February 2, 2007

Dr. Michael D. Shelby
CERHR Director, NIEHS
P.O. Box 12233
MD ED-32
Research Triangle Park, NC 27709

Sent as e-mail attachment to: Shelby@niehs.nih.gov

Re: Comments on the draft CERHR report of December 2006 on the reproductive and developmental toxicity of bisphenol A

Dear Dr. Shelby:

I am pleased to provide these comments on behalf of the Polycarbonate/BPA Global Group in regard to the draft CERHR report on the reproductive and developmental toxicity of bisphenol A. The Polycarbonate/BPA Global Group represents the leading global manufacturers of bisphenol A and polycarbonate plastic, who for many years have supported and conducted scientific research to understand whether bisphenol A has the potential to cause health or environmental effects and to support scientifically sound public policy.

We have also supported critical reviews by scientific experts of the many studies relevant to whether bisphenol A has the potential to cause health or environmental effects. When properly conducted, such reviews can be of high value to support public policy, guide future research and reduce controversy. Consequently, our comments are provided in the spirit of promoting a well-conducted and scientifically sound expert panel evaluation of the scientific evidence regarding the potential reproductive and developmental toxicity of bisphenol A.

Our attached comments are divided into three sections along with a series of supporting tables. The first section describes the need for a weight of evidence approach to the CERHR evaluation of bisphenol A and discusses several overarching aspects of the evaluation. The second section discusses several considerations specific to bisphenol A that are critical for a scientifically sound assessment. The third section provides detailed comments on all sections of the draft CERHR report.

Given the extremely large volume of research on bisphenol A, we recognize that this evaluation has already required an enormous effort from the CERHR staff and the expert panel. However, there is quite a bit of work left to be done, some of which we believe must be done in advance of the expert panel meeting at which conclusions will be drawn. In particular, as outlined in our comments, there are two areas that will require work in advance of the panel meeting. First, a

number of studies are missing from the draft report that should be included. Although the missing studies may not substantially change the weight of evidence or conclusions that are drawn, they should be considered for inclusion in the interest of completeness.

Second, a number of studies included in the draft report have been inconsistently evaluated in regard to their utility and it is not apparent if uniform criteria were used to evaluate all relevant studies. Regardless of the reason, it is critical that individual studies be consistently evaluated with uniform criteria in order to reach sound and defensible conclusions.

We appreciate the effort from all involved in this evaluation and look forward to a successful conclusion. Please do not hesitate to contact me if I can be of further assistance to clarify any comments or if additional information is needed. I can be reached at (703) 741-5588 or by e-mail at steve_hentges@plastics.org.

Regards,

Steven G. Hentges, Ph.D.
Executive Director
Polycarbonate/BPA Global Group

Attachments

Comments of the Polycarbonate/BPA Global Group
on the December 2006 Draft CERHR Report on the
Reproductive and Developmental Toxicity of Bisphenol A

February 2, 2007

1. General Considerations for a Sound Scientific Evaluation of Bisphenol A
   - **a.** A weight of evidence approach is needed for the evaluation
   - **b.** Weight of evidence framework
   - **c.** Clear criteria are needed for review of individual studies with consistent application across all studies
   - **d.** Guideline studies and use of Good Laboratory Practices should be given additional weight
   - **e.** All positive and negative findings from relevant studies should be included in the weight of evidence

2. Considerations Specific to Bisphenol A that are Critical for a Scientifically Sound Evaluation
   - **a.** Human exposure
   - **b.** Route of exposure
   - **c.** Sub-mammalian studies

3. Detailed Comments on Draft CERHR Report
   - **a.** Section 1
   - **b.** Section 2
   - **c.** Sections 3 and 4

1. General Considerations for a Sound Scientific Evaluation of Bisphenol A

a. A weight of evidence approach is needed for the evaluation

A complete evaluation of the potential risks of bisphenol A (BPA) to reproduction and development requires, at a minimum, detailed knowledge of several areas of research:

- Metabolism and pharmacokinetics
- Human exposure
- Reproductive and developmental toxicity

The draft CERHR report appropriately includes sections on each of these key areas and, as is evident from the report, bisphenol A is one of the best studied of all substances. Each of these sections includes a large number of published studies, which vary substantially in size, scope, quality and relevance to human health.

Many types of studies are included in the draft CERHR report, ranging from *in vitro* studies, to small-scale *in vivo* studies with limited scope, to large-scale multi-generational studies with broad scope. Likewise, the endpoints and findings reported in these studies cover a wide range. To reach sound conclusions, the panel must distinguish between toxicological effects and other physiological responses that do not result in an adverse effect. The panel must also determine what data are most relevant for judging the potential of bisphenol A to present a risk to human reproductive health.

It is not clear from the draft CERHR report how the panel will use the large amount of relevant information available on bisphenol A to reach conclusions. Given the large number and diversity of studies, it is of particular importance for the panel to apply a weight of evidence approach to systematically review and critically analyze the studies, leading to transparent conclusions based on all of the relevant evidence.

The need for a rigorous weight of evidence approach was highlighted in a November 2005 report from the EU Scientific Committee on Health and Environmental Risks, which noted that reported bisphenol A low-dose effects have not been reproduced.[1] The committee recommended that "assessment of such effects requires a rigorous and science based weight-of-the-evidence approach, which needs to consider that the findings at low doses represent changes without, or, at best, unknown toxicological significance."

Very recently a scientific panel of the European Food Safety Authority conducted a weight of evidence assessment of bisphenol A and established a Tolerable Daily Intake.[2] That assessment is incorporated in its entirety in these comments as Attachment 1.

---

[1] Scientific Committee on Health and Environmental Risks. Opinion on Endocrine Disrupting Chemicals: a Non-animal Testing Approach. November 25, 2005. Available on the Internet at http://europa.eu.int/comm/health/ph_risk/committees/04_scher/docs/scher_o_015.pdf.

[2] European Food Safety Authority. Opinion of the Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food on a request from the Commission related to 2,2-bis(4-hydroxyphenyl)propane (Bisphenol A). Adopted 29 November 2006.

We strongly encourage the CERHR panel to apply a weight of evidence approach to reach sound conclusions based on the many studies summarized in the draft report. Specific aspects that should be considered and applied are included in the following parts of this section.

b. Weight of evidence framework

A suitable weight of evidence framework includes the analytical elements described below.[3]

Elements Focused on Internal Validity

- Rigor – Studies should be evaluated for their proper conduct and analysis. Greater weight should be given to better-conducted studies. Some studies may have been performed so poorly that their results should be substantially or entirely discounted.

- Power – The statistical power of an experimental design should be examined for its ability to detect effects of a given magnitude. Studies of higher statistical power should be given higher weight over studies of lower statistical power that are otherwise comparable.

- Corroboration – The replication of findings among similar studies and the observation of similar effects under relevant conditions may increase the confidence that the findings represent a real effect in experimental animals as replication is a fundamental principle of the scientific process. Conversely, lack of corroboration across many studies with similar experimental conditions is grounds to doubt the validity of an experimental result reported in one or a few studies. In a multi-generational study, an exposure-related effect should appear across generations.

Elements Focused on External Validity

- Universality – The degree to which a finding (i.e., either the presence or absence of an effect) is consistently reproduced in validated study designs and test systems increases the confidence that it is valid and may apply to humans. In contrast, if an effect is restricted to a certain species, strain, or route of administration, the ability to generalize the response to other species or routes becomes more questionable.

- Proximity – When clearly established effects have been shown in a species with greater similarity to humans or at a dose level similar to that expected in humans, such results weigh more heavily than those in dissimilar species, by inappropriate routes, or at markedly different dose levels. In this regard, the significant route dependency and species dependency in the metabolism and pharmacokinetics of bisphenol A must be taken into account. In humans, the oral route of exposure is most relevant. By this

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

route in humans, bisphenol A is subject to virtually complete pre-systemic clearance by first-pass metabolism to a primary metabolite (BPA-glucuronide) that is then rapidly eliminated in urine. The primary metabolite has not known biological activity and, in particular, has been shown to have no estrogenic activity.

- Relevance – From what is known about the underlying biological basis for a toxic response in animals, it may be possible to judge (based on knowledge of animal and human physiology) whether similar metabolism, mechanisms of damage and their repair, and molecular targets of action should be expected to operate in humans. Confidence in applicability to humans can increase or decrease accordingly.

- Cohesion – The extent to which all of the data are consistent and are subject to a single, biologically plausible explanation increases weight compared to a situation where inconsistencies require ad hoc explanations and exceptions to general patterns. A common hypothesis is that bisphenol A may act via an estrogenic mode of action at low doses to cause adverse effects. The weight of evidence analysis should therefore examine the data to determine if there is a replicable pattern of estrogenic effects within and across the many studies. Lack of a consistent pattern of effects significantly reduces the biological plausibility of the hypothesis that bisphenol A acts via an estrogenic mechanism to cause reproductive and developmental effects at low doses.

c. Clear criteria are needed for review of studies with consistent application across all studies

Studies used in the evaluation must have both suitable quality and relevance

As a first step in a weight of evidence evaluation, each study within the scope of the evaluation must first be individually reviewed to determine its utility for the evaluation. This step must be completed consistently based on clear criteria for utility of studies within the context of the evaluation.

The utility of a study is a function of two principal aspects: (a) the quality and reliability of its data and (b) the applicability and relevance of those data to addressing a particular question about potential human hazard or risk.

The first aspect – study quality – has to do with the rigor of its design, the adequacy of its sample size, the absence of potential confounding factors, the accuracy of the measurement of endpoints, and the sufficiency of data presentation in the publication. These qualities affect how reliable the data may be judged to be for whatever application, and a study with poor quality will have low utility for any purpose since its outcomes are of uncertain validity and cannot be relied upon.

The second aspect – study applicability – addresses how well data from even a well conducted and reliable study addresses the question at hand. A published definition for utility of studies in CERHR evaluations[4] includes the critical aspect of applicability: "utility (applicability to the purpose of drawing conclusions on whether or not a chemical adversely affects reproduction)."

A study may have high utility for some questions and lower utility for others. Direct demonstrations of a phenomenon of interest, for example the generation of reproductive dysfunction in live animals dosed at levels and by a route applicable to human experience, will have more applicability than studies that only indirectly address the primary hazard question. For instance, studies in ovariectomized animals or studies dosing animals by injection may have uses in illuminating modes of action, but they do not themselves demonstrate a hazard relevant to any real human circumstances.

For a study to have high utility, its results must be both reliable and relevant to human exposure scenarios. The panel must keep in mind that some studies may be judged reliable but not relevant to the purpose of this evaluation, which is whether bisphenol A is likely to be a potential reproductive or developmental hazard or risk to humans.

CERHR should apply uniform criteria for review of studies

Guidelines for this step have been published[4] and are also available on the CERHR website. Section III of either document lists the following evaluation and quality criteria:

[4] Shelby, M. D. National Toxicology Program Center for the Evaluation of Risks to Human Reproduction: Guidelines for CERHR expert panel members. Birth Defects Research (Part B) 2005; 74: 9-16.

Evaluation Criteria

• Numbers of animals in each treated and control group • Species, strain, and sex • Age at beginning and end of treatment • Route of administration • Purity of substance • Solvent or vehicle • Controls (untreated, solvent, positive) • Doses • Dosing schedule • Basis for dose selection	• Duration of treatment • Endpoints observed • Method of examination • Age at observation • Number of animals and/or litters observed • Statistical methods utilized • Statistical significance • Author's conclusions supported by the data • GLP study

Quality Criteria

- Was an appropriate number of animals used?
- Were they randomly assigned to experimental groups?
- Was the test chemical defined, source and purity stated?
- Was there chemical verification of dosing preparations?
- Were age of animals and duration of exposure appropriate?
- Were appropriate endpoints observed?
- Were endpoints observed at appropriate life stages?
- Were data reported in appropriate detail?
- Were appropriate statistics employed?

Although the CERHR criteria name most of the important factors to consider, they do not emphasize the quality/relevance distinction. For the relevance question, the CERHR criteria name appropriate factors, but there is little indication about what values for those factors render a study more or less useful for the question at hand, which is whether bisphenol A is likely to be a potential reproductive or developmental hazard or risk to humans.

Study review criteria have been inconsistently applied

Sections 3 and 4 of the draft CERHR report summarize a large number of studies, identify some of their strengths and weaknesses, and assign utility rankings to each study.

The study review criteria, however, seem to have been applied inconsistently. A notable example where this is readily apparent is the inconsistent treatment of route of exposure. For some studies, subcutaneous exposure is identified (correctly) as a weakness of the study and provides part of the basis for a judgment of low utility. For other studies, subcutaneous exposure is identified (incorrectly) as a strength of the study and provides part of the basis for a judgment of high utility. Because of its importance for bisphenol A, route of exposure is discussed in more detail in Section 2 of these comments.

Making it more difficult to assess whether study review criteria have been applied consistently are the wide range of utility descriptors that are used in the utility paragraphs for studies in Sections 3 and 4 of the draft report. Descriptors used include at least the following:

• Adequate • Adequate and useful • Marginally adequate • Barely adequate • Inadequate • Not adequate • Inadequate by itself	• Moderate utility • Minimal utility • Limited utility • Little utility • Marginal utility • No utility	• Very useful • Moderately useful • Useful • Slightly useful • Marginally useful • Relevant and useful • Not useful
• Suitable • Suitable but has minimal utility	• Valuable	

The basis for the descriptors is sometimes explained, sometimes not; sometimes further qualified with text that implies a lower utility ranking; and in some cases text is provided with no clear indication of utility.

Overall, it is difficult to assess whether the utility descriptions are based on one or both of the two critical components – quality and relevance – that determine the utility of a study for this evaluation. This makes the report less transparent and, more importantly, will make it difficult for the panel to appropriately weigh the evidence and reach sound and defensible conclusions.

Utility descriptors and criteria should be standardized and uniformly applied

The utility descriptors and criteria should be standardized as described below and consistently applied to all studies in Sections 3 and 4 of the draft CERHR report. This should be completed and a revised report prepared before the public panel meeting to facilitate a productive meeting.

Our proposed criteria for utility rankings applicable to the bisphenol A evaluation are shown in the table below.

Proposed Utility Ranks for Bisphenol A Studies Evaluated by CERHR	
Rank	**Criteria**
High utility	Wide dose range, including low doses.* Large number of relevant endpoints.** Sufficient number of animals per dose group. Large sample size. Appropriate statistical analysis. Oral route of exposure.
Adequate	Some relevant endpoints and doses. Sufficient number of animals per dose group. Appropriate statistical analysis. Sufficient number of dose groups to demonstrate dose-response trend Oral route of exposure.
Limited Utility	Small dose range or number of doses. Non-oral route of exposure. No dose-response trend for reported effects. Inconsistent pattern of effects for reported effects. Some important data not reported. Limited statistical analyses. Inappropriate study design. Reproductive and developmental effects not primary endpoint.
Inadequate	No relevant endpoints. Number of animals per dose group too low to draw conclusions. Much important, relevant data not provided. No or inappropriate statistical analyses. Animals with co-exposures to other compounds. Ovariectomized or orchiectomized animals.

** The relevant dose range is defined in guideline studies by the respective test guideline. In the case of bisphenol A, a wide dose range including low doses is of highest utility for the weight of evidence evaluation to clarify open questions due to published studies using low doses.*

*** Relevant endpoints include reproductive organ weights, perinatal characteristics, pubertal characteristics, teratogenic effects, and effects on reproductive function. Endocrine activity and changes in gene expression should not be considered because these do not necessarily indicate adverse effects.*

Considerations for evaluating the particular criteria listed above include the following:

Dose range. With a wider dose range, more information can be gleaned regarding the nature of effects and whether there is a dose-response relationship. In addition, a NOAEL or LOAEL is more likely to be established with a wide dose range. Thus, a study with a wide dose range is considered to be of high utility, and a study with a small dose range is of limited utility. The relevant dose range is defined in guideline studies by the respective test guideline.

In the case of bisphenol A, a wide dose range including low doses is of highest utility for the weight of evidence evaluation to clarify open questions due to published studies using low doses and because humans are potentially exposed to only very low doses of bisphenol A.

Sample size/animals per dose group. A higher number of animals in a dose-group means the study has more power to detect a difference among dose-groups. Therefore, a study with a large sample size is of higher utility. If the number of animals per dose group is too small to draw conclusions, the study is inadequate.

Large number of relevant endpoints. The more developmental and reproductive endpoints measured in a study, the more useful the study is. Relevant endpoints include reproductive organ weights, perinatal characteristics, pubertal characteristics, teratogenic effects, and effects on reproductive function, including validated guideline study endpoints. Some measurements of endocrine activity (e.g., uterotrophic effects) or changes in gene expression should not be considered as relevant endpoints because these do not necessarily indicate adverse effects, but they may be useful for determining mode of action.

Oral route of exposure. The potential for exposure of humans to bisphenol A is principally *via* ingestion. Bisphenol A undergoes extensive first pass metabolism in humans, thus the amount reaching general circulation is much lower than that which was ingested and essentially negligible at lower doses due to extensive capacity for pre-systemic clearance. Thus, non-oral exposures, which bypass metabolic pathways, are of limited utility for determining effects in humans.

Dose-response trend. The presence of a dose-response trend is likely indicative of a valid treatment-related effect, while a lack of a dose-response trend implies that observed effects might not be treatment-related. For example, studies with only one dose group or studies in which effects are observed only at a single, intermediate dose group are of limited utility.

Inconsistent patterns of effects. Inconsistent patterns of effect (*i.e.*, such as effects in one of two paired organs) suggests that observed effects are likely not treatment-related. Studies with inconsistent patterns of effects are of limited utility.

Important data not reported. In certain cases, a study does not provide sufficient information to properly interpret results. Studies with missing data will be of limited utility or inadequate, depending on how much data is missing.

Limited statistical analyses. A proper statistical analysis is crucial to determine if there are treatment-related effects. Studies with limited analyses are of limited utility, and studies with inappropriate statistical analyses are inadequate for determining whether there are treatment-related effects.

Inappropriate study design. If the study design is not sound, then this leaves the interpretation of results in question. Studies with inappropriate study designs are of limited utility.

Reproductive and developmental effects not primary endpoint. If a study was not designed to examine reproductive and developmental effects, a finding of an effect is unlikely to be definitive evidence for a treatment-related effect. Thus, studies in which reproductive and developmental effects are not the primary endpoints will typically be of limited utility.

Animals with co-exposures to other compounds. If animals are co-exposed to other compounds, it is not possible to determine whether an effect was due to bisphenol A alone. Thus, these studies are inadequate for determining effects of bisphenol A in humans.

Ovariectomized or orchiectomized animals. Studies of ovariectomized or orchiectomized animals can be useful in some contexts for determining mode of action; however, because the hormonal milieus of these animals are altered, these studies are inadequate for predicting effects in humans.

In general, all of the criteria in the "high utility" ranking are necessary to merit this ranking. Otherwise, if at least one criteria is met for the other categories, that is where a study should fall.

Application of standardized descriptors and criteria

In Section 3 of these comments, the standardized criteria and descriptors are applied to the studies summarized in Sections 3 and 4 of the draft CERHR report. We strongly encourage the CERHR panel to adopt this approach to improve the transparency of the CERHR report and, more importantly, to facilitate the ability of the panel to reach sound and defensible conclusions based on a weight of evidence evaluation.

We have evaluated studies first by quality, judging lower quality studies to be of low utility, and then by relevance, recognizing that even studies that are well conducted in their own terms may not be the most relevant to the human reproductive and developmental hazard characterization of bisphenol A. For quality judgments, additional weight was given to studies conducted under GLP although conduct of a study under GLP is not absolutely necessary, since well conducted but non-GLP studies may have highly reliable data (see following section for additional comments on the value of GLP). Studies that fail to show internal consistency – for example, by having effects only at some isolated time points but not others, or having effects in one member of a paired set of organs but not the other, or showing no regularity in the pattern of response over dose levels – are also rated as of lower quality, since the internal inconsistency suggests that individual data points may be unreliable indicators of any general phenomenon.

In our recommended utility judgments, the highest utility goes to studies that are *in vivo* evaluations of reproductive and developmental effects in intact mammals (generally rats and mice) dosed in a manner that is relevant to human exposures (i.e., orally). A wide range of doses increases utility because it provides information relevant to low-dose effects, if any, and because it provides an opportunity better to characterize the dose-response relationship. As noted elsewhere, oral exposure is important because bisphenol A has substantial first-pass metabolism, with conjugated compound being inactive, and so non-oral routes may lead to much higher systemic exposures (and different metabolites) than comparisons of the nominal

dose rates would indicate. Studies examining intermediate endpoints or reporting non-pathological alteration of tissues are given lesser utility ratings, because such outcomes have only indirect bearing on the ability of bisphenol A actually to cause manifestations of reproductive or developmental dysfunction. We recognize, however, that such studies can have uses for other, ancillary, questions.

We recommend that studies be segregated by the time of exposure, according to whether exposure is preconception, during gestation, during gestation and nursing, only during nursing, or directly to young or mature animals. These exposure modes are relevant to different parts of the larger hazard identification question, and results may not be comparable across dosing time regimes.

We have not judged studies by whether the authors' conclusions are supported by their data. We feel that the Panel ought to be drawing its own conclusions, and that it can do so better than any individual study author by virtue of seeing the whole array of information across studies. Individual study authors may have had a more limited context without the benefit of the latest results when they drew their conclusions.

As stated elsewhere, interpreting a body of data consists of much more than simply examining studies individually. A full weight-of-evidence characterization entails comparing results across studies that have examined similar endpoints over similar dose ranges, seeking evidence of consistency or contradiction. Such comparisons also affect the interpretation of individual studies in that, when a reported result has not been repeated in several other studies, this increases the likelihood that the reported phenomenon may be an artifact or statistical anomaly rather than a genuine demonstration of a biological response. A full examination of this consistency across studies is a task for the future Panel meeting, but we have noted in our utility ratings when results in particular studies are at odds with those of similar studies.

d. Guideline studies and use of Good Laboratory Practices should be given additional weight

Several of the studies reviewed in the draft CERHR report were conducted according to the requirements of standardized guidelines that have been designed for evaluation of reproductive and developmental toxicity. By design, these studies will satisfy the CERHR evaluation and quality criteria for review of individual studies (discussed above) and results from these studies should accordingly be given additional weight in the overall evaluation.

Additional weight should also be given to studies conducted under Good Laboratory Practice (GLP) regulations or principles. Extensive monitoring, auditing, and quality assurance are integral parts of the GLP process, resulting in increased reliability of data from GLP studies. Numerous other aspects of GLP should be part of any good scientific practice (e.g., written protocol with justification for the study design and test methods, standard operating procedures, data collection records, data retention procedures, identification and stability of the test substance, confirmation of doses administered, independent quality assurance review). These aspects may not exist or may not be readily determined from published non-GLP studies.

e. All positive and negative findings from relevant studies should be included in the weight of evidence

As noted above in Section 1(b) of these comments, replication of findings among similar studies and the observation of similar effects under relevant conditions is a fundamental principle of the scientific process. Conversely, lack of replication is grounds to doubt the validity of single experimental results.

In this regard, it will be important for the panel to fully consider all positive and negative findings from studies that are included in the evaluation. The draft CERHR report includes a large number of studies, many of which report positive findings on specific endpoints. However, many of these same studies also report negative findings on other endpoints. To fully test for corroboration of the positive findings, the panel must equally consider all of the relevant positive and negative findings in these many studies.

The draft CERHR report currently includes summary tables, organized by study, with data from studies deemed to have sufficient utility to include in the overall evaluation (e.g., Tables 97-100). These tables should include all positive and negative data for relevant endpoints from each listed study. It would also be helpful, if not preferred, to include tables that are organized by endpoint to more easily compare all relevant positive and negative findings.

The utility of a comprehensive comparison of low-dose bisphenol A studies was published by Ashby et al.[5] The authors describe observations, together with a reanalysis of the data

[5] Ashby, J., Tinwell, H., Odum, J. and Lefevre, P. Natural variability and the influence of concurrent control values on the detection and interpretation of low-dose or weak endocrine toxicities. Environmental Health Perspectives 2004; 112: 847-853.

associated with several reported instances of weak or low-dose endocrine effects that have proven difficult to confirm in independent laboratories and propose recommendations on experimental study design for examination of weak or low-dose effects.

Examples of the utility of this approach are two recent weight of evidence evaluations of bisphenol A that have been published in the peer reviewed literature (Goodman et al.,[6] CERHR reference 55; Gray et al.,[7] not referenced in the draft CERHR report).

[6] Goodman, J. E., McConnell, E. E., Sipes, I. G., Witorsch, R. J.., Slayton, T. M., Yu, C. J., Lewis, A. S. and Rhomberg, L. R. An updated weight of the evidence evaluation of reproductive and developmental effects of low doses of bisphenol A. Critical Reviews in Toxicology 2006; 36: 387-457.

[7] Gray, G. M., Cohen, J. T., Cunha, G., Hughes, C., McConnell, E. E., Rhomber, L., Sipes, I. G. and Mattison, D. Weight of the evidence evaluation of low-dose reproductive and developmental effects of bisphenol A. Human and Ecological Risk Assessment 2004; 10: 875-921.

2. Considerations Specific to Bisphenol A that are Critical for a Scientifically Sound Evaluation

a. Human exposure

An understanding of exposure is necessary in any risk evaluation and is particularly important for the CERHR evaluation of bisphenol A. Many of the reproductive and developmental studies that have been reviewed by the panel involve "low-dose" effects and in many of these studies, the doses tested are described as "environmentally relevant." These terms have been applied to a wide range of doses with little apparent consideration of actual human exposure. As a result, much of the "information" on human exposure to bisphenol A is better characterized as misinformation. The best assessment of human exposure to bisphenol A (but not the internal dose or actual bioavailability as discussed below) is provided by an examination of the human urinalysis data.

To properly assess the reproductive and developmental studies and to reach defensible risk conclusions, the panel will need to carefully assess actual human exposure to understand what truly is a "low" or "environmentally relevant" exposure. Reliable estimates of human exposure can be derived from the substantial amount of relevant information that is available. That same information also allows for outliers and misinformation to be identified.

A clear distinction should be made between exposure and bioavailability

In the draft CERHR report, the term *exposure* is not given a precise definition. The report cites some of the published literature as estimating "exposure" from measurements of bisphenol A in food or the environment with various assumptions on the percent absorption (example: page 17, lines 12, 22, 35). A clear distinction should be made between the term *exposure* (which should be defined as the contact of an individual with a chemical arising from any source and by any potential route of entry into the body) and internal dose or the *bioavailability* of a compound (which is defined as the entry of the compound as parent compound into the systemic circulation).

It should be recognized that exposure to a chemical may not result in the delivery of the chemical from the site of exposure to other areas of the body. In the case of bisphenol A, oral ingestion by mammalian species results in poor bioavailability due to the extensive presystemic clearance through first pass intestinal tract and hepatic metabolism. This presystemic clearance through metabolism is especially pronounced in humans, with rapid urinary excretion of BPA-glucuronide resulting in essentially 100% excretion within 24 hrs of ingestion (Völkel et al., 2002, CERHR reference 91; Völkel et al., 2005, CERHR reference 62).

Urine biomonitoring provides the most reliable estimate of human exposure

Regarding human exposure, the CERHR guidelines[4] state "Direct measurement in human tissues or fluids is particularly important and should be included where possible." While this comment may be generally valid, it is particularly appropriate for bisphenol A.

As noted above and in the draft CERHR report, the metabolism and pharmacokinetics of bisphenol A in humans have been characterized in several controlled human pharmacokinetic studies. After oral ingestion, which is the only significant route of human exposure, bisphenol A is efficiently metabolized to BPA-glucuronide, essentially 100% of which is excreted into urine within 24 hours of ingestion. This information means that measurement of bisphenol A in urine after complete hydrolysis of conjugates will provide a reliable estimate of bisphenol A exposure (i.e., defined as the contact with bisphenol A that occurred, but not the internal dose or bioavailability).

In particular, studies that have collected urine over a 24-hour interval are likely to provide the best estimates of daily exposure (see Tsukioka et al., 2004, CERHR reference 72; Arakawa et al., 2004; CERHR reference 76; both cited in CERHR Table 12, page 25) since analysis of a 24-hour sample effectively is a direct measure of exposure over the previous 24-hour period. These studies indicate that mean or median human exposure is in the range of 0.02-0.03 µg/kg/day (20-30 nanograms/kg/day).

Many other studies that report bisphenol A concentrations in spot samples of urine are summarized in the draft CERHR report and additional studies are provided in Section 3 of these comments. Although spot sample concentrations do not directly measure daily intake, they can be used to estimate daily intake based on typical daily urine excretion rates or daily creatinine excretion rates (in cases where bisphenol A concentrations are reported as a ratio against creatinine concentration). Estimates of daily intake from the many spot sample studies are very consistent with the daily intake measurements from 24-hour urine samples. Notably for the US population, the median daily intake estimated from a recent CDC urinary biomonitoring study of 394 individuals (Calafat et al., 2005, CERHR reference 58) is 0.026 µg/kg/day.

Collectively, the many urine biomonitoring studies consistently indicate that mean or median human exposure to bisphenol A is in the range of 0.02-0.03 µg/kg/day.

Daily intake estimated from biomonitoring data is supported by exposure estimates from other methodologies.

