, Rubin MA, Humphrey PA, Sundberg JP, Rozengurt N, Barrios R, Ward JM, Cardiff RD. Prostate pathology of genetically engineered mice: definitions and classification. The consensus report from the Bar Harbor meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. Cancer Res. 2004 Mar 15;64(6):2270-305)**

Page 200: Ishido et al. 2004 **This study used ethanol/olive oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**
Page 201: Masuo et al. 2004 **This study used olive oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**

Page 202: Masuo et al. 2004 **What vehicle was used? Results were removed with no explanation.**
?

Page 203: Ishido et al. 2005 **This study used ethanol/olive oil vehicle without testing for background estrogenicity. No information was provided on housing conditions, and the sample size for each endpoint was not included. There were no statistics performed. Results were removed with no explanation.**

Page 204: Patisaul et al. 2006 **These animals were fed a phytoestrogen-free diet in the last week of gestation, but previous diet is not stated Early developmental exposures to estrogens can have significant effects on developmental endpoints and can confound other results. This study used sesame oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**
Page 205: Patisaul et al. 2006 **This study used oil vehicle without testing for background estrogenicity. Results were removed with no explanation.**

Page 206: Shikimi et al. 2004 **This study used sesame oil vehicle without testing for background estrogenicity. It was written that estradiol benzoate was used as a positive control, but then the use of 17β-estradiol was included as a strength.**

Page 207: Zsarnovsky et al. 2005 **Housing and feed information was not provided. What was the vehicle used? Results were removed with no explanation.**

Page 208: Morrissey et al. 1987 **As written for the rat section of this study, it was stated that this study was performed according to GLP, however, there is no mention of the**

housing conditions, caging, food type, water source, bedding type or type of vehicle used.

Page 209: vom Saal et al. 1998 **The comment about these mice being maintained as** "an outbred stock in a closed colony" **is without context. What information does the panel have about how any of the commercial suppliers maintain an outbred stock?**

Page 210: Nagel et al. 1997 **Again, comments about the mouse colony are out of context. Also, considering the feed and bedding to be a weakness has NOT been applied to most of the other studies listed, especially because many studies do not give any information about their housing conditions.**

Page 211: Cagen et al. 1999 **It was written that this study** "attempted to duplicate the findings by vom Saal et al. and Nagel et al. by repeating their procedures" **but then there is an acknowledgement of significant deviations from those earlier studies including a different endpoint tested (PND90 vs. PND180), different feed, an approximation of gestational age in some animals, single vs. group housing, and the commercial source of mice. The panel failed to mention the weakness of using corn cob bedding, as was mentioned for the previous study. The group had an inability to duplicate control breeding levels at the fecundity level of the commercial supplier. The authors noted non-dose-related effects in several endpoints and therefore** "concluded that the effects were not treatment related" **although no other explanation was given. The positive control DES had no significant effects indicating that this experiment failed (in spite of it using GLP), yet this study was not considered inadequate. Also on this topic, the panel wrote that** "only 1 study has shown effects at this dose" **but this is not true. It was stated that** "the study authors noted that the diethylstilbestrol dose was considered the 'maximum effect' oral dose by vom Saal". **Also, within this review, Gupta et al. show effects of DES at 0.1 µg/kg. There are many other examples in the DES literature.**

Page 212: Ashby et al. 1999 **It was stated that this study was an attempt to replicate vom Saal and Nagel's study. However, this study also used several protocol deviations including a different diet, diet changes during gestation, an approximation of gestational age in some animals, and individual vs. group housing conditions. Information was not listed regarding caging material, water bottles, or bedding. It was written that** "vaginal opening was delayed in the diethylstilbestrol-treated group and in the naïve control group" **indicating a flaw in study design, a lack of effect in the positive-control group, or contamination of the negative control group. The panel wrote that** "the use of small samples and the lack of effect of the positive control is a weakness, but it is unclear why this dose of diethylstilbestrol was expected to give a positive response." **While it is good that the expert panel's recommendation for this study has changed from** "very useful in the evaluation" **to** "marginally useful" **it is still unclear why it is not considered inadequate.**

Page 214: Howdeshell et al. 1999 **This study used oil vehicle without testing for background estrogenicity. Comments from the panel include:** "although the authors identified a litter-based analysis, it was not always clear that this applied to all analyses (in study figure 1, the n values exceed the number of dams, suggesting that some of the data were analyzed on a per pup basis." **Because this study was examining animals at several different intrauterine positions, more than one pup per litter was used. Also, the panel included the endpoints as a weakness, saying:** "the use of time from vaginal

opening to first estrus is not a standard endpoint for assessing puberty in mice." **The panel does not appear to be aware that time of vaginal opening is considered similar to time of thelarche in humans, and time of first estrus is comparable to time of menarche in humans.**

Page 215: Gupta et al. 2000 **This study used corn oil/ethanol vehicle without testing for background estrogenicity. The panel criticized the use of more than one animal per litter for some endpoints, and then alternatively criticized the use of one animal per litter for other endpoints:** "Questions were raised about whether sampling 1 pup/litter on PND60 provided a reliable estimate, especially for highly variable endpoints such as anogenital distance, which can be affected by sex of the adjacent fetuses in the uterus."

Page 217: Iida et al. 2002 **This study used corn oil vehicle without testing for background estrogenicity. Why were the results removed? A weakness was** "inadequate methods for histopathological preservation." **Is the panel suggesting that the fixative of paraformaldehyde is inappropriate?**

Page 217: Timms et al. 2005 **It was stated that this study used 0.1 µg ethinyl estradiol/kg/day as a positive control, but then the panel commented that a strength of the study was the use of 17β-estradiol as a positive control.**

Page 218: Palanza et al. 2002 **It was stated that** "It is unusual that pre- and postnatal exposure had effects but not the combination of pre- and postnatal exposure, and failure to explain this finding is a weakness." **But the authors propose a mechanism for how this is occurring: "An unexpected finding here is the absence of an effect on maternal behavior in females first exposed to BPA during fetal development and then again in adulthood during late pregnancy. One hypothesis is that fetal exposure to BPA results in permanent changes in systems that maintain homeostasis. This shift in homeostatic mechanisms may alter the subsequent response to chemical exposure at a later life stage relative to the response that would occur with no prior exposure to the chemical. There has been speculation that short-term exposure to chemicals, such as BPA, could lead to different outcomes relative to long-term exposure, such as in a multigenerational study (*23*). Our findings are thus intriguing, but a much broader study involving administration of EDCs at different times in life, which includes physiological measures as well as behavior, will be required to answer this question."**

Page 220: Nishizawa et al. 2003 **This study used olive oil vehicle without testing for background estrogenicity. The panel did not provide information about the number of dams or offspring examined. The actual results were not included either. Why was this study was considered adequate despite one weakness being** "lack of clarity on sample origins and sizes for each assay."

Page 220: Nishizawa et al. 2005 **This study used olive oil vehicle without testing for background estrogenicity. What were the results? Again, why was this study considered adequate if one weakness is** "lack of clarity on sample origins and sizes for each assay."

Page 222: Nishizawa et al. 2005 **This study used olive oil vehicle without testing for background estrogenicity. Why was this study considered adequate if one weakness**

is “lack of clarity on sample origins, uncertainty about statistical analyses, and sizes for each assay.”

Page 224: Imanishi et al. 2003 **This study used olive oil vehicle without testing for background estrogenicity. Why were the results removed?**

Page 225: Yoshino et al. 2004 **This study used ethanol/corn oil vehicle without testing for background estrogenicity. Why were the results removed?**

Page 227: Berger et al. 2007 **This study used peanut oil vehicle, and fed animals a diet containing soy without testing for background estrogenicity.**

Page 228: Kawai et al. 2003 **This study used corn oil vehicle without testing for background estrogenicity. This study tested some animals a total of 3 times, but the panel did not comment on the lack of a repeated measures statistical test.**

Page 229: Kawai et al. **This study used corn oil vehicle without testing for background estrogenicity. It is unclear, based on the experiments conducted, why the panel expected the authors to perform statistical procedures regarding** “nested factors associated with repeated measurements.”

