COMMENTS ON THE INTERIM CERHR REPORT ON BISPHENOL A (APRIL, 2007)

Frederick S. vom Saal, Ph.D.
Professor of Biology
Division of Biological Sciences
University of Missouri
Columbia, MO 65211

GENERAL COMMENTS

It is disturbing that I submitted comments after the posting of the initial draft of this report identifying factual errors in the report, and these errors were not corrected.

My overall assessment of this report is that the initial draft, written by Sciences International (SI), contained numerous factual errors. Importantly, these errors were not evenly distributed in that they were focused on discrediting government-funded studies conducted primarily in academic institutions by recognized authorities in a variety of disciplines that reported significant effects of bisphenol A (BPA). The present interim draft has done nothing to rectify this blatant pro-industry bias. This bias is consistent with a statement uncovered in a letter (posted at EWG.org) from the former president of SI, Elizabeth Anderson, to a client (RJR Tobacco Company); she wrote that the association with government agencies gave SI "unique credibility to negotiate with regulators on behalf of our private sector clients". This report is the "smoking gun" demonstrating that SI engaged in a systematic distortion of the science to benefit its corporate clients that manufacture millions of pounds per year of this chemical.

The decision by the CERHR to continue to use the flawed report initially prepared by SI after deciding to fire SI due to undisclosed conflicts of interest requires investigation by an independent agency as opposed to an internal review conducted by the NTP. My view is that an independent investigation of this report will confirm that it contains inaccurate information and that the inaccuracies are aimed at discrediting government-funded research. The idea that a group of scientists, including a number working for corporations, without a background working with BPA (a stated criterion for being chosen to serve on this panel) could go through over 700 articles on BPA and rectify the myriad of errors made by SI is not logical. Also, what will be apparent to any knowledgeable reviewer is the glaring failure to "connect the dots", with the result that studies that should be looked at as a "package" of information are treated as independent, unrelated observations (I provide examples below). This, of course, reflects the lack of expertise with this specific subject (BPA) by the members of this panel. The errors and lack of attention to critical issues are a serious concern and demonstrate that selection of at least a few panel members directly involved in conducting research on BPA, and with prior familiarity with the published BPA literature, was needed. It is clearly time that the policy of the NTP that having demonstrated expertise disqualifies someone from serving on a CERHR panel be re-evaluated by an outside agency.

That an agency (the NTP) within NIH would allow this tainted review process to continue is disturbing. There should thus also be concern with the other reports (about 20) that were

prepared by SI, and again an independent agency, not the NTP, should be charged with re-examining these reports to determine whether factual errors and bias are also found in these other reports. Government agencies charged with self-examination (which is what is now occurring with regard to the prior CERHR reports prepared by SI) are prone to conclude that all prior actions were correct.

If one were seeking to establish a mechanism that would be virtually certain to underestimate the potential for harm to be caused by a chemical, the CERHR mechanism is exactly the process that they would want to establish to achieve that objective.

EXAMPLES OF ERRORS AND EVIDENCE OF BIAS IN THE REPORT

ISSUE 1: TYPE OF FEED

On P 210 the panel stated: **"The Purina 5001 chow has high and variable levels of soy phytoestrogens, and the corn cob bedding is known to be problematic due to antiestrogenic constituents." On P 219** regarding the Palanza et al 2002 EHP study, the panel stated**: "The use of a diet high in soy isoflavones is an additional weakness.".** These statements are clearly intended to reduce the credibility of the findings we reported. The panel member(s) responsible for this statement should consider the following:

In all research conducted in my laboratory a combination of Purina 5008 pregnancy and lactation chow and Purina 5001 maintenance chow have been used. Neither of these rodent feeds is as variable as Purina 5002, which was used in the Tyl et al. 2003 study as well as in the Cagen et al. 1999 study (the panel identified this study by Cagen as a "replication" of the Nagel et al. 1997 study conducted in my laboratory, which is inaccurate given that they did not use the same feed). Thigpen et al. (2003 Comp Med) reported that some batches of Purina 5002 feed could interfere with the detection of DES effects, accounting for the failure of both the Cagen and Tyl studies to find effects of BPA, and the failure of Cagen to also find any effects of the positive control, DES.