In addition to urine biomonitoring data, the draft CERHR report also summarizes data and exposure estimates from a variety of other sources and methodologies. Although most of these methodologies do not measure human exposure directly, these exposure estimates are generally consistent with the more direct urine biomonitoring estimates and provide further support that the urine biomonitoring exposure estimates are valid.

Notable examples include exposure estimates from

- Canned foods and beverages: mean daily intake = 0.0083 µg/kg/day (Thomson and Grounds, 2005, CERHR reference 39)
- Total diet: daily intake = 0.00195 µg/kg/day for adults, 0.00475 µg/kg/day for children (Miyakawa et al., 2004,[8] incorporated via CERHR reference 30)
- Hospital meals: average daily intake = 0.0084 µg/kg/day (Higuchi et al., 2004[9]
- Hospital meals: mean daily intake = 0.02 µg/kg/day (Imanaka (2001) as cited in Miyamoto and Kotake, CERHR reference 30)
- Children's aggregate exposure: range = 0.018 – 0.071 µg/kg/day (Wilson et al., 2003, CERHR reference 11)
- Children's aggregate exposure: median daily intake = 0.0714 and 0.0608 µg/kg/day for two locations (Wilson et al., 2006, CERHR reference 12)

Doses described as environmentally relevant may not be

Many reproductive and developmental toxicity studies of bisphenol A report that the doses tested are "environmentally relevant." If explained at all, these studies frequently cite studies that do not provide authoritative exposure data to substantiate the claim of environmental relevance. As a result, much information on human exposure to bisphenol A is better characterized as misinformation. Several examples that are specifically mentioned in the draft CERHR report are listed in the attached Table 1.

In light of the human exposure estimates above, it is apparent that doses characterized as "environmentally relevant" are usually not. In general, studies that examine "low" doses use doses that are several orders of magnitude or more above actual human exposure levels. Thus, in reaching risk conclusions, the panel must take care to refer to actual human exposure estimates and not accept unsupported characterizations of doses as environmentally relevant.

Blood biomonitoring data is not suitable for human exposure estimates

There is now evidence that under certain conditions, hydrolysis of the glucuronide metabolite of bisphenol A may occur during specimen sampling or during analysis of biological specimens (Waechter et al., 2007[10]). Hence the measurement of bisphenol A itself in human tissues or fluids may not be an accurate reflection of the true bioavailability of parent compound unless the investigators have clearly demonstrated that in the process of specimen sampling and analysis that hydrolysis of conjugates has been precluded.

A cursory examination of the existing human blood data available on bisphenol A may lead to the conclusion that *"Bisphenol A is absorbed in humans as indicated by the detection of bisphenol A in blood from the general population (Section 1) and in maternal and fetal*

[8] Miyakawa, H., Shimamura, Y., Suzuki, K., Ibe, A. and Saito, K. Determination of bisphenol A in total diet study samples by GC/MS. Tokyo-to-Kenko Anzen Kenkyu Senta Kenkyu Nenpo 2004; 55: 157-161.
[9] Higuchi, M., Miyata, D., Kawamura, S., Ueda, E., Imanaka, M. and Tonogai, Y. Estimation of daily intake of phenols in hospital meal samples. Shokuhin Eiseigaku Zasshi 2005; 45: 339-343.
[10] Waechter, J., Domoradzki, J., Thornton, C. and Markham, D. Factors affecting the accuracy of bisphenol A and bisphenol A-monoglucuronide estimates in mammalian tissues and urine samples. Toxicology Mechanisms and Methods 2007; 17: 13-24.

fluids." (draft CERHR report page 31, lines 32-33). As absorption is defined as the entry of parent compound in the systemic blood circulation, it is more accurate to say that bisphenol A is only poorly absorbed following oral administration to humans. As noted in detailed comments in Section 3, reports of bisphenol A concentrations in human blood in the range of 1 to 2 ng/ml (or higher) in individuals without known prior exposure to bisphenol A challenges credulity based on the known human pharmacokinetics and urinary biomonitoring. Therefore, the blood data from the general population should be regarded with great caution. To produce blood concentrations in the ranges reported would likely require the individual to be ingesting daily or multiple daily doses in the range administered by Völkel et al., 2002 (CERHR reference 91) or higher, which is not supported by or consistent with the extensive urine biomonitoring data on bisphenol A, or the exposure estimates from food consumption or other sources.

b. Route of exposure

As discussed in Section 1 of these comments, a key aspect that determines the utility of an individual study is the relevance of the study to the question at hand. The CERHR guidelines[4] provide some guidance on how the relevance of experimental animal data should be assessed:

> Relevant = human data, or animal data for which pharmacokinetic and mechanism information is adequate to demonstrate a particular similarity to humans.
> Assumed relevant = no information available to modify the assumption that the data are relevant.
> Irrelevant = pharmacokinetic or mechanistic features of the animal models are known and demonstrated to be inconsistent with human exposure or response.

This guidance appropriately highlights the importance of pharmacokinetic information in assessing the relevance of experimental animal data. As discussed in detail in this section, an important experimental parameter for which pharmacokinetic information is critical in assessing relevance is route of exposure.

Draft CERHR report is inconsistent in assessing relevance of studies with respect to route of exposure

The draft CERHR report summarizes reproductive and developmental toxicity studies that involve a variety of oral and non-oral routes of exposure. For many of these studies, route of exposure is specifically highlighted in the strengths/weaknesses and or utility sections of the study summaries. However, it is readily apparent that CERHR has not consistently assessed the relevance of studies with respect to route of exposure.

The attached Table 1 lists studies for which route of exposure is specifically highlighted in the strengths/weaknesses or utility sections. For many of these studies subcutaneous routes of exposure (e.g., subcutaneous injection, osmotic pumps, silastic capsules) are described as a weakness and the studies are ranked as having limited or no utility, at least partly on the basis

of the route of exposure. However, in other cases, subcutaneous routes of exposure are described as a strength and the studies are ranked as adequate or higher.

As described below, non-oral routes of exposure, including subcutaneous routes, are not relevant for human exposure and, thus, animal studies that use non-oral routes of exposure should uniformly be ranked as having limited or no utility for the CERHR evaluation.

Studies involving non-oral routes of exposure have limited utility

As discussed in Section 2 of the draft CERHR report and above in these comments, humans are orally exposed to very low levels of bisphenol A through the diet. Significant differences between the oral route of exposure and other routes, in particular subcutaneous routes, have been demonstrated in both pharmacokinetic studies and in studies that examine the potential endocrine activity of bisphenol A.

The metabolism and pharmacokinetics of bisphenol A in humans have been studied in several controlled dose human studies (Völkel et al., 2002, CERHR reference 91; Völkel et al., 2005, CERHR reference 62; Tsukioka et al., 2004, CERHR reference number 72). These studies demonstrate that, by the oral route of exposure, bisphenol A is cleanly metabolized to BPA-glucuronide, which is then rapidly cleared by excretion into urine with a half-life of less than 6 hours. No parent bisphenol A was detected in blood or urine after an oral exposure of 5 mg/person (~83 micrograms/kg/day), which is well above typical human exposure levels.

The primary metabolite BPA-glucuronide has no known biological activity and, in particular, has been shown to have no estrogenic activity (Matthews et al., 2001, CERHR reference 138). Although some studies have suggested that BPA-sulfate may also form to a lesser extent in humans, this metabolite also has been shown to have no estrogenic activity (Shimizu et al., 2002, CERHR reference 143).

Thus, by the relevant oral route of exposure in humans, the bioavailability of bisphenol A is extremely low and there is no indication of biologically active metabolites, in particular estrogenic metabolites.

In contrast, animal studies have shown that the bioavailability of bisphenol A by different routes of exposure are substantially different with non-oral routes showing much higher bioavailability.

For example, Pottenger et al. (2000, CERHR reference 93) investigated pharmacokinetic parameters in Fischer 344 rats after oral dosing (gavage), intraperitoneal (i.p.) or subcutaneous (s.c.) injection. This study is partially evaluated in the draft CERHR report and the following data should be additionally emphasized.

Pottenger et al. (2000) investigated single doses of 10 or 100 mg/kg ^{14}C-labelled bisphenol A by oral gavage, i.p. or s.c. injection. Parent bisphenol A levels and radioactivity were determined in blood and plasma samples. Radioactivity was determined in urine and fecal samples. At 7 days post-dosing, animals were sacrificed and selected tissues (brain, liver,

kidneys, peri-renal fat, gonads, uterus, and skin) and remaining carcass were analyzed for radioactivity. Examination of metabolic products was also conducted by HPLC on selected urine and fecal samples. The authors concluded:

- "The **relative bioavailability of BPA was highly route-dependent**…Estimation of BPA AUC for s.c. administration resulted in values from 7-fold (high dose females) to 245-fold (high does males) greater than for oral administration"

- "Oral administration resulted in **Cmax** values that were from **1 to 2 orders of magnitude lower** than those found following i.p. or s.c. administration"

- **Three additional unidentified metabolites** were observed following i.p. or s.c. but not oral exposure.

- "Both i.p. and s.c. administration resulted in larger fractions of plasma radioactivity comprised of unchanged parent compound,…**Unchanged parent [BPA]** comprised **27-51% and 65-76%** of total plasma radioactivity following **i.p. and s.c.** administration, respectively, whereas **oral administration resulted in only 2-8%** of total plasma radioactivity comprised of unchanged parent compound"

The substantially greater values for unchanged parent bisphenol A observed following i.p. or s.c. administration indicate that essentially complete first-pass metabolism occurred following the oral administration of *high* doses to rats. Any circulating levels of the parent compound were substantially and rapidly further reduced by the extensive conjugation of bisphenol A by hepatic glucuronyl transferases. In the rat, some deconjugation of BPA-glucuronide by intestinal microflora occurred for the conjugate excreted via the bile, and with reabsorption resulted in enterohepatic circulation. This phenomena is clearly absent in the metabolism and kinetics of bisphenol A in humans as the biliary route of excretion is not predominant.

Accordingly, oral and non-oral doses cannot be directly compared and it is not clear that non-oral exposures should be considered true "low" doses. Any effects seen with non-oral exposures should be interpreted with caution from the point of view of predicting consequences of oral exposure because of the severe differences observed in relative bioavailability and metabolism.

Consistent with the significant differences in bioavailability from oral and non-oral routes of exposure, significant differences in effect levels have also been observed with oral and non-oral exposures. For example, in an immature rat uterotrophic assay (Yamasaki et al., 2000, CERHR reference 99), female rats were given daily s.c. injections of 0, 8, 40 or 160 mg/kg bisphenol A, or daily oral (gavage) doses of 0, 40, 160 or 800 mg/kg bisphenol A, on postnatal days 18-20. Animals were sacrificed 24 hours after the final dose and uteri removed and weighed. A repeat study using the same experimental protocol was conducted. Statistically significant increases in the relative wet and dry uterine weights were seen at 160 mg/kg and above after oral dosing and at 8 mg/kg (lowest dose investigated) and above for s.c. dosing. In the first study, plasma concentrations of bisphenol A were detected in all groups given bisphenol A sc and in groups given 160 and 800 mg/kg bisphenol A orally, with a

dose–response effect. Comparing plasma concentrations for the same dose between sc and oral routes, bisphenol A values were much higher in the 160 mg/kg bisphenol A sc group than in the group given the same dose of bisphenol A orally.

Overall, oral exposure results in much lower tissue levels of bisphenol A than would occur by other routes of exposure and BPA-glucuronide is the only metabolites identified in humans after oral exposure to bisphenol A. Non-oral routes of exposure result in substantially higher bioavailability and different metabolism.

Therefore, pharmacokinetic features of *in vivo* experiments involving non-oral routes of exposure in laboratory animals are "known and demonstrated to be inconsistent with human exposures or response," and data from such studies should be considered "irrelevant to consideration of human risk" according to the CERHR guidelines.[4]

As noted above, animal studies that use non-oral routes of exposure should uniformly be ranked as having limited or no utility for the CERHR evaluation.

c. Sub-mammalian studies

Sections 3 and 4 of the draft CERHR report include summaries of a number of studies on a variety of submammalian species (e.g., fish, frogs). The utility of these studies for the CERHR evaluation is variously described as "not clear," "useful", "slightly useful," "limited utility," or "not useful," but it is not apparent how these studies could have any utility.

Although in Section 3 of these comments we have noted some additional strengths and weaknesses of these studies, our primary comment is that these studies should all be ranked as having no utility for the CERHR evaluation. As noted in certain cases in the draft CERHR report, these studies do have utility for assessing ecotoxicological hazards, but are of no utility for assessing hazards and risks to human reproduction and development.

For these studies to be of any utility, several major hurdles would need to be overcome. The most prominent hurdle is the lack of information on mode of action in the submammalian species to assess whether observed effects could have any relevance at all to humans. A second hurdle is the difficulty in correlating aquatic concentrations of bisphenol A from studies on aquatic species to an oral dose equivalent for humans. Related to both of these issues is the lack of information on metabolism and pharmacokinetics in the submammalian species to assess whether there are any similarities with humans.

The extrapolation of experimental data from animals to humans has been, and continues to be, a dynamic topic in the risk assessment process. The state of the art is far from a well-defined and universally accepted approach. The use of substantial uncertainty factors (100x or 1000x) for extrapolation of rodent data to human hazard and risk remains a key element for regulatory agencies around the world. The ability to extrapolate from other species, even if considered useful, would require uncertainty factors so large as to render the evaluations scientifically unjustified.

bisphenol A was evaluated in a Case Study by the Environmental Protection Agency to determine the usefulness and validity of cross-phylum to cross-species evaluation of reproductive toxicity (EPA, 2005[11]). While the report concluded that such evaluations hold promise for elucidating the Mechanism of Action (MOA) for bisphenol A and other chemicals, it also identified a number of major gaps that prohibit direct comparisons of current data.

[11] US Environmental Protection Agency. A cross-species mode of action information assessment: A case study of bisphenol A. EPA/600/R-05/044F. April 2005.

3. Detailed comments on draft CERHR report

Detailed comments on the draft CERHR report are presented in the form of the following attached tables.

a. Section 1 of the draft CERHR report

Comments on Section 1 are provided in Table 2.

b. Section 2 of the draft CERHR report

Comments on Section 2.1 are provided in Table 3.

Comments on Sections 2.2-2.6 are provided in Table 4.

c. Section 3 and 4 of the draft CERHR report

Comments on human and mammalian reproductive and developmental studies in Sections 3 and 4 (excluding neurodevelopmental studies) are provided in Table 5. This table includes comments on studies that are already included in Sections 3 and 4 of the draft CERHR report. Several notable features of the table include:

- To make the table more useable as a stand-alone document, each study is briefly summarized with key technical details, the strengths and weaknesses highlighted by the CERHR panel, and the CERHR utility ranking. The "Additional Considerations" column includes additional strengths and weaknesses that were not highlighted by the CERHR panel. Studies highlighted in blue were evaluated by Goodman et al. (2006)[6], which can be referred to for additional information on the evaluation on these studies.

- All of the strengths and weaknesses along with key technical details were then evaluated using the critieria outlined in Section 1(c) of these comments to derive a utility ranking, which appears in the column on the right side of the table. Studies with proposed rankings that differ significantly from the CERHR panel ranking are highlighted in yellow.

- Note that a number of studies are listed in Table 5 with duplicate lines (e.g., Cagen et al., 1999, CERHR reference 294). Duplicate lines indicate that the study includes both reproductive and developmental endpoints. CERHR should consider these studies within the scope of both Sections 3 and 4 so ensure that all relevant endpoints are included in the weight of evidence evalutions. It should be noted that for some studies with both reproductive and developmental endpoints (e.g., Ema et al., 2001, CERHR reference 292), CERHR has already included the studies in both Sections 3 and 4.

- As described above in Section 1(c), we encourage CERHR to reassess these studies with a uniform set of criteria and consistently assign standardized utility rankings before the public panel meeting. We offer our evaluations as a guide.

Comments in Table 6 are in regard to reproductive and developmental studies that are not included in Sections 3 or 4 of the draft CERHR report but should be included. Studies that are italicized are included elsewhere in the draft CERHR report but should also be included in Sections 3 or 4 as appropriate to ensure that the endpoints relevant to these sections are considered. Each study has been evaluated and presented as in Table 5 described above. We encourage CERHR to consider these studies for inclusion in the report for completeness.

Comments on neurodevelopmental studies in Section 3 are provided in Table 7. This table includes comments on studies that are already included in Section 3 of the draft CERHR report. It should be noted that some studies with neurodevelopmental endpoints are described in other parts of Section 3, presumably for other developmental endpoints, but are not included in the neurodevelopmental part of Section 3. These studies are noted in the table and should be included in the neurodevelopmental part of Section 3 to be sure that all relevant data is considered. Several notable features of the table include:

- Similar to Table 5, each study in this table is briefly summarized with key technical details, the study author's conclusions, the strengths and weaknesses highlighted by the CERHR panel, and the CERHR utility ranking. Our additional comments and utility ranking added for each study.
- In cases where a study has both neurodevelopmental endpoints as well as other reproductive and developmental endpoints, the study will be listed in both Tables 5 and 7. In a few of these cases, the utility rankings in the two tables may differ since the study may have higher or lower utility for one type of endpoint versus another.

- As for the studies in Table 5, we encourage CERHR to reassess these studies with a uniform set of criteria and consistently assign standardized utility rankings before the public panel meeting. We offer our evaluations as a guide.

Comments in Table 8 are in regard to studies that are not included in the neurodevelopmental parts of Section 3 of the draft CERHR report but should be included. Each study has been evaluated and presented as in Table 7 described above. We encourage CERHR to consider these studies for inclusion in the report for completeness.

Comments on the sub-mammalian studies in Sections 3 and 4 are provided in Table 9. As noted above in Section 2(c) of these comments, our primary comment is that all of these studies should be designated as inadequate for the CERHR evaluation.

Table 1

Dose and Route of Exposure
Examples from CERHR Draft Report of December 2006
February 2, 2007

Page	Lines	Dose(s)	Route of Administration[1]	Author (A) or CERHR (C) Description	Comments
130	20-21	0.025 mg/kg/day	sc osmotic pump	A: "environmentally relevant doses"	Subcutaneous route of exposure is not relevant for human risk assessment and the single dose tested is several orders of magnitude higher than typical human exposure
148 149	14-15 5	0.040 and 0.400 mg/kg/day	oral	A: "the doses selected were relevant to human exposures from can linings and dental sealants" C: "Strengths of this study are…used biologically-relevant concentrations"	The doses are several orders of magnitude higher than typical human exposure
170	9	0.01 mg/kg	sc silastic capsules	C: "The paper has many strengths, from the use of multiple biologically relevant doses"	The single dose tested is well above typical human exposure and the route of exposure is not relevant for human risk assessment

[1] The oral route of exposure includes several experimental methods including exposure through the diet, drinking water, gavage and micropipette. These are identified in the CERHR report summaries for each study but are generically categorized in this table as "oral."

179	34-35	0.0024 mg/kg/day	oral	C: "Strengths are the oral route of exposure, the use of environmentally-relevant dose level"	The single dose tested is well above typical human exposure
193	20-21	0.000025 and 0.00025 mg/kg/day	sc osmotic pump	C: "The administration of very low doses by subcutaneous pump…are strengths of this study"	Subcutaneous route of exposure is not relevant for human risk assessment
194	31-32	0.025 and 0.25 mg/kg/day	sc osmotic pump	C: "The administration of very low, environmentally relevant doses by subcutaneous pump…are strengths of this study"	Subcutaneous route of exposure is not relevant for human risk assessment and the doses are several orders of magnitude higher than typical human exposure
195	6-7	0.00025 mg/kg/day	sc osmotic pump	A: "The bisphenol dose was selected because it was thought to be environmentally relevant"	Subcutaneous route of exposure is not relevant for human risk assessment
209 210	39 3-4	0.000025 and 0.00025 mg/kg/day	sc osmotic pump	A: "environmentally relevant doses" C: "The use of sc pumps to deliver low doses…is strength"	Subcutaneous route of exposure is not relevant for human risk assessment
211	21	0.000025 and 0.00025 mg/kg/day	sc osmotic pumps	C: "relevant doses"	Subcutaneous route of exposure is not relevant for human risk assessment
227	41-42	0.002, 0.020, 0.200, 2 and 200 mg/kg/day	oral	C: "Additional strengths are…the oral route of administration"	

263	30-32	11, 78, 158 and 250 mg/kg/day	sc injection	C: "the relevancy of the model for human risk assessment is limited because the route of administration/dosing paradigm was not relevant"	
271	18-19	10, 50 and 250 mg/kg/day	sc injection	C: "route of administration…severely limits the utility of this study for human risk assessment"	
276	23-24	235, 466 and 950 mg/kg/day	oral	C: "a relevant route of administration"	
277	46	0.0000002-2 mg/kg/day	oral	C: "a relevant route of administration"	
280	21-22	0.3 and 3 mg/kg/day	sc injection	C: "SC is not a relevant route of exposure"	
283	6	0.1 and 100 mg/kg	sc injection	C: "The sc route of exposure…is not relevant to human exposure"	
285	22	3 mg/kg/day	sc injection	C: "route of administration was not relevant for human risk assessment"	
288	28-29	0.002, 0.02 and 0.2 mg/kg/day	oral	C: "a relevant route of exposure"	

290	25-26	0.05, 0.5 and 5 mg/kg	ip injection	C: "route of administration was not relevant for human risk assessment"	
291	37	0.0024 mg/kg/day	sc injection	C: "route of administration was not relevant"	
292	18-19	150 mg/kg/day	ip injection	C: "route of administration is not relevant for human risk assessment"	
308	19-26	2.4, 4.2 and 8.1 mg/kg/day	silastic implants	C: "When compared to studies that used the oral route of exposure, 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. However, the administration of bisphenol A by silastic implants makes the extrapolation for human risk assessment difficult in the absence of an improved pharmacokinetic understanding." and "an irrelevant route of exposure, which makes extrapolation for human risk assessment difficult"	

Table 2

Polycarbonate/BPA Global Group Comments on
CERHR Draft Report of December 2006
Section 1, Pages 1-30

February 2, 2007

Page	Lines	Comment
2	4-5	Additional information on trace impurities in commercial BPA is available in Terasaki et al. (2004)[1] and Terasaki et al. (2005)[2].
2	12-13	Bisphenol A is manufactured by the acid catalyzed condensation of phenol and acetone, not with an alkaline catalyst as stated (see CERHR reference 6).
2	15-18	Current manufacturers of BPA in the US are Bayer MaterialScience, Dow Chemical Company, General Electric, Hexion Specialty Chemicals, and Sunoco Chemicals (see CERHR reference 6, company names updated to current corporate identities).
2	19	The referenced information was provided by the American Plastics Council.
2	25-26	The referenced 1991 production volume is not realistic. See CERHR reference 6 for a more realistic value.

[1] Terasaki, M., Nomachi, M., Edmonds, J. S. and Morita, M. Impurities in industrial grade 4,4'isopropylidene diphenol (bisphenol A): Possible implications for estrogenic activity. Chemosphere 2004; 55: 927-931.
[2] Terasaki, M., Shiraishi, F., Nishikawa, T., Edmonds, J. S., Morita, M. and Makino, M. Estrogenic activity of impurities in industrial grade bisphenol A. Environmental Science and Technology 2005; 39: 3703-3707.

3	6-9	A search of the Household Products Database for all products containing CAS No. 80-05-7 revealed only three epoxy adhesives that contained BPA as an ingredient in the adhesive formulation. These products are described as for hobby/craft or home use. The broader list of products referenced in the CERHR draft report apparently resulted from a search for "bisphenol A," which primarily identifies resin products that are made from BPA but do not contain BPA as an ingredient. The broader list of products is thus not relevant for the CERHR evaluation.
3	11-17	Additional FDA authorized uses of polymers or other substances manufactured from BPA include: 21CFR177.1440, 21CFR 177.1580, 21CFR 177.1585, 21CFR 177.1655, 21CFR 177.2280, 21CFR 177.2420, 21CFR 177.2600 and 21CFR 178.2010.
3	21	The referenced information was provided by the American Plastics Council.
3, 4	47-51, 1-2	More recent research has shown that BPA is likely to photo-degrade in natural surface waters containing dissolved organic matter, which acts as a photo-sensitizer and is ubiquitous in surface waters. For example, Chin et al. (2004)[3] concluded that photosensitized degradation of BPA may be as important as biodegradation.
4	24-43	It should be noted that the BPA in air was collected as particular matter. It is not known whether any of the detected BPA is absorbed or bioavailable after inhalation of the particulates.
4	24-43	When available, mean or median values should be reported along with the range in the studies discussed in this paragraph.
4	41-43	The level of BPA in dust is incorrectly reported.

[3] Chin, Y., Miller,, P. L., Zeng, L., Cawley, K. and Weavers, L. K. Photosensitized degradation of bisphenol A by dissolved organic matter. Environmental Science and Technology 2004; 38: 5888-5894.

4	48-49	It should be clarified that the median value reported (0.14µg/L) applies only to the samples in which BPA was detected above the reporting limit of 0.09 µg/L. BPA was not detected in 58.8% of the samples (i.e., <0.09 µg/L). Thus, the true median value is below the limit of detection and may be well below the reported value based on censored data.
5	17-20	It should be noted that the high migration values reported in one study (~192 µg/L and 654 µg/L; CERHR reference 18) resulted from treatment of the bottles to high temperatures (70 and 100 degrees Centigrade) for 240 hours. These extreme test conditions are not relevant to real-life use of baby bottles and the resulting data is thus of no relevance for human exposure assessment.
5	29	The correct CERHR reference is (35).
6	---	Mountfort et al. table entry: The technical report[4] that supports the Mountfort et al. publication (CERHR reference 19) should also be cited since some of the procedure details and results appear to be from the report rather than the publication.
7	---	CSL table entry: The results are inconsistent with those in the CSL report. Migration into 3% acetic acid after 20 washes should be ND to 0.51 ppb, and migration into 3% acetic acid after 50 washes should be ND to 0.7 ppb.
8	---	Biles et al. table entry: The units in the Biles et al. study were incorrectly published. The units expressed in Biles et al. as ng/L (converted to µg/L in the CERHR report) should have been expressed as ng/mL (personal communication from the authors). The correct CERHR reference number is (35).

[4] Mountfort, K. A. Investigations into the potential degradation of polycarbonate baby bottles during sterilization with consequent release of bisphenol A. Central Science Laboratory 1997 Report FD 97/08.

9	13-16	Consistent with the use of data from measurement of BPA in actual canned foods instead of migration into food simulants (see page 9, lines 5-8), the simulation study described in lines 13-16 should be replaced by the Brenn-Struckhofova and Cichna-Markl study (CERHR reference 46 and last entry in Table 4). As a minimum, the results described in CERHR reference 46 should be discussed in the text.
9	---	1st table entry: The level of BPA in diluted infant formula is incorrectly calculated/reported. The correct CERHR reference is (23), not (35).
9	---	2nd table entry: The limit of detection for Goodson et al. (<0.002 mg/kg) is incorrectly reported as <0.002 µg/kg.
14	37-40	Additional information on overestimation of BPA levels in biological samples with the ELISA method is provided in Inoue et al. (CERHR reference 90) and should be discussed here.
14	---	Section 1.2.3.4 There are many additional studies that have been omitted from this section, most notably the Arakawa et al. study that followed a protocol for the collection of urine over a 24 hour period (CERHR reference 76, listed in Table 12). These investigators employed GC/MS/MS for the measurement of BPA in urine specimens collected over 24 hours from 36 male subjects. The analysis for BPA was carried out following the enzymatic hydrolysis of BPA-glucuronide with β-glucuronidase and the detection limit for BPA in urine was 0.38 ng/ml. No information was given as to the sulfatase activity present in the β-glucuronidase, although the enzyme was derived from *E. coli,* which is suggested to be low in arylsulfatase activity (Ye et al., 2005). BPA found in the urine post-hydrolysis ranged from <0.21 to 14 µg/day (median value 1.2 µg/day), corresponding to estimated exposures (external doses) of <0.003 to 0.23 µg/kg/day (median value 0.02 µg/kg/day) assuming an average body weight of 70 kg. This study along with the Tsukioka et al. study (CERHR reference 72, listed in Table 12) are very important in that they both employed 24 hour urine collection. This protocol provides data that provide a more accurate assessment of daily exposure. This is because estimates of daily exposure from "spot" specimens assume the concentration of BPA in urine is constant throughout the day and also must be corrected on the basis of estimated daily urine volume or creatinine measurements (if available). Neither assumption nor correction on the basis of volume or creatinine is needed if urine is collected over a 24 hour period.