Page 231: Markey et al. 2001 **In this case, the panel considers the** “administration of very low doses by subcutaneous pump” **to be a strength. This statement by the panel is odd:** “An additional weakness is that the proliferative changes reported in mammary tissues in virgin mice have not been satisfactorily established as precursors of breast cancer” **since this was not claimed by the study, and the point of the study was to illustrate the effects of exposure on tissue organization.**

Page 232: Markey et al. 2003 **Again, it is odd that, in this case, the panel considers the** “administration of very low doses by subcutaneous pump” **to be a strength. Why were the results of this study removed?**

Page 233: Vandenberg et al. 2007 **The report states:** “Strengths of this paper are the rigor with which the measurements were made, and the fact that the authors were trying to quantify endpoints that are difficult to measure (e.g., the relationship of the stroma to the epithelium). The relevance of the endpoints is a strength as is the low dose used.” This is inconsistent with the statement following the previous one “A weakness is inappropriate statistical analysis of a complex study design that may have produced too many positive findings”. **What does this mean:** “a weakness is inappropriate statistical analysis of a complex study design that may have produced too many positive findings”**? First, the study design was not complex: dams were exposed to either control or BPA. Single offspring were chosen for each morphological measurement so that litter effects were not needed. Each endpoint was analyzed separate from others. Also, the panel wrote that a weakness was** “the possible influence of the vehicle (50% DMSO) on the uptake and distribution of the bisphenol A.” **Where is the evidence for this statement? This is a vehicle recommended by the pump manufacturer. The differences observed between controls and BPA-exposed offspring negates this unwarranted comment.**

Page 235: Honma et al. 2002 **This study used sesame oil vehicle without testing for background estrogenicity. This study examined age of vaginal opening and age of first estrus, and the comment from the panel was that it is** “one of the few studies that

appropriately examines the onset of puberty in the mouse as an endpoint" **but it is unclear how these methods were appropriate when previous studies criticized by the panel used the same methods. Also, the panel concluded that** "the study is considered supplemental because of the subcutaneous route of exposure" **yet many other studies used this same route, and the same conclusion was not made.**

Page 236: Iwasaki and Totsukawa 2003 **Results were removed with no explanation.**

Page 236: Nikaido et al. 2004 **Results were removed with no explanation.**
Page 237: Park et al. 2005 **This study used corn oil vehicle without testing for background estrogenicity. This review is based on a translation provided by the American Plastics Council, which is concerning.**

Page 238: Park et al. 2005 **This study used corn oil vehicle without testing for background estrogenicity. Again, this review is based on a translation provided by the American Plastics Council, which is concerning.**

Page 238: Sato et al. 2001 **This study used olive oil vehicle without testing for background estrogenicity.**

Page 239: Rubin et al. 2006 **Why were the results removed? Why is it claimed that** "the uncertainty about sample size and relationship to litter" **are considered a weakness when it is clearly stated that one male and one female per litter were analyzed per time point and the size is above the arbitrary number the panel set as an acceptable one?**

Page 240: Toyama et al. 2005 **This study used olive oil vehicle without testing for background estrogenicity.**

Page 241: Berger et al. 2007 **This study fed BPA in a peanut butter supplement and fed animals a soy-containing food without testing for background estrogenicity. It is unclear why a Chi Square test would be an appropriate statistical test for body weight.**

Page 243: Kabuto et al. 2004 **What type of water bottles were used to deliver BPA? Why were the results removed?**

Page 244: Takao et al. 2003 **What type of water bottles were used to deliver BPA?**

Page 245: Matsumoto et al. 2004 **The study used a sample size of 3 for some endpoints. Why were the results removed?**

Page 245: Suzuki et al. 2003 **The sample size was 3 or 6 for some endpoints.**

Page 246: Tando et al. 2007 **Why were the results removed?**

Page 247: Mizuo et al. 2004 **Why were the results removed?**

Page 248: Miyatake et al. 2006 **This study used olive oil vehicle without testing for background estrogenicity.**

Page 248: Ryan and Vandenbergh 2006 **The number of dams per group was not indicated. Age of puberty was defined as** "the first day on which cornified cells were detected in 4-7 females/group" **and the panel referred to a strength of this work as** "the appropriate evaluation of pubertal onset" **even though this is an odd way to measure puberty, especially because the total number of females examined per group for this endpoint was not described. It was also stated that** "Bisphenol A 200 µg/kg bw/day resulted in a decrease in errors on earlier trials than the control in the radial arm maze, but this effect was not characterized by the authors as providing strong evidence of an alteration in spatial memory." **Was any other explanation given?**

Page 250: Tyl et al. 2006 **Several changes were referred to as having non dose-related effects and were therefore disregarded as** "unlikely to be of biological significance." **This explanation is unsatisfactory and may indicate non-monotonic dose response curves (see earlier). There are also 2 conflicting statements:** "strengths include… the thoroughness of the histologic evaluation" **but** "it did not appear that histopathological data were statistically analyzed" **.**

Page 251: Suzuki et al. 2002 **This study used sesame oil vehicle without testing for background estrogenicity. Also, it says that** "data tables list the sample size as 8-11/group/time period" **and that** "2 or 3 pups per litter were used in each analysis" **This implies that fewer than 6 litters were used for each group/time period.**

Page 253: Nikaido et al. 2005 **These animals were given water from polycarbonate bottles. Results were removed with no explanation.**
Page 254: Markey et al. 2005 **The panel questioned whether the litter or offspring was considered the statistical unit, but it clearly says that 1 pup was used per litter. Results were removed with no explanation.**

Page 255: Munoz de Toro et al. 2005 **Again, the panel questioned whether the litter or offspring was considered the statistical unit, but it clearly says that 1 pup was used per litter. Results were removed with no explanation. It also says that** "the statistics appear to be inappropriate in not accounting for repeated measures" **but this statistical test is not required for these endpoints.**

Page 256: Nakahashi et al. 2001 **This study used sesame oil or sesame oil/DMSO as a vehicle without testing for background estrogenicity.Results were removed with no explanation.**
Page 257: Aikawa et al. 2004 **This study used sesame oil vehicle without testing for background estrogenicity. Results were removed with no explanation.** Page 258: Toyama and Yuasa 2004 **This study used DMSO and olive oil vehicle without testing for background estrogenicity. There was a sample size of 5 animals/dose/timepoint for BPA exposed animals and 3-4 vehicle control mice, which the panel referred to as** "critically small sample size." **Results were removed with no explanation.** Page 259: Evans et al. 2004 **This study used corn oil/alcohol vehicle without testing for background estrogenicity.**

Page 260: Morrison et al. 2003 **This study used corn oil/alcohol vehicle without testing for background estrogenicity. There were significant issues with the statistics and a reduced sample size of 5, but this study was still considered adequate.**

Page 260: Savabieasfahani et al. 2006 **This study used cottonseed oil vehicle without testing for background estrogenicity. The small sample size was considered a weakness, but there were 11/group. Also, the panel said:** "the single time point for bisphenol A plasma determination at an unknown time relative to sc injection is a weakness." **However, it said that pregnant ewes were** "injected sc on GD 30-90" **and that** "maternal blood samples were taken on GD 50, 70 and 90" **so this is a confusing statement.**

Pages 262 to 275 included studies of invertebrates, frogs, fish, reptiles and birds. These studies were not considered relevant for human risk assessment. However, there were some panel comments that were inaccurate or confusing, so they are listed below.