Using the combination of 5008/5001 we have reported that an increase in free serum eestradiol of 0.1 pg/ml (parts per trillion delivered via the mother by Silastic capsule s.c.) in fetal CF-1 mice significantly increased the size of the fetal prostate and caused malformation of the urethra; this slight increase in fetal serum estradiol also permanently increased prostate size and increased prostate androgen receptors measured in adulthood. We showed the same effect with oral administration to pregnant CF-1 mice (via feeding using a pipetter to reduce stress associated with gavage) DES at 0.02, 0.2 and 2 micrograms/kg/day maternal body weight (vom Saal et al. 1997, PNAS). In addition, we showed the same effects on the fetal prostate (additional gland formation, hyperplasia of the basal epithelial cells in the developing glands, and gross urethral malformations) of a 0.1 microgram/kg/day maternal body weight dose of either DES or ethinylestradiol as well as a 10 microgram/kg/day dose of BPA fed to pregnant CD-1 mice (Timms et al. 2005, PNAS). We then removed the prostatic region of the embryonic urogenital sinus in CD-1 male mice and cultured the mesenchyme cells (these are the estrogen and androgen-responsive cells) and showed that a dose of estradiol (0.27 pg/ml; parts per trillion) that was

exactly the free concentration that stimulated the fetal prostate in our prior *in vivo* experiment (vom Saal et al. 1997, PNAS) stimulated a significant increase in androgen receptor gene activity (Richter et al. 2007, EHP), thus ruling out any interference in the response of this tissue to estradiol by the phytoestrogens in 5001 feed. In this same experiment, the concentration of BPA that significantly stimulated both androgen and estrogen receptor gene activity was 0.23 ng/ml (parts per billion) which is exactly in the range of human fetal exposure to unconjugated BPA (Schonfelder et al. 2002, EHP). Using the same Purina feeds we have reported that fetal exposure to a maternal oral dose of BPA or 2.4 micrograms/kg/day significantly increased postnatal weight gain and advanced the age at first estrus in female CF-1 mice (Howdeshell et al. 1999, Nature). Feeding pregnant CD-1 mice 10 micrograms/kg/day BPA also led to impaired maternal behavior (Palanza et al. 2002, EHP). Finally, a large number of studies conducted over a 20-year period with these feeds showed differences in postnatal phenotype due to very small differences in fetal testosterone and estradiol levels based on an animal's intrauterine position relative to siblings of the same or opposite sex. How could anyone aware of these findings propose that the use of the 5008 and 5001 feeds was "problematic"? Clearly, over the last 30 years of conducting research on mice with these feeds we have provided extensive evidence for effects of low doses of estrogen.

Every major finding discussed above has been replicated: Gupta (2000, PSEBM), a recognized authority on male reproductive system development, replicated our prostate findings (while neither Ashby et al. 1999 or Cagen et al. 1999 could replicate this finding, none of the authors on either of these industry-funded papers had ever conducted an experiment related to the male reproductive system, and in each case, they also reported no effect of their positive control, DES). The intrauterine position effects have also been shown by Dr. John Vandenbergh in numerous studies with mice, and these effects have also been reported by a large number of other investigators. Clearly the feeds we used did not disrupt our ability to detect these very low dose effects caused by endogenous hormones, since these findings were replicated by researchers using other species and mouse strains and different feeds.

The above set of studies, which represent a small subset of the large number of studies concerning effects of low doses of endogenous hormones, hormonally active drugs, and endocrine disrupting chemicals conducted in my laboratory using these same feeds over the last 30 years provides an example of "connecting the dots". The panel failed to integrate the information from these different studies and thus made a criticism regarding the feed that was inconsistent with 30 years of published research as the basis for reducing the importance of studies reporting adverse effects of BPA, and this is also the case in other portions of the report regarding other issues. With specific regard to the critical issues relating to BPA, this is not surprising, given that none of the panel members came into this process with a prior knowledge of the massive (> 700 published articles) BPA literature. There are independent academic researchers who have extensive knowledge of this large literature on BPA, and they were all explicitly excluded from membership on this panel.

ISSUE 2: ROUTE OF ADMINISTRATION AS THE BASIS FOR ELIMINATING STUDIES FROM CONSIDERATION BY THE PANEL

On P 130 the panel states:

"the Panel carefully considered the value of studies where Bisphenol A was administered anywhere other than to the mouth or stomach of the experimental animal. 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."