		Another notable study that should be discussed in this section and included in Tables 5 and 6 is Fukata et al. (CERHR reference 56), which separately measured BPA as parent and conjugate in 52 matched urine and blood specimens with a sensitive and reliable LC/MS/MS based analytical method. Low levels of BPA, essentially all in the form of BPA-glucuronide, were found in all of the urine samples. The reported levels are comparable to the levels reported in many other urine biomonitoring studies, indicating that the test subjects had a typical level of exposure to BPA. However, neither parent BPA nor BPA-glucuronide was found in any of the blood samples with a detection limit of 0.2 ng/ml. These results confirm the expectation from controlled human pharmacokinetic studies that at typical human exposure levels, BPA in any form should not be detected in blood with a detection limit in the range of the sensitive limit established in this study. Because BPA-glucuronide levels in blood and urine are tightly linked by pharmacokinetics, one implication of this study is that reports of measurable levels of BPA in blood, in the form of parent BPA or BPA-glucuronide, should be confirmed with analysis of matched urine samples, which should show correspondingly high levels of BPA-glucuronide in urine. Other published urine biomonitoring studies that were not cited or reviewed include Brock et al. (CERHR reference 415, mentioned on page 262); Hanaoka et al. (CERHR reference 82, discussed on page 261); Matsumoto et al.[5], Kawaguchi et al.[6], Liu et al.[7], Yang et al.[8], and Yang et al.[9] These additional studies further support the conclusion made in the CERHR draft report that urinary concentrations are very low, in range of 1-2 μg/L. Collectively the data from these studies are consistent with the data from controlled human pharmacokinetic studies

[5] Matsumoto, A., Kunugita, N., Kitagawa, K., Isse, T., Oyama, T., Fouremen, G. L., Morita, M. and Kawamoto, T. Bisphenol A levels in human urine. Environmental Health Perspectives 2003; 111: 101-104.

[6] Kawaguchi, M., Inoue, K., Yoshimura, M., Ito, R., Sakui, N., Okanouchi, N. and Nakazawa, H. Determination of bisphenol A in river water and body fluid samples by stir bar sorptive extraction with in situ derivitzation and thermal desorption-gas chromatography-mass spectrometry. Journal of Chromatography B 2004; 805: 41-48.

[7] Liu, Z., Wolff, M. S. and Moline, J. Analysis of environmental biomarkers in urine using an electrochemical detector. Journal of Chromatography B 2005; 819: 155-159.

[8] Yang, M., Kim, S., Chang, S., Lee, I. and Kawamoto, T. Urinary concentrations of bisphenol A in relation to biomarkers of sensitivity and effect and endocrine-related health effects. Environmental and Molecular Mutagenesis 2006; 47: 571-578.

[9] Yang, M., Park, M. S. and Lee, H. S. Endocrine disrupting chemicals: Human exposure and health risks. Journal of Environmental Science and Health Part C 2006; 24: 183-224.

		(Völkel et al., CERHR reference 91; Tsukioka et al., CERHR reference 72; Völkel et al., CERHR reference 62) in that most of the biomonitoring studies that attempted to measure parent BPA in urine detected either none: Ouchi and Watanabe (CERHR reference 71), LOD of 0.2 ng/ml, storage conditions unspecified, n = 48; Völkel et al. (CERHR reference 62), storage at -20º C, n = 19; or only trace amounts (Tsukioka et al. CERHR reference 72), 0.1 to 0.27 ng/ml, LOD not reported, n = 91 with 80 out of 91 below the LOD; Ye et al. (CERHR reference 69), 0 to 0.6 ng/ml, LOD of 0.3 ng/ml, storage conditions unspecified, n = 30 with 27 out of 30 below the LOD. Exceptions are the studies of Matsumoto et al.[5], for both 1992 and 1999 data the range is estimated at 0 to 7 µg BPA/g creatinine or 4.2 ng/ml, LOD of 1.7 ng/ml, storage at -80º C, n = 106 with 88 of 106 below LOD; and the study of Kim et al. (CERHR reference 70) who reported that free BPA in urine ranged from 0.068 to 2.36 ng/ml. Based on what is known about the pharmacokinetics of BPA in humans, it is unlikely that the presence of other than trace levels of parent BPA in urine were due to the actual excretion of parent BPA. It is more likely that parent BPA was present in urine due to the unintentional hydrolysis of BPA-glucuronide or BPA-sulfate during sample handling and analysis.
15	2-6	*"Two studies (60, 61) reported urinary bisphenol A levels that were orders of magnitude higher than commonly observed levels.* It is accurate that the total BPA concentrations in human urine (measured typically post-hydrolysis of BPA-glucuronide and BPA-sulfate) were remarkably consistent across most of these studies despite the fact that these analyses were conducted in several different laboratories, using different sampling paradigms and analytical methods. An exception to the other studies is the study of Yang et al. (CERHR reference 61) where based on the total BPA in 73 spot urine specimens (after hydrolysis of BPA conjugates) the range of BPA concentrations was 0.68 to 586 µg/L. However, the geometric mean of the BPA urinary concentration (9.54 µg/L) was calculated by assigning a value of half of minimum value of detected urinary BPA (0.34 µg/L) to all specimens where BPA was not detected. BPA was not detected in 25% of the specimens and the analytical limit of detection was reported as 0.012 µg/L, well below the value of 0.34 µg/L assigned to these specimens. The high assigned value likely contributed to the higher geometric means reported, along with specimens that were "outliers" on the upper end of these concentrations. More recently, Yang and coworkers[8] have reported urinary BPA concentrations (after hydrolysis of conjugates) in 172 spot urine specimens. With a reported detection limit of 0.026 µg/L, BPA was detected in 97.5% of the specimens with a median value of 7.86 µg/L. The few samples with non-detected BPA were assigned a value of half of the detection limit and geometric means were separately reported for males (6.88 µg/L) and females (5.01 µg/L), both of

		which may have been influenced by outliers on the upper end of the concentration range. In a recent review article by Yang et al.[9], the authors report additional unpublished data from 2001 (median urinary BPA concentration of 4.20 μg/L in adults) and 2002 (median urinary concentration of 0.97 μg/L in children). Of particular importance for interpreting the several sets of data from Yang et al. is the comment that development of analytical methods for BPA is ongoing and that the apparent decrease in BPA concentrations over time is mainly due to the analytical techniques rather than a true decrease. In other words, the more recent lower values, which are more consistent with other values measured worldwide, are more likely to be representative of actual concentrations and the higher values reported in earlier studies may be overestimated. In the case of Mao et al. (CERHR reference 60), the authors also report the concentrations of 17-β-estradiol excreted in urine of the female human subjects studied to be in the mg/L range (Table 3 of reference 60) whereas typical 17-β-estradiol concentrations are known to be in the μg/L range for adult females. We agree with the analysis of Goodman et al., (CERHR reference 55) that this is likely a 1000-fold laboratory reporting error (mentioned on lines 4-6, page 15 of the draft CERHR report). Corrected for this apparent reporting error, the reported BPA concentrations are in the range typically reported in the many other studies reviewed in the draft CERHR report.
15	---	Table 5. The important human pharmacokinetic study of Völkel et al. (CERHR reference 91) should also be included in the table. These investigators found that following a direct oral dose of 5 mg BPA, no parent BPA was detected in blood at an LOD of 2.3 ng/ml and the peak plasma concentration of BPA-glucuronide was 176 ng/ml. Assuming that even 2% of the dose had escaped first pass metabolism (i.e., 100 μg) and was absorbed into the bloodstream, with an average plasma volume of 3000 ml it would have produced initial concentrations of about 33 ng/ml (100 μg/3000 ml blood). Thus, Völkel et al. would have had to sample blood within the first few minutes post-dosing (~50 minutes was their first sampling time) in order to have a possibility to detect any free BPA. This is because of the further rapid metabolism of BPA that would have occurred as the circulating blood reentered the liver, along with distribution to other tissues. Thus, reports that parent BPA concentrations are in the range of 1 to 2 ng/ml (or higher) in individuals without known prior exposure to BPA challenges credulity. To produce blood concentrations in this range would presumably require the individual to be ingesting daily or multiple daily doses in the range administered by Völkel et

		al. or substantially higher. Thus studies to provide closer examination and validation of the analytical methods used in many of these trace analyses is recommended. The entry for the study of Kuroda et al., should be corrected. This publication states that 9 sets of maternal and cord blood were collected from healthy pregnant women (not sterile patients) and 21 sets of peripheral blood serum and ascitic fluid was collected from patients. The data in this paper are reported as 0.46 mean; 0.43 median; range 0.21-0.79. The Methods column should be reviewed for accuracy. LC/MS (not HPLC) was often used.
16	---	Table 6 The designation of <LOD is not used consistently in this table. In some lines <MDL is used, in others a "<" symbol is used in conjunction with a concentration either in µg/L or nM. The free BPA concentrations for the Ouchi and Watanabe study should be <0.2 not ≤0.2. The number of subjects for the study of Tsukioka (CERHR reference 72) for the data reported was 34 females and 57 males. The data under the "Free" column for these subjects is for the 24-hour specimens collected from 11 males and 11 females shown in the next line. The mean reported for these subjects was 0.08 µg/L, range 0.01-0.27 µg/L. The footnote c on the data shown in the Total column for the Fujimaki et al. study is incorrect; the β-glucuronidase used also contained sulfatase. In general, data generated using ELISA methods should not be considered to be reliable measurements. ELISA methods for BPA in biological matrices have been shown to be non-specific for BPA and hence will overestimate BPA concentrations (Inoue et al., CERHR reference 90; Fukata et al., CERHR reference 56). ELISA is therefore inappropriate for use in biomonitoring studies. Völkel et al. table entry: The reported limit of detection for BPA-glucuronide in urine is 25 pmol/ml (10.1 µg/L) and the reported limit of quantitation is 65 pmol/ml (26.3µg/L).

Table 3

Polycarbonate/BPA Global Group Comments on
CERHR Draft Report of December 2006
Section 2.1, Pages 31-66

February 2, 2007

Page	Lines	Comment
31	4-5	*"The studies in this section demonstrate that bisphenol A is absorbed in humans and experimental animals following oral exposure."* This statement needs more qualification. Absorption is typically defined as the entry of parent compound in the systemic blood circulation. Later in this section (lines 13-14) it is correctly stated: "*In humans and experimental animals, most of a bisphenol A dose is metabolized to bisphenol A glucuronide prior to absorption.* This is based on the work of Völkel et al. (CERHR reference 91), which reported that no parent BPA was detected in human blood at an LOD of 2.3 ng/ml following oral administration of a 5 mg (5000 μg) dose of BPA in a controlled human pharmacokinetic study. A more accurate statement would be: "*The studies in this section demonstrate that bisphenol A is* poorly *absorbed in humans and experimental animals following oral exposure.*
31	14-15	*"In rats, there is evidence that glucuronidation does not occur in fetal liver…"*. To the contrary, there is evidence that glucuronidation of BPA in fetal liver may occur as Domoradzki et al. (CERHR reference 100, discussed on CERHR draft report pages 39-41) found that BPA-glucuronide was present in the fetus of Sprague-Dawley rats following the oral administration of BPA to the dams on gestation day 16.
31	32-33	*"Bisphenol A is absorbed in humans as indicated by the detection of bisphenol A in blood from the general population (Section 1) and in maternal and fetal fluids."* As absorption is typically defined as the entry of parent compound in the systemic blood circulation, it is more accurate to say that bisphenol is poorly absorbed following oral administration. As noted above, the blood data from the general population should be regarded with great caution. To produce blood concentrations in the ranges reported would likely require the individual to be ingesting daily or multiple daily doses in the range administered by Völkel et al. (CERHR reference 91) or higher. Thus, studies to

		provide closer examination and validation of the analytical methods used in many of these trace analyses is recommended. Much of the data in Table 18 was generated using the ELISA methods. As noted above, data generated using ELISA methods should not be considered as reliable measurements of BPA concentrations. ELISA methods for BPA in biological matrices have been shown to be non-specific for BPA and hence likely overestimate BPA concentrations (Inoue et al., CERHR reference 90; Fukata et al., CERHR reference 56). ELISA is therefore inappropriate for use in biomonitoring studies. The limitations of the ELISA methods should be clearly identified in this section as well as in section 1.0.
33	4-5	*"In humans, bisphenol A was measured in cord blood and amniotic fluid, demonstrating distribution to the embryo or fetus."* This statement is based on the data presented in Table 18. Most of these data were generated with ELISA methods, which have been shown to be non-specific for BPA. Therefore these data likely do not represent BPA at all but may represent the metabolite BPA-glucuronide or, even more likely, other unrelated substances such as phytoestrogens that are commonly present in humans. Data generated with other methods (e.g., GC/MS, Schönfelder et al., CERHR reference 84) are inconsistent with the known human pharmacokinetics information generated by Völkel (CERHR reference 91) as well as the many consistent estimates of human exposure from urinary biomonitoring data and food sources. Overall, the weight of evidence from the existing data do not support that the embryo/fetus is exposed to BPA as the parent compound.
34	28-30	*"They noted that according to unpublished data from their laboratory, the percentage of glucuronidated bisphenol A in mid-term amniotic fluid was ~34%, which is much lower than reported values for other human fluids (>90%).* Sample handling, extraction, concentration, chemical derivatization, or instrument injection can cause BPA metabolites (BPA-glucuronide or BPA-sulfate) to chemically degrade or be cleaved enzymatically, changing the actual ratio of BPA to BPA-glucuronide (Waechter et al.[1], this reference should be included and discussed in the CERHR draft report, for example in this section as well as in section 1.0). Without appropriate analytical controls it might be easily concluded that parent BPA was present in the specimen when only a BPA conjugate was actually present.

[1] Waechter, J., Domoradzki, J., Thornton, C. and Markham, D. Factors affecting the accuracy of bisphenol A and bisphenol A-monoglucuronide estimates in mammalian tissues and urine samples. Toxicology Mechanisms and Methods 2007; 17: 13-24.

34	---	General comment on Section 2.1.1.2 Most of the papers discussed in this section are lacking many details on analytical methods, which if present might provide the reader with greater (or lesser) confidence in the data generated. The following are critical considerations in the development and execution of analytical methods for BPA biomonitoring studies, most of which are lacking in the studies cited. Sufficient specimen(s) should be collected over time from each individual to allow for more robust interpretation of analytical findings (e.g., 24-hour urine collection or multiple blood collections over 24 hours). Each sample of a biological specimen should be analyzed at least in duplicate (triplicates recommended) along with: • External standards (defined as solutions containing known concentrations of BPA). These are the standards that define the BPA concentration - detector response curve. • Blanks (blanks are defined as all the solvents and reagents used in the analysis and are analyzed without an aliquot of the biological specimen added). These determine the background or baseline response of the detector (if any). • Controls (defined as samples of the same type of biological specimen as the specimens being taken for biomonitoring that are known to contain either no BPA or contain a known background concentration of BPA). These samples determine the background or baseline response for the biological specimens of interest. • Fortified controls (defined as samples of the same type of biological specimen being analyzed that are known to contain no BPA or a pre-determined concentration of BPA that have also had a precisely known amount of BPA added to a precisely known amount of specimen). These controls allow for extraction efficiency and detector response in the matrix of interest to be determined.

		• Samples (defined as biological specimens of interest for biomonitoring having an unknown concentration of BPA [if any]). • Fortified samples (biological specimens and solvent(s) used for external standards that have had a precisely known amount of BPA added to a precisely known amount of the specimen *at the time of collection*). These allow for the determination of BPA stability during sample collection, transport (if any), storage and analysis. The method should be validated over the entire concentration range reported using a fortified control matrix. Method validation should include a mass balance experiment to determine the absolute method recovery. Fortified controls with a minimum of 3 points or concentrations (a 5 point calibration is recommended) should be prepared and analyzed with each set of samples. These should encompass the full range of concentrations over which BPA is being quantified, and also include the appropriate blanks, controls, and isotopic crossover samples (if appropriate).
36	31-40	The relatively high percentages of parent BPA reported to be found in urine in the studies cited are unlikely to represent what was actually excreted based on the controlled human pharmacokinetics studies of Völkel et al.(CERHR references 62 and 91) where virtually no free BPA was present in the urine of humans dosed orally with BPA. As noted above, hydrolysis of the BPA conjugates present may occur to release parent BPA if specimens are not handled carefully. Alternately, dividing specimens and using field spikes is one method to account for any potential hydrolysis of BPA conjugates during handling and analysis.
37	27-42	These data clearly support the concept that parent BPA is only poorly absorbed and should be used to modify the statement about absorption at the beginning of Section 2.1. An example is found in the publication of Pottenger et al. (CERHR reference 93) who demonstrated that at a relatively high dose of 100 mg/kg, the peak (and transient) blood concentration achieved in female rats was only 2.29 µg/g. If the entire amount administered to these rats (~25 mg) had been "instantaneously" absorbed into an estimated blood volume of 12 ml, the peak concentration would have been in the range of 2 mg/g.

38	4-12	Fecal excretion data are not a definitive measure of the extent of oral absorption. The analysis of parent compound in blood across time is the definitive measure for the determination of bioavailability following oral administration. The fact that more than 50% of fecal elimination occurred at 24 hours post-dosing in the study of Pottenger et al. (CERHR reference 93) is only indicative of enterohepatic circulation that occurs in rats. Inoue et al. (CERHR reference 119) demonstrated that in an everted small intestine rat model, 83% of the BPA is conjugated by the small intestine as it is absorbed. In a separate study, which should be included in the CERHR report, the same laboratory reported 65% hepatic glucuronidation occurred on first pass through the isolated perfused livers of rats.[2] These and other data support the conclusion that BPA undergoes extensive presystemic clearance by both the intestinal tract and liver resulting in poor oral bioavailability.
38	20-29	The oral bolus dose of 1000 mg/kg administered by Takahashi and Oishi in this study was approximately one third to one fourth of the oral LD_{50} in rats. At this dose, saturation of metabolism was likely and the data generated on fetal distribution are not relevant to human exposures, which are many orders of magnitude lower.
42	23-24	It should be noted that the data in Figure 3 of Miyakoda et al. (CERHR reference 102) were sufficiently variable to preclude the conclusions on the relative half-life of BPA in fetuses vs. maternal blood made by the authors and repeated in the CERHR draft report.
45 46	30-40 1-24	It is noted on Line 17 and 18 of page 46 that the authors (Zalko et al.) stated that no differences in metabolites were observed in mice between the oral and subcutaneous routes of administration (data not shown). This is surprising in that there were several unique metabolites of BPA observed following administration to rats by parenteral routes (Pottenger et al., CERHR reference 93) that were not observed following oral administration. Nevertheless, in both a rat study (Pottenger et al.) and a mouse study (Zalko et al, CERHR reference108), there were route-dependent quantitative differences in the metabolites formed that demonstrate the lack of relevance for risk assessment of toxicology studies conducted using these parenteral routes of administration.

[2] Inoue, H., Yokota, H., Makino, T., Yuasa, A. and Kato, S. Bisphenol A glucuronide, a major metabolite in rat bile after perfusion. Drug Metabolism and Disposition 2001; 29: 1084-1087.

47	20-21	The study cited (Uchida et al., CERHR reference 109) used only the subcutaneous route of administration, a route of administration not relevant for human exposure. The conclusion of the study was that the placental barrier does not protect the fetus from bisphenol A. While this conclusion may be accurate for the data at hand, the presystemic clearance of BPA following oral exposure was virtually complete, a biological mechanism that is highly effective in precluding fetal exposure to bisphenol A.
59	47-49	A conclusion is made by the authors that bisphenol A glucuronide was not transferred from the dams to the fetuses based on the finding that there were no differences in the concentration of BPA after treatment of the fetal homogenates with β-glucuronidase. This conclusion may be incorrect as BPA glucuronide has been shown to be readily hydrolyzed in rat fetal tissue.[1] As the use of internal standards was not reported it is possible that hydrolysis of any BPA glucuronide that present occurred prior to treatment with β-glucuronidase.
60	43-46	Many of the BPA concentrations used in the *in vitro* metabolism study by Jaeg et al. (CERHR reference 123), were very high relative to the concentrations of BPA in tissues following *in vivo* administration. Therefore the metabolites found are more an artifact of the system used, rather than a representation of expected *in vivo* metabolites.
61	7-13	Studies in non-mammalian species are not relevant in an assessment of potential risk to human health.

Table 4

Polycarbonate/BPA Global Group Comments on
CERHR Draft Report of December 2006
Sections 2.2-2.6, Pages 66-110

February 2, 2007

Page	Lines	Comment
67	12-15	*"The overall conclusion of the European Union was that is [sic] was somewhat unclear if bisphenol A induces orthodox skin sensitization, photosensitization, or responses in individuals previously sensitized to another substance, such as epoxy resins."* Subsequent to the European Union conclusion, a Local Lymph Node Assay in mice was conducted in 2002 according to OECD testing guideline 429 and following Good Laboratory Practices regulations. This study demonstrated that bisphenol A has neither an irritating nor a sensitizing potential in mice after dermal application. (Bayer AG. Bisphenol A Local Lymph Node Assay in Mice (LLNA/IMDS), Report no. AT00155. H.-W. Vohr. 2002). In 2003, a second Local Lymph Node Assay was performed to investigate the photoreactive potential of bisphenol A in mice according to OECD testing guideline 429 and following Good Laboratory Practices regulations. This test showed that bisphenol A has no photoreactive potential in mice after epicutaneous application and UV-A irradiation. No indication of substance specific UV-dependent activation of the cells of the immune system after dermal application was found by the method used. (Bayer AG. Bisphenol A Study of Photoreactive Potential in Mice (LLNA/IMDS), Report no. AT00413. H.-W. Vohr. 2003).
67 67	19-21 32-33	The recently conducted bisphenol A 2-generation study in mice (CERHR reference 376, discussed in detail in CERHR draft report sections 3 and 4) should be cited in this section to support liver as the target organ not only in rats but also in mice.

67	24-27	It should be mentioned that the European Union concluded that cecal enlargement is not regarded as a toxicologically significant observation of relevance to humans. See the following conclusion in the EU Risk Assessment (identified in CERHR reference 2), section 4.1.2.6.3, page 185, second paragraph: "The caecal enlargement was observed at 25 mg/kg and above and was without any associated histological abnormalities. In addition, it was not observed in a 2-year study at doses up to about 140 mg/kg or a multigeneration study at doses up to 500 mg/kg/day. Consequently, this is not regarded as a toxicologically significant observation of relevance to humans. A NOAEL of 74 mg/kg has been established for rats from a 2-year study.")
67	37-39	It should be mentioned that the European Union concluded that this observation is of doubtful toxicological significance. See the following conclusion in the EU Risk Assessment (identified in CERHR reference 2), section 4.1.2.6.3, page 185, fourth paragraph: "In a 90-day dietary study in dogs, a no effect level of approximately 80 mg/kg was identified, with increases in relative liver weight being the only other finding observed at approximately 270 mg/kg: in the absence of histopathology this finding is of doubtful toxicological significance.")
69	18-20	*"The study authors concluded that change in estrous cyclicity was the only useful endpoint for evaluating the endocrine-mediated effects of bisphenol A."* This sentence might be misleading. The authors used ethinylestradiol (EE) in an independent experiment to investigate if the parameters included in the "enhanced OECD test guideline 407" might be useful to detect endocrine-mediated effects. The authors did observe specific endocrine effects for EE, but did not observe specific endocrine-mediated effects of bisphenol A (except continued diestrus stage over 4 days in 3/9 high dose females). Yamasaki et al. (CERHR reference128) discussed the BPA data as follows: "These findings demonstrate that the estrous cycling, organ weight change and histological changes were useful parameters to detect the endocrine-disrupting effects of EE. On the other hand, the highest dose of BPA was an apparently adverse-effect dose because death or decreased body weight gain was observed in this group, but the endocrine-mediated effect was only detected in the estrous cycling, and other parameters did not change in the highest dose group. Although the dose giving the endocrine-mediated effects brought about the other adverse toxicity changes in this study, the enhanced TG 407 using EE or BPA was considered to be a suitable screening test for the detection of endocrine-mediated effects because endocrine related changes were detected in this assay."

69	---	Table 47: The 7% decrease in male terminal body weight in the high dose group should be 17% (control males 422.5 g, high dose males 352.7 g).
72	---	Figure 2 (and also page 74/Table 48 and page 81/Table 49): • Comparison of estrogen-like activity based on "molar potency/comparator" is error prone and not comprehensible and, consequently, clear indication that these data should be interpreted with caution is warranted. Data provided by Vivacqua et al. (CERHR reference 47) are discussed in detail on page 72 and demonstrate that similar experimental data might lead to substantially different interpretations. For example, BPA might be 1000-fold or 2-fold less active than 17β-estradiol based on the method used. In Table 48, single values were reported for each study and it is not indicated that alternative interpretations are possible. It might be helpful to the reader, to indicate that the relative potency values (half-maximal response or effects at similar molecular concentrations) and/or to indicate the range of molar potency/comparator values calculated by different methods (e.g., Vivacqua et al.: 0.5 – 1000) are dependent on the method used to derive each value. In addition, conventional units such as EC50 (or IC50) or NOAEL could be indicated. • Fish data are not relevant for human risk assessment and should be deleted. • The conclusions that can be drawn from this table are limited because data from juvenile and adult animals, oral and subcutaneous application, different strains and different experimental parameters are all combined together. This point should be stated clearly. Overall, the conclusion on p. 79: "The assertions of some investigators notwithstanding, the Expert Panel notes that oral bisphenol A does not consistently produce estrogenic responses and, when see, estrogenic effects after oral treatment occur at high dose levels" is appropriate.
74	20-21	*" ...and a study using MCF-7 cell proliferation found 17β-estradiol and bisphenol A to have synergistic effects (142)."* It should be mentioned that the study has a number of limitations and weaknesses, and it should be interpreted with caution. For example, BPA and 17β-estradiol seem to have antagonistic activity at low BPA concentrations [compare Table 1 line 1 (17β-estradiol control) and BPA 10 and 40 nM data], and non-monotonic dose response was observed at high BPA concentrations leading to maximal cell proliferation at intermediate BPA concentrations.

79	22	*"These authors treated immature CD-1 mice with bisphenol A and evaluated…"* The route of administration (subcutaneous implantation of osmotic pumps) should be indicated.
81	---	Table 49 • The study-selection criteria are not clear. For example, Hong et al. (CERHR references 220 and 221) investigated pregnant and lactating rats, all other studies in the rat-uterus section investigated immature or ovariectomized animals. Other studies on uterus weight are not mentioned.[1,2] The selection of Laws et al. (CERHR reference 202) as the only study on vaginal opening should be supplemented by other relevant studies, for example 2 studies by Tyl et al. (CERHR references 376 and 411). • The studies on fish and invertebrates are not relevant to human risk assessment and the list of studies does not appear to be comprehensive. • Only uterine weights are listed in most studies on "rat uterus" although other parameters were measured in some studies; additional parameters were listed for some studies but not for others (e.g., clabindin D9k mRNA, cfos expression,…). • The calculation of "molar potency/comparator" is not readily comprehensible [see Comment above].
89	2-4	*"This model showed an increase in ER activity after a single sc bisphenol A dose of 25 µg/kg bw (P = 0.052), with further increases in activity after 0.8 and 25 mg/kg bw."* It should be mentioned that the biological significance of the observed ER activity in transgenic mice at 25 µg/kg is difficult to judge because data were reported only in one figure (Fig. 6) in which the authors did not include data on the control group (oil-control). Further, the number of aminals investigated per dosegroup (n=6-7) is low and no information on the history of use with the animal model were provided.
91	29-30	Change "estrogenic or antiestrogenic" to "androgenic or antiandrogenic."

[1] Central Toxicology Laboratory. Bisphenol-A: Uterotrophic Assay in Immature Rats (Oral Dosing). CTL Unpublished Report No: CTL/P/6029. 1999.
[2] Central Toxicology Laboratory. Bisphenol-A: Uterotrophic Assay in Immature Rats (Subcutaneous Dosing). CTL Unpublished Report No: CTL/P/5943. 1999.

92	21	*"…with bisphenol A* ***[purity not indicated]****…"* Nishino et al. used BPA provided by Bayer: (PtNr. 97.001/Prod.Nr. 04111095). The purity of the material was 99.9% and the analytical material balance can be provided upon request.
93-94	30-33 1-6	The Hunt et al. study (CERHR reference 262) has a significant number of limitations and weaknesses, and caution is warranted in drawing conclusions from it. • The principle effect of BPA measured was congression failure, a misalignment of chromosomes during metaphase stages of meiosis. While one might expect that a misalignment of chromosomes during meiotic metaphase would be associated with improper chromosome segregation, this has not been conclusively demonstrated for meiotic cells. The alignment of chromosomes at metaphase is a dynamic process. Chromosomes that appear to be misaligned at one point might quickly become properly aligned and segregate properly. As a result, chromosome misalignment and aberrant congression represent cellular observations that might eventually lead to aneuploidy but are not considered definitive effects. • Overall, results have been generated by pooling animals of different ages, strains/genotypes, and derived from different breeding stocks (e.g., see Figure 3). It is not appropriate to pool and compare the hyperploid results in 8-12 month old mice with those that are 4 weeks of age. • Because of animal-to-animal variability, the experimental unit used for the statistical analysis of *in vivo* experiments should be the frequency of abnormal events per animal. Hunt et al. used the number of cells as statistical unit and aberrant cells appear to have been pooled from several animals and analyzed without accounting for animal-to–animal variability. Notably, there is no indication of the number of animals used in the various experiments in the report. • To judge whether the results are truly significant increases in congression failure resulting from treatment with BPA, it would be important to compare the treatment results with historical control rates. • It would be important to confirm the observations in an independent experiment before drawing conclusions from the observations. Attia et al (CERHR reference 263) conducted an independent experiment and concluded that the aneuploidy predicted by Hunt et al. could not be demonstrated. Since this study (Attia et al., 263) is published only as an abstract it should be evaluated in detail when more information becomes available. • Significant effects on fertility would be expected if the reported high frequencies of congression failure resulted in high frequencies of aneuploidy since most of the aneuploid events would be lethal to the developing embryo.