Page 265: Oka et al. 2003 **20 µM BPA induced apoptosis in the nervous system of frog embryos, but 10 µM 17β-estradiol did not have the same effect, so the study authors concluded that** "the effects did not appear to occur through an estrogenic mechanism" **and the panel members wrote that** "the use of 17β-estradiol exposure to suggest a non-estrogenic mechanism of bisphenol A toxicity is a strength." **However, this description does not support the conclusion reached. No anti-estrogenic compounds such as ICI were tested.**

Page 267: Yang et al. 2005 **The panel stated that the cultureware was not discussed, but it clearly says that these experiments took place in tanks. Also, the panel's statement that** "the lack of attention to statistical analysis is a weakness and makes the authors' conclusions unreliable" **is a statement that could be applied to many other studies.**

Page 268: Kishida et al. 2001 **It says that** "Expression of the CYP450 aromatase gene was determined in 50 embryos/treatment group using an RT-PCR/Southern blot technique" **and also that** "the Southern blot analysis revealed a ~3-fold increase in the band intensity of CYP450 aromatase at the high concentration (10 µM) of bisphenol A." **There appears to be some mistake here. Southern blotting is used to analyze DNA content. Are the authors/panelists suggesting that BPA treatment changes the genetic content of the exposed fish? Also, it was stated that** "a weakness of this paper for the current evaluation is the lack of morphometric data." **This statement is unwarranted, since the authors intended to examine steroidal endpoints. Also, it was stated that** "The significance of the observed change in aromatase is not clear." **Altered aromatase expression will alter the hormal profile of the animal. If, in fact, expression of aromatase is increased by BPA treatment, this would indicate a possible shift in testosterone/estrogen ratio.**

Page 269: Segner et al. 2003 **The statement from the expert panel** "this study is useful in showing a lack of effect on fertilization at environmentally relevant concentrations of bisphenol A" **is confusing. First, the study examined doses of 94, 188, 375, 750 and 1500 µg/L; the Roepke study (page 262) examined doses of 250, 500, 750 and 1000 µg/L, which were described as** "levels exceeding environmentally relevant concentrations." **The null hypothesis for Segner's experiments is that BPA has no effect on the studied endpoints. If a statistically significant difference between vehicle and BPA-exposed animals is found, the null hypothesis is rejected (and, thus, BPA is described as having an effect.) If the null hypothesis is not rejected, it does not make the null hypothesis correct.**

Page 272: Berg et al. 2001 **It was stated that** "culture ware not discussed" **but this study used injection directly into the egg yolk, indicating that the egg itself was the cultureware.**

Page 273: Panzica et al. 2005 **It was stated that** "culture ware not discussed" **but this study used injection directly into the egg yolk, indicating that the egg itself was the cultureware.**

Page 274: Furuya et al. 2002 **This study used corn oil vehicle without testing for background estrogenicity.**

Page 274: Sashihara et al. 2006 **Why were birds provided continuous lighting? This study used ethanol and sesame oil vehicle without testing for background estrogenicity.** "This provides a vacuum for the interpretation of the dose-related increase in vocalizations that were reported." **This statement is unclear.**

Page 275: Furuya et al. 2006 **This study used alcohol/corn oil vehicle.**

Pages 276 to 280 are in vitro studies. These studies were not considered relevant for human risk assessment. However, there were some panel comments that were inaccurate or confusing, so they are listed below.

Page 276: Takai et al. 2000 **First, it is unclear how this study is not important for human risk assessment. It examines mouse embryos in a tissue culture dish. Is this not relevant for IVF? Second, these researchers are finding effects at the low and high end of the doses studied, indicating a biphasic (non-monotonic) dose response curve. Finally, the panel stated that** "the use of serum-free and phenol red-free media is an appropriate way to avoid estrogenic contamination" **but this is inaccurate. Many tissue culture products (tubes, dishes, etc.) are made from polycarbonate plastics. Many others contain estrogenic contaminants or ingredients. Using a sensitive assay for estrogens such as the ESCREEN can detect these estrogenic activities. Also, the panel states that this estrogen-free media** "is an artificial environment compared to the estrogen-rich milieu in which preimplantation embryos normally develop." **With the number of births originating from IVF, this is not necessarily true any longer.**

Page 276: Takai et al. 2001 **The same statements made above about estrogen contamination and the artificial environment are repeated for this study. It was also stated that** "this study did not evaluate the effect of exogenous bisphenol A under physiologic conditions." **What is meant by the "physiologic" condition, BPA is not physiological in any way.**

Pages 280 to 288, is the summary for the developmental toxicity data. This section is filled with inconsistencies and inaccurate statements.

Page 281: **It is stated that** "No increase in malformations was observed in mice with oral gavage of bisphenol A at doses of ≤1250 mg/kg bw/day (Morrissey 1987)." **The lowest dose studied in this report was 160mg/kg/day. This is not a dose relevant to humans.**

Page 283: **It is stated that** "Although some sporadic effects were reported for anogenital distance in male and female rats, study authors concluded that the endpoint was not affected by prenatal, lactational, and/or post-weaning exposure to bisphenol A." **The rat study by Rubin et al was cited here, but this conclusion was not taken from it. .**

Page 283: **It is stated that** "There were some indications that bisphenol A exposure may affect serum LH levels in male rats after exposure to ≤1.2 mg/kg bw/day administered during gestational or postnatal periods, but the biological significance of the effect was uncertain because of questions regarding exposure characterization, lack of dose response relationships, and reproducibility of the effect (Rubin 2001, Akingbemi 2004)." **What is meant by questions regarding exposure characterization? Is this even the result from Rubin et al study quoted here? In it, we only measured LH in long-term ovariectomized females.**

Page 283: **It is stated that** "no effects on prostate or sperm production were observed in more robust studies with multiple dose levels and larger group sizes." **There are no references listed here. I can only assume they are referring to the Cagen and Ashby studies. Both of these studies are flawed in that they did not replicate the vom Saal and Nagel studies, and both had positive controls that did not give effects, indicating that these studies were flawed in some way.**

Page 284: With regard to the Ramos 2001 study, it says "lack of dose response relationships were noted" **but the panel actually means that no linear/monotonic dose responses were noted.**

Page 284-5: "PIN- a lesion considered by consensus to be a prostate cancer precursor in humans, and of unclear significance in rodents (Ho, 2006)." **This statement is contradicted by data showing that rodent PIN could undergo neoplastic progression (Kim MJ, Bhatia-Gaur R, Banach-Petrosky WA, Desai N, Wang Y, Hayward SW, Cunha GR, Cardiff RD, Shen MM, Abate-Shen C. Nkx3.1 mutant mice recapitulate early stages of prostate carcinogenesis. Cancer Res. 2002 Jun 1;62):2999-3004).**

Page 285: "Hyperplastic or neoplastic changes (alone or in response to N-nitroso-N-methyl urea) were seen in mammary tissue of rats the dams of which were treated subcutaneously with bisphenol A 0.0025 or 0.025 mg/kg bw/day during gestation {Durando, 2007 #2450; Murray, 2007 #2451}. One single dose level study (using 50% DMSO vehicle in subcutaneous Silastic implants) reported that gestational exposure to 0.000250 mg/kg/d in mice caused quickened maturation of the mammary fat pad which was associated with increased ductal area and extension, decreased epithelial cell size, and altered collagen density in pubertal and adult mice." **In the first sentence, this studiy actually reported their presence in glands treated with 0.0025, 0.025, 0.25 and 1mg/kg/day. Silastic implants we not used in this study. A"quickened" maturation of the mammary fat pad-was not seen i.e. we did not see maturation at an earlier time; instead, a more advanced development was observed. Finally, the effects described (increased ductal area & extension, decreased epithelial cell size, altered collagen density) were all observed in fetal mammary glands. It was proposed that these alterations in tissue organization may lead to the changes we've observed at puberty and in adult life.**

Page 285: "At doses between 0.020 and 2.5 mg/kg bw/day, either no effect or accelerated vaginal opening was reported (Markey 2003, Iwasaki 2003)." **Markey used doses of 0.000025 and 0.00025 mg/kg/day, much lower than the doses described here.**

Page 286: “Changes in estrus cyclicity in mice were reported following gestational exposure to ≥0.02 mg/kg bw/day (Markey 2003, Honma 2002, Nikaido 2004) but the effects were not always dose-related.” **The Markey study used doses of 0.000025 and 0.00025mg/kg/day, much lower than the cited dose. Were the results not dose related, or were they not monotonic?**