MY COMMENTS: The BPA panel eliminated from consideration any study that did not involve oral administration. However, human blood and urine levels of BPA are greater than any current model of exposure predicts should be found. Because we cannot account for the high levels of BPA found in virtually all humans (median values in the parts per billion), it is likely that there are multiple routes of exposure to this ubiquitous environmental contaminant that is produced in excess of 6 billion pounds per year worldwide. The only way that blood and urine levels can be in the part per billion range in virtually everyone examined is that there is continuous exposure to BPA via multiple routes, a prediction which directly contradicts that of the panel. The decision to eliminate from consideration any animal study that did not administer BPA orally is thus inappropriate and inconsistent with the data from a large number of biomonitoring studies. This prediction of exposure from multiple sources is consistent with information presented in the report on **P 4. Section** "*1.2.3.1 Environmental fate and bisphenol A levels in environment* Bisphenol A may be present in the environment as a result of direct releases from manufacturing or processing facilities, fugitive emission during processing and handling, or release of unreacted monomer from products {European-Union, 2003 #2146}. According to the Toxics Release Inventory database, total environmental release of bisphenol A in 2004 was 181,768 pounds, with releases of 132,256 pounds to air, 3533 pounds to water, 172 pounds to underground injection, and 45,807 pounds to land {TRI, 2004 #2251}." Given the unexplained high levels of BPA in humans, these are likely underestimates of the actual environmental contamination that is occurring as BPA-containing items are discarded into terrestrial and aquatic ecosystems.

In addition, the members of the panel did not take into account that while there are data showing that route of administration impacts the rate of clearance of BPA in adults, due to more rapid metabolism of orally administered BPA, no such data exist for newborn rodent pups. Newborn pups have been shown to have limited capacity to metabolize BPA, and the differences in rate of metabolism seen in adults associated with different routes of administration would thus not be expected to occur in pups. In summary, particularly during development, all routes of administration provide valuable information regarding the potential health outcomes of exposure to BPA and should be considered in evaluating the potential for adverse health effects due to human exposure to BPA.

Finally, below I will discuss in more detail the confusion that obviously exists among panel

members about the issue of "false positive" effects noted in their statement above. Regardless of route of administration, it is astonishing that the panel would conclude that administration of BPA to pregnant and lactating rodents via an infusion pump s.c. at a dose 2-million times lower than the current published LOAEL (see studies from Ana Soto and Beverley Rubin laboratories showing effects on offspring of BPA at a maternal dose of 0.025 microgram/kg/day), would produce "false positive" effects because there was continuous exposure rather than exposure one time per day via oral administration. First, if the findings are reliable, and many have been repeated in multiple experiments and are thus clearly reliable, then they cannot be categorized as "false positives".

The decision to completely ignore all non-oral administration BPA research is a demonstration of a lack of understanding by the panel regarding the importance of explaining the high exposure to BPA that has been reported for humans in many studies, a lack of understanding that metabolic studies in adults are irrelevant for fetuses and newborn babies (babies are not little adults is the maxim in pediatric medicine), and a lack of understanding of what constitutes a "false positive" effect.

In summary, route of administration is an invalid basis for eliminating studies from consideration, and all such studies should be considered in evaluating the health effects of BPA.

ISSUE 3: INACCURATE DESCRIPTION OF THE CONTENT OF AN ARTICLE

On P 209-210 the panel reviewed vom Saal et al. (1998 Tox Ind Sci) and deemed it of limited use due to statistical issues. The discussion of the statistical issues is **entirely inaccurate.** These inaccurate statements were identified in detail in the review I previously submitted in response to the initial draft but have been ignored by the panel.

Regarding my published study (vom Saal et al 1998), I have to wonder how carefully this report was prepared if the lack of clarity about the strain of mouse used and the method of assigning males to groups could be included as a criticism. The beginning of the "Animals" section of this publication (see p 243) states: "CF-1 mice were purchased from Charles River Laboratories…". Regarding the selection of males, it was stated that they were "randomly selected" and that the males were those used to report the prostate data in the Nagel et al study, where we also clearly identified that only one randomly selected male per litter was used. Testis weight and the analysis of it (by ANCOVA) are extensively discussed on p 246-7, and the testis weight data are presented in Table 2. It appears that the reviewer failed to note that the daily sperm production data were adjusted for testis weight (DSP per gram of testis). The statement about the need for knowledge of testis weight is thus absolutely correct, but if this was recognized as being so important, how could the reviewer fail to know that we had taken this into account. Even a superficial examination of the paper will reveal that this criticism is invalid based on looking at the title of Figure 1 from this publication, which states "Daily Sperm Production Per Gram Testis". This appears to be a blatant attempt to decrease the impact of these findings by making up false criticisms.