		Effects on fertility have not been found in several multi-generation studies in mice and rats at doses lower than, similar to, or substantially higher than the doses in this study (see CERHR references 292, 293, and 376).
108	---	Similar to the Hunt et al. study above, a new study from the same researchers (Susiarjo et al.[3]) could be included in this section. The results of the study are limited by the use of only one dose group treated with bisphenol A by an implanted capsule, a route of administration that is not relevant for human risk assessment. The number of animals in the study is low and it is not clear if statistical evaluations are based on the litter/dam, as would be appropriate, or on another basis. As with the Hunt et al. study above, significant effects on fertility would be expected in the second generation but have not been observed in several multi-generation studies in mice and rats at doses lower than, similar to, or substantially higher than the single doses in this study (see CERHR references 292, 293, 376 and 411).
108	17-18	*"Findings regarding sensitization potential were not clear."* As noted above in comments on page 67/lines 12-15, new studies indicate no sensitizing potential after dermal application.
108	20-23	*"Possible target organs or systems of toxicity identified in repeat-dose animal studies with oral dosing include cecum, liver, kidney, and male and female reproductive systems."* As noted above in comments on page 67/lines 24-27, the European Union concluded that cecal enlargement is not toxicologically significant and defined a NOAEL of 74 mg/kg. Comprehensive studies in rats and mice (Tyl et al. CERHR references 293, 376 and 411) indicate that liver and kidney are the primary target organs. Male and female reproductive systems should not be listed as primary target organs as effects on these organs were observed only at substantially higher doses.

[3] Susiarjo, M., Hassold, T. J., Freeman, E. and Hunt, P. A. Bisphenol A exposure in utero disrupts early oogenesis in the mouse. PLoS Genetic 2007; 3: 1-8.

109	20-21	*"Anti-androgenic activity was demonstrated in in vivo systems using cells transfected with androgen receptor reporting systems (Table 49)."* "In vivo" is somewhat misleading and should be changed to "cell based." The correct table is Table 51.

Table 5

Polycarbonate/BPA Global Group Comments on
CERHR Draft Report of December 2006
Sections 3 and 4 (Human and mammalian reproductive and developmental studies included in the draft CERHR report, excluding neurodevelopmental studies)

2-Feb-07

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Aikawa et al., 2004	*381*	*212*	Y	M	M	Neonatal	SC	Y	*Strengths:* Estradiol used as positive control. *Weaknesses* : PND18 not optimal time to examine histological changes.	*Weaknesses* : The Materials and Methods section reports that there are 10-20 mice per treatment group, but 5-14 animals per treatment group were described in the Results section. It is unclear which animals received which diets. The diet effects are as prominent as the BPA effects. The BPA dose was not corrected for body weight. Significant reduction in body weight was observed in BPA-treated dams and high variation in uterine weights in the BPA-treated groups. SC not a relevant route of exposure.	Slightly useful	Limited utility
Akingbemi et al., 2004	*306*	*140*	Y	M	R	Prenatal/neonatal	Oral gavage	Y	*Weaknesses:* Questions about adequacy of the sample size with respect to the number of litters represented and number of offspring from each litter.	*Weaknesses:* Sample size per treatment group sometimes unclear. One dose level. High inter-experimental variation between the control groups from two experiments. Litters were not considered in randomizing animals to treatment groups. Data were pooled from two experiments and the number of dams treated were not given.	Adequate	Limited utility
Akingbemi et al., 2004	*306*	*140*	Y	M	R	Puberty/Adult	Oral gavage	Y	*Weaknesses* : Inadequate number of animals for analyses of hormonal changes. No histopathological evaluation. No estrogenic positive control.	*Weaknesses* : Sample size per treatment group sometimes unclear. One dose level. High inter-experimental variation between the control groups from two experiments. Litters were not considered in randomizing animals to treatment groups.	Adequate	Limited utility
Akingbemi et al., 2004	*306*	*140*	Y	M	R	Puberty	Oral gavage	N	*Strengths:* Helpful examination of postnatal effects following adolescent exposure.	*Weaknesses:* No relevant endpoints.	Adequate	Inadequate
Al-Hiyasat et al., 2002	*442*	*286*	Y	M	M	Adult	Oral gavage	Y	*Weaknesses:* Small n/group. No method presented for randomization. No dose-response for many endpoints. Fertility in controls is suspect. Males killed shortly after mating, which could have influenced sperm in epididymis. Inappropriate statistics. Statistical significance is suspect.	*Weaknesses* : There was a lack of consistency between the paired testis weight and the left testis weight both qualitatively and quantitatively. There was no clear dose-response pattern with growth. Only relative organ weights reported.	Minimal utility	Inadequate
Al-Hiyasat et al., 2004	*422*	*268*	Y	F	M	Adult	Oral gavage	Y	*Weaknesses* : Small n (10) for fertility endpoint. No confirmation of mating. Fertility rates in low and mid-dose groups result of variability. Body and organ weights measured in 5 mice/dose level.	*Strengths* : Authors noted that their findings were consistent with those of another group that reported dose-dependent effects on the rat uterus. Results were consistent with a dose-dependent effect. *Weaknesses* : Body, ovary and uterus weights were measured in only five (of 15) animals per dose group. Only relative organ weights were provided. Inappropriate statistics.	Limited utility	Limited utility
Anahara et al., 2006	*444*	*291*	Y	M	M	Adult	SC	N	*Weaknesses* : Non-oral. One dose level. No information on BPA purity, type of feed, bedding, or caging. No correlating adverse outcome assessed.	*Weaknesses:* No relevant endpoints.	Limited utility	Inadequate
Ashby et al., 1999	*343*	*177*	Y	B	M	Prenatal	Oral	Y	*Strengths* : Close replication of vom Saal and Nagel studies (only diet differed). Used solo and group cages. Study supported by NTP statistics subpanel. DES used as positive control. *Weaknesses* : DES showed no response, unclear why.	*Strengths* : Data analyzed by litter. The lack of effects with orally administered DES is not a weakness of this study. The dose of DES administered was three orders of magnitude below the LOEL for DES in other studies (McLachlan, 1981; Newbold, 1995). Orally administered DES, like BPA, is extensively glucuronidated via first pass metabolism (Metzler, 1981) and DES-glucuronide has been shown to not have estrogenic activity (Waechter et al., 2001). The alleged effects of low doses of orally administered DES on prostate weight (Vom Saal et al., 1997) have only been claimed by one laboratory and like BPA these findings have not been reliably reproduced. *Weaknesses* : Diet fed to animal included soy. Control prostatic weight may have been artifactually high.	Very useful	High utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Ashby et al., 2003	*433*	*278*	Y	M	R	Adult	Oral gavage	Y	*Strengths* : Very well-conducted. Comprehensive. Studied potential variables that might account for discrepancies in other studies.	*Strengths* : Stainless steel cages. Adult exposure over short time period. Three dose levels. This study was conducted under GLP standards. *Weaknesses* : Number of animals/group unclear; 3 animals per cage at start, 10 animals altogether. Inter-experiment variability among control groups for body and epididymis weight and sperm and DSP parameters; authors concluded this was due to chance or intrinsic study-to-study variability. Body weight was measured on a daily basis, but values were given only at start and end of studies. Ashby et al. (2003) compared their control values for daily sperm production with the mean of 34 control values taken from the literature. They found that their control values and those of Ashby et al. (1997) and Tinwell et al. (2002) were quite similar to each other and to the literature average, while control values from Sakaue et al. (1999), Ohsako et al. (2001) and Sakaue et al. (2001), in particular, were higher. Ashby et al. (2003) noted that higher than normal control values could have led to apparent decreases in BPA-treated animals that were seen by Sakaue et al. (2001).	Highly useful	High utility
Atanassova et al., 2000	*327*	*162*	N	M	R	Neonatal	SC	Y	*Strengths:* BPA compared with other estrogenic compound. Group known to have extensive expertise in testis biology and male fertility in general. *Weaknesses* : Only one dose level of BPA used, this varied on a weight basis.	*Weaknesses* : High dose only. SC not a relevant route of exposure.	Suitable	Limited utility
Cagen et al., 1999	*294*	*131*	Y	M	R	Prenatal/postnatal	Drinking water	Y	*Strengths* : Large number of dose levels. Large number of animals per dose level. Study performed with technical care. Use of positive control and two negative control groups. *Weaknesses* : Little effect with diethylstilbestrol treatment.	*Strengths* : Large dose range used. Large dose groups. Data were analyzed with the litter as the experimental unit. This study was conducted under GLP standards and used two methods for sperm counting. The lack of effects with orally administered DES is not a weakness of this study. The dose of DES administered was three orders of magnitude below the LOEL for DES in other studies (McLachlan, 1981; Newbold, 1995). Orally administered DES, like BPA, is extensively glucuronidated via first pass metabolism (Metzler, 1981) and DES-glucuronide has been shown to not have estrogenic activity (Waechter et al., 2001). The alleged effects of low doses of orally administered DES on prostate weight (Vom Saal et al., 1997) have only been claimed by one laboratory and like BPA these findings have not been reliably reproduced.	Adequate (with qualifications)	High utility
Cagen et al., 1999	*294*	*131*	Y	F	R	Adult	Drinking water	Y	*Strengths:* Large number of dose levels. Large number of animals per dose level. Study performed with technical care. Use of positive control and two negative control groups. *Weaknesses* : Little effect with diethylstilbestrol treatment.	*Strengths* : Large dose range used. Data were analyzed with the litter as the experimental unit. This study was conducted under GLP standards and use of two methods for sperm count. The lack of effects with orally administered DES is not a weakness of this study. The dose of DES administered was three orders of magnitude below the LOEL for DES in other studies (McLachlan, 1981; Newbold, 1995). Orally administered DES, like BPA, is extensively glucuronidated via first pass metabolism (Metzler, 1981) and DES-glucuronide has been shown to not have estrogenic activity (Waechter et al., 2001). The alleged effects of low doses of orally administered DES on prostate weight (Vom Saal et al., 1997) have only been claimed by one laboratory and like BPA these findings have not been reliably reproduced.	Adequate (with qualifications)	High utility
Cagen et al., 1999	*342*	*176*	Y	F	M	Adult	Oral pipette	Y	*Strengths:* Use of litter analysis. Large sample size. NTP subpanel agreed with authors' conclusions. *Weaknesses* : Strain in this study differs from those in Vom Saal and Nagel studies. Termination at 90 d (vs. 180 d). Solo housing. Lack of response to DES, though this could have been an inappropriate dose.	*Strengths* : Multiple doses of BPA used. Analyses of BPA concentration in dose solutions. The use of a different strain should not be considered a 'weakness'; as the same strain (CF-1 mice) was used and has been shown to be sensitive to estrogen induced effects (Tyl et al, 2006). Solo housing should not be considered as a weakness, but a strength as solo housing removes the confounding effect of group housing on male reproductive organ weight (Bartos and Brain, 1993). In addition, this strain was sensitive to estrogenic effects in a study using individual housing, (Tyl et al., 2006). This study was conducted under GLP standards.	Useful	High utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Cagen et al., 1999	*342*	*176*	Y	M	M	Prenatal	Oral pipette	Y	*Strengths:* Use of litter analysis. Large sample size. NTP subpanel agreed with authors' conclusions. *Weaknesses* : Strain in this study differs from those in Vom Saal and Nagel studies. Termination at 90 d (vs. 180 d). Solo housing. Lack of response to DES, though this could have been an inappropriate dose.	Strengths: Multiple doses of BPA used. Analyses of BPA concentration in dose solutions. The use of a different strain should not be considered a 'weakness'; as the same strain (CF-1 mice) was used and has been shown to be sensitive to estrogen induced effects (Tyl et al, 2006). Solo housing should not be considered as a weakness, but a strength as solo housing removes the confounding effect of group housing on male reproductive organ weight (Bartos and Brain, 1993). In addition, this strain was sensitive to estrogenic effects in a study using individual housing, (Tyl et al., 2006). This study was conducted under GLP standards.	Useful	High utility
Chitra et al. 2003	*436*	*282*	Y	M	R	Adult	Oral gavage	Y	*Weaknesses* : Small sample size per group. Low confidence in control values (~95% motile sperm, literature suggests 60-85%). Apparent decrease in motile sperm in BPA-treated animals consistent with published controls (maybe species/strain differences). Olive oil could affect BPA stability. ROS may interact with BPA. Data could be the same as in Chitra et al. (435)	*Weaknesses:* Graphs are difficult to interpret. Control values differ in magnitude and variance from other studies, particularly those conducted by NTP (NTP, 2001). No statistical test was conducted to determine whether reduction in sperm motility and count was dose-dependent. It was unclear how many animals were used for measuring sperm. The following were not reported for epididymis histology: how many sections were examined, what percentage of the sections examined were abnormal for either treated or control animals, and what scoring method was used to determine "degeneration." No data were provided on epididymis weight, which was of note because Chitra et al. (2003a) found epididymis weight was reduced under similar conditions. Finding stands in contrast to the lack of effect reported in other studies, including the much higher injection doses of Takahashi and Oishi (2003). Chitra et al. (2003b) showed only one photomicrograph in their publication for each dose for each region of the epididymis. The results for sperm count and motility are almost identical in the separate studies with 45 and 60 d dosing (Chitra 435 and 436), and the reported error bars are exceptionally small.	Limited utility	Inadequate
Chitra et al., 2003	*435*	*281*	Y	M	R	Adult	Oral gavage	Y	*Strengths:* Relatively well-conducted study. Consistent dose-response for testis and epididymis weights and sperm parameters. Consistent with Sakaue et al. *Weaknesses:* Unclear how olive oil vehicle affected BPA stability. Small sample size per group.	*Weaknesses* : It was unclear how many animals were used for measuring sperm. Sperm count and motility data are shown in figures (values not shown) and are difficult to interpret. The major objective of the study was to examine oxidative stress. Control values differ in magnitude and variance from other studies, particularly those conducted by NTP (NTP, 2001). No statistical test was conducted to determine whether reduction in sperm motility and count was dose-dependent. The effect on testis weight was small (decrease of 2.6 – 3.5%). Epididymis and testis weights were decreased while prostate weights were increased at all doses. If one hypothesizes that the prostate and epididymis are affected by BPA via an endocrine disruptive mechanism, it is not unreasonable to expect that both organs, being testosterone targets, respond to the xenobiotic in the same direction (i.e., increase or decrease in size) in a dose-related fashion. The results for sperm count and motility are almost identical in the separate studies with 45 and 60 d dosing (Chitra 435 and 436), and the reported error bars are exceptionally small in view of the nature of the endpoints in question, particularly so for sperm motility.	Limited utility	Inadequate
Della Seta et al., 2005	*420*	*267*	Y	F	R	Adult	Oral pipette	Y	*Weaknesses* : Low pregnancy rate in controls. One dose level (no dose-response). Not stated whether analysts blinded to treatment. Residual BPA in oral cavity may have altered taste perception. This was a behavioral study, not designed to test fertility.	*Weaknesses:* There was no explanation as to why animals were cross-fostered.	Limited utility	Limited utility
Della Seta et al., 2006	*323*	*159*	Y	M	R	Puberty	Oral gavage	N	To be added.	*Strengths:* Appropriate statistics. *Weaknesses:* Primary endpoints neurological and behavioral. Body weight only developmental endpoint. Only one BPA dose level.	To be added	Inadequate
Durando et al., 2007	*291*	*129*	Y	F	R	Prenatal	SC pump	Y	To be added	*Weaknesses* : No information on the exact number of female offspring examined for various endpoints. SC not a relevant route of exposure. One dose level. High variability was noted for control group AGD from this study compared to work cited in Murray et al., 2007. The litter did not appear to be used as the statistical unit of analysis. Mammary gland development studies in rodents may not represent useful models for humans due to significant differences in the hormonal milieu.	To be added	Limited utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Elswick et al., 2000	*296*	*132*	Y	M	R	Prenatal/neonatal	Drinking water/Oral gavage	Y	*Strengths* : "Interesting" approach to examination of background variance and litter effects. Data argues for multiple pup/litter sampling. Significant effects noted in only 1 block, raising questions about technician experience/training. *Weaknesses:* Does not report original data.	*Strengths* : Large dose range. *Weaknesses:* Only 1-2 animals per litter sampled. Authors noted large intralitter variability in prostate weights. Authors state that the article was not intended to be a report on the biological effects of BPA. The NTP low dose workshop 2001) statistics subpanel concluded that multiple pups/litter increases statistical power and hence will reduce the false negative rate (does not affect the false positive rate). Hence, for studies that measured prostate weight for multiple pups/litter, this should be considered a strength.	Adequate (with qualification)	Inadequate
Ema et al., 2001	*292*	*130 300*	Y	B	R	2-generation	Oral gavage	Y	*Strengths* : This well-designed comprehensive low-dose assessment of potential bisphenol A-related effects on multiple generations of rats examined a wide variety of hormonally sensitive endpoints. The study had appropriate power with an appropriate number of rats per group. Oral. The concentrations of the dosing solutions were verified. *Weaknesses* : It would have been helpful if a dose level that caused maternal toxicity was also used; however, given the objective of this study it is a minor point.	*Strengths:* Low doses. Multi-generation study. Examined multiple reproductive/developmental endpoints. Large number of animals used. Considered relevance of changes that were not consistent across generations. Large dose range. This study was conducted under GLP regulations and was compliant with the US EPA testing guidelines (U.S. EPA, OPPTS, 837.3800, 1998) with the exception that the highest dose did not produce evidence of toxicity.	Adequate/Highly useful	High utility
Evans et al., 2004	*382*	*214*	Y	F	S	Puberty	IM injection	Y	*Strengths:* Unique animal model. *Weaknesses* : High dose.	*Weaknesses:* IM injection is not a relevant route of exposure. Non-rodent model.	Useful	Limited utility
Fisher et al., 1999	*324*	*159*	N	M	R	Neonatal	SC	Y	*Weaknesses* : Single dose of BPA used. No tissues other than the testis examined.	*Weaknesses:* High dose only. SC not a relevant route of exposure.	Suitable	Limited utility
Fukumori et al., 2003	*332*	*165*	Y	M	R	Neonatal	SC	Y	*Weaknesses* : Animals sacrificed at young age, possibly before prostatic development was complete. Original article in Japanese, figures not translated to English.	*Weaknesses* : No information on statistics given. Study is unclear regarding the number of rats used. Incomplete information on methodology. Tables not in English. SC not a relevant route of exposure.	Suitable	Inadequate
Funabashi et al., 2001	*416*	*263*	Y	F	R	Prenatal and neonatal	SC	N	*Strengths* : Design and statistics appear to be appropriate. *Weaknesses:* Sample size was small. Control for litter effects was not clear.	*Weaknesses:* No relevant endpoints. SC not a relevant route of exposure.	Not clearly stated	Inadequate
Funabashi et al., 2003	*418*	*265*	N	F	R	Adult	SC	N	*Strengths* : Sufficient sample size. Well-conducted study. *Weaknesses:* Non-oral.	*Weaknesses:* No relevant endpoints. High dose only.	Limited utility	Inadequate
Funabashi et al., 2004	*419*	*266*	N	F	R	Adult	SC	N	*Weaknesses* : Non-oral. High dose. Single dose. No functional/physiological correlate.	*Weaknesses* : No relevant endpoints.	Limited utility	Inadequate
Goloubkova et al., 2000	*216*	*262*	N	F	R	Adult	SC	Y	*Weaknesses:* Non-oral. High dose. Single dose. No functional/physiological correlate.	None.	Not clearly stated	Limited utility
Gupta, 2000	*347*	*179*	Y	M	M	Prenatal	Oral	Y	*Strengths* : Oral. Low dose. DES used as positive control. Prostate measurements at 3 postnatal timepoints. In vitro study to support in vivo results. *Weaknesses:* Statistical analysis could be weakness, but response satisfactory. Lack of attention to litter effects and effects on AGD.	*Weaknesses* : Only one BPA dose used. Used only one animal per litter at each timepoint (not relevant for all endpoints, just organ weights). This study not conducted under GLP standards. The raw data from this study were not evaluated by the NTP low dose peer review workshop participants, therefore, a rigorous peer review has not been conducted.	Very useful	Limited utility
Hanaoka et al., 2002	*82*	*261*	Y	M	H	NA	BADGE spray	N	*Strengths:* Blood draws and urine samples time standardized. Measured BPA with HPLC. Appropriate statistics. Confounders (age, alcohol use, smoking) considered. *Weaknesses* : Small sample size led to limited power. Multiplicity of exposures.	*Weaknesses:* Exposure to BADGE, not BPA. Conjugated BPA likely included in the BPA measurement.	Not clearly stated	Inadequate
Herath et al., 2004	*439*	*284*	Y	M	R	Adult	SC	Y	*Weaknesses* : Non-oral. Inconsistency in hormone data.	*Weaknesses* : Animals were exposed to luteinizing hormone releasing hormone (LHRH) in addition to BPA. Only one dose was tested.	No utility	Inadequate
Ho et al. 2006	*336*	*169*	Y	M	R	Neonatal	SC	Y	*Strengths* : "The paper has many strengths, from the use of multiple, biologically relevant doses of bisphenol A…" Search done to identify molecular mechanisms for observations made.	*Weaknesses* : Small number of animals used. One BPA dose group. SC not a relevant route of exposure. Half of the animals treated with estradiol and testosterone capsules. One reproductive/developmental endpoint. Statistics not described but appear to be inappropriate.	Suitable	Limited utility
Honma et al., 2002	*364*	*195*	Y	F	M	Prenatal	SC	Y	*Strengths* : Low doses.	*Strengths* : Statistical analyses were performed using litter-based values. *Weaknesses* : SC not a relevant route of exposure. Changes observed within a dose group were not consistent over time for several endpoints. The findings are inconsistent with an estrogenic effect, as one would predict an estrogenic compound to cause a decrease in anogenital distance.	Useful	Limited utility
Howdeshell et al., 1999 (Howdeshell and vom Saal, 2000)	*345*	*178*	Y	F	M	Prenatal	Oral	Y	*Strengths* : Oral. *Weaknesses* : Assessment of puberty onset using vaginal smears. Omission of description of husbandry conditions. Lack of positive control. Time from vaginal opening to first estrous is not a standard endpoint. Authors say statistics are conducted on a per litter basis, but some statistics appear to be on a per-pup basis.	*Weaknesses* : Only one dose level. The use of intrauterine position identification is not a validated practice in guideline studies and is not relevant to the human situation.	Useful (with some exceptions)	Limited utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Ichihara et al., 2003	*304*	*137*	Y	M	R	Prenatal/neonatal	Oral gavage	Y	*Strengths* : Wide dose range. Several endpoints evaluated. Design reasonable for endpoints measured. Good screening study. *Weaknesses* : Sample sizes inadequate for prostate cancer endpoint and hormonal endpoints.	*Weaknesses:* Unclear if animals from different litters were distributed among treatment groups. Number of animals unclear (number in the figures does not correspond to the text).	Inadequate (alone, but possibly useful when considered with other studies)	Limited utility
Ichihara et al., 2003	*304*	*137*	Y	F	R	Adult	Oral gavage	Y	*Strengths* : Wide dose range. Several endpoints evaluated. Design reasonable for endpoints measured. Good screening study. *Weaknesses* : Sample sizes inadequate for prostate cancer endpoint and hormonal endpoints.	*Weaknesses* : Unclear if animals from different litters were distributed among treatment groups. Number of animals unclear (number in the figures does not correspond to the text).	Inadequate (alone, but possibly useful when considered with other studies)	Limited utility
Iida et al., 2002	*350*	*181*	Y	M	M	Prenatal	Oral gavage	Y	*Strengths* : Oral exposure. Good dose-response trend. *Weaknesses* : Lack of information on husbandry. Small sample size. Lack of adjustment for litter effects.	*Weaknesses:* There was an apparent dose-response trend, with a greater frequency of morphological changes with increasing doses; however, no statistical tests were performed to confirm this. Data were reported for the randomly selected animals used in histopathological exams.	Not useful by itself	Limited utility
Iwasaki and Totsukawa, 2003	*365*	*196*	Y	F	M	Prenatal	SC	Y	*Strengths* : Three dose levels. Estradiol used as positive control.	*Weaknesses:* SC not a relevant route of exposure. The day of vaginal opening was noted to be delayed at the lowest dose, unchanged at the middle dose, and advanced at the highest dose. Uterine weight decreased in the lower dose group and increased in higher dose group after treatment with estradiol. Pup survival and body weight were affected only in the lowest dose group. One pup per litter used for the uterotrophic assay, all other endpoints examined the litter as a whole.	Moderately useful	Limited utility
Kabuto et al., 2004	*107*	*200*	Y	M	M	Prenatal/neonatal	Oral (drink)	Y	*Strengths:* Low doses. Oral exposure. *Weaknesses:* Use of males only.	*Weaknesses* : The main objective of this study was to investigate the modifications in endogenous antioxidant capacity and oxidative damage in the brain, liver, kidney and testis. The magnitude of the decrease in testis weight was greater for the lower dose group. The experimental unit appears to be the individual offspring, not the litter. There was no positive control. The number of offspring was not indicated. Only 8-13 offspring per dose group were included in body and organ weight analyses. Lack of dose-response for some endpoints.	Useful	Limited utility
Kato et al., 2003	*333*	*166*	N	F	R	Neonatal	SC	Y	*Strengths* : Study was carefully performed and documented. *Weaknesses* : Only high dose of BPA used.	*Weaknesses:* SC not a relevant route of exposure.	Useful	Limited utility
Kato et al., 2006	*335*	*168*	Y	M	R	Neonatal	SC	Y	*Strengths:* Wide range of BPA doses used. Endpoints relevant to overall topic.	*Weaknesses* : SC not a relevant route of exposure. "The present data revealed that neonatal exposure [sp] to BPA caused no adverse effects on males and estrogen-mediated response in the testis. In contrast E2 given neonatally elicited incomplete preputial separation, lower copulatory rate, decrease in reproductive organ weight and number of sperm, degeneration of germinal epithelium, and alteration of mRNA expression in the testis. In the males given E2, however, no changes were found in the body weight, observation of fetuses, sperm analysis or serum testosterone levels."	Suitable	Limited utility
Kawai et al., 2003	*358*	*191*	Y	M	M	Prenatal	Oral pipette	Y	*Strengths:* Two low dose levels. Oral exposure. *Weaknesses* : Lack of husbandry information. Lack of consideration of litter effects.	*Weaknesses:* The main objective of this study was to examine the influence of BPA on aggression during the maturation process in male mice. The magnitude of the decrease in relative testis weight was greater for the offspring of the dams exposed to 0.002 mg/kg-d (the lower of the two dose groups). Alterations did not persist and were not evident at 16 weeks. Data analyzed were obtained from randomly selected animals, not all animals. In most cases, specific data were missing, with outcomes shown only in figures. The experimental unit appears to be the individual offspring, not the litter.	Moderately useful	Inadequate
Khurana et al., 2000	*331*	*165*	N	B	R	Neonatal	SC	N	*Strengths* : "Moderate" dose levels used. Male and female animals assessed. *Weaknesses* : Lower dose showed more marked effects than the higher dose.	*Weaknesses* : High dose only. SC not a relevant route of exposure. No relevant endpoints.	Suitable	Inadequate
Kim et al. 2002	*125*	*280*	Y	M	R	Puberty/Adult	Oral (drink)	Y	*Weaknesses* : Absence of information on animal groups size and other study design features (BPA purity, feed, caging, etc.).	*Strengths:* Large dose range. *Weaknesses:* Article written in Korean. The number of samples per animal analyzed for sperm counts and histological examinations was not stated.	Not useful	Limited utility
Kim et al., 2001	*276*	*112*	N	B	R	Prenatal	Oral gavage	Y	*Strengths* : Standard embryo-fetal developmental toxicity study. Incorporates no effect dose and high maternally toxic dose. Measures AGD. Observed fetal effects only at level that caused maternal toxicity. *Weaknesses* : No information about birth process. Postnatal viability/function. Gross visceral exam may be insensitive to certain abnormalities.	*Strengths* : Good sample size. *Weaknesses* : High doses only.	Adequate	Adequate