Page 286: “Decreased uterine lamina propria volume, increased BrdU incorporation by uterine epithelial cells, and increased expression of progesterone receptor (not dose related) and ER by uterine epithelial cells was observed in mice at a bisphenol A dose of ≤0.000250 mg/kg bw/day (Markey 2005).” **The actual Markey 2005 paper reads: “The percentage of epithelial cells expressing detectable levels of PR was increased by approximately 14-fold in the 25 ng BPA/kg-BW per day group (89.9 ± 6.9; *P* < 0.01), and by approximately 13-fold in the 250 ng BPA/kg-BW per day group (82.6 ± 15.6; *P* < 0.01) relative to the control group (6.6 ± 2.3; Fig. 3A). As depicted in Figure 3B, females exposed to 25 and 250 ng BPA/kg-BW per day exhibited a 125-fold (74.5 ± 12.0; *P* < 0.01) and 124-fold (75.0 ± 15.7; *P* < 0.01) increase, respectively, in the number of cells showing high intensity (++) staining for PR relative to control females (0.6 ± 0.5).” showing that this is not a correct description of the findings.**

Page 286: “In mouse studies examining the effects of parenteral gestational exposure on the mammary gland, changes in the development of mammary structures, BrdU incorporation, and progesterone receptor expression by mammary epithelial cells were observed at a bisphenol A dose of ≥0.000025 mg/kg bw/day {Markey, 2001 #455;Markey, 2003 #2117;Muñoz-de-Toro, 2005 #644}, but the results were not always dose-related.” **Again, the panel is pointing out the non-monotonic effects of BPA exposure.**

Page 287: “Treatment of pregnant mice with bisphenol A 0.000250 mg/kg bw/day using sc minipumps resulted in loss of sexual dimorphism in number of tyrosine hydroxylase-positive neurons in the anteroventral periventricular preoptic area and loss of sexual dimorphism in open-field testing {Rubin, 2006 #2432}.” **Loss of sexual dimorphism in brain structures and behavior was also seen at 0.000025mg/kg/day.**

Page 287: “Two studies suggest no effect on thyroid function following oral exposure of rat dams during the gestation and lactation periods. A non-dose related increase in serum thyroxine levels was only observed on 1 of 4 evaluation periods during postnatal development at ≥1 mg/kg bw/day orally {Zoeller, 2005 #698}.” **This contrasts with the abstract from Zoeller’s study: “We now report that dietary exposure to BPA of Sprague Dawley rats during pregnancy and lactation causes an increase in serum total T4 in pups on postnatal d 15, but serum TSH was not different from controls. The expression of the TH-responsive gene RC3/neurogranin, measured by in situ hybridization, was significantly up-regulated by BPA in the dentate gyrus. These findings suggest that BPA acts as a TH antagonist on the beta-TR, which mediates the negative feedback effect of TH on the pituitary gland, but that BPA is less effective at antagonizing TH on the alpha-TR, leaving TRalpha-mediated events to respond to elevated T4.”**

Page 288: “Following oral exposure of mice to bisphenol A during gestation, changes 1 were observed for mRNA expression of arylhydrocarbon receptors, receptor repressor, or nuclear translocator and retinoic acid and retinoid X receptors in brain, testes, and/or ovary at 0.00002–20 mg/kg bw/day, {Nishizawa, 2005 #2226;Nishizawa, 2005 #665;Nishizawa, 2003 #760}. The strongest effects were found at the lowest doses following exposures during organogenesis (GD 6.5-13.5 or 6.5-17.5).” **First, a basic understanding of developmental**

biology shows that this is obvious. However, organogenesis of many organs including the brain continues postnatally, outside the windows described.

Detailed Comments on Section 4.0 Reproductive Toxicity

Page 315: Takeuchi et al. 2002 “ELISA has not been standardized for human sera, and may over-estimate bisphenol A due to non-specific binding (see Section 1.1.4), although the authors cited a 0.97 correlation between this assay and the better quality HPLC analysis.” **With a correlation of 0.97, it appears that ELISA replicates the biochemical assay quite well.**

Page 316: Takeuchi et al. 2004 **Same comments as above.**

Page 317-9: Sugiura-Ogasawara et al. 2005 **Under the section for utility, it states:** “Because of the design and analysis flaws in this work, the only conclusion that can be drawn is that for some reason, very small numbers of women (maybe 4 or 5) with repeated spontaneous abortions have elevated levels of bisphenol A and that a few of these also have elevated antinuclear antibody levels. The reasons for those few individuals having elevated levels are not clear.” **The tone of this statement is inappropriate.**

Page 319: Yang et al. 2006 **A proper description of BPA levels in urine would include the enzymatic treatments that were used to examine the urine to determine what percent was parent BPA, BPA-gluc and BPA-sulf. A careful examination of this paper shows that these authors treated samples with glucunonidase, so the detected BPA is likely BPA-gluc (+ any parent BPA).**

Page 320: Luconi et al. 2001 “While a novel approach, this study used a concentration of bisphenol A that was lower than some of the 17β-estradiol concentrations, despite a significant literature showing that bisphenol A is less efficient at binding the ER than is 17β-estradiol…” **First, the panel writes that the concentrations used in this study were:** “1 µM bisphenol A, 1 µM 17β-estradiol, 10 µM progesterone” **indicating that this study used the same concentration of bisphenol A and estradiol. Additionally, while BPA is less efficient at binding to nuclear ER than estradiol, it binds to membrane ER at the same or higher affinity. The authors did not use any method to prevent binding to potential ER in the membrane.**

Page 320: Hanaoka et al. 2002 “A plausible (P=0.07) correlation between bisphenol A and decreasing FSH was reported. The authors took care to note that all levels were within the clinical normal range.” **I assume this statement is about the normal range for FSH and not BPA, which has no clinical normal range as it is an exogenous contaminant. Also, it was stated that** “According to Brock et al. (2001), urine glucuronides of bisphenol A are a longer-lived (12-48h) biomarker, so the sampling appears o have been appropriate.” **Because the cited study was not a metabolism study, where a dose of BPA was administered and then followed through the excretory pathway, and instead was a study sampling urine and testing it for BPA concentrations, this statement is inappropriate.**

Pages 322 through 390 address reproductive toxicity studies in animals treated with BPA during sexual maturity. (Studies examining exposure during fetal development, neonatal development, and puberty were assessed in Section 3.2.)

Page 322: Goloubkova et al. 2000 **This study used very small sample sizes (3-4 for some endpoints, 6-8 for others) but this was not mentioned as a weakness. Why were the results removed?**

Page 323: Funabashi et al. 2001 **This study used sesame oil vehicle without testing for background estrogenicity. The sample size was 6 dams/group (not mentioned as a weakness by the panel) and the dose administered was not adjusted for animal weight (also not mentioned as a weakness). The lack of info on husbandry was not included as a weakness.**

Page 323: Yamasaki et al. 2002 **The funding source is not listed, but an examination of papers from this groups shows they work for "**Chemicals Evaluation and Research Institute of Japan" **a group hired by industry to perform tests on chemicals. This study used olive oil vehicle without testing for background estrogenicity. No info was reported on bedding (but not mentioned by the panel.) It says that 3 doses of ethinyl estradiol were tested, but then notes from the panel include:** "… females from the mid- and/or high-dose 17β-estradiol group experienced alterations…" **The panel did not mention the very high doses as a weakness. It also states:** "The study authors concluded that change in estrus cyclicity was the only useful endpoint for evaluating the endocrine-mediated effects of bisphenol A." **Does this mean that the authors noted changes in estrus cyclicity, or are they suggesting that changes in the weight and histopathology of the uterus, vagina and ovary are not reliable endpoints for assessing endocrine-mediated effects.**

Page 324: Spencer et al. 2002 **This study used a sample size of 5 for at least some endpoints, but that was not considered a weakness by the panel. The lack of info on husbandry was not included as a weakness.**