On P 283 it states: "One group of investigators reported decreased sperm production efficiency (LOAEL 0.020 mg/kg bw/day) {Vom Saal, 1998 #187} **and increased prostate weight at 0.002**

but not 0.020 mg/kg bw/day {Nagel, 1997 #6;Vom Saal, 1998 #187} in offspring of mouse dams exposed during pregnancy.**"**

The underlined statement is inaccurate. We reported statistically significant effects on the prostate of BPA at both 0.002 and 0.020 mg/kg/day.

On P 98 [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 members of the panel who wrote this obviously did not read the Nagel et al. 1997 study, specifically the discussion on p 74 regarding the implications for estrogen responsiveness of exposure to xenoestrogens such as BPA due to of the lack of binding to high affinity plasma proteins; this is also discussed in the companion paper (vom Saal et al. 1998) describing the effects on other male reproductive organs in addition to the prostate. Our prediction was not focused just on the prostate but on all estrogen-responsive endpoints, in direct contradiction to the conclusion reached by the panel.

On P 214 the review of Howdeshell et al. (1999 Nature) demonstrates confusion by the reviewer(s). We examined female siblings identified as developing between 2 males (2M females), between a male and a female fetus (1M females) and not next to a male (0M females), and found that the effects of fetal exposure to a very low dose of BPA on their postnatal rate of growth and timing of peripubertal events was greatest in the females with the highest fetal levels of estradiol (0M females) and lowest in the females with the lowest fetal levels of endogenous esttradiol, suggesting additive effects. There is a huge literature on intrauterine position effects. The statistical criticisms used to reach the conclusion that this study had "marginally utility" for evaluation purposes are: "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."

The person who wrote this had no understanding of this paper and that siblings within litters from different intrauterine positions were being compared. In addition, the reviewer appears to be unaware that he/she was reading a brief communication in *Nature*, which does not allow for inclusion of all of the details deemed essential to the reviewer. Thus, an article deemed important enough to be accepted for publication in the most prestigious journal in the world is dismissed by this panel. The statement that we did not provide the NIH Low Dose Panel with the data from this experiment is inaccurate. We were told not to submit these data, since the statistics sub-panel was overwhelmed with too many other studies to analyze, and the data from this experiment were not in an Excel spread sheet.

On P 215 in the analysis of the Gupta (2000 PSEBM) article showing that fetal exposure to BPA via feeding the mother stimulated an increase in prostate weight at 3 times during postnatal life, the following criticism was made by the panel: "[It was not clear if the offspring or litter was considered the statistical unit.]". On page 63 in the article by Gupta it states: "For determining organ weight and biochemical analyses 15 offspring (1 from each of the 15 litters) were used for

each time point." What could be clearer than that statement by Gupta. In the "strengths and weaknesses" section a criticism is: "the apparent lack of attention to possible litter effects."

On P 216, regarding the criticisms of the Gupta 2000 paper, the entire section about the criticism by Elswick et al. should be deleted. This criticism was based on a model proposed by Elswick et al. 2000 that was deemed "misleading", "illogical" and "flawed" by the NIH Low Dose Review Panel. The CERHR BPA panel has very selectively quoted from the prior Low Dose Peer Review Panel report. It is thus noteworthy that the Elswick et al 2000 study, which was deemed to have been essentially a fraudulent misrepresentation of statistically significant findings of a increase in prostate size caused by developmental exposure to bisphenol A but was presented as being not statistically stignificant by Elswick, led to pages of scathing criticism of the statistics by the Low Dose Peer Review Panel. A pathetic attempt by Elswick to create a statistical model suggesting that their results were really false positive findings, which was shown to be complete nonsense in the Low Dose Panel Review report, somehow escaped the notice of this CERHR panel, and, in fact, ended up being used as the basis for criticizing another well conducted study. See the review of the Elswick paper, which forms the basis for the criticism of the Gupta article, on p A86 - A91 of the Endocrine Disruptors Low Dose Peer Review at: http://ntp.niehs.nih.gov/index.cfm?objectid=06F5CE98-E82F-8182-7FA81C02D3690D47

This is another example of bias in the initial report prepared by SI. It should be noted that the Elswick et al 2000 study was conducted at the Chemical Industry Institute of Toxicology (CIIT) in Research Triangle Park, NC, and is consistent with the conclusions of all other industry-funded studies, 100% of which report that BPA causes no significant effects at low doses. The report prepared by SI, which has corporate clients who manufacture BPA, is thus consistent with all other publications about BPA from organizations associated with or receiving funds from industry.