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Kim et al., 2001	*276*	*112*	N	F	R	Adult	Oral gavage	Y	*Strengths:* Standard embryo-fetal developmental toxicity study. Incorporates no effect dose and high maternally toxic dose. Measures AGD. Observed fetal effects only at level that caused maternal toxicity. *Weaknesses* : No information about birth process or postnatal viability/function. Gross visceral exam may be insensitive to certain abnormalities.	*Strengths* : Good sample size. Some dose-response related effects seen. *Weaknesses:* High doses only.	Adequate	Adequate
Kim et al., 2003	*105*	*113*	Y	B	R	Prenatal	Oral gavage	Y	*Strengths:* Good verification of dosing solution. *Weaknesses:* Number of animals used unclear. Unclear whether fetal data was analyzed with the litter as a unit. Unclear when dams were killed and analyzed. Unclear what was actually done in study. No dose-related effects. Small sample size.	*Strengths* : Large dose range. *Weaknesses:* Unclear whether data analyzed by animal or by litter. Groups are not statistically equivalent at the beginning (e.g. body weight at day 0 in 200 µg group increased compared to control). Inconsistencies between text and figures. Values for BPA concentration in the serum similar at different doses.	Inadequate	Inadequate
Kim et al., 2003	*105*	*113*	Y	F	R	Adult	Oral gavage	Y	*Strengths:* Good verification of dosing solution. *Weaknesses* : Unclear what was actually done in study. Number of animals used unclear. Unclear whether fetal data was analyzed with the litter as a unit. Unclear when dams were killed and analyzed. No dose-related effects. Small sample size.	*Strengths* : Large dose range. *Weaknesses:* Unclear whether data analyzed by animal or by litter. Groups are not statistically equivalent at the beginning (e.g. body weight at day 0 in 200 µg group increased compared to control). Inconsistencies between text and figures. Values for BPA concentration in the serum similar at different doses.	Inadequate	Inadequate
Kobayashi et al., 2002	*300*	*136*	Y	B	R	Prenatal/neonatal	Oral gavage	Y	*Strengths* : Study was better able to address maternal toxicity than offspring outcomes. *Weaknesses* : Sample sizes too small to judge postnatal endpoints.	*Weaknesses* : Unexpected high mortality in the 400 mg/kg-d dose group.	Inadequate	Limited utility
Kobayashi et al., 2005	*302*	*143*	Y	B	R	Prenatal/neonatal	Oral gavage	N	*Strengths:* Wide range of doses used in the study. *Weaknesses:* Dose levels high. Limited endpoints addressed. Small number of animals used.	*Weaknesses:* No relevant reproductive endpoints.	Minimal utility	Inadequate
Kubo et al., 2001	*310*	*145*	Y	B	R	Prenatal/neonatal	Oral (drinking water)	Y	*Strengths:* Variety of biological and behavioral endpoints assessed. *Weaknesses:* Lack of experimental detail reported for the study. Difficult to determine the amount of BPA the animals received.	*Weaknesses* : Primary endpoints are neurological and behavioral effects. Only one dose level.	Limited utility	Limited utility
Kubo et al., 2003	*311*	*146*	Y	B	R	Prenatal/neonatal	Oral (drinking water)	Y	*Weaknesses:* Failure to describe how much BPA dams received during pregnancy. No effects noted on reproductive tract.	*Strengths:* Data from the suckling period was examined by litter. *Weaknesses* : Data from post-weaning was analyzed for individual animals. Number of offspring used for tests is unclear and appears to differ depending on endpoint examined.	Adequate	Limited utility
Kwon et al., 2000	*297*	*144*	Y	B	R	Prenatal/neonatal	Oral gavage	Y	*Strengths* : Study was well performed and presented. Wide dose range. Positive control used. Good number of reproductive organs and endpoints evaluated. *Weaknesses:* Limited analysis of reproductive organs. Did not determine pup exposure during lactation.	*Strengths:* Good variety of endpoints examined. *Weaknesses* : Only one dose below 5 mg/kg-d.	Adequate	Adequate
Kwon et al., 2000	*297*	*144*	Y	F	R	Adult	Oral gavage	Y	*Strengths* : Study was well performed and presented. Wide dose range. Positive control used. Good number of reproductive organs and endpoints evaluated. *Weaknesses* : Limited analysis of reproductive organs. Did not determine pup exposure during lactation.	*Weaknesses* : Only one dose below 5 mg/kg-d.	Adequate	Adequate
Laviola et al., 2005	*359*	*192*	Y	B	M	Prenatal	Oral pipette	Y	*Weaknesses* : One dose level. Small sample size.	*Strengths:* Data for prenatal treatment was analyzed for between-litter effects. *Weaknesses* : The main objective of this study was to measure behavioral effects from treatment with BPA and/or methoxychlor. Data not shown for body weight or sex ratio at birth. All data besides prenatal treatment analyzed for within-litter effects.	Slightly useful	Limited utility
Luconi et al., 2001	*414*	*260*	Y	M	H	NA	NA	N	*Weaknesses:* Limited information on spermatozoa samples (i.e., number of donors or samples per donor). BPA used as a tool to examine receptor behavior. BPA concentration too low. One BPA concentration used. No assessment of whether effects altered sperm function or behavior.	*Weaknesses: In vitro* exposure. No relevant endpoints.	Not useful	Inadequate
Markey et al., 2001	*360*	*192*	Y	F	M	Prenatal	SC pump	Y	*Strengths* : Use of SC pump. Examined mammary gland. Low dose levels.	*Weaknesses* : Non-oral exposure. Low number of animals (6-10) and only 1 pup/litter selected. The use of s.c. injections using an osmotic pump should be considered a weakness (not a strength) as this is a route of exposure not relevant to potential human exposures. It is unclear if the positive controls were run concurrently with the BPA treated groups. No dose-response for many changes reported.	Useful	Limited utility
Markey et al., 2003	*361*	*193*	Y	F	M	Prenatal	SC pump	Y	*Strengths* : Use of SC pump. Low doses. Multiple measures of ovarian cycle and tissues.	*Weaknesses:* The experimental unit appears to be the individual offspring, not the litter. The number of animals per treatment group in the different experiments is not clear. The reported data were incomplete for several experiments. The use of s.c. injections using an osmotic pump should be considered a weakness (not a strength) as this is a route of exposure not relevant to potential human exposures. The effect of the pump implantation was not investigated. Same comment as above for use of osmotic pump.	Useful	Limited utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Markey et al., 2005	*379*	*209*	Y	F	M	Prenatal/neonatal	SC pump	Y	*Strengths:* Use of SC pump.	*Weaknesses* : The Methods section indicates that there was one animal examined per litter per treatment group (10 litters per group), but the Results section describes 11 animals in the low-dose BPA group. The number of animals used differed between tests. The use of s.c. injections using an osmotic pump should be considered a weakness (not a strength) as this is a route of exposure not relevant to potential human exposures. Same comment as above for use of osmotic pump. No dose-response for many changes reported or evaluation of samples for only one dose level reported.	Very useful	Limited utility
Matsumoto et al., 2004	*371*	*202*	N	B	M	Prenatal/neonatal	Oral	Y	*Weaknesses* : Difficulty calculating BPA intake. High BPA exposure. Lack of information on dam number and husbandry. High level of pup body weight decrement and mortality.	*Weaknesses* : Only one dose level used. Body weight is only relevant developmental endpoint.	Not useful	Limited utility
Matsumoto et al., 2004	*371*	*269*	N	F	M	Adult	Oral	Y	*Weaknesses* : Single high dose that likely induced maternal toxicity. Difficult to dilineate if findings in mouse pups are the result of BPA-related effects of maternal toxicity or effects on pups.	*Weaknesses* : One dose level used.	No utility	Limited utility
Miyatake et al., 2006	*374*	*204*	Y	M	M	Prenatal/postnatal	Oral	N	*Weaknesses* : Only two dose levels.	*Weaknesses:* No relevant reproductive endpoints.	Moderately useful	Inadequate
Mizuo et al., 2004	*373*	*203*	Y	M	M	Prenatal/neonatal	Oral	N	*Strengths* : Wide dose range. *Weaknesses* : No information on number of dams per dose group, purity of BPA, feed, caging, bedding, ages of testing, or sex of offspring tested.	*Weaknesses:* No relevant reproductive endpoints.	Not useful	Inadequate
Moon et al., 2001	*445*	*291*	N	M	B	Adult	IP injection	N	*Weaknesses* : Non-oral exposure. No concurrent positive control. High dose.	*Weaknesses:* No relevant endpoints.	No utility	Inadequate
Morrison et al., 2003	*383*	*215*	Y	F	S	Juvenile	IM injection	Y	*Strengths: Useful in combination with Evans et al., 2004. Follow-up study of Evans et al.* *Weaknesses: No information on feed or bedding composition or caging materials. BPA purity not reported.*	*Weaknesses:* Small sample size. Short fixation period for histopathology. Animals were ovariectomized during dosing. One BPA-treated animal (of six) was dropped from statistical analyses because it was more than three standard errors from the mean of the endometrial to myometrial ratios within the BPA-treatment group.	Limited utility	Inadequate
Morrissey et al., 1987	*273*	*111*	N	B	R	Prenatal	Oral gavage	Y	*Strengths* : Adequate sample sizes. Verification of dosing solutions. Maternal toxicity associated with lethality. Traditional embryofetal development study. Appropriate conclusion about teratogenicity (based on lack of fetal endpoints in presence of maternal toxicity). *Weaknesses* : Absence of data from high-dose group. Absence of no-effect dose. No information about birth process, postnatal viability/function. Gross visceral exam likely insensitive to certain abnormalities.	*Weaknesses* : High doses only.	Adequate	Adequate
Morrissey et al., 1987	*273*	*111*	N	F	R	Adult	Oral gavage	Y	*Strengths:* Adequate sample sizes. Verification of dosing solutions. Maternal toxicity associated with lethality. Traditional embryofetal development study. Appropriate conclusion about teratogenicity (based on lack of fetal endpoints in presence of maternal toxicity). *Weaknesses:* Absence of data from high-dose group. Absence of no-effect dose. No information about birth process, postnatal viability/function. Gross visceral exam likely insensitive to certain abnormalities.	*Weaknesses* : High doses only.	Adequate	Adequate
Morrissey et al., 1987	*273*	*173*	N	B	M	Prenatal	Oral gavage	Y	*Strengths:* Oral exposure route. *Weaknesses* : High doses only.	None. Comment: CERHR states 'the oral route of exposure is a strength'. This be listed as a strength for all oral exposure studies. Also, shouldn't all studies not using the oral route be identified as a weakness?	Moderate utility	Adequate
Morrissey et al., 1987	*273*	*173*	N	F	M	Adult	Oral gavage	Y	*Strengths* : Oral exposure route. *Weaknesses* : High doses only.	None. Comment: CERHR states 'the oral route of exposure is a strength'. This be listed as a strength for all oral exposure studies. Also, shouldn't all studies not using the oral route be identified as a weakness?	Moderate utility	Adequate
Munoz de Toro et al., 2005	*363*	*210*	Y	F	M	Prenatal and neonatal	SC pump	Y	*Strengths* : Relevant doses. Long-term perinatal exposure. Tested similar doses as Markey et al., 2005 (379).	*Strengths* : Food, bedding, and cages tested negligible for estrogenicity by E-SCREEN assay. Water supplied in glass bottles. *Weaknesses:* SC not a relevant route of exposure. One pup per litter used for each experiment. No data is shown for four-month-old animals. Absence of dose-response for many changes reported.	Very useful	Limited utility
Murray et al. 2007	*290*	*128*	Y	B	R	Prenatal	SC pump	Y	To be added.	*Weaknesses:* Number of dams or pups used in the study not indicated. No dose-response seen with increased incidence of hyperplastic ducts. SC not a relevant route of exposure. High variability was noted for control group AGD from this study compared to work cited in Durando, 2006.	To be added	Inadequate
Naciff et al. 2005	*288*	*127*	Y	F	R	Adult	SC	Y	*Strengths* : Relevant endpoints. Good strategy used. Adequate numbers of animals for gene expression. *Weaknesses* : Small number of animals for histopathology. Concern over the use of neat DMSO as vehicle.	*Strengths* : Large dose range. *Weaknesses* : SC not a relevant route of exposure. Few functional reproductive endpoints.	Adequate	Limited utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Naciff et al., 2002	*246*	*126*	N	B	R	Prenatal	SC	Y	*Weaknesses* : Sample size small but adequate for analyses. Endpoints appropriate but of limited nature. Concern about the use of neat DMSO as vehicle.	*Weaknesses* : High doses only. SC not a relevant route of exposure.	Adequate (with qualification)	Limited utility
Naciff et al., 2005	*288*	*127*	Y	B	R	Prenatal	SC	Y	*Strengths* : Relevant endpoints. Good strategy used. Adequate numbers of animals for gene expression. *Weaknesses:* Small number of animals for histopathology. Concern over the use of neat DMSO as vehicle.	*Strengths* : Large dose range. *Weaknesses:* Non-oral exposure. Few functional reproductive endpoints.	Adequate	Limited utility
Nagao et al. 2002	*369*	*199*	Y	M	M	Prenatal	Oral gavage	Y	*Strengths* : Oral exposure. Three low dose levels. Careful description of methods. Use of low-phytoestrogen diet. Confirmation that mice strain was estrogen-sensitive.	*Strengths:* Doses determined based on daily body weights. Good dose range.	Very useful	High utility
Nagao et al., 1999	*325*	*160*	N	B	R	Neonatal	SC	Y	*Strengths* : Study well-performed and documented. *Weaknesses* : Only one single high dose used for BPA. Exposure period may have excluded more sensitive time periods.	*Weaknesses* : Non-oral exposure. High dose level.	Suitable	Limited utility
Nagao et al., 2002	*369*	*199*	Y	M	M	Puberty	Oral gavage	Y	*Strengths* : Oral exposure. Three low dose levels. Careful description of methods. Use of low-phytoestrogen diet. Confirmation that mice strain was estrogen-sensitive.	*Strengths* : Doses determined based on daily body weights. Good dose range.	Very useful	High utility
Nagao et al., 2002	*369*	*288*	Y	M	M	Adult	Oral gavage	Y	*Strengths* : Extremely well-conducted. Appropriate number of mice per dose group. Two strains used. Positive control used. Sperm data presented in light of historical data.	*Strengths* : Doses determined based on daily body weights. Good dose range.	Highly useful	High utility
Nagel et al., 1997	*205*	*175*	Y	M	M	Prenatal	Oral gavage	Y	*Strengths* : Use of the same methods as vom Saal et al. (341). Use of dose levels in the range of human exposure. Independent confirmation of the data analysis by the NTP Statistics Subpanel. *Weaknesses* : Lack of clarity on the mouse strain that was used. The Purina 5001 chow has high and variable levels of soy phytoestrogens, and the corn cob bedding may be problematic due to antiestrogenic constituents. The method of selection of males is not clear, and it appears that litter of origin was not considered. This study did not use a positive control, although there are earlier reports from this laboratory using diethylstilbestrol.	*Weaknesses* : Only one male per litter used in the study. See additional considerations for vom Saal et al. 1998 (this is a companion paper). Comment: Unlike the vom Saal (1998) citation which CERHR was divided as to it's utility, this study was determined by CERHR to be 'useful in the evaluation'. This study was essentially the same experiment so it is difficult to see how it is deemed useful. All of the weaknesses identified under vom Saal et al. 1998, appear to be relevant for this citation.	Useful	Adequate
Nakahashi et al., 2001	*380*	*212*	Y	M	M	Neonatal	Injections	Y	*Weaknesses* : Difficult to calculate BPA doses. Lack of husbandry and statistical information. Injection route not indicated. BPA purity not reported. No information on caging and bedding materials. Numbers of litters not indicated.	*Weaknesses* : Non-oral exposure.	Slightly useful	Limited utility
Narita et al., 2006	*357*	*190*	Y	B	M	Development	Oral	N	*Weaknesses* : Poorly written.	*Weaknesses* : No relevant endpoints	Inadequate	Inadequate
Negishi et al., 2003	*315*	*151*	Y	B	R	Prenatal/neonatal	Oral gavage	Y	*Weaknesses* : Doses were high enough to produce gross body weight changes. Analysis not litter-based. No positive control.	*Weaknesses:* Unclear whether data used for analysis was by animal or by litter.	Adequate	Limited utility
Negishi et al., 2003	*315*	*151*	Y	F	R	Adult	Oral gavage	Y	*Weaknesses* : Doses were high enough to produce gross body weight changes. Analysis not litter-based. No positive control.	None.	Adequate	Inadequate
Negishi et al., 2004	*316*	*152*	Y	M	R	Prenatal/neonatal	Oral gavage	Y	*Weaknesses* : Single dose level used.	*Weaknesses* : Organ weight data not shown. Only one dose level used, data not reported (authors just stated that there were no effects) for relevant reproductive/developmental endpoints.	Adequate	Inadequate
Negishi et al., 2004	*316*	*152*	Y	F	R	Adult	Oral gavage	Y	*Weaknesses* : Single dose level used.	*Weaknesses* : Data not shown for relevant reproductive endpoints. Only one dose level used, data not reported (authors just stated that there were no effects) for relevant reproductive/developmental endpoints.	Adequate	Inadequate
Nieminen et al., 2002	*423*	*270 292*	N	B	P	Adult	Oral	N	*Weaknesses:* High doses only. No reproductive endpoints. Small sample size.	None.	No utility	Inadequate
Nieminen et al., 2002	*424*	*270 292*	N	B	V	Adult	SC	N	*Weaknesses* : Small sample size. Non-oral exposure. Lack of similar studies in literature with this species.	*Weaknesses* : High doses only. No relevant reproductive endpoints.	Severely limited (F)/no utility (M)	Inadequate
Nikaido et al., 2004	*366*	*197*	Y	F	M	Prenatal	SC	Y	*Weaknesses:* Lack of clarity regarding sample size. Weak description of histopathogy findings.	*Weaknesses* : SC not a relevant route of exposure. The experimental unit appears to be the individual offspring, not the litter. The number of dams treated is not clear. The number of offspring per litter and per dose group is not clear. A statistical evaluation of BPA effects on the mammary gland was not performed. Effects on mammary gland and estrous cycle are inconsistent with the results reported in Nikaido et al. 2005.	Moderately useful	Limited utility
Nikaido et al., 2005	*378*	*208*	N	F	M	Puberty	SC	Y	*Strengths* : DES used as positive control. *Weaknesses:* Lack of information on sample size or results from later in life.	*Weaknesses:* SC not a relevant route of exposure. High doses only.	Not useful	Limited utility
Nishizawa et al., 2005a	*353*	*184*	Y	F	M	Adult	Oral	N	*Strengths* : Oral exposure. Low dose exposures. Exposures at different time periods.	*Weaknesses:* Gene expression study. No adverse developmental or reproductive effects measured.	Slightly useful	Inadequate
Nishizawa et al., 2005b	*354*	*185*	Y	F	M	Adult	Oral	N	*Strengths* : Wide dose range. Oral exposure. *Weaknesses* : Did not specify method.	*Weaknesses:* Gene expression study. No adverse developmental or reproductive effects measured.	Moderately useful	Inadequate

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Nishizawa et al., 2005c	*355*	*187*	Y	F	M	Adult	Oral	N	*Strengths* : Wide dose range. Oral exposure. *Weaknesses* : Did not specify method.	*Weaknesses:* Gene expression study. No adverse developmental or reproductive effects measured.	Useful	Inadequate
NTP, 1985, 1989	*456 457*	*307*	Y	B	M	Adult	SC pump	Y	*Strengths* : Well-conducted study. *Weaknesses* : Non-oral exposure.	*Weaknesses* : High doses only. Comment: CERHR notes that silastic implants are an 'irrelevant route of exposure, which makes extrapolation for human risk assessment difficult.' Shouldn't this also be the case for the s.c. studies published by Markey et al?	Extrapolation difficult	Limited utility
NTP, 1989	*457 458*	*308*	N	B	M	2-generation	Oral	Y	*Strengths* : Large number of animals. Multiple endpoints. *Weaknesses:* Direct exposure to BPA did not occur during cohabitation, therefore direct exposure to BPA was absent during sperm maturation, capacitation and ovulation.	*Weaknesses* : High doses only.	Highly useful	Adequate
Palanza et al., 2002	*352*	*183*	Y	F	M	Adult	Oral pipette	N	*Strengths* : Oral exposure. Low dose. Exploration of effects on complex on maternal behaviors. *Weaknesses* : Pre- and post-natal exposures had effects alone but not in combination, no explanation of this finding. Diet high in soy isoflavones.	*Weaknesses:* One dose. No dose-response trend. No adverse developmental or reproductive effects measured.	Very useful	Inadequate
Park et al., 2004	*421*	*268 290*	Y	B	R	Adult	IP injection	Y	*Weaknesses* : Frequency (5 times) and route of administration not relevant to humans.	*Weaknesses* : No statistical tests for dose-response trends were conducted.	Minimal value/Not useful	Inadequate
Park et al., 2005a	*367*	*198*	Y	B	M	Prenatal	IP injection	Y	*Strengths* : Three doses. *Weaknesses* : Lack of information on husbandry. Non-oral exposure. Poor presentation of histopathology.	*Weaknesses* : This study is written in Korean with the abstract, figures and tables in English. The number of dams per dose group is not specified.	Marginal utility	Limited utility
Park et al., 2005a	*367*	*198*	Y	F	M	Adult	IP injection	Y	*Strengths* : Three doses. *Weaknesses* : Lack of information on husbandry. Non-oral exposure. Poor presentation of histopathology.	*Weaknesses:* This study is written in Korean with the abstract, figures and tables in English. The experimental unit appears to be the individual offspring, not the litter. The number of dams per dose group is not specified.	Marginal utility	Limited utility
Park et al., 2005b	*368*	*198*	Y	F	M	Adult	IP injection	Y	*Weaknesses* : Inadequate description of methods. Non-oral exposure. Poor presentation of histology.	*Weaknesses* : This study is written in Korean with the abstract, figures and tables in English. The number of dams per dose group is not specified, although the number of offspring examined is.	Marginal utility	Limited utility
Park et al., 2005b	*368*	*198*	Y	B	M	Prenatal	IP injection	Y	*Weaknesses* : Inadequate description of methods. Non-oral exposure. Poor presentation of histology.	*Weaknesses* : This study is written in Korean with the abstract, figures and tables in English. The number of dams per dose group is not specified, although the number of offspring examined is.	Marginal utility	Limited utility
Peknicova et al., 2002	*443*	*288*	Y	M	M	3-generation	Oral (drink)	Y	*Weaknesses* : Very few experimental details provided: no information on bedding, caging, BPA purity, number of animals per dose group, age during treatment, duration of treatment, mating procedures, or whether females were treated.	*Weaknesses* : The incidence and severity of findings of testicular histopathology not reported. Some reported information is inconsistent. Age at sperm sampling are not described.	No utility	Inadequate
Ramos et al., 2001	*286*	*124*	Y	B	R	Prenatal	SC pump	Y	*Strengths* : "Interesting" design with respect to choice of endpoints. *Weaknesses:* Certain design aspects unclear. Small sample size. Considerable uncertainty about numbers and litter effects. Concern with the use of neat DMSO as solution vehicle.	*Weaknesses* : SC not a relevant route of exposure. Only relevant endpoint was AGD, for which data is not reported, just mentioned in passing by the authors.	Marginally adequate (with qualification)	Inadequate
Ramos et al., 2003	*287*	*125*	Y	B	R	Prenatal	SC pump	Y	*Weaknesses* : Very small sample size for most measures. Uncertainty about litter origin and representation in each necropsy group. Altered values given without reporting normal range of variation or likely functional significance of changes.	*Weaknesses:* SC not a relevant route of exposure.	Inadequate	Limited utility
Razzoli et al., 2005	*425*	*271*	Y	F	G	Adult	Oral	N	*Strengths* : Well-conducted study. *Weaknesses* : No reproductive endpoints.	*Weaknesses* : No relevant endpoints.	Limited utility	Inadequate
Rivas et al., 2002	*329*	*163*	Y	M	R	Neonatal	SC	Y	*Weaknesses* : Only one BPA dose, this was high and varied depending on rat weight.	*Weaknesses:* SC not a relevant route of exposure.	Suitable	Limited utility
Rubin et al., 2001	*217*	*133*	Y	B	R	Prenatal/neonatal	Drinking water	Y	*Strengths* : Study incorporates range of basic developmental and gross functional reproductive endpoints. Estrous cycle changes are plausible. *Weaknesses* : Sample sizes small, does not use litter as the unit. Exposures poorly defined.	*Weaknesses* : Dose levels estimated based on drinking water levels and the assumption that all water was consumed. Exposure may have been underestimated due to an assumed low water consumption.	Barely adequate	Inadequate

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Ryan and Vandenbergh, 2006	*375*	*204*	Y	F	M	Prenatal and neonatal	Oral gavage	Y	*Weaknesses:* No discussion of developmental or reproductive effects (only neurodevelopmental effects).	*Weaknesses* : The main objective of this study was to examine anxiety and spatial memory. All animals were kept in polycarbonate cages, which were monitored for signs of deterioration or damage. The experimental unit appears to have been the individual offspring, not the litter. Number of dams dosed was unclear. Small number of female offspring per group used to determine onset of puberty.	Useful	Limited utility
Saito et al., 2003	*289*	*127*	Y	M	R	Prenatal	SC	Y	*Strengths* : Examination of testosterone levels at 13 weeks of age. *Weaknesses:* Sample size is too small for drawing firm conclusions.	*Weaknesses* : SC not a relevant route of exposure. Only one dose level used.	Inadequate (alone, but possibly useful when considered with other studies)	Limited utility
Saito et al., 2003	*437*	*282*	Y	M	R	Puberty	SC	Y	*Weaknesses:* Non-oral exposure. Experimental design concerns. Data not shown for several endpoints.	*Weaknesses* : The main objective of this study was to determine effects on testosterone and estradiol levels. Doses were not adjusted for changes in body weight. Non-oral dosing.	Not useful	Limited utility
Sakaue et al., 2001	*432*	*277*	Y	M	R	Adult	Oral gavage	Y	*Weaknesses:* Variability in control daily sperm production between experiments. Small sample (5 animals per dose group). No histopathology presented. The decrease in daily sperm is unlikely to affect fertility.	*Weaknesses* : No dose response trend. Control variability. A more comprehensive study conducted by Ashby et al. (2003) did not replicate results, Ashby et al. (2003) found control values for daily sperm production from Sakaue et al. (1999), Ohsako et al. (2001) and Sakaue et al. (2001), in particular, were higher than literature average (n = 34 studies). Ashby et al. (2003) noted that higher than normal control values could have led to apparent decreases in BPA-treated animals that were seen by Sakaue et al. (2001). In addition, Ashby et al. (2003) showed that in the Sakaue et al. (2001) study, there were higher standard deviations for control and low-dose group daily sperm production values than for the higher dose groups. These authors commented that the high inter-experiment variability in the Ashby et al. (2003) study and the high variability in the no- and low-dose groups in the Sakaue et al. (2001) study suggest that daily sperm production might not be reliable enough to test effects of low doses of BPA.	Limited utility	Inadequate
Schonfelder et al., 2002	*279*	*118*	Y	B	R	Prenatal	Oral gavage	Y	*Weaknesses* : Unclear methodology. Uncertainty of group sizes. Uncertainty of number of offspring examined.	*Weaknesses* : Number of animals sacrificed in estrus was larger than the number of animals sacrificed in diestrus. Dose groups (negative control, EE group, and BPA groups) were not run concurrently. This is a serious design flaw as any potential effects may be confounded by time differences. The vehicle used for the positive control (peanut oil) was different than what was used for the negative control and BPA groups (2% mondamin). This study was not conducted under GLP standards.	Inadequate	Limited utility
Schonfelder et al., 2004	*280*	*119*	Y	B	R	Prenatal	Oral gavage	Y	*Weaknesses* : Small number of dams in study. Uncertain number of litters examined.	None. Dose groups (negative control, EE group, and BPA groups) were not run concurrently. This is a serious design flaw as any potential effects may be confounded by time differences. The vehicle used for the positive control (peanut oil) was different than what was used for the negative control and BPA groups (2% mondamin). This study was not conducted under GLP standards.	Inadequate	Limited utility
Sharpe et al., 2003	*330*	*164*	N	M	R	Neonatal	SC	Y	*Weaknesses* : Single high, variable dose of BPA used.	*Weaknesses:* SC not a relevant route of exposure.	Suitable	Limited utility
Spencer et al., 2002	*417*	*264*	N	F	R	Adult	SC	Y	*Weaknesses* : Non-oral exposure. One high dose level.	None.	Limited value	Limited utility
Stoker et al., 1999	*326*	*161*	N	M	R	Puberty	SC	Y	*Strengths* : Experiments appropriately performed. Data dependable. *Weaknesses* : Single, high-dose of BPA.	*Weaknesses:* SC not a relevant route of exposure.	Suitable	Limited utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Sugiura-Ogasawara et al., 2005	*67*	*258*	Y	F	H	NA	NA	Y	*Weaknesses* : Timing and number of blood samples not clearly defined. Inappropriate statistical analyses. Cases and controls not comparable because of differences in occupation and unknown fertility in controls. Controls not evaluated for hypothyroidism and systemic lupus erythematosus. ELISA can overestimate BPA due to cross-reactivity with other chemicals. No QA information. Confounders and effect modifiers not effectively managed.	*Weaknesses* : Inconsistencies among study results: no association between BPA and future miscarriages, but association with past miscarriages. Conjugated BPA likely included in the BPA measurement.	Not clearly stated	Inadequate
Suzuki et al., 2002	*377*	*207*	N	F	M	Prenatal	SC	Y	*Strengths:* DES used as positive control. *Weaknesses* : High dose. Non-oral exposure.	None.	Useful	Limited utility
Suzuki et al., 2002	*377*	*207*	N	F	M	Neonatal	SC	Y	*Strengths:* DES used as positive control. *Weaknesses* : High dose. Non-oral exposure.	None.	Useful	Limited utility
Suzuki et al., 2003	*372*	*202*	Y	B	M	Prenatal/neonatal	Oral	N	*Weaknesses* : Inadequate description of methods. Non-oral exposure. Poor presentation of histology.	*Weaknesses:* No relevant reproductive endpoints.	Not useful	Inadequate
Takagi et al., 2004	*305*	*139*	N	B	R	Prenatal/neonatal	Diet	Y	*Strengths:* Range of endpoints used. Estradiol group included. Complete statistical evaluation. Range of endpoints was better than average. *Weaknesses* : Small sample sizes used.	*Strengths* : Good reporting of statistics. Large dose range. Large number of endpoints studied. *Weaknesses* : High doses only. No dose-response.	Barely adequate	Limited utility
Takahashi and Oishi, 2003	*438*	*283*	N	M	R	Adult	Oral	Y	*Strengths* : Comprehensive and well-conducted. Compares strains and exposure route on BPA effects. *Weaknesses:* High doses.	*Weaknesses* : Relative organ weights were not shown, but analyses were described in the text.	Useful	Limited utility
Takahashi and Oishi, 2003	*438*	*283*	N	M	R	Adult	SC, IP	Y	*Strengths* : Comprehensive and well-conducted. Compares strains and exposure route on BPA effects. *Weaknesses:* High doses.	*Weaknesses:* SC and IP not relevant routes of exposure. Relative organ weights were not shown, but analyses were described in the text.	Useful	Limited utility
Takahashi and Oishi, 2003	*438*	*289*	N	M	M	Adult	Oral	Y	*Strengths* : Comprehensive and well-conducted. Compares strains and exposure route on BPA effects. *Weaknesses:* High doses.	*Weaknesses* : Relative organ weights were not shown, but analyses were described in the text.	Marginal utility	Limited utility
Takahashi and Oishi, 2005	*431*	*275*	N	M	R	Adult	Oral	Y	*Strengths* : Relatively well-conducted study. *Weaknesses* : Testis histopathology data of limited value because of fixation techniques.	*Strengths* : Some dose-response related effects seen. *Weaknesses:* High doses only.	Limited utility (for histopath); Adequate for other	Adequate
Takao et al., 1999	*441*	*286*	Y	M	M	Pubertal	Oral (drink)	Y	*Strengths:* Well-conducted study. *Weaknesses* : No reproductive endpoints. Paucity of experimental details (i.e., sample size per dose group, purity of BPA, histopathology evaluation).	None.	Limited utility	Adequate
Takao et al., 2003	*370*	*201*	Y	M	M	Pubertal	Oral (drink)	Y	*Weaknesses* : Lack of information on age of sacrifice. Stability of BPA not determined.	*Weaknesses* : The main objective of this study was to examine effects of BPA on estrogen receptors in testes. Small number of animals per group. Inadequate reporting of exposure conditions.	Marginally useful	Limited utility
Takashima et al., 2001	*299*	*134*	N	B	R	Prenatal/neonatal	Diet	Y	*Strengths* : Good size and duration of the study. *Weaknesses:* Study saw small effects despite high exposure level. Effects on body weight and thyroid-stimulating hormone levels are discounted. Sample size inadequate to address neoplasm endpoints. Information insufficient to judge appropriateness of statistical analyses and reliability of findings.	*Weaknesses* : High dose only. One dose level used. 10-11 dams/group. Main endpoint of the study was carcinogenesis, not reproductive/developmental effects.	Adequate (but limited)	Limited utility
Takashima et al., 2001	*299*	*134*	N	F	R	Adult	Diet	Y	*Strengths* : Good size and duration of the study. Study saw small effects despite high exposure level. *Weaknesses* : Effects on body weight and thyroid-stimulating hormone levels are discounted. Sample size inadequate to address neoplasm endpoints. Information insufficient to judge appropriateness of statistical analyses and reliability of findings.	*Weaknesses* : High dose only. One dose level used. Main endpoint of the study was carcinogenesis, not reproductive/developmental effects.	Adequate (but limited)	Limited utility
Takeuchi and Tsutsumi, 2002	*65*	*257*	Y	B	H	NA	NA	Y	*Weaknesses* : Small sample size. Crude study design. Inadequate analyses. No time standardization for serum draws. ELISA method has been shown to overestimate BPA levels. Little information provided on selection of comparison group, recruitment methods, and participation rates. Lack of diagnostic criteria for polycystic ovary syndrome. Confounders not accounted for. No information regarding distribution of data.	*Weaknesses* : One measurement at one time point per person. Conjugated BPA likely included in the BPA measurement.	Gives some insight	Inadequate
Takeuchi et al., 2004	*64*	*258*	Y	B	H	NA	NA	Y	*Weaknesses:* Small sample size. Blood draws time-standardized. Confounders/effect modifiers not accounted for. Inappropriate statistical analyses. ELISA method can overestimate BPA. Unclear if same study subjects in Takeuchi and Tsutsumi (2002) study.	*Weaknesses* : One measurement at one time point per person. Conjugated BPA likely included in the BPA measurement.	Not clearly stated	Inadequate