Page 325: Funabashi et al. 2003 **This study used sesame oil vehicle without testing for background estrogenicity. Very small sample sizes were used for some experiments (n=3-5 for immunohistochemistry, n=3-4 for dose response curve, n=5-8 for behavior experiments). Notes from the panel state** "the number of animals per group is sufficient given the nature of this study design" **but does not make sense according to the criteria established by the panel itself, and a basic understanding of the variability in histological and behavioral endpoints. It was also stated that** "brains were removed and fixed in 2% paraformaldehyde." **This is probably an insufficient fixation for an adult rat brain; histological examinations of adult brains normally require perfusion for proper assessments. The lack of info on husbandry was not included as a weakness.**

Page 326: Funabashi et al. 2004 **This study used sesame oil vehicle without testing for background estrogenicity. Small sample sizes (n=6/group or n=5-6/group) were not noted as a weakness. It was stated that** "the study authors concluded that bisphenol A can alter the neocortical function through the progesterone receptor in adult rats, but the physiological significance of the effect is not known." **The panel should consider that changes from normal physiology are always significant. It was also noted that a weakness was** "only one dose level administered at a single time point" **but this study assessed the effects of BPA exposure 0, 6, 12 and 24 hours after injection. If the panel means that this study is flawed because only adult animals were used, then all studies assessed in Section 4.1 are equally flawed. Finally, it was seen as a weakness** "the failure to examine any physiological or functional endpoints" **but it was noted that BPA exposure leads to** "the induction of progesterone receptor mRNA, an estrogenic response."

Is the panel suggesting that this is not a functional endpoint? The lack of info on husbandry was not included as a weakness.

Page 326: Della Seta et al. 2005 **This study used peanut oil vehicle without testing for background estrogenicity. The actual sample size assessed in the behavior experiments was 7-9, a fairly small sample size for behavioral parameters. Also, it was not stated whether the statistics accounted for multiple samples from the same dams.**

Page 327: Park et al. 2004 **This study used ethanol/corn oil vehicle without testing for background estrogenicity. Additional animals were examined that were not injected with vehicle or BPA, but the results from these animals were not discussed, so it cannot be determined if the vehicle was having some effect. The lack of info on husbandry was not included as a weakness.**

Page 328: Berger et al. 2007 **This study used peanut oil vehicle without testing for background estrogenicity. It was stated that a weakness was** "faulty methodology" **but this was not explained.**

Page 328: Al-Hiyasat et al. 2004 **This study was criticized for a relatively small sample size for some endpoints, but is still considered marginally adequate. It was also stated that** "confirmation of mating was not performed (cohabitation was for 10 days; if the mice mated on day 10, the necropsy would have been performed on GD7." **This statement is confusing. Proper examination of a uterus 7 days after conception would allow for confirmation of mating.**

Page 330: Matsumoto et al. 2004 **Based on the results discussed by the panel, it is not clear why this study appears in this section in addition to section 3. The lack of a defined sample size and no discussion of litter effect statistics for offspring endpoints were not noted as weaknesses by the panel. The lack of info on husbandry was also not included as a weakness.**

Page 330: Nieminen et al. 2002 **It is stated that animals were exposed to BPA** "in feed at concentrations providing doses of 0, 10, 50, or 250 mg/kg bw/day for 2 weeks." **If exposure is occurring through feed, the exposure dose must be approximate. This should be stated. The type of feed used and other housing conditions were not described. This study was also criticized for the** "absence of reproductive endpoints." **If the panel believes this is the case, why is it included in a section on "Reproductive Toxicity".**

Page 331: Nieminen et al. 2002 **It was stated that** "mortality was significantly increased by bisphenol A treatment, with incidences of 18, 36, and 20% in the low- to high-dose groups. No mortality was observed in the control group." **These percentages do not make sense considering the sample sizes for each group of 7, 5 and 8. Also,** "the study authors concluded that wild mammals such as field voles could be more susceptible to bisphenol A-induced toxicity than laboratory rodents." **There were no differences noted in pooled LH levels (pools from males + females) but a basic knowledge of endocrinology would indicate that pooling male and female sera for this assay is inappropriate. Finally, the panel stated that** "this study provides evidence that voles are more sensitive (based on mortality) to sc bisphenol A administration than rats or mice." **To make this statement, the same study would have to test and compare all 3 species, an experiment that has not been performed.**

Page 332: Razzoli et al. 2005 **This study used corn oil vehicle without testing for background estrogenicity. It was stated that females received ethinyl estradiol as a positive control, but under study strengths, the positive control was listed as 17β-estradiol. In this study, cohabitation occurred for 21 days. It states that** "during the cohabitation period, social behavior (e.g. agonism, social investigation, huddling, and nest sharing) was observed and body weights of females were measured." **How often were the behavioral parameters examined, and in which light cycle?**

Page 333: Oehlmann et al. 2000 **It is stated by the panel that** "bisphenol A has stimulatory (17β-estradiol-like) effects on the spawning masses and eggs of snails. These changes occurred in the absence of a histological correlate." **What histological correlate to increased egg production would be suggested?**

Page 333: Forbes et al. 2007 **There was no mention of cultureware, and the panel did not point this out as in previous studies. What was the vehicle used to dissolve BPA? It was unclear if the authors used repeated measures statistics or some type of litter effect statistic.**

Page 334: Schirling et al. 2006 **There appears to be an error in units here: these snails were not exposed to 50 or 100 g BPA/L. It was stated that** "ethinyl estradiol treatment significantly decreased embryo heart rate compared to the water control group but not compared to the DMSO control" **but then the authors were criticized for** "the comparison of ethinyl estradiol to the water control instead of the DMSO control."

Page 334: Xu et al. 2002 **The cultureware and type of media used were not described, so it is not possible to determine if other estrogenic contaminants were present. The section for strengths/weaknesses is full of statements that are unclear. What is meant by:** "Because this study used in vitro study [sic] PMSG-stimulated murine cells, metabolism is likely to have been minimal (if present at all)" **Metabolism of BPA or metabolism of nutrients in the culture medium?** "…in vitro dosimetry of bisphenol A is difficult to extrapolate to in vivo dosimetry." "Based on the data presented the mechanism by which bisphenol A may be inducing cell cytotoxicity/apoptosis is likely not "endocrine disruptor" mediated." **What is endocrine disruptor mediated? In addition, no evidence has been presented that this effect is occurring through a non-endocrine mechanism (nor has any evidence been presented showing that it is working through an endocrine mechanism, other than the fact that BPA is known to have estrogenic activity in vitro and in vivo.**

Page 335: Mlynarcikova et al. 2005 **Again, the section for strengths/weaknesses has many confusing statements. First,** "Potential estrogenic effects were observed at 10^{-5} M bisphenol A. Decreased in responses observed at the 10^{-4} M concentration are likely due to nonspecific cytotoxicity" **or these data may indicate the presence of a non-monotonic dose response.** "There was no mention of whether phenol red-free media were used or whether fetal bovine serum was charcoal-stripped." **There was no mention of the type of cultureware used; as stated earlier, many plastics used in tissue culture are contaminated or engineered with estrogenic chemicals.** "With exception of the highest dose level, there was no dose response (inconsistent trends)" **The two parts of this statement are contradictory.** "The statistical flags are potentially due to random chance." **All findings are potentially due to random chance (i.e. with a p value=0.05, there is a 1 in 20 probability that the effect is due to random chance alone.)**

Page 336: Mohri & Yoshida 2005 **It is stated that** "estrogens may affect the oocyte by regulating calcium oscillations and that bisphenol A could affect oocyte maturation" **but** "It is unclear if calcium oscillations play a role in oocyte maturation in other species, including humans." **Perhaps the panel is not familiar with the importance of calcium oscillations in oocyte development in response to fertilization. See a recent review by Ducibella, Schultz & Ozil (2006).**

Page 336: Yamasaki et al. 2002 **This study used olive oil vehicle without testing for background estrogenicity. Again, it was stated that** "change in estrus cyclicity was the only useful endpoint for evaluating the endocrine-mediated effects of bisphenol A." **This is confusing for 2 reasons: first, this section is about male rats (which do not have estrus cycles); second, there are obvious effects reported on prostate and testis weight, 2 other endocrine-related endpoints.**