ISSUE 4: STATISTICAL REASONS FOR REJECTING ARTICLES FOR EVALUATION
The panel identified on **P 130 "an insufficient number of animals for rigorous statistical analysis… (in the case of a small n) that insufficient animals were analyzed to provide confidence in the results."** as a basis for eliminating studies from consideration. There is a fundamental statistical fallacy here, namely, that a specific number of animals is required to have "reliable" (repeatable) results. I attended the first panel meeting and listened to members of the panel discuss this issue, and was stunned to find that there were members of the panel who believed that it was appropriate to set establish numbers of animals that they would accept without regard to what one should learn in introductory statistics: the ability to find statistically significant effects is based on precision (error variance) and magnitude of the effect. As a member of the University of Missouri Animal Care and Use Committee (ACUC), we required investigators to justify the number of animals to be used based on conducting a power analysis, and experienced investigators typically had prior data that provides information about the number of animals required to achieve statistical significance based on estimates of variance and the expected magnitude of effect. If an investigator had used the criteria established by this panel as the basis for establishing numbers of animals per group, the proposal would be rejected by any ACUC as not in compliance with the NIH guide for the care and use of animals. This rationale by

the panel to eliminate studies from consideration is clearly inappropriate in that it is inconsistent with NIH policy to use the fewest number of animals required to achieve statistical significance.

A number of studies conducted in my laboratory were deemed of marginal utility due to statistical issues. The criticisms demonstrate a serious flaw in the panel report in that statistically significant findings (such as those in Nagel et al 1997 EHP) were deemed inadequate (presumably false positive findings) due to what the panel considered to be a small sample size, even though the findings were statistically significant and subsequently replicated in other studies conducted in our laboratory. This makes absolutely no sense at all. The fact that other investigators conducting what they were reporting as replications of our experiments, but which in fact were conducted using different procedures, is not a legitimate basis for discounting our findings, particularly when most of the "negative" results were from studies funded by chemical companies or their trade organizations.

ISSUE 5: FURTHER EVIDENCE OF BIAS IN THE INITIAL SI REPORT AND IN THE APPLICATION OF STUDY REJECTION CRITERIA BY THE PANEL

On **P 213** an example of clear bias in the application of the issue of group size is the study by Ashby et al 1999 in which adult female mice were reported to not differ in their uterine weight as a result of prenatal exposure to BPA or DES. In the review written by SI and not challenged or commented on by members of the panel, SI stated: "In female offspring from the bisphenol A groups, there were no significant effects on body weight or organ weights, including cervix, uterus, vagina, and ovary." Anyone examining the literature would know that in intact female mice, uterine weight varies over 4-5 fold during the estrous cycle. We found a significant (30%) increase in uterine weight in adult female mouse offspring in response to exposure to a low 0.1 microgram/kg/day DES dose administered on gestation day 11-17 (this dose also stimulated an increase in prostate weight in males). But, to find this we ovariectomized the adult females and (based on extensive pilot data) administered a dose of estradiol via Silastic capsule to the ovariectomized females that stimulated uterine growth, but not maximally. We thus revealed a prenatal DES effect by conducting an experiment in which small differences could be revealed and background variance was very low (Alworth et al. Tox Appl Pharm 183:10, 2002). In sharp contrast, Ashby et al 1999 examined 36 control gonadally intact females from 8 litters for uterine weight and compared BPA and DES treated females to these control females. The lowest uterine weight value in the control group was 67 mg and the largest uterus was 290 mg (range = 223 mg). The mean for the 36 females was 133 mg and the SD was 52 mg (this is based on data provided by Ashby to the NIH Low Dose Peer Review Panel that met in 2000). Power analysis conducted on these data demonstrate insufficient power to find the predicted magnitude of effect, given that there was no attempt to control for variability in uterine weight by controlling stage of estrous cycle or by ovariectomy and estrogen treatment. This flaw should have been noted and the study should have been deemed unusable by the panel.

SUMMARY

The lack of attention to gross design and statistical flaws in industry-sponsored studies provides evidence of bias in the initial report prepared by SI staff. The lack of awareness or concern with these flaws by the members of the CERHR BPA panel means that this report

cannot be trusted to provide regulatory agencies with an accurate, unbiased assessment of the potential human health effects of bisphenol A.

The flaws and bias identified above are just examples chosen to demonstrate that this report is unfit for use by the CERHR. There is no way for any panel to correct all of the errors in this report.