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Talsness et al., 2000	*277*	*115*	Y	B	R	Prenatal	Oral gavage	Y	*Strengths* : Good postnatal evaluation of various endpoints from pup to adulthood. Verification of dosing solution concentrations. *Weaknesses* : Did not use litter as unit of analysis. Only two dose levels examined. Vaginal opening data for controls outside normal range for SD rats. Unclear analysis of estrous cycle data. Incorrect analysis of F1 data. Magnitudes unreliable for data on developmental disruptions. Authors' comments on dose-response must be viewed with caution.	*Weaknesses* : Most data appears to have been analyzed on a per-animal basis instead of on a per-litter basis. No dose-response trends observed. Dose groups (negative control, EE group, and BPA groups) were not run concurrently. This is a serious design flaw as any potential effects may be confounded by time differences. The vehicle used for the positive control (peanut oil) was different than what was used for the negative control and BPA groups (2% mondamin). This study was not conducted under GLP standards.	Not adequate	Inadequate
Tan et al., 2003	*307*	*142*	N	M	R	Puberty	Oral gavage	Y	*Strengths:* Study well performed. Endpoints tested were documented. *Weaknesses:* Only one high dose of BPA used. Exposure period avoids early development. No histology performed on seminal vesicles or coagulating gland tissue.	None	Moderate utility	Limited utility
Thuillier et al., 2003	*282*	*121*	Y	M	R	Prenatal	Oral gavage	N	*Strengths* : Endpoints are a strength. *Weaknesses* : Inadequate methodological detail for any informed judgment of study quality.	*Weaknesses:* No relevant reproductive endpoints.	Inadequate	Inadequate
Timms et al., 2005	*351*	*182*	Y	M	M	Prenatal	Oral pipette	Y	*Strengths:* Oral exposure. Low dose. DES and estradiol used as positive controls. Sophisticated measures applied to prostate.	*Strengths:* The authors described effects as "virtually identical" to the effects observed after low doses of DES. *Weaknesses* : One dose. No dose-response trends observed. One pup per litter used for analysis, this pup was a male fetus that developed in utero between another male and a female fetus. Is the removal of pups to determine intra-uterine position add an element of stress to the experiment. This is not relevant to the human situation. This study was not conducted under GLP standards. Absolute prostate weights were not given.	Very useful	Limited utility
Tinwell et al., 2002	*278*	*117*	Y	B	R	Prenatal	Oral gavage	Y	*Strengths:* Good dose range. Appropriate measures selected. Four dose levels. Two rat strains. Dosing solution verified. Positive control used. Data appropriately analyzed with litter as experimental unit. *Weaknesses* : Small sample size.	None. Strengths: This study was conducted under GLP regulations.	Adequate	Adequate
Tohei et al., 2001	*434*	*279*	Y	M	R	Adult	SC	N	*Weaknesses* : Non-oral exposure. Small sample size. No functional consequences of hormone alterations shown	*Weaknesses* : No relevant reproductive endpoints.	Minimal utility	Inadequate
Toyama and Yuasa, 2004	*334*	*167*	Y	M	R	Neonatal	SC	Y	*Strengths:* Well performed and documented. Multiple doses. Use of both rats and mice. *Weaknesses* : Selective data presentation. Failure to examine sperm morphology in the fertile 15 week old animals to determine whether the changes in sperm maturation seen at earlier time points had resolved or whether the animals were fertile in the face of such abnormalities.	*Weaknesses* : SC not a relevant route of exposure. Study was observational. No statistical analyses were performed. Effects were reversible. Authors noted there was no dose-response relationship for abnormal sperm.	Suitable	Limited utility
Toyama and Yuasa, 2004	*334*	*213*	Y	M	M	Neonatal	SC	Y	*Strengths* : Well performed and documented. Multiple doses. Use of both rats and mice. *Weaknesses* : Selective data presentation. Failure to examine sperm morphology in the fertile 15 week old animals to determine whether the changes in sperm maturation seen at earlier time points had resolved or whether the animals were fertile in the face of such abnormalities.	*Weaknesses* : SC not a relevant route of exposure. Effects on sperm probably don't represent adverse effects in view of the lack of impact on fertility. Study was observational, no statistical analyses were performed. No clear dose-response relationship observed.	Suitable	Limited utility
Toyama et al., 2004	*440*	*285*	Y	M	R	Adult	SC	N	*Weaknesses* : Non-oral exposure. No statistical analyses. Limited data set. Fertility assessment not meaningful because 2 animals per dose group. Background incidence not discussed.	*Weaknesses* : Effects noted were reversible. Observational study. Organ weight data not shown.	Not useful	Inadequate
Toyama et al., 2004	*440*	*290*	Y	M	M	Adult	SC	N	*Weaknesses* : Non-oral exposure. No statistical analyses. Limited data set. Fertility assessment not meaningful because 2 animals per dose group. Background incidence not discussed.	*Weaknesses* : Effects noted were reversible. Observational study. Organ weight data not shown.	Not useful	Inadequate
Tyl et al., 2002	*293 411*	*130 302*	Y	B	R	3-generation	Oral gavage	Y	*Strengths* : Well-designed and comprehensive. Large number of rats per dose group (30). Multiple endpoints examined. Oral exposure. Concentration of BPA in the test diet was verified. Maternal and paternal toxicity was identified. Wide dose range.	*Strengths:* This study was conducted under GLP standards and a large historical control database exists for rats in this laboratory.	Adequate/highly valuable	High utility
Tyl et al., 2006	*376*	*205 312*	Y	B	M	2-generation	Oral	Y	*Strengths* : Large number of animals. Multiple endpoints. Large dose range. Rigorous study. Concurrent positive control. Thorough histological evaluation. *Weaknesses* : Statistics not performed on some histopathology data.	*Strengths:* This study was conducted under GLP standards.	Very useful	High utility
Vandenberg et al., 2007	*362*	*195*	Y	F	M	Prenatal	Implanted osmotic pump	Y	Not stated	*Weaknesses* : SC not a relevant route of exposure. One dose level.	Not stated	Limited utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
vom Saal et al., 1998	*341*	*174*	Y	M	M	Prenatal	Oral	Y	*Strengths* : Oral exposure. Low doses. 36% increase in preputial weight at 0.002 mg/kg bw/day seems robust. *Weaknesses:* Lack of clarity concerning the strain of mouse. Failure to weight-adjust the maternal dose daily. Lack of litter of origin in randomly selected males. Lack of information on testis weight (which is needed for consideration of daily sperm production). Questions about the statistical analysis. NTP statistics subpanel did not confirm statistical significance.	*Strengths:* Vom Saal did report testis weights. The association between BPA and prostate weight remained significant in the NTP panel's reanalysis, although the p value increased from <0.01 to <0.05. Prostate and testis weight changes were significant after control for body weight. *Weaknesses:* Mouse strain used by vom Saal et al. in this study no longer exists, and may have been more sensitive to estrogenic effects than mice used by Ashby et al. Use of one pup from each litter could increase the likelihood of a false negative, but would probably not increase the false positive probability (based on Haseman et al. 2001). Use of only 5 total mice for sperm parameters and 7 for prostate weight. Only 1 pup/litter was selected and it is not clear that technicians did so randomly. There is no published literature to suggest that switching housing conditions (from group to individual at 5 months of age) eliminates dominant-submissive effects established from group housing. This study was not conducted under GLP conditions.	Panel divided	Limited utility
Wang et al., 2004	*283*	*121*	Y	M	R	Prenatal	Oral gavage	Y	*Strengths* : Generally well-conceived study. *Weaknesses* : Small sample size.	*Strengths* : Large dose range. Doses adjusted based on maternal body weight. *Weaknesses* : Only one relevant reproductive endpoint (number of spermatagonia/tubule or testis). Effects described are likely adaptive rather than adverse.	Inadequate (alone, but possibly useful when considered with other studies)	Limited utility
Williams et al., 2001	*328*	*162*	N	M	R	Neonatal	SC	Y	*Strengths:* Group known to have expertise. Experiments well performed. *Weaknesses* : Only one dose of BPA used, this was variable because of changing body weights of animals.	*Weaknesses:* High dose only. SC not a relevant route of exposure.	Suitable	Limited utility
Wistuba et al., 2003	*281*	*120*	Y	M	R	Prenatal	Oral gavage	Y	*Strengths:* Good focus on male reproductive tract/function. *Weaknesses* : Too few animals to provide reliable data.	*Weaknesses* : Authors noted "remarkable" intralitter variation. Dose groups (negative control, EE group, and BPA groups) were not run concurrently. This is a serious design flaw as any potential effects may be confounded by time differences. The vehicle used for the positive control (peanut oil) was different than what was used for the negative control and BPA groups (2% mondamin). This study was not conducted under GLP standards.	Inadequate (alone, but possibly useful when considered with other studies)	Inadequate
Yamasaki et al., 2002	*128*	*264 275*	N	B	R	Adult	Oral gavage	Y	*Strengths* : Well-conducted GLP study. *Weaknesses* : High dose. Increase in testis weight after BPA administration is likely an artifact of decreased terminal body weight.	*Strengths* : Multiple doses used. Good description of statistical analyses. Male and female rats used.	Adequate	Adequate
Yoshida et al., 2004	*104*	*138*	Y	F	R	Prenatal/postnatal	Oral gavage	Y	*Strengths* : BPA determinations made. Animal exposure levels anchored to human exposures. Good range of endpoints measured. Study is a good screening study. *Weaknesses* : Number of animals sacrificed at endpoints was small. Numbers too small for cancer evaluation and for definitive conclusions to be reached for adult reproductive endpoints. Insufficient description of statistics to determine how data from multiple sampling points were evaluated.	*Weaknesses* : BPA found in the serum, milk, and liver of control animals as well as in the drinking water and pellet diet fed to all animals.	Inadequate (alone, but possibly useful when considered with other studies)	Limited utility
Yoshida et al., 2004	*104*	*138*	Y	F	R	Adult	Oral gavage	Y	*Strengths* : BPA determinations made. Animal exposure levels anchored to human exposures. Good range of endpoints measured. Study is a good screening study. *Weaknesses* : Number of animals sacrificed at endpoints was small. Numbers too small for cancer evaluation and for definitive conclusions to be reached for adult reproductive endpoints. Insufficient description of statistics to determine how data from multiple sampling points were evaluated.	*Weaknesses:* BPA found in the serum, milk, and liver of control animals as well as in the drinking water and pellet diet fed to all animals.	Inadequate (alone, but possibly useful when considered with other studies)	Limited utility
Yoshino et al., 2002	*303*	*136*	N	M	R	Prenatal/neonatal	Oral	Y	*Weaknesses:* Small number of male offspring used at various timepoints to determine organ endpoints. Report resembles screening study rather than definitive study.	*Strengths:* Study done twice to check unexpected finding in first study. *Weaknesses:* High doses only. Only two dose levels used. Data reported on a per-pup basis, rather than by litters, small number of pups used at each timepoint.	Inadequate	Limited utility
Yoshino et al., 2002	*303*	*137*	N	F	R	Adult	Oral	Y	*Weaknesses* : Report resembles screening study rather than definitive study.	*Weaknesses* : High doses only. Only two dose levels used.	Inadequate	Limited utility

Author	CERHR Ref.	CERHR pg	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro/Devel Endpt	CERHR Strengths/Weaknesses	Additional Considerations	CERHR Ranking	Proposed Revised Ranking
Yoshino et al., 2004	*356*	*189*	Y	B	M	Prenatal	Oral gavage	Y	*Strengths:* Oral exposure. Wide dose range.	*Weaknesses* : The main objective of this study was to examine effects of BPA on the immune system. There were inconsistencies in the reporting of the duration of dosing (17 days in Materials and Methods, 18 days in the Results). Offspring body weights shown only for males, not females. Body weights were expressed as averages of five randomly selected male littermates. It is unclear whether offspring from more than one litter per dose were weighed.	Moderately useful	Inadequate
Zoeller et al., 2005	*308*	*143*	Y	B	R	Prenatal/neonatal	Oral (diet)	Y	*Strengths:* Wide dose range. *Weaknesses:* No litter-based analysis. No positive control.	*Weaknesses:* Study unclear as to whether data was examined by animal or by litter; many of the reported results are for the dams. Few reproductive endpoints. Primary endpoints are effects on thyroid hormones and brain development.	Limited utility	Limited utility
Zoeller et al., 2005	*308*	*143*	Y	F	R	Adult	Oral (diet)	Y	*Strengths* : Wide dose range. *Weaknesses* : No litter-based analysis. No positive control.	*Weaknesses* : Study unclear as to whether data was examined by animal or by litter; many of the reported results are for the dams. Few reproductive endpoints. Primary endpoints are effects on thyroid hormones and brain development.	Limited utility	Limited utility

Previously reviewed in Goodman et al.
Strongly disagree with CERHR assessment

Abbreviations: BPA, Bisphenol A; Y, yes; N, no; F, female; M, male; B, both male and female; NA, not applicable; BADGE, BPA diglycidyl ether
Species Abbreviations: B, rabbit; G, gerbil; H, human; M, mouse; P, polecat; R, rat; S, sheep; V, vole

Table 6

Polycarbonate/BPA Global Group Comments on CERHR Draft Report of December 2006

Studies Missing From The Draft CERHR Report That Should Be Included

2-Feb-07

Author	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro Endpt	Strengths/Weaknesses	Proj Ran
Ashby and Odum, 2004	Y	F	R	Juvenile	Oral gavage	Y	*Weaknesses* : Large dose range. No statistical test for dose-response trends. Animals dosed for very short duration. Only 1 relevant endpoint considered.	Adequa
Berger et al., 2006	N	F	M	Adult	SC	Y	*Weaknesses* : SC not a relevant route of exposure. High doses only. Dose groups varied in size. Inappropriate statistics.	Limited
Berger et al., 2006	N	F	M	Adult	Dietary	Y	*Weaknesses:* High doses only. Number of animals per group low. Inappropriate statistics.	Limited
Colerangle and Roy, 1997	Y	F	R	Puberty	SC pump	Y	*Strengths:* Appropriate statistics. *Weaknesses* : One developmental endpoint examined (mammary gland development). SC not a relevant route of exposure. One low BPA dose level.	Limited
Diel et al., 2004	Y	F	R	Adult	Oral gavage	Y	*Strengths* : Multiple strains of rats used. Appropriate statistics. *Weaknesses* : Only one low dose (5 mg/kg-d), Data are shown in figures only, no precise values given.	Adequa
Gui et al., 2005	Y	M	R	Adult	IP	Y	*Weaknesses* : Article in Chinese, translation unavailable. Two doses, one below 5 mg/kg-d. No information regarding at what dose effects were observed. IP not a relevant route of exposure. No information regarding statistical methods.	Inadequ

Author	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	F'n'l Repro Endpt	Strengths/Weaknesses	Proj Ran
Hiroi et al., 2004	Y	F	H	NA	NA	Y	*Weaknesses* : Small subject numbers. One non-fasting serum BPA measurement per person. BPA measured by ELISA. Confounders and effect-modifiers not accounted for.	Inadequ
Koda et al., 2005	Y	F	R	Adult	SC	Y	*Strengths:* Appropriate statistics. *Weaknesses:* BPA was the positive control in this study. Only one relevant endpoint considered. Large dose range. SC not a relevant route of exposure. Rats were ovariectomized.	Inadequ
Long et al., 2000	Y	F	R	Adult	IP	N	*Strengths:* Appropriate statistics. *Weaknesses:* IP not a relevant route of exposure. No relevant reproductive endpoints. Ovariectomized rats.	Inadequ
Mehmood et al., 2000	Y	F	M	Puberty	SC	Y	*Weaknesses* : SC not a relevant route of exposure. Only 4 animals used per group. Large dose range. Limited number of endpoints examined. 3 days of dosing. Statistical methods not fully described.	Limited
Nishino et al., 2006	Y	M	R	Puberty	Oral gavage	Y	*Weaknesses:* BPA dosed via oral gavage, while the positive control group was dosed subcutaneously. Rats were orchiectomized. Inappropriate statistics.	Inadequ
Noda et al., 2005	Y	B	R	Neonatal	SC	Y	*Strengths:* Appropriate statistical methods. *Weaknesses:* Number of animals dosed unclear. SC not a relevant route of exposure. Dosing for 5 days. Low number of offspring and litters. Authors noted that shortened AGD in female rats did not originate from estrogenic activities of BPA.	Limited
Rubin et al., 2006	Y	F	M	Adult	SC	Y	*Strengths:* Appropriate statistics. *Weaknesses:* Number of dams dosed not specified. SC not a relevant route of exposure. Primary endpoints neurodevelopmental. Authors state, "Only litters with normal distributions of males and females were included in these studies." It is not known if (and how many) litters were excluded from the study.	Inadequ

Author	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	F'n'l Repro Endpt	Strengths/Weaknesses	Proj Rar
Savabieasfahani et al., 2006	Y	F	S	Prenatal	SC	Y	*Strengths* : Sufficient number of animals. Appropriate statistical methods. *Weaknesses:* One BPA dose level. SC not a relevant route of exposure.	Limited
Seidlova-Wuttke et al., 2004	Y	F	R	Adult	Dietary	Y	*Strengths:* Appropriate statistics. *Weaknesses:* Only four (of 12) animals per group were used for the histopathology examinations. Only two dose levels used and the reported doses of 33 and 333 micrograms/kg are not clear from the description of the study. Rats were ovariectomized. Primary endpoints were effects on ER.	Inadequ
Steinmetz et al., 1998	Y	F	R	Adult	IP injection or Implanted capsule	Y	*Strengths* : Appropriate statistics. *Weaknesses* : Number of animals dosed not reported. One dose level of BPA for some experiments. IP and SC not relevants routes of exposure.	Limited
Tinwell et al., 2000	Y	F	M	Neonatal	Oral	Y	*Strengths:* Large dose range. Appropriate statistics.	Adequa
Tinwell et al., 2000	Y	F	M	Neonatal	SC	Y	*Strengths:* Appropriate statistics. *Weaknesses:* SC not a relevant route of exposure. No dose-response trend observed.	Inadequ
Toyama, 2005	Y	F	M	Adult	SC	Y	*Weaknesses:* Article not in English, translation available from APC. SC not a relevant route of exposure. Large dose range. Statistics not appropriate.	Limited
Yamada et al., 2002	Y	F	H	NA	NA	Y	*Strengths:* Appropriate statistics. *Weaknesses:* Participation rates not stated. Samples taken over a ten-year period. BPA measured by ELISA. Conjugated BPA likely included in the BPA measurement. Confounders and effect-modifiers not considered.	Limited
Yamasaki et al., 2002a	Y	F	R	Adult	SC	Y	*Strengths:* Appropriate statistics. *Weaknesses:* SC not a relevant route of exposure. Primary objective to investigate effects of phytoestrogen content in food on uterotrophic response.	Limited
Yamasaki et al., 2002b	N	B	R	Adult	Oral	Y	*Strengths:* Appropriate statistics. *Weaknesses:* High doses only.	Adequa

Author	Low dose	Sex	Species	Lifestage during dosing	Exposure Route	Fn'l Repro Endpt	Strengths/Weaknesses	Prop Ran
Yang et al., 2006	Y	B	H	NA	NA	Y	*Strengths* : Appropriate selection of study subjects. Appropriate statistics. BPA measured by HPLC. Confounders and effect-modifiers considered. *Weaknesses:* No association found in primary analysis, only marginal association in sub-analysis. Participation rates not stated. Unclear how subset of subjects were selected for sister chromatid exchange analyses. Blood draws not time-standardized.	Adequa

Abbreviations: BPA, Bisphenol A; Y, yes; N, no; F, female; M, male; B, both male and female; NA, not applicable; BADGE, BPA digly
Species Abbreviations: H, human; M, mouse; R, rat; S, sheep

Table 7
Polycarbonate/BPA Global Group Comments on
CERHR Draft Report of December 2006
February 2, 2007

Below each study table are additional considerations that were used to rank the utility of each study. This review and the utility ratings focus primarily on neurodevelopmental results reported in the study. Some of the studies were not found in the CERHR neurodevelopmental sections, but were evaluated elsewhere in the CERHR review. The rating criteria used for this table, defined at the end of this table, are adapted from the criteria used to rank studies for reproductive and developmental endpoints other than neurodevelopmental endpoints.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Adriani (2003)	320 (p.156)	Sprague -Dawley rats	**Prenatal Postnatal** GD0-PND25	Oral (in oil); micropipette	0.04	1 rat/sex/litter; 9 litters Novelty preference (n=1 rat/sex/litter, age=PND 30-45), impulsivity (n=1 rat/sex/litter, age ≥PND 70), open field with pharmacologic challenge (n=1 rat/sex/litter however 4 litters used for saline and 5 for amphetamine, age ≥PND 70)	Y	N	BPA increased novelty induced stress and reduced impulsivity in both sexes. BPA decreased restlessness profile in males to look more like females. BPA reduced effects of amphetamine to increase activity in males. BPA produced neophobia in adolescent females but not in males.	Well-performed using well-established protocols. Use of single exposure level. Conclusion that bisphenol A (BPA) causes demasculinization not supported by lack of male-female difference in behavior in controls. Questioning of labeling figures.	Suitable	Limited utility

Additional considerations:

Inconsistency of text and tables was addressed by the authors in an erratum (Adriani et al., Environmental Health Perspectives, 113:A368, 2005.

The authors' conclusion that BPA caused demasculinization of behavior on impulsivity is not supported by experimental evidence. For example, there is no evidence that the change in nose poking during delay is due to an organizational effect on the brain. Therefore, CERHR should not state that the data "reinforce the idea of demasculinizing effects." The behavior that was "demasculinized" has no scheduled consequence (i.e., neither punished nor rewarded), and could be more prone to chance group differences. Therefore, more caution is needed in interpreting results of a single experiment with one dose level using an endpoint not demonstrated to be hormonally mediated. The primary behavior that was under schedule control was increased preference for large reinforcement (decreased "impulsivity") in both males and females following BPA exposure. In this and other behaviors, the effects of BPA were not consistent with the hypothesis that BPA causes demasculinization of behavior either because there were no gender differences in the control, or changes were not in a direction consistent with "demasculinization". CERHR should not state that this paper is consistent with many others showing similar effects. There are no other studies evaluating the operant behavior and novelty preference paradigm used in this study. It is possible that CERHR made this statement based on the conclusion of Adriani et al. (2003) that their results on the operant behavior suggest demasculinization consistent with Farabollini et al. (2002). Closer comparison indicates that Farabollini et al. (2002): a) measured sexual behaviors that are not comparable with Adriani (2003), b) saw that males had slight impairment of sexual performance but on the whole, authors conclude the important measures of male sexual behavior were not disrupted by BPA, and c) point out that "demasculinization" is not congruent with the hypothesis that estrogen or aromatizable androgens can masculinize the brain.

Aloisi (2002)	314 (p 150)	Sprague -Dawley rats	**Prenatal Postnatal** GD0-PND21	Oral (in oil); micropipette	0.04	9-16 rats/gender/treatment group, 7 litters Pain test (formalin injection), open field, hormone levels (n=11 male and 9 female pups, age=22 weeks for all endpoints) Animals were cross-fostered and randomly assigned to sham or formalin-treated groups	N	N	Perinatal exposure to BPA alters nociceptive responses in male and female rats. BPA was associated with opposite modifications in males and females with locomotor and exploratory activities increasing in females and decreasing in males, although the change was not statistically significant.	Added dimension being pain response. Inconsistency in data. Lack of some methodologic detail. Study separates antenatal and postnatal exposure but not a group of both exposures. Use of single dose level.	Barely adequate or useful	Limited utility

Additional considerations:

A utility rating based on the authors' conclusion that BPA causes differential effects on males and females is not supported by the experimental data. Also, this hypothetical increase in difference between male and female behavior is inconsistent with other papers that claim BPA decreases difference between males and females for locomotor behaviors.