Page 337: Takahashi & Oishi 2001 **The panel's description of the methods are unclear. 4-week old animals were fed BPA in their diets for 44 days, and** "rats were killed when mean body weight of controls reached ~200 g." **So, how much time elapsed between the end of the dosing period and the time of kill? Also, it was stated that** "food intake was said to be slightly decreased according to dose" **so were the estimated intake values corrected for these differences?**

Page 338: Sakaue et al. 2001 **This study used ethanol/corn oil vehicle without testing for background estrogenicity. The panel noted the small sample sizes (n=5 or n=8 for different endpoints) but this study was still considered adequate. What is meant by the statement that** "no histopathological correlate was presented"**? If a human demonstrated altered sperm count, would this be insignificant without a testis biopsy?**

Page 339: Ashby et al. 2003 **It is stated that the support for this study is not indicated but these authors are clearly associated with the chemical industry. This study used ethanol/corn oil vehicle without testing for background estrogenicity. It was noted that:** "Subtle genetic differences in the rats were suggested as a possible reason for differences in results between the 2 studies." **Does this indicate that the rats were purchased from 2 different vendors, i.e. Ashby used the Charles River strain, a rat that has been suggested to be less responsive to estrogens? Finally, a study that finds no positive results and fails to replicate a previous study, but does not contain a positive control (a weakness not pointed out by the panel) may indicate a technical error.**

Page 341: Tohei et al. 2001 **This study used sesame oil vehicle without testing for background estrogenicity. Comments made by the expert panel include:** "SC is not a relevant route of exposure" **a fact that has been repeated throughout this document but is not true.** "Blood collection via decapitation is not appropriate, because decapitation stress affects plasma prolactin and LH secretion." **Any method requiring the handling of animals at time of kill could potentially affect plasma prolactin and LH secretion.** "No mention is made of the order of killing. If controls were killed first and the guillotine was not cleaned between uses (and animals were not in separate rooms), there may be serious confounding of the data." **This statement is inappropriate and unwarranted.** "Because rat plasma testosterone levels are normally highly variable, the low degree of variability in this study, given the small sample size, is remarkable." **Again, the tone of this statement is inappropriate.**

Page 342: Kim et al. 2002 **This study was translated by the American Plastics Council, making its presentation questionable. It was stated that** "parts of organs were preserved in formalin and examined histologically" **and** "no histopathological alterations were reported for the testis, epididymis, seminal vesicle, prostate, spleen, or brain." **In a previous review, the panel asserted that fixation of testis with formalin is inappropriate (Takahashi & Oishi 2001). Additionally, histology on adult brains that were not perfused is not accurate either.**

Page 343: Chitra et al. 2003 **This study used olive oil vehicle without testing for background estrogenicity. Interestingly, the comments from the panel include this vehicle as a possible problem:** "A potential [sic] significant concern in this study is the use of olive oil as the vehicle. The stability/reactivity of bisphenol A was not determined and it is possible that bisphenol A interacted with olive oil, resulting in the observed findings." **It is odd for the panel to make this statement here. A) The stability/reactivity of BPA has not been demonstrated in any of the other oils used, either. B) 10 studies cited in Section 3 (and 2 earlier studies in Section 4) used olive oil vehicle, but this point was not brought up. C) The panel is still not acknowledging the fact that many oils are contaminated with estrogenic substances. Additionally, this study had a sample size of 6 but was still considered adequate. No information was provided about cages or bedding, but this was not considered a weakness by the panel.**

Page 344: Chitra et al. 2003 **This study used olive oil vehicle without testing for background estrogenicity (the problem addressed in the previous study was not mentioned here.) Why were the results removed?**

Page 345: Saito et al. 2003 **This study used corn oil vehicle without testing for background estrogenicity. The lack of information on husbandry was not considered a weakness by the panel. Doses given were not adjusted for weight of the animals. It was noted that a problem was** "the use of an inappropriate method of anesthesia when measuring hormone levels" **but no details were provided to verify this statement.**

Page 345: Takahashi & Oishi 2003 **Propylene glycol was used as a vehicle in this study. The sample sizes for different endpoints were 5-6 and 8, but the panel seems to regard the use of small numbers as good:** "This paper reports a comprehensive study comparing 2 mouse and 2 rat strains using minimal numbers of animals per group." **Some animals were given BPA in their diets, but the type of chow used was not described so it is unknown how much estrogenic activity it contained. It was stated that** "the study authors concluded that bisphenol A is more toxic through sc and ip exposure routes than by oral exposures in the diet." **The panel concludes that the results from this study** "suggest that systemic exposure is necessary for bisphenol A estrogenic activity to be exhibited and strongly indicate that route of administration (oral vs. ip) is an important consideration." **First, this conclusion could only apply to animals exposed in adulthood, as was done in this study. However, this conclusion cannot be supported by this study because different doses were used in the different modes of exposure (and only a single dose was examined for oral exposure) The animals injected (ip and sc) were only injected 4 days/week while the animals exposed in diet were exposed constantly, making these two modes of exposure difficult to compare.**

Page 347: Herath et al. 2004 **This study used a soy-containing commercial feed without testing for background estrogenicity, and this feed expected to have estrogenic**

activity, but this was not included as a weakness by the panel. A single dose was examined, but this was not considered a weakness. It was stated that "a major weakness of this paper is the inconsistency in the hormone data (control data after 2 weeks were dramatically different than after 5 weeks even though both are from sexually mature rats." **However, treatment began at 50 days of age, which is around the time of puberty in the rat.**

Page 348: Toyama et al. 2004 **This study used DMSO/olive oil vehicle without testing for background estrogenicity (DMSO considered a weakness by the panel). The lack of information on husbandry was not considered a weakness.**

Page 349: Takao et al. 1999 **This study exposed animals to BPA in their drinking water, but it was stated that** "water intake was significantly reduced [by 8%] in the high-dose group exposed for 4 weeks" **indicating that the exposure levels did not remain constant throughout the experiment (and actual exposure concentrations are unknown).**

Page 349: Al-Hiyasat et al. 2002 **This study used sample sizes of 10 mice/group, a number greater than several studies that were considered adequate in this section. However, the panel's comments include:** "number of animals per group was too small for a definitive assessment of study endpoints" **and** "weaknesses include small sample sizes for endpoints." **The panel commented that** "there was no indication that mating was confirmed by checking for sperm in the vagina" **but in mice, actual sperm are not normally assessed; the presence of a copulatory plug is seen at the vaginal opening. The panel also commented that** "the method of randomization (or initial body weights) was not presented." **First, this information is lacking from most of the studies cited in this section. Second, the only acceptable way to truly randomize animals is using a random number generator, not distributing them to different groups by body weight (as some previous studies describe.) The panel noted** "an absence of a dose response in several of the endpoints assessed" **when, in fact, some of these endpoints may have a non-monotonic dose response curve. The panel also noted that** "given that mice usually have poor (relative to rats) fertility rates, collectively the controls in this study are suspect." **It's unsure whether the panelists are referring to actual studies that show this or their own personal experience, but this is in direct conflict with our own experiences and observations (i.e. within 4-5 days of pairing 2 females per male, generally >80% demonstrate copulatory plugs.)**

Page 351: Nagao et al. 2002 **Under strengths/weaknesses, the panel states that a strength was the** "examination of 2 strains (1 which demonstrated a greater sensitivity to 17β-estradiol)" **but, in fact, this study did not examine the effects of BPA on 2 different strains. As stated earlier,** "an initial experiment… found that C57BL/6N mice were more sensitive to 17β-estradiol than ICR mice, and the study authors therefore used C57BL/6N mice to examine the effects of bisphenol A." **Additionally, a strength was listed as** "inclusion of 17β-estradiol as a positive control" **but as per the above statement, it was not included in the BPA studies as a positive control. Either the panelists did not summarize this study properly, or they did not understand the experiments conducted.**

Page 351: Peknicova et al. 2002 **It is stated that mice were exposed through drinking water to** "doses of 0.000002 and 0.000020 mg '/animal's weight/day'." **What is meant by this? It appears that the dose was not changed to match the animal's weight, and the actual dose must be approximate if it is administered in drinking water. The vehicle**

used to dose the water was not listed. The lack of information on husbandry was not considered a weakness by the panel. There was no discussion of the type of fixative used for histopathological examination of the testis, as in earlier studies.