Strength/weakness section should include that there was no clear operational definition of the subjective behaviors measured and observations were not conducted blind. In addition, the author's conclusion that BPA differentially affects male and female behavior is not supported by the data. The number of litters was low and the litter was not the unit for statistical evaluation. The relevance of the reported effects to humans requires additional evaluation. Since this study is an exploratory study, inclusion of a positive control group would be helpful to evaluate the test system in more detail. "Added dimension being pain response" is not a strength of the study, but simply a description of the study.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Carr (2003)	322 (p 158)	Fischer 344 rats	**Postnatal** PND 1–14	Oral; gavage to pups	0.1, 0.25	At least 1/sex/litter for n=10, unknown # litters Body weight (age=weighed daily), swim ability (age=PND 33, 1 day before acquisition), acquisition (age=PND 34 for 4 days), memory (age=PND 38) Within litter design such that male and female pups within same litter were assigned to different treatment groups, but not limited to one pup per sex per litter to achieve 10 per treatment group.	N	Y	There were no treatment-related effects on swimming ability or motivation or on acquisition of maze solution. Treatment with E2 and low BPA disrupted normal gender-dependent pattern of acquisition, while treatment with high BPA did not. In a probe trial, females treated with high BPA spent significantly less time in the escape quadrant.	Additional behavioral dimensions captured. Use of positive control. High doses (not damaging). Limited number of endpoints investigated.	Adequate and useful	Limited utility

Additional considerations:

Two dose levels within same order of magnitude should not be considered "low" and "high" doses, and therefore raise the question of why a response was seen at one dose but not the other. There is the possibility of cross-contamination of BPA doses and E2 in this within-litter design. The E2 dose group can not be described as a positive control since no effect was observed in E2 treated animals except for an increase in acquisition time at day 3, but no effect at days 1, 2, or 4. The statistical method was not clear and non-significant observations should not be taken into consideration without replication.

The authors state that the "high" BPA dose led to decreased time spent in the escape quadrant during the probe trial without any effects on acquisition based on comparing males and females. However, there was a tendency for increased acquisition in females at the "high" dose, which could have led to the decreased time spent in the escape quadrant during the probe trial. The direction of this effect is inconsistent with the authors' conclusions that BPA decreased sex difference because: a) control males tend to perform better than females, and b) "high" BPA causes females to perform worse (so less like males). The positive control (E2) had no effects on the probe trial.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Della Seta (2005)	420 (p 267)	Sprague -Dawley rats	**Prenatal postnatal** GD0-PND21	Oral; micropipette	0.04	Maternal behavior elements: 7-9 Dams/behavioral group; 17-23 Dams exposed per treatment group; 9/17 BPA and 12/23 controls gave birth and litters were cross-fostered with 4 males and 4 females to each dam.	N	N	Treatment of mothers with BPA reduced licking-grooming, and marginally reduced frequency of ano-genital licking and duration of arched-back posture, irrespective of gender of pups and period of observations. The weak behavioral effects observed here do not clarify whether exposure to BPA differentially affects maternal behavior toward male and female pups.	Suggests a low, oral dose of BPA (0.04 mg/kg bw/day) affects pregnancy and maternal behavior. Unusually low pregnancy rate observed in controls (18/32), raising concern about study design. Only one dose level, so dose-response relationship could not be established. Not stated that analysts were blinded to treatment. Because BPA/peanut oil was “fed” to mice, residual BPA may have been retained in oral cavity of dam resulting in altered grooming via altered taste perception.	Limited	Limited utility

Additional considerations:

The CERHR evaluation identifies important limitations of the paper. In addition, it should be noted that the changes in frequency, and duration of behavioral elements were considered by the author to be weak effects and that there is no evidence that duration or frequency of normal maternal behaviors have any adverse impact on nurturing. It is not clear if the time of recording behaviors was balanced across the different treatment groups since testing would have occurred over several hours and two days. These are important controls that would be expected to be reported in a well-conducted behavioral study of this type.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Della Seta (2006)	323 (p 159)	Sprague Dawley rats	**Postnatal (juvenile)** PND 23–30	Oral (in oil); micropipette	0.04	78 males from 16 dams, numbers of treatment/ control not given Weight gain (n=?, age=PND 21-31, taken every 2 days for a total of 6 times, assume all animals), juvenile behavior (play, exploring, object -directed; n=12, age=PND 45), adult socio-sexual behavior (n=9-10, age=>90), estradiol and testosterone levels (n=5-8, age=PND 37, PND 105) Weanlings were housed in groups of four with one animal per litter but distribution to different treatment groups not stated	N.A.	Y	EE permanently altered sexual behavior of adult males. BPA's effects were much weaker and fewer in number but were in the same direction as ethinylestradiol indicating an estrogenic mechanism. On the whole, there is an indication that exposure to estrogenic substances at early puberty has a short-term effect on immature behavior of male rats, with a decrease in the exploratory drive directed to environment and conspecifics; the effects found for BPA, weaker but in the same direction of EE, indicate to be due to its estrogenic activity.	To be added	Adequate and useful	Limited utility

Additional considerations:

The authors state that the behavioral measures were adapted from a previous paper so it is not possible to evaluate the rigor with which objective operational definitions were established for different behaviors. Four animals from each treatment group were observed at one time for 9 minutes which is greater than the 2-minute period evaluated in Dessi-Fulgheri et al., 2002. An important limitation of this study is the use of only one dose, and that the actual frequency for all the dependent variables measured in both males and females for the juvenile behavior portion of the study were not presented. This would allow one to have a better perspective of the overall pattern of behaviors for both genders. Instead, only the largest correlation coefficients from the principal component analysis (PCA) used to group different behavioral measurements using statistical criteria were presented. These groupings are called “factors”. The results of ANOVA applied to each factor, using individual PCA factor scores as variables and including both animals exposed to ethinylestradiol and bisphenol A were presented. Thus, effects on EE alone can contribute to statistical significance of the ANOVA. Only the factors with statistical significance were reported. Based on inspection of limited data provided in Figure 1 for juvenile behaviors and the post-hoc analyses, BPA had lower incidence of sniffing an object and possibly biting the object (although frequency of biting is very low in controls indicating very little contribution to overall behaviors). Given the lack of findings in pouncing, chasing, crawl over, jumping withdrawing, exploring, rearing, the biological significance of decrease in sniffing an object is uncertain given all the other behaviors the animals could be engaged in with 3 other animals in the same arena. Evaluation of the pattern of behavior across all these related behavioral measures including those that were not statistically significant by ANOVA is important and cannot be done properly without reporting all the data, not just statistically significant data. It should also be noted that duration of the animals’ behavior was not included, although this type of data was presented for sexual behaviors in this paper and for social behavior in other papers from this group. This measure is useful to compare with frequency because an animal can have less number of initiated behaviors (so lower frequency) but have same total time engaged in this behavior. The adult socio-sexual behavior is limited because the animals were not very active, so that a lower criterion of at least 2 mounts was used to consider animals sexually active and included in the evaluation. In terms of the most significant sexual behaviors, namely intromission and intromission/mount, BPA was similar to controls and EE was greater than both oil and BPA. Given the overall weak effect of BPA, these findings are of uncertain biological significance. It is premature to suggest that these results are indicative of a hormonal effect on behavior.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Dessi-Fulgheri (2002)	319 (p.155)	Sprague Dawley rats	**Premating, prenatal neonatal** Higher dose: GD14-PND6 Lower dose: 10 day premating – PND 21	Oral (in oil); micropipette	0.04, 0.4;	12 or 15 F and M from 9-11 litters Play toward males or females, nonsocial exploration, defensive, low-intensity mating, sociosexual, ground exploration, social interest (n=15 pups/sex/dose control and high dose, 12 pups/sex/dose low dose, age=PND 35, 45, 55 for all endpoints) For each treatment group, the animals were randomly chosen from different litters	N	N	Sex differences were seen in four behavioral categories defined by principal component factors (play with females, low-intensity sexual behavior, sociosexual exploration, ground exploration. BPA masculinized female behaviors in play with females and sociosexual exploration. BPA intensified male behavior in play with females. (Play with females [factor 2] is composed of pounce, chase, nape and withdraw from females; sociosexual exploration [factor 6] is defined as sniffing female genitalia and body and self grooming.)	Evaluated socio-sexual consequences of exposure at young age. Two controls not performed (prolonged high-dose and short low-dose exposure) limiting degree of data interpretation.	Marginally useful	Limited utility

Additional considerations:

Additional important limitations of study not mentioned by CERHR are: 1) only behaviors displayed during minute 2 and minute 3 were evaluated, which may not be sufficiently representative of the behaviors measured; 2) only results of composite principal component factors were analyzed and presented without tabular or graphical presentation of the specific measurements evaluated that make up these factors; 3) no positive control to help validate/support the grouping of behaviors into different factors based on gender differences and to support conclusions that BPA masculinized two female behaviors and intensified male behavior; 4) the paper claims effects on play behavior are due to estrogenic mode of action, but play behavior has been linked to androgenic effects; (5) the terms used to define statistically derived factors are named "social interest," "sociosexual exploration," and "defensive behavior" which are completely subjective and almost arbitrary and present a danger of misleading the reader that these behaviors are directly relevant to humans; and 6) the conclusions of this paper depend on multiple layers of statistical assumptions that ANOVA of pooled factors from principle component analysis is appropriate. Also, the use of different dosing schedules limits comparisons between study doses. This paper more specifically represents two different studies of a single dose each than one study using two doses that can be compared.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Ema (2001)	292 (p.130)	Sprague Dawley rats	**Prenatal, postnatal** Prior to mating through PND 22 for two generations	Oral; gavage (dam and selected pups from each generation)	0.0002, 0.002, 0.02, 0.2;	25 F and M from 25 litters each generation All F1 and F2 pups from 25 litters/ dose level were observed for pinna detachment, incisor eruption, eye opening, testes descent. Developmental landmarks (n=1 pup/sex/litter, age=daily beginning at PND6 [surface righting], PND 7 [negative geotaxis] and PND 13 [mid-air righting]), open field [ambulation, rearing, grooming, urination, defecation; n=1 pup/sex/litter, age=5-6 week], Biel multiple T-maze (n=6 F1rats/sex/group from different litters, age=6-7 weeks), brain weight (n=1 pup/sex/litter, age=not clear but appears to be around weaning [F1 and F2 weaning], males after mating, females after weaning pups [F1 adult]). (Other endpoints not relevant to neurological assessment are not included in this list.)	Y	N	Focused on behavioral endpoints: No compound-related changes in surface righting reflex, negative geotaxis, mid-air righting reflex in F1 and F2 generations, or behavior in the open field or water filled multiple T-maze in the F1 generation. No effects on functional reproductive measures in any generation.	Thorough evaluation, size of dose range, large number of animals, litter-based analysis, and verification of the dosing solution are good. Failure to replicate previous findings (at 0.002 and 0.02 mg/kg bw/day) and lack of positive control group.	Adequate	Adequate

Additional considerations:

The large range of dose levels and the large number of litters make this a very useful study. Open field observations and ontogeny of behavior were evaluated using large sample size and the litter was the statistical unit of analyses. However, the criteria used to measure rearing, grooming and ambulation during 3-minute open-field evaluation were not described. The methods for and results of the multiple T-maze were not adequately described and a much smaller sample size was used (6) than for other endpoints. It was not clearly described how these animals were selected. CERHR cites lack of a positive control group as a weakness because it "leaves persistent questions about the ability of this group of rats to respond." CERHR should consider adding the following for a more balanced discussion of this study: *"However, many BPA studies evaluating effects on neurobehavior and sexual behaviors were conducted with Sprague-Dawley rats (Adriani, 2003, Aloisi, 2002, Dessi-Fulgheri, 2002, Della Seta, 2006, Farabollini, 1999, 2002). Therefore, the large sample size and use of the litter as the experimental unit makes this study useful to compare reproductive outcome, open field ambulation, rearing, grooming, multiple t-maze learning, reproductive indices with other studies claiming effects on similar endpoints in the same strain and species."* CERHR should be consistent in the way it cites lack of positive control as a deficiency for all studies, including studies reporting effects of BPA. The primary limitation of this study for neurobehavioral assessment is that the criteria used to measure rearing, grooming, and ambulation during the 3-minute open-field evaluation are not described.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Facciolo (2002)	312 (p.148)	Sprague Dawley rats	**Premating, prenatal, postnatal** 10 days premating to weaning (44 days)	Oral (in oil); method not specified	0.04, 0.4;	M and F pups (not specified) from treated dams Ligand-binding assay (n=5-6), autoradiograms (n=4-6, age=PND 10, PND 23), binding activity to sst 2 receptors (n=10, 11, age=PND 10, PND 23) competition assay (n=5), regional change in binding of sst 2 receptors (high dose, n=12, 13; unknown N for controls, age=PND 10, PND 23 respectively) Pups were cross-fostered to reduce litter effects	N	N	Somatostatin type 2 receptors may represent an important mediating point for promoting estrogenic mechanisms of BPA, especially during early developmental phases.	Carefully performed. Used biologically relevant concentrations with specific neuropeptide receptors. Suggest that GABA system could mediate some of the xenoestrogenic effects of bisphenol A. Lack of some specific experimental details. Random assignment of 1 pup/litter within treatment groups is a weakness.	Useful	Limited utility

Additional considerations:

We agree with CERHR that this exploratory study is well designed and investigates effects at the cellular or receptor level. However, the biological significance of these changes on neurodevelopment still needs to be determined. Therefore, while the work contributes towards furthering understanding of possible mechanisms of action, this study has only limited use for risk assessment at this time. We note that in general the Facciolo papers are much more rigorously conducted than the Funabashi papers discussed below.

CERHR comments that "The random assignment of 1 pup/litter within treatment groups is a weakness." However, the authors appeared to have done this in order to reduce maternal influence on genetic siblings which would provide an experimental strength. CERHR may want to reconsider this comment. The study is well conducted in terms of ligand binding studies, which substantiate functionality of the receptor alterations. The use of multiple endpoints to test the presence and activity of the sst 2 receptor strengthens the conclusions. The use of high and low doses of BPA and oral dosing of the dams also provides strength to the study. However, the study is limited in that the data were not analyzed with respect to gender and there was no positive control for estrogenic mode of action in this study.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Facciolo (2005)	313 (p.149)	Sprague Dawley rats	**Premating, prenatal, postnatal** 8 days premating until weaning (42 days)	Oral (in oil); micropipette	0.04, 0.4;	Female offspring of treated dams (n=12 litters per treatment, 8 control litters) Somatostain receptor type 3 mRNA levels by quantitative autoradiography and in situ hybridization and sst_3–α $GABA_A$ subunit interaction(n=4/treatment/age, age=PND7 and PND55 for each technique) Pups were cross-fostered to reduce litter effects	N	N	sst3 receptor mRNA decreased in layer V of the adult frontoparietal cortex from exposure to BPA during gestation and lactation. In 7-day old rats sst3 receptor mRNA was reduced in the hypothalamic periventricular nucleus and increased in the ventromedial nuclei. Further changes involved an increase and decrease in sst3 receptor mRNA levels in the presence of αGABA (α1, 5) receptor subtype agonists. Therefore there appears to be a modulatory influence of BPA on neural circuits with potential influence on neuroendocrine/sociosexual behaviors.	Well controlled work examining effects of antenatal and lactational exposure provided orally to dam on the expression profile of somatostatin receptor subtype 3 and the role of GABA in its expression profile. Rigorous study performance and nature of endpoints were strengths. No major weaknesses.	Valuable	Limited utility

Additional considerations:

This paper presents mechanistic studies evaluating receptor alterations. The doses investigated are orders of magnitude higher than typical human exposures and should not be described as "relevant." The functional consequence of these changes to neurodevelopment requires further study before these findings, by themselves, can be used for risk assessment purposes. Additional limitations are that only 4 animals per group were used for GABA receptor subtype influence on SST3 receptor mRNA levels and a lack of a positive control such as estradiol. This study is also limited in that the mRNA endpoints are not substantiated with protein detection and quantitation, and the lack of neurobehavioral studies to support the authors' hypothesis. This work should be followed by other studies that study the functional relevance of these findings experimentally.

In adults given 400 μg/kg BPA, strong effects were seen in the frontoparietal cortex, while adults given 40 μg/kg had results in the ventromedial hypothalamic nucleus and hippocampal stratum radiatium. The authors do not address how effects are seen in different tissues at doses separated by only 10 fold. CERHR Table 76 (pg 150) had errors (i.e., PND 55 at 0.400 radiatium hippocampal reported increase while paper reported decrease. Also of note, one Facciolo figure (Fig 6) shows bars denoting distance from control and some very small bars were labeled highly significantly different from control, creating a question of correct labeling in the paper, or of statistical errors. These questions should be addressed before the study is used for further assessment. Overall, these are early studies that have been well performed, but are limited in their value for extrapolation to humans until they can be followed up with SST3 receptor binding assays, and studies that link the functional changes more directly with these molecular/biochemical changes.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Farabollini (1999)	317 (p 153)	Sprague Dawley rats	**Prenatal; postnatal** lower dose: 10 days premating until weaning; higher dose: GD 14-PND6	Oral (in oil); micropipette	0.04, 0.4;	12 or 15 F and M from 11 litters Holeboard test, plus maze, elevated plus-maze test (n= 15 [control], 15 [low], and 12 [high], age= 85 days for all endpoints) For each dose level, animals were randomized, but paper does not clearly explain how the 12–15 males and females per dose group were selected for testing out of the 9–11 litters of pups per dose group	N	N	Various aspects of nonsocial behavior were affected by BPA, differently in males and females. Contrary to our expectation, a clear masculinization of females was not observed. In general, the factor analysis indicated that BPA treated males had reduced motivation to explore and reduced anxiety. BPA females had reduced motor activity and reduced motivation to explore. There were no substantial differences between prolonged treatment with low dose and shorter exposure to high dose.	Adequately and rigorously performed. Lack of two controls (prolonged high-dose and short low-dose exposure). Behavioral effects induced by the two protocols was similar.	Adequate and useful	Limited utility

Additional considerations:

CERHR states in the rating of this study that this study "raises concern that early exposure, even to low doses of BPA, can give rise to permanent behavioral changes," and included Table 78 showing percentage change from control for only those statistically significant behavioral tests. This table does not report hole board (self grooming [f,d], rearing [f,d] % internal/total crosses) and plus maze (percent time in closed arms, rearing [f,d], self grooming[d], head-dips[f] measures) which were not affected by BPA at either low or high dose levels. The use of percentages in this table makes it difficult for the reader to see that differences in magnitude in some measures are small, such as number of entries into various sections (2.5 v. 0.9; 2.5 v. 4.2), number of stretched-attend postures (2.6 vs. 4.6), and frequency of head-dipping (2.4 vs. 5.8). This is a largely negative study, considering the multiple comparisons made on different behavioral endpoints, some of which are closely related (frequency and duration of same behavior), or ratios of behavioral endpoints. As the author stated, the direction of the effects on females is not consistent with the hypothesis that environmental estrogens masculinize the female brain. Therefore, the CERHR overstates the significance of this paper to raising concerns about permanent behavioral changes. Also, the use of different dosing schedules limits comparisons between study doses. This paper more specifically represents two different studies of a single dose each than one study using two doses that can be compared.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Farabollini (2002)	318 (p 154)	Sprague Dawley rats	**Prenatal only, postnatal only** Mating to birth Birth to weaning	Oral (in oil); micropipette	0.04	12 F and M from 7 treatment and 13 control litters; pups were cross fostered so that pups exposed prenatally were nursed by vehicle dams, or born to vehicle dams but nursed by BPA-treated dams, and control. Intruder test (n=12 pups/sex/group [prenatal-only, post-natal only, and control groups], age=100 days), sexual orientation (n=12 pups/sex/group [prenatal-only, post-natal only, and control groups], age=100 days and 1 week) and activity (n=12 pups/sex/group [prenatal-only, post-natal only, and control groups], age=100 days and 1 week [males], and a second 1-week interval [females]	N	N	Males but not females increased defensive behavior to an intruder. BPA had no effect on male sexual orientation toward a stimulus to females, but a slight decrease in latency and frequency of intromissions. BPA produced a small increase in sexual motivation and receptive behavior. The direction of the effect of the observed effect is not in line with expectation of a masculinization/ defeminization of the brain.	Carefully performed. Addresses aggressive/defensive behavior and sexual performance with interest in both male and female offspring. Use of single dose level of BPA. Does not address underlying biological mechanisms.	Serious contribution and suitable	Limited utility

Additional considerations:

The study is of limited rating because of only one dose level, and because sexual behavior was analyzed by pooling the pre- and post-groups to test for possible effects – there was no significance when analyzed separately.

CERHR states that this study is a serious contribution and the observations of this study suggest a potentiation of female behavior and a decrease in masculinity in adults resulting from perinatal exposure to low doses of BPA. Although this study is well conducted, the CERHR overstates the significance of this study and does not discuss the large number of negative results and relatively weak positive results. The paper is a largely negative study even though the authors conducted multiple comparisons of the different behaviors as well as of ratios and percentage of behaviors without correction for multiple comparisons. The few effects on individual parameters were described by the authors to be "slight," "weak," and/or "marginally significant" and not in the direction consistent with an estrogenic mode of action that should increase masculinity. There were no statistically significant effects on any of the directly measured behaviors in the intruder test for males or females. The only statistically significant effect noted was a slight increase in the ratio of defensive behavior to total agonistic behavior in females but not males. There were no BPA effects on sexual orientation for males and females. There were only marginally significant effects on lordosis posture and exit latency that became more significant when animals from the two BPA groups (treated either prenatally or postnatally) were pooled together in an *a posteriori* analyses. There were no effects on lordosis quotient (number of lordosis/number of mounts) and proceptive behavior in female rats. Although there were slight increases in number of intromissions and intromission latency and increased duration of genital sniffing, the important measures of sexual activity such as latency of ejaculation and refractory period were not significantly affected. There is no justification to single this study out as a serious contribution compared to other studies.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Fujimoto (2006)	285 (p 123)	Wistar rats	**Prenatal** GD13-PND0	Oral; drinking water	0.015 Estimated by authors from 0.1 ppm water	20–24 F and M from 6 litters Open field test (n=22–24)passive avoidance learning (n=20–24) elevated plus maze (n=21–23) forced swimming test (n=24) all at 6-9 weeks of age	N	N	Prenatal BPA exposure impairs sexual differentiation in rearing and struggling behavior and facilitates depression-like behavior.	Good methods to examine functional disruptions in sexually dimorphic behaviors. Lack of clarity about the nature of disruption of sexually dimorphic behavior patterns indicated by author's conclusions. Small sample size. Use of singe dose level (not confirmed). Lack of clarity of statistical methods regarding litter.	Adequate	Limited utility
Additional considerations: This study is rated as limited utility not only because of the use of a single low dose, but also because of (a) low number of litters and lack of litter as the experimental unit, and (b) lack of a positive control to demonstrate the effect of an estrogen analog and support the conclusions of this study. Also missing is data indicating the stability of BPA in the water and frequency drinking water solutions were freshly made up. Rearing and struggling behavior has not been rigorously demonstrated to be a sexually dimorphic behavior. Furthermore, there were no treatment-related effects on the other closely related open field behaviors such as total distance and time in center even though there were gender differences shown in these behaviors. There is no clear evidence of a hormonal basis for these endpoints, so it is overstating the evidence to use this data as evidence that BPA 'impairs sexual differentiation'.												
Funabashi (2001)	416 (p 263)	Wistar rats	**Adult** ovariect-omized	SC single dose	Single dose 10mg BPA (40 mg/kg bw) or 17β estradiol, 10µg	7-8 week old ovariectomized rats N=6/group 3 groups, 10 days post ovariectomy, rats were given: oil vehicle or BPA 10 mg, or 10 µg 17β estradiol. Rats were sacrificed 24 hours later and mRNA was extracted from pre-optic area, medial basal hypothalamus, and anterior pituitary for examination by Northern Blot for changes in progesterone receptor, neurotensin and preproenkephalin.	N	Y	The study authors concluded that BPA increases the expression of progesterone receptor mRNA in the pre-optic area and anterior pituitary of adult ovariectomized rats.	Demonstrated that BPA induced progesterone receptor mRNA in the pre-optic area and anterior pituitary, similar to 17β-estradiol. Potential concomitant changes in progesterone receptor protein were not examined. Single dose and route was sc.	Limited	Inadequate
Additional considerations: One high dose using a subcutaneous route of exposure makes this study inadequate for risk assessment purposes. In addition, the molecular work is not sufficiently rigorous to support the conclusions made by the authors and of uncertain biological significance. CERHR should consider including the following discussion in the strengths/weaknesses section. This study provides evidence that BPA has an effect on stimulating the transcription of progesterone receptor mRNA, which suggests increases will follow in protein. However, no follow up protein data were provided to rigorously support this conclusion. The authors bring up a very relevant point to consider at the end of the paper by commenting that endocrine disruptors such as BPA have been shown to behave in an inverted U-dose response so that very low doses of BPA could have more of an effect than higher doses, yet in this study and the ones that follow by the same authors, we see that BPA in fact does not behave this way but it takes a relatively high dose, one that would not be achieved in humans under any realistic circumstances, to promote the induction of progesterone receptor transcripts. Also, BPA did not mimic the estrogenic effect that induced preproenkaphlin mRNA within the same study.												

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Funabashi (2003)	418 (p 265)	Wistar rats	**Adult** ovariect-omized	SC single dose	approx. 0.004, 0.04, 0.4, 4, 40 From author, 10 mg/day/ animal equals approximat ely 40 mg/kg	Exp 1. 3-5/group, 3 groups: sesame oil sham, 10 mg BPA (40 mg/kg), 10 μg 17β estradiol (40 μg/kg bw). Injected 14 days post ovariectomy. Rats sacrificed the day following dosing and brains fixed in paraformaldehyde for localization and detection of PR protein Exp 2. 3-4/group, 4 groups .004, .040, .4, 4 mg/kg BPA, rats sacrificed the next day and brains processed as above. Exp 3. 4 groups, n=5-8 per group: rats were injected with oil, 10mg BPA, 10 μg estradiol and then next day with 1mg progesterone (P), or oil. Groups: (oil +P), (BPA+ oil), (BPA+P) (Est+P). Lordosis and rejection responses were measured 5-7 hours after last injection.	N.A.	Y	The study authors concluded that the findings suggest that BPA influences sexual behavior by altering the progesterone receptor system in the hypothalamus.	Well conducted study that demonstrated rats injected with BPA exhibit 17β-estradiol-like responses. Consistent observation with estrogenic activity of BPA. Number of animals per group is sufficient. Weakness is sc route of administration.	Limited	Inadequate

Additional considerations:

Both the use of ovariectomized animals and a single high subcutaneous dose makes the first experiment of this study inadequate for risk assessment purposes, as do the small group sizes. This study is more important for its exploration of a possible mechanism of action than as a study for risk assessment at this time. This study demonstrates that ovariectomized adult rats injected with 10 mg BPA have an increased number of progesterone receptors as detected by immunoreactivity in the pre-optic area and the ventromedial hypothalamic nucleus, and that this increase is directly dose dependent above 0.1 mg BPA. This study does NOT show estradiol-like responses in rats treated with BPA because the main behavioral test used to test this functionally, lordosis, was negative. The results from the second test, rejection, were also not supportive of an estrogen-like response because rats in the positive control group that were given 17β-estradiol, did not show statistically different rejection behavior than the oil alone group or BPA+ oil group. Therefore, the comparisons of the rejection response of rats given BPA+Progesterone are not relevant to a claim of consistency as an estrogen agonist. One statistically different response from the control does not make an “estrogen-like” response if the estrogen treated group does not respond similarly with statistical significance. The authors overstate the significance of their study, and CERHR should be careful to document the internal inconsistencies of their results with the authors’ conclusions.