Page 352: Takahashi & Oishi 2003 **Mice used in this study were 4 weeks old at the beginning of exposure; this is pubertal development in mice. It was stated that** "relative testis weight was not significantly affected, but when the value from 1 mouse with a high relative testis weight was deleted, the effect attained statistical significance." **Was this mouse identified as a statistical outlier using the 2-sigma statistical approach, or was it just deleted arbitrarily? Removing an animal because it doesn't 'fit' is not statistically appropriate. It was also stated that** "the study authors concluded that the testicular toxicity of bisphenol A is 'relatively weak,' based on the co-occurrence of liver and kidney toxicity at exposure levels causing testicular effects." **This statement is not scientifically accurate or when considering that BPA exposure (like exposure to other hormones) is likely to have effects at multiple levels of organization, i.e. gene expression, cellular endpoints, tissue/organ and whole body.**

Page 353: Park et al. 2004 **This study used ethanol/corn oil vehicle without testing for background estrogenicity, but a second group of control animals were not injected with vehicle. (Results for this second group of controls were not presented, so it is unknown if the vehicle was having some effect.) The lack of information on husbandry was not considered a weakness by the panel.**

Page 353: Toyama et al. 2004 **This study used DMSO/olive oil vehicle without testing for background estrogenicity (DMSO considered a weakness by the panel). No information was provided about fixation methods for testes. The lack of information on husbandry was not considered a weakness by the panel.**

Page 354: Anahara et al. 2006 **This study used corn oil vehicle without testing for background estrogenicity. Included as a weakness:** "There were no apparent differences in levels of protein expression between various estrogenic agents/treatments." **Unexpected results are not automatically wrong or flawed. It was also stated that** "no adverse outcomes of the changes in cortactin were explored" **but it was previously stated that** "cortactin is one of several proteins that control spermatid development." **The potential issues associated with a change in this protein are clear. The lack of information on husbandry was not considered a weakness by the panel.**

Page 355: Moon et al. 2001 **This study used corn oil vehicle without testing for background estrogenicity. In the strengths/weaknesses section, it was stated that** "there is no evidence that bisphenol A has any effect on the ability to attain an erection resulting in copulation in mice or rats." **Is the panel referring to a particular study that assessed this endpoint, or are they assuming that because BPA-exposed animals are fertile that they do not have problems with penile erections? Fertility is not necessarily indicative of whether or not animals have problems achieving erections at some points.**

Page 355: Nieminen et al. 2002 **Large amounts of experimental details should be included in this section. It was also stated:** "This study provides evidence that the bisphenol A administered to polecats increases GST and UDPGT activity. Since these findings were dose-related it appears that in the polecat bisphenol A increases phase 2 metabolism but has minimal effects on hormone levels." **But then,** "Due to the limited

number of animals and the absence of a dose-response relationship, the hormonal changes in this study are difficult to interpret." **These two statements are contradictory. It was also stated that a weakness was the** "absence of reproductive endpoints" **but then why was this study included in this section?**

Page 356: Nieminen et al. 2002 **This whole section is directly repeated from the earlier citation for this study. In fact, even the references to "males" were mistakenly left and not switched to "females".**

Page 357: Kinnberg & Toft 2003 **What dose of estradiol was used as a positive control? The panel writes** "There was no apparent low-dose effect" **does this mean there was no observed low dose effect for the endpoint studied (gross malformations of the testes).**

Page 358: Nikula et al. 1999 **The panel fails to recognize that plastic cultureware could potentially contaminate these experiments. It was stated that** "this study provides compelling evidence that the actions of bisphenol A maybe [sic] non-estrogen mediated." **First, it is assumed that the panel meant non-estrogen receptor mediated. Second, this experiment does not support this conclusion. The effects of a single dose of BPA were not the same as the effects of a single dose of estrogen. This does not mean that the effects of BPA are not ER-mediated. The use of ICI would have helped support this statement.**

Page 359: Murono et al. 2001 **The description of the results and the authors' conclusions do not match. BPA was found to have no effect on any endpoint described by the panel (cell viability, basal or hCG-induced testosterone production, hCG binding to LH receptors, or conversion of hydroxycholesterol to testosterone.) However,** "the study authors noted the similarity of effect between bisphenol A and 17β-estradiol." **This is confusing. Was stripped serum used in this experiment (not mentioned by the panel)?**

Page 359: Akingbemi et al. 2004 **What doses were used for the positive controls? Again, the panel refers to a study being** "well-conducted by a respected lab" **indicating that the panel is making judgments based on investigator reputation and not necessarily on the robustness of the cited study. The results from this study clearly demonstrate that a non-monotonic dose response curve is present for the assessed endpoints. In fact, the panel writes that** "these data are highly relevant and suggest the occurrence of selective low dose effects in the absence of high dose effects."

Page 360: Song et al. 2002 **What is meant by** "methods used in this study are described in conjunction with the results." **Is this a factor of the journal style? It was stated that** "This study appears to have been well conducted and links in vitro bisphenol A administration to dose-related (classic, not inverted) activation of *Nur77* and subsequent downstream signal transducing proteins. Various confirmatory experiments supported this relationship." **This is not true, as the subsequent experiments only examined a single dose and therefore were not able to confirm the prevalence of a monotonic dose response curve at additional endpoints. It was also stated that** "the toxicological implications of these findings were not addressed" **but were the endocrinological implications (or some other implication) addressed instead?**

Page 361: Hughes et al. 2000 **This study examined doses of 100, 300 or 600 μM BPA. Why weren't lower (and higher) doses studied? In cell culture studies, it is generally accepted that doses should span several log phases (i.e. should test 1×10^{-10} through 1×10^{-4}). It was stated that** "bisphenol A inhibited calcium ATPase activity in rat testis microsomes." **At which doses was this effect observed?**

Page 362: Tabuchi et al. 2002 **Again, this study uses a very narrow range of BPA doses for an in vitro study. This study performed a microarray on cells exposed to vehicle or 200 μM BPA, even though this dose of BPA reduced cell viability by 33%. Of the 1081 genes examined, 3 were downregulated and 10 were upregulated by exposure (although, with a p-value of 0.05, 50 genes are expected to be changed by chance. So, what p-value was used to select these genes? It was also stated that** "the study authors concluded that microarray analysis is a useful tool for investing [sic] molecular mechanisms of bisphenol A-induced toxicity in testicular cells." **However, the genes were not followed through to demonstrate that these changes are relevant, so this conclusion is not supported by the study performed.**

Page 362: Tabuchi & Kondo 2003 **What is meant by "mock cells"?**

Page 363: Tabuchi et al. 2006 **How do the authors explain that** "cell viability was decreased" **but** "there was no evidence of apoptosis"**? The panel includes as a strength the use of** "state-of-the-art technology". **Are they referring to the use of microarrays, which has been used for many years? What is meant by the statement:** "It is not surprising that given this 'high' bisphenol A concentration, 'novel' and likely non-specific gene changes were noted"**?**

Page 364: General Electric unpublished studies **These studies have not been published and are not peer-reviewed. Thus, they do not belong in this review.**

Page 364: Ema et al. 2001 **Large portions of this section are directly repeated from Section 3 and only the effects on the adult F0 generation are needed in this section. At the end of the first generation, it was stated that** "1 or 2 F1 weanlings/litter/sex (25/sex/group) were selected to continue in the study." **How did the authors deal with the uneven contributions (1 vs 2) from different litters? It was stated that there were** "non-dose-related decreases in percentages of females with normal estrus cycles (76 versus 96% in controls)." **At which dose(s) was this observed? A strength was that** "the concentrations of the dosing solutions were verified (both prior and after.)" **How were the concentrations verified after dosing occurred?**

Page 366: Tyl et al. 2002 **Again, large portions of this section are directly repeated from Section 3 and only the effects on the adult F0 generation are needed in this section. It is important to note how different the actual intakes of BPA were from the target intakes, but also how much variability there was in intake within each dosing group. It says that** "30 F1 offspring/sex/group were randomly selected." **Was this selection 1 animal per litter or how were litter effects accounted for? It was also stated that** "up to 3 F1 weanlings/sex/litter were killed for organ weight measurements" **but this indicates a variable number used from different litters. How were litter effects accounted for? It was stated that** "Some small (~5%) decreases in pup body weight during the lactation period at lower doses were apparently not considered treatment-related by study authors." **Why not? These authors found essentially no effects of BPA exposure (**"the study authors concluded hat bisphenol A should not be considered a

selective reproductive toxicant”**); any study that finds no positive results and also did not use a positive control is suspicious. With regard to the CERHR’s review of this paper, it was said that:** “the unnecessary use of ANOVA with Dunnett test was noted. Some possible outliers and 10-fold errors in data points that could have effected conclusions were observed.” **Yet it was also stated that** “statistical methods were well thought out and appropriate.” **These two statements are contradictory.**

Page 372: NTP 1984 & Morrissey et al. 1989 **The lowest dose examined in this study was estimated at 6.25mg (the panel did not state if this is 6.25mg/kg/day or 6.25mg/mouse/day.) In either case, not a single dose approaching relevant human exposures was examined. It was stated that** “the study authors noted that reproductive tract weight in the high dose group was greater [by 52%] than in the control group but statistical significance was not achieved because of high variability.” **Were these animals killed in proestrus to decrease potential variability? On page 372 line 43, this should state** “task 3”**. It was stated that** “Dosing was continued throughout the breeding and separation periods. However implants were often expelled through cutaneous lesions or the incision site.” **This indicates that the surgical implantation was performed improperly! Additionally, “**When animals expelled their implant, a new one was inserted but pregnant mice were allowed to complete their pregnancy before insertion of the new implant. Therefore dosing was not uniform.” **This non-uniform dosing is concerning. Animals that lose their implants should be removed from the study. Finally, it was stated that** “this study provides evidence that the manifestation of maternal toxicity is dependent on the route of administration and that route-dependent metabolism may be important for toxicity.” **The animals in this study were only dosed s.c. so this conclusion is unsupported.**

Page 373: NTP 1985 & Morrissey et al. 1989 **This study used polycarbonate cages. It was stated that animals were randomly assigned to treatment groups** “according to body weight” **but this indicates that assignments were not random. It is highly concerning that experiments pairing BPA-exposed animals with non-exposed animals for mating were not exposed to BPA during the mating period of 21 days. However, the conclusion from this study was that** “the cross-over test revealed no effect on mating or fertility in either males or females exposed to bisphenol A” **and it was later stated that** “therefore, direct exposure to bisphenol A was minimal or nonexistent during sperm maturation, capacitation, and ovulation.” **Was there some indirect exposure instead?**

Page 377: Tyl et al. 2002 **Again, this entire summary is copied from Section 3. The same issues apply here as well.**

Page 379: Tyl et al. 2006 **Again, this summary is repeated from Section 3. The same issues apply here as well.**

Page 387: Kwak et al. 2001 **It was stated that** “reproductive impairment occurs in a concentration-dependent manner” **but the panel does not describe how reproductive impairment was assessed. The use of** “only 3 concentrations” **was considered a weakness but then the panel states:** “of note is the classic dose response obtained in this apparently sensitive model.”

Page 387: Sohoni et al. 2001 **The list of strengths and weaknesses for this study were very confusing:** “General toxicity” was identified and good histology was used. The conclusions regarding weak estrogenic activity were appropriate at 160 μg/L and higher. Other effects were likely due to general toxicity.”

Page 388: Kang et al. 2002 **It was stated that** "this study exhibited classic dose responses in the affected endpoint" **but only 2 doses of BPA were examined.**

Page 390: Utility of Reproductive Toxicity Data **It is stated that** "there are 2 studies that measured serum bisphenol A in healthy women, women with polycystic ovary syndrome, and healthy men and evaluated correlations with serum gonadotropins, prolactin, testosterone, and other androgens." **The panel should be aware that there are other studies that measured BPA levels as well as sex hormones in healthy individuals including Takeuchi & Tsutsumi 2002 and Takeuchi et al 2004.**

Page 390: Summary of Reproductive Toxicity Data **The section with reference to the Charles River SD rat is lacking citations, so it is impossible to determine whether the panel is correctly interpreting the results or conclusions of these studies.**

Page 391: "Human reproductive studies are summarized in ." **This sentence is incomplete.**

Page 391: "Sugiura-Ogasawara et al. (2005) attempted to examine an association first observed in laboratory animal studies…" **What does the panel mean, "attempted to examine"? The tone is inappropriate.**

Page 393: "One study reported estrus cycle alterations in offspring of rats given 1.2mg/kg bw/day bisphenol A in drinking water from GD 6 through the lactation period (Rubin 2001). Estrous cycle alterations were not reported in other rat oral exposure studies covering a wide range of dose [sic] (<1-475 mg/kg bw/day) administered during all or part of the gestational or lactational periods." **How did these other studies assess estrous cycle abnormalities?**

Page 393: "There were some indications that bisphenol A exposure may affect serum LH levels in male rats after exposure to ≤1.2 mg/kg bw/day administered during gestational or postnatal periods, but the biological significance of the effect was uncertain because of questions regarding exposure characterization, lack of dose response relationships, and reproducibility of the effect {Rubin, 2001 #521;Akingbemi, 2004 #2104}." **Again, we would like to reiterate that the Rubin study examined female rats after long-term ovariectomy while the Akingbemi study examined male rats. These sexes are not expected to have the same LH levels. Additionally, the Rubin study did not have a "lack of dose response" because significant effects were only observed in the highest dose administered.**

Page 393: **With regard to the "irreproducibility" of the Nagel and vom Saal prostate studies:** "No effects on prostate or sperm production were observed in more robust studies with multiple dose levels and larger group sizes." **What the panel fails to mention here is that these "more robust" studies were not performed properly because their positive controls did not give positive results.**

Page 394: "Changes in estrous cyclicity were reported following gestational exposure to ≥0.002 mg/kg bw/day (Markey 2003)" **but the Markey reference used doses of 0.000025 and 0.00025 mg/kg bw/day.**

Page 394: A decrease in the number of mice with corpora lutea was observed following gestational exposure to 10 mg/kg bw/day and increases in polyovular follicles were observed following neonatal exposure to 100 mg/kg bw/day {Suzuki, 2002 #556}." **The Markey 2003 paper also demonstrated that exposure to 0.00025 mg/kg/day during gestation led to a significant increase in the percentage of ovarian tissue occupied by antral follicles in 3-month-old mice.**

Page 394: "In mouse studies examining the effects of parenteral gestational exposure on the mammary gland, changes in the development of mammary structures, BrdU incorporation, and progesterone receptor expression by mammary epithelial cells were observed at a bisphenol A dose of ≥0.000025 {Markey, 2001 #455;Markey, 2003 #2117;Muñoz-de-Toro, 2005 #644}, but the results were not always dose-related and there was no consistency of response at different evaluation time periods." **With regard to the results** "not always dose-related" **the panel is reminded of the extensive literature in the field of endocrinology demonstrating the presence of non-monotonic dose response curves. With regard to the statement that** "there was no consistency of response at different evaluation time periods" **this statement appeared in the first draft in reference to the different levels of BrdU incorporation at different ages. Differences in BrdU labeling have been noted based on age of the animal, but not dependent on the time of evaluation. Basic understanding of the mammary gland reveals that differences in BrdU incorporation at different points of development are expected; the mammary gland is an organ undergoing constant reorganization and remodeling**.