Funabashi (2004) (J Neuroendocrinol)	419 (p 266)	Wistar rats	**Adult** ovariect-omized	SC single dose	Single dose 10 mg (approx. 40 mg/kg bw)	Exp 1 n=6, 2 groups one sham one experiment. Animals ovariectomized at 7 weeks, injected once sc 10 days later with 10 mg bisphenol A (age is approximate, not well described). Sacrificed at 24hrs post injection. Frontal, parietal and temporal cortex examined for progesterone receptor (PR) mRNA by Northern Blot analysis. Exp. 2 n=5-6 rats/group, 4 groups, 0, 6, 12, 24h post injection. frontal, temporal, and occipital cortex examined for PR mRNA by Northern Blot	N.A.	N	Bisphenol A may affect sexual neurobehavioral endpoints in adults because it has the ability to alter progesterone receptor mRNA in the frontal cortex past the neurodevelopmental stages.	Study links single high dose 10 mg (40 mg/kg) sc bisphenol A to the induction of progesterone receptor mRNA in the frontal cortex, an estrogenic response. A positive control is absent, as is data indicating whether the corresponding protein is increased. There is only one sc dose, bypassing potential first-pass metabolism. The authors imply that an effect on the progesterone receptor may have an effect on behavioral endpoints but do not show studies to ascertain that relationship.	Limited	Inadequate

Additional considerations:

This study is inadequate primarily because of the single high subcutaneous dose to ovariectomized female rats. This, together with the limitations of the molecular work leads us to rate this study as inadequate. The study conclusions are not well-supported by the experimental data because of the lack of data on actual progesterone receptor protein levels or function. There was no positive control for maximal estrogen induced progesterone receptor mRNA. The authors attempt to link neurobehavioral outcome with mRNA induction of progesterone receptor without measuring receptor protein levels to validate mRNA increases and show translated protein. There were also no activity assays (receptor binding studies) and physiologic response data on sexual behavior. The authors state that they have done the localization of the PR protein in the frontal cortex with immunohistochemistry and this supports their mRNA data (Northern) but have not published the results---this is not a sufficient explanation especially since their 2003 paper has experiments that are published with this endpoint in other regions of the brain.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Funabashi (2004) (Psychoneuronendocrinology)	284 (p 122)	Wistar rats	**Prenatal; postnatal** dams treated until weaning (dosing start time is unknown)	Oral; drinking water	2.5; Estimated by authors from 10 mg/L	8-11 M and F pups from 8-11 litters Immunocytochemistry for corticotropin-releasing hormone-staining neurons in preoptic area, and anterior and posterior bed of the stria terminalis (n=8–11, age 4-7 months) The number tested is consistent with 1/sex/litter, but is not stated	N	N	Exposure to BPA during gestation and lactation reduces sex difference in the bed of the stria terminalis and the pre-optic area.	Appropriate design and statistics. Relevance and subtleties of endpoints measured. Uncertain number of animals examined and duration of dosing period. Results suggest disruption of normal patter of sexually dimorphic neurons, important to concerns relevant to reproductive function and sexual dimorphic behavior. Control for litter effects not clear.	Adequate	Limited utility
Additional considerations: The route of exposure is relevant but only one dose is used. The presence of a change in protein levels is an exploration of a possible mechanism of action, not demonstration of a functional effect. Molecular information is important to guide further research, but is not sufficient to use for risk assessment purposes. The conclusions of this study are based on one experimental parameter (immunocytochemistry) that is a qualitative rather than a quantitative technique, usually used to show trends or presence/absence of a protein. The expression of a protein is better quantified with real time PCR and Western blots with densitometry. The authors speculate about specific functional changes, but do not provide the data to support this hypothesis. The authors conclude that both female and male animals have altered numbers of neurons in the anterior and posterior bed of the stria terminalis to eliminate the gender difference seen in the control animals. These results would be stronger with stereology to control for total cell count and area dimensions, as opposed to immunohistochemistry. A strength of this study is that the time of sacrifice of the females was controlled for estrus cycle. A weakness is that the authors do not state that they have tested the stability of the BPA/water mixture or how often they replace the treatment water with a fresh batch.												
Halldin (2005) This paper reports data based on Halldin 2001	402 (p225)	Japan-ese quail	**Prenatal**	Injection into egg yolk	67, 200 μg/g egg	N=5-7 females, 4-7 males Male endpoints: sexual behavior, testosterone analysis, testis weight, body weight. Female endpoints: egg laying, oviduct morphology, body weight	N	Y	BPA did not cause any significant estrogen-like effects on behavior, testosterone, gondao-somatic index, or testis weight in males, or on egg laying in females. In females, exposure to BPA showed a tendency toward retention of the right oviduct. We conclude that the risk for adverse reproductive toxicity in avian wildlife is probably low.	2 positive controls and attention to sexual behavior are strengths. The expression of exposure level as μg per egg is difficult to compare to human exposure, and the lack of details about the reporting of methods and results, and the lack of apparent statistical analysis, are weaknesses.	Limited Utility	Inadequate
Additional considerations: CERHR reviewed the review paper. Our rating takes into account the original 2001 paper. This study is inadequate for human health risk assessment purposes because of the lack of ability to compare exposure level as mg/kg/day and understand relevance to humans without demonstration of similar responses in another mammalian system. This work might be relevant for wildlife risk assessment and possibly for interest as a mechanistic study, but is not appropriate for human risk assessment at this time. The original 2001 paper was not reviewed, so the weaknesses noted by CERHR are because the 2005 paper is a review. In the original 2001 paper, details about the data, methods, and statistics were complete and appropriate. Halldin 2005 is a review article using data from three papers by the same first author. The BPA data is from Halldin 2001. The authors' premise is that previous work in Japanese quail indicated that tetrabromobisphenol A affected reproductive parameters in quail and chickens, so BPA, TBBPA and DES were more completely examined in this paper. The laboratory has a history with this model and performed appropriate controls, such as documentation of the spread of radiolabeled BPA within the egg yolk for embryonic uptake and distribution. The endpoints were appropriate but not extensive. The statistics were appropriate. The Kurskall-Wallis test was used for comparing groups, and Chi-squared for frequency of retained oviducts in females. The number of eggs was low; increasing the number of eggs exposed might bring the frequency of females retaining the right oviduct into statistical significance.												

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Ishido (2004)	337 (p 170)	Wistar rats	**Postnatal** PND5	Intracisternal injection to pup	0.02, 0.2, 2, 20;	6 M for motor activity, unknown number for neuro-chemistry, from unknown # litters (probably 10) Motor activity (n=6, age=4-5 weeks), DNA microarray (n=?, age=4 or 8 weeks), brain TH+ immunohistochemistry (n=?, age=8 weeks)	N	Y 6OH-DA	BPA affected central dopaminergic activity, resulting in hyperactivity possibly from reduction of tyrosine hydroxylase.	Good range of concentrations of BPA and correlation of behavior patterns it induced with expression of specific dopamine receptor sets. Inability to correlate doses of BPA given by intracisternal injection with s.c. or oral routes. Behavioral data are stronger than cursory molecular study where one possibly relevant receptor was chosen from a microarray set with minimal follow-up.	Suitable	Inadequate

Additional considerations:

We recommend that CERHR change rating for this study due to inadequate methods and incomplete reporting of data. Although the 24 hour motor activity test is superior to other methods for measuring motor activity, the behavioral evaluation is severely limited by relatively low sample size, and lack of information on control of confounders for motor activity such as balancing the time of testing across treatments. There was no description of methods used to measure degree of immunoreactivity of tyrosine hydroxylase. There was no reporting of number of animals selected for DNA microarray and incomplete reporting of doses causing decrease in gene expression. The use of intracisternal injection to the brain at PND 5 limits the use of this study for risk assessment purposes, although such approaches can provide insight into mode of action if pharmacokinetic data are available to relate intracisternal injection to oral routes of exposure. This would have to include a demonstration of metabolism of parent BPA within brain tissue to make up for the lack of first-pass metabolism that occurs with oral dosing.

It should be noted that there is a monotonically increasing dose response for motor activity following increasingly higher intracisternal doses of BPA. This is in contrast to the overall lack of effects on motor activity following oral doses seen in other studies of motor activity. The conclusions drawn on the basis of the molecular data presented are unsubstantiated and speculative far in excess of the data presented. It is important to note that although the authors use a positive control, it was not for its estrogenic mode of action. Instead, 6-OHDA was used as a positive control for degeneration of dopaminergic neurons. Therefore, this study does not get a higher rating because of inclusion of positive control. Furthermore, it is unfortunate that they did not include the 6-OHDA animals in the motor activity test. Instead, the authors cite unpublished observations (from another study?) that 6-OHDA caused hyperactivity comparable to bisphenol A.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Kawai (2003)	358 (p 191)	CD-1 mice	**Prenatal** GD11-17	Oral (in oil); micropipette	0.002, 0.02;	8–16 M for testes weight, 10–32 M for contact time from 7 litters Aggression (more precisely contact time) (n=4-5 male pups, 30 control, 32 low BPA, and 26 high BPA, age=PND 8,12, and 16) testes weight (n=8-14 mice/group [litter does not appear to be unit], age PND 9, 13, and 17, 1 week after aggression test), testosterone (n=8–14 mice/group [litter does not appear to be unit], age PND9, 13, 17, 1 week after aggression test)	N	N	BPA increased aggression scores at 8 weeks but not at 12 weeks. BPA decreased relative testis weight but had no effect on serum testosterone concentration. These results demonstrate that BPA temporarily activated aggressive behavior in mice and interfered with normal development of reproductive organs.	Good use of 2 low dose levels and oral route of administration. Lack of husbandry info and consideration of possible litter effects.	Moderately useful	Limited

Additional considerations:

The CERHR statement that the study showed increased aggression should be removed from the utility section because the study only measured contact time and the author's own text clearly states that there was no indication of any attacking behavior. CERHR should distinguish between what the author concludes and what the study actually measured. More critical analysis is needed when discussing results of studies in the sections on utility and strength and weaknesses because the presumption is that the reviewer critically reviewed the study and agrees with authors statements. If there is no data to support the authors' conclusions, then these sections should avoid stating them in these sections.

CERHR should add the following strengths and weaknesses. Strength: Testosterone levels and testes weight were measured and correlation between contact time and testosterone concentration was statistically analyzed. Weaknesses: The authors overstate their results such that CERHR reviewer misinterpreted that BPA caused increase in aggression. It would be more accurate to state that BPA increased contact time during which mice sniffed each other but did not show any attacking behavior.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Kubo (2001)	310 (p 145)	Wistar rats	**Prenatal, postnatal** GD1-PND 21	Oral; drinking water	1.5 (approx, calculated from water)	11–14 F and M from 5 litters Open field (n=11–14 pups/group, age=6 weeks) passive avoidance (n=11–14 pups, age=7 weeks), size of locus coeruleus (n=6–7 pups/ group, age=20 weeks) and SDN-POA (n=6-7 pups/group, age=20 weeks), brain weight (n=6–7 pups/group, age=20 weeks), reproductive organ weights (n=12–14, age=12 weeks), hormone levels (n=5–10, age=12 weeks)	N	N	Control female offspring showed higher activity, lower passive avoidance memory, larger locus coeruleus and smaller SDN-POA than male controls while the BPA-exposed group did not show any sexual dimorphism inl measurements except SDN-POA.	Variety of biological and behavioral endpoints assessed. Lack of experimental detail including dose of BPA received by animals.	Limited utility	Inadequate

Additional considerations:

CERHR should downgrade the rating of this study from limited to inadequate because there was inadequate discussion of methods, there were only 5 litters/group, the litter was not considered the experimental unit, and there was inadequate discussion of how animals were selected for the different endpoints with respect to representing different litters. There was inadequate discussion of the statistical analyses conducted. Based on the figures it appears that the statistical analyses focused on differences between males and females in any one treatment group. It is not clear whether there was NO statistical difference between BPA treated animals and controls or whether the authors failed to conduct this analysis, although it would be surprising if they did not. Rearing, cell volume and cell density measurement were not conducted blind to treatment level. Taken together, these flaws in the methods and reporting of the data make this study inadequate.

CERHR is correct that the major weakness of the study is the lack of experimental detail, but only comments on difficulty in determining how much BPA was received by the animals is cited by CERHR. Although this is an important limitation, it is not the most serious deficiency of the study. CERHR did not mention additional important weaknesses of this study described above under “rating”. In -trial passive avoidance behavior as a test for cognition can be confounded by motor activity because higher activity can result in decreased latency to cross to the dark side where the shock is. Therefore, gender differences in passive avoidance may simply be a reflection of gender differences in activity level (females > males).

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Kubo (2003)	311 (p.146) .	Wistar rats	**Prenatal, postnatal** GD1-PND21	Oral; drinking water	0.03, 0.3 (approx, calculated by authors);	20–24 F and M for open field, 7–8 F and M for anatomy, from 6 litters Open field (n=20–24 offspring/ sex/group, age=6 weeks), locus coeruleus and SDN-POA (n=7–8 offspring/sex/group, age=14 weeks), reproductive system parameters (including serum hormone levels including 17βestradiol and testosterone) (n=13–15 offspring/ sex/group, age=12 weeks [females killed at proestrus stage on day after diestrus day]), sexual behavior (n=7–13 offspring/sex/group, age=11–12 weeks) For measurements made prior to weaning: The individual animal was the experimental unit for behaviors measured after weaning	Y	Y	BPA disrupts or inverts the normal sexual differentiation in both brain structure and behavior without any adverse effect on the reproductive system. DES but not BPA and RVT influenced female sexual development but not in males.	Failure to describe methods to determine dose received by dam. Good subtlety and relevance of neurologic endpoints assessed. Finding related to brain development and size of the locus coeruleus along with possibly related behavioral changes (no effects on reproductive tract noted). Explored low dose exposures to BPA	Adequate	Limited utility

Additional considerations:

CERHR should point out that the litter was not the experimental unit and the total number of litters in this study was only 6. For comparison, EPA guidelines require 1 male or female from each litter from at least 20 litters, hence the study is of limited utility. However, although the previous paper had important flaws, similar effects were measured in both studies. It is puzzling that DES did not produce expected effects on the SDN-POA. It is also puzzling that the there appears to be a reversal in gender size rather than an increase in the difference. The authors provide no explanation for the reversal of gender differences for some parameters and this inconsistency also limits the utility of the study. According to the authors, the primary hypothesis being tested is that the brain is female type in mammals and the brain is only masculinized in males by 17beta-estradiol converted from testosterone via aromatase in the neurons in the perinatal critical period. The female brains are unaffected. The authors also cite evidence that neonatal injection of testosterone into females led to LC size similar to normal male LC.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Kwon (2000)	297 (p.144)	Sprague Dawley rats	**Prenatal, postnatal** GD11–PND 20	Oral; gavage	3.2, 32, 320;	F and M, mean values used from 8 litters (in other words, greater than 1 male and 1 female per litter, but litter was always the statistical unit) SDN-POA size (mean of 6–8 litters), pubertal development (mean of 7–8 litters), estrous cyclicity (mean of 7–8 litters), lordosis behavior (n=7–9 animals). Other endpoints not relevant to neurological assessment are not included in this list.	Y	Y	Pre- and post-natal exposure to BPA did not have any apparent effect on female rat pubertal development and reproductive functions.	Well performed and presented. Wide dose range (across three logs). Good use of diethylstilbestrol as a positive control. Good number of reproductive organs and endpoints evaluated. Limited analysis of reproductive organs and lack of determination of pup exposure during lactation.	Adequate	Adequate
Additional considerations: Agree with CERHR that lack of histology on organs is a weakness, but overall study is good and deserves the rating of adequate. The data is strengthened by use of greater than 1 animal/sex/litter and analyses of data using litter as the experimental unit.												
Laviola (2005)	359 (p.192)	CD-1 mice	**Prenatal** GD11-18	Oral (in oil); micropipette	0.01	3 F and M from each litter, 3–4 litters assumed Conditioned place preference, amphetamine-induced motor activity (n= 3/sex/litter, age= 60 days for each endpoint)	Y	N	Prenatal effects were sex-dependent, and no changes were seen in males whereas BPA eliminated the place conditioning seen in control females. Locomotion was not affected. As a whole, (BPA) affected some steps in the organization of the brain dopaminergic system in females leading to long-term alternations in neurobehavioral function.	Use of only one dose level. Small sample size.	Slightly useful	Limited utility.
Additional considerations: None.												
Mizuo (2004)	373 (p.203)	ddY mice	**Prenatal, postnatal** Mating to weaning	Oral; diet	0.03, 7.5, 30;	6–10 M from an unknown # of litters Conditioned place preference (n=6–10 mice/group, age not reported), morphine-induced activity (n=9–10 mice, age not reported), RNA (μ-opioid receptor/GAPDH), G-protein activation.	N	N	BPA produces supersensitivity of the morphine-induced rewarding effect and hyperlocomotion without direct changes to the u-opiod receptor, but that BPA does affect development of the central dopaminergic system.	Wide dose range. Essential information is not included.	Not useful	Inadequate
Additional considerations: This paper does not appear in the CERHR neurodevelopmental sections, but is discussed on p. 203 and included in a table on p. 305. Descriptions are lacking detail that would allow replication or good evaluation. These studies suggest that there is an enhancement of the reward effect and hyperlocomotion induced by morphine in the presence of BPA although the molecular mechanism for this is unclear. The authors speculate that this is due to dopamine D1 receptor changes in the limbic forebrain (previous paper) rather than the lower midbrain since the μ-opiod receptor did not change in this region at the message or protein level. The authors hypothesize that the supersensitivity to morphine is a result of BPA's effects on the dopaminergic receptor. The biochemical/molecular work is preliminary at best.												

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Miyatake (2006)	374 (p 204 and 230)	DdY mice	**Prenatal and neonatal**	Oral gavage	0.003, 200	7 M from an unknown # of litters	N	Y	BPA enhanced the dopamine-dependent rewarding effect induced by morphine	2 doses used, 1 low and 1 very high. Both had similar effects.(p 204)	Moderately useful	Limited utility
		ICR mice	**Neonatally**-derived cells	8 days after cell plating	10 fM – 1 μM, dilutions of 10, inclusive	Unknown # of M and F brains from unknown # of litters, obtained at postnatal day 1, placed in neuron-glia co-culture or as purified astrocytes Immunohistochemistry for GFAP and caspase-3, Ca++ imaging	N.A.	Y	Treating purified astrocyte cultures or neuron/glia co-cultures with BPA activated astrocytes (stellate morphology and increase in GFAP) in a biphasic response, and enhanced the Ca++ response to dopamine. High BPA also induced the activation of caspase-3. 17β-estradiol did not have these effects.	Multiple concentrations of BPA over a wide range, evaluation of multiple endpoints, comparison to known receptor ligands are good.	Not useful	Inadequate
Additional considerations: Although the authors balance order of drug injection and compartment for receiving treatment across subjects, and time spent in compartments was scored by infrared sensor, both good design components, the use of extremely different doses is not useful for a dose-response. Given the biphasic response seen in vitro by the same authors, the use of one low and one very high dose is difficult to interpret since both doses showed a similar response. The cell culture methodology is good and well-described, although lacking functional endpoints relevant to human risk-assessment. These studies provide preliminary mechanistic data that may be useful for developing hypotheses on possible modes of action, but are not useful for human health risk assessment at this time.												
Nagao (1999)	325 (p.160)	Sprague Dawley rats	**Postnatal** PND 1-5	s.c. injections	300	15 M from an unknown # of litters Viability and growth (n=22–31, age=0-9 weeks), reproductive performance (n=22–25), male sexual behavior (n=10, age-14 weeks), weight and pathology of reproductive organs and brain (n=12–15, age=PND21 and 14 weeks), size of SDN-POA (n=?, 14 weeks)	N	Y	Neonatal exposure to estradiol benzoate affects reproductive function and the SDN-POA, but high dose BPA does not.	Well performed and documented study comparing effects of BPA and estradiol benzoate. Good documentation of behavioral (mating) and biological (genital tract development) endpoints in both male and female rats. Use of single high dose of BPA and choice of PND 1-5 for exposure are weaknesses.	Suitable	Limited utility
Additional considerations: The study is rated limited utility only because only a single very high dose is used, and is given s.c., eliminating the first-pass metabolism of oral doses. Also, the authors do not state that the behavioral studies are monitored by an observer blind to treatment. Additionally, the males are presented with an ovariectomized female that has been given hormones to bring her into receptivity, but there is no redundancy to determine whether a male that does not respond to a given female would otherwise respond to an intact female or that the same ovariectomized female would simulate response in males that had previously responded to ovariectomized females. However, the negative results for BPA on the SDN-POA and on reproductive parameters including sexual behavior at this extremely high dose is reassuring. The observation that there is no change in the SDN-POA is consistent with lack of effect of BPA on SDN-POA seen at lower doses in other studies.												
Narita (2006)	357 (p 190)	ddY mice	**Prenatal, postnatal** Mating to weaning	Oral; diet	0.006, 0.06, 0.6, 100, 400 (calculated by CERHR);	Unknown total number of litters. See endpoints for sample size. Conditioned place preference (n=6-14, age=7 weeks), morphine-induced motor activity (n=5-15, age=7 weeks), dopamine binding assay (n= 3, age=7 weeks)	N	N	Pre- and neo-natal exposure to BPA potentiates the central dopamine receptor-dependent functions, resulting in supersensitivity of morphine-induced hyperlocomotion and reward in mice.	Poorly written and difficult to understand. Inability to pass its message to reader.	Inadequate	Inadequate
Additional considerations: Agree with CERHR review that methods are poorly described. Additional considerations to add to strengths and weaknesses of this study are: A biphasic response is seen at extreme ends of a dose-response for drug-paired place, total activity after morphine, and GTP binding. These results would need to be repeated and much better described to carry weight, but could contribute towards better understanding of possible effects of BPA to increase sensitivity to morphine.												

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Negishi (2003)	315 (p.151)	Fischer 344 rats	**Prenatal, postnatal** GD10–PND 20	Oral (in oil), instrument not specified (dam only)	4, 40, 400;	9–36 F and M from 8–9 litters Body weight (8–9 litters/group [8 pups/litter], age= PND 7, 14, 21, 28, 56, 84), motor activity (n=12–27 offspring/sex/group, age=4 weeks), active avoidance (8–9 offspring/ sex/group, age=4 and 8 weeks), open field (n=not reported, age=8 weeks)	N	N	BPA increased immobile time during dark phase in female offspring; altered the active avoidance responses in male offspring (improved in the mid- and high-dose group at 4 weeks but decreased in the low-dose group at 8 weeks of age); and increased grooming at the low dose. Based on these findings, the author concludes that there are sex differences in the behavioral sensitivity to BPA even though there is no sex difference in behavioral score.	Doses sufficiently high to produce gross body weight changes. Three different measures of behavior collected and organ weight at necropsy from same animal. Lack of evaluation of hormone-dependent behaviors and assessment of more hormone-dependent tissues or processes. Analysis was not litter based and there was no positive control.	Adequate	Limited utility
Additional considerations: The lack of low doses and inconsistent pattern of effects makes this study of limited utility. There were no clear criteria established to define "grooming," "stretching," and "other" behaviors. The conclusions overstated the significance of the results. The behavioral effects noted were not very robust effects based on overall pattern of effects. The increase in “immobile time” measured at 4 weeks has uncertain biological significance given that there were no effects on 12-hour motor activity. The increase in avoidance response in BPA treated animals at 4-weeks is consistent with improved performance. The decrease at 4 mg/kg/day at 8 weeks is difficult to consider as evidence of a U-shaped dose-response curve in the absence of such an effect at 4 weeks of exposure and lack of effect during the second and third block of 50 trials and lack of significant interaction between treatment x gender and increase in grooming but not other parameters during the open field evaluation.												
Negishi (2004)	316 (p.152)	Fischer 344 rats	**Prenatal, postnatal** GD3– PND20	Oral (in oil); micropipette	0.1	9-10 M from an unknown number of litters Open field (n=1 pup/litter [9–10], age=8 weeks), motor activity (n=1 pup/litter [9–10], age=12 weeks), passive avoidance (n=1 pup/litter [9–10], age=13 weeks), elevated plus maze (n=1 pup/litter [9–10], age=14 weeks), active avoidance (n=1 pup/litter [9–10], age=15 weeks), tranylcypromine-induced activity (n=1 pup/litter [9–10], age=22–24 weeks)	Y	N	BPA had no effects on open field test, spontaneous motor activity, or elevated plus maze test. There were no effects on passive avoidance test but a tendency towards remaining in the light environment which is consistent with improved performance. BPA decreased active avoidance of shock. BPA attenuated the activity increasing effects of a single dose of tranylcypromine.	Variety of endpoints to provide data. Use of single dose level which point to effects that are not gross structural changes but subtle behavioral effects	Adequate	Limited utility
Additional considerations: The active avoidance test may have been confounded by the passive avoidance test that was performed in the same animals. The passive avoidance test has a one trial test that did not have an acquisition compound. Only males were tested limiting ability to compare results of this study with other studies claiming disruption of gender effect. Limited utility due to only one dose tested.												

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Palanza (2002)	352 (p.183)	CD-1 mice	**Prenatal, postnatal** GD14-18 (in utero) and 2–2.5 months (as adults) of F1 females	Oral (in oil): micropipette (dams and selected "pups" as 2nd generation of dams)	0.01	15–20 F and M pups (M for postnatal development) from 9 F0 and 31 F1 litters Maternal behavior elements (n=dams 20 oil-oil, 15 oil-BPA, 15 BPA-oil, 15 BPA-BPA, age= lactation days 2–15), birth measurements (n=litter offspring 20 oil-oil, 15 oil-BPA, 15 BPA-oil, 15 BPA-BPA, age=PND 1), postnatal developmental landmarks (n=8 litters/group [all 10 pups/litter evaluated], age=PND 3, 5, 7, 9 [also 13 for body weight])	Y	N	Dams exposed to BPA either as fetuses or in adulthood spent less time nursing their pups and more time out of the nest compared with the control group. Females exposed to BPA both as fetuses and in adulthood did not significantly differ from controls. No alterations in postnatal reflex development were observed in the offspring of the females exposed to BPA.	Good oral route of administration, low dose of BPA, and exploration of effects on complex maternal behavior. Failure to explain that pre- and postnatal exposure had effects but not the combination of the two. 1 in 6 maternal behaviors was affected in exposed mice during both time periods. Use of diet high in soy isoflavones is poor.	Very useful	Limited utility

Additional considerations:

We disagree with CERHR rating this study as very useful because of the single dose level used. As CERHR noted, the authors did not adequately explain why pre- and postnatal exposure had effects but not the combination of the two, which also makes the study of limited utility. Other important limitations are (a) there was no mention of whether the observations and behavioral tests were conducted without knowledge of treatment group, and (b) there was no mention of whether the time of evaluation was balanced across treatment level. The strengths were that there were operational definitions for the normal maternal behaviors recorded, and postnatal developmental landmarks and growth were evaluated in offspring to better evaluate potential impact of changes in maternal behavior elements.

This paper does not appear in the neurodevelopmental section. CERHR should consider adding the following points in their review. The methods section is unclear regarding exactly how long animals were evaluated each of the 30 times they were evaluated. The methods section state that dams were observed once every 4 minutes, but it does not appear that animals were observed for the entire 4 minutes because only behaviors displayed "at the moment of observation" are recorded and the maximum frequency possible is 30. [Four minutes is a very long period of time during which the frequency of different behavioral measures would be much greater than 1 per observations period.] It is important to report precisely how long each animal was observed because that provides greater assurance that objective evaluation of each animal was conducted in identical manner. Therefore, Figure 1 is mislabeled as "average percent time" spent on maternal behavior variables, when it should be "average percent frequency". Although, this did not impact our rating of the study, the study design might have been improved by evaluating maternal behavior for an extended period of time (rather than snap shots throughout the day), so that a more integrated pattern of maternal behavior might be analyzed. Since the maternal behaviors were measured as independent snap shots, it would also have been helpful to include assessments of whether pups had milk in their stomach or whether the dam already established a nest for the litter to put into perspective the time spent nursing or building a nest. Statistically significant differences in time spent in these normal behaviors were assumed to be adverse and were not discussed in relation to whether these changes have biologically meaningful effects on function given that BPA had no effects on "in-nest" behavior, "licking", or "forced nursing behaviors, body weight, cliff-drop aversion reflex. The non-statistically significant tendency towards decreased righting reflex occurred only in the BPA-OIL offspring, but not the OIL-BPA or BPA-BPA offspring. Thus, there are relatively few effects that occur in a consistent pattern. There is no data to support the author's suggestion that BPA's effects on maternal behavior are a result of a direct effect on the neuroendocrine substrates underlying initiation of maternal behavior. There is no evidence that the changes in measures are indicative of an adverse effect on nurturing that impacts development. It should be noted that the statistical analyses are not conducted on the actual scores of the maternal behavior observations but on converted percentages of the maximum frequency possible (30) for each observational period.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Patisaul (2006)	338 (p.171)	Sprague Dawley rats	**Neonatal** PND1-2	Twice daily s.c. injection to pups	500 μg /day 83 (our approximation based on birth weight of 6 grams from Charles River)	F and M, 5–8 of each gender from 5 litters Size of the AVPV and number of TH+ and ERα cells within (age= PND 19 for all endpoints) Litters were cross-fostered	N	Y	Acute exposure to endocrine-active compounds during a critical period alters AVPV development; also, BPA demasculinized TH immunoreactivity in the male AVPV and defeminized females.	Good use of 17β-estradiol as a positive control and measurement of ERα receptors. Relatively high dose level of BPA and use of injection route of exposure of newborn pups is poor	Moderately useful	Limited utility

Additional considerations:

This paper should be rated lower than it was by CERHR because of the relatively high dose given twice daily as subcutaneous injections directly to pups makes this study limited for risk assessment purposes. The normal route of exposure for a pup of this age would be through the dam's milk.

CERHR should note that the paper does not clearly state if the borders of the nuclei were drawn without knowledge of treatment level. Reference 5 of this paper indicates that sexual dimorphism of the AVPV does not occur until PND 30–90 while these animals were sacrificed at PND 19. The total number of cells TH+ and ERα+ was quite low, approx 12–20 in females and 6–8 in males; is such a small number biologically relevant and able to be evaluated? These data should be considered preliminary. In addition, the author's postulation that changes in the AVPV neonatally may "ultimately affect the estrous cycles of adult females" is not supported by the results of several multi-generation studies at any dose.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Porrini (2005)	321 (p 157)	Sprague Dawley rats	**Prenatal, postnatal** Mating to PND 21	Oral (in oil); micropipette	0.04	18 M and F from 12 litters Social and non-social exploration, defensive toward males, play with males or females, low-intensity mating, social grooming (n= ?, 35, 45, 55 days of age for all endpoints) Pups were cross-fostered to reduce litter effects	N	N	BPA does not induce a clear masculinization of female behavior but is able to defeminize some aspects of female behavior.	Well-performed study with poorly researched endpoints (juvenile play behavior) that have implications for reproductive behavior later in life. Data are objective (blinded). Use of single dose level is poor. Fostering pups within treatment groups prevents evaluation of intrauterine effects. Evaluation of play is questionable (only rough and tumble play is sexually dimorphic). Behavior is organized by androgens, not estrogens, decreasing the biologic plausibility of the conclusions	Suitable	Limited utility

Additional considerations:

This study has only one dose level and does not have a positive control to support conclusions made by the authors. Also, the results do not present the actual behavioral measurements.

CERHR should indicate that this study only reports the composite data from factor analyses and does not report results (means/s.d.) of actual measures as has been reported in other studies by this group. It is very important to evaluate the results of the actual measurements made because the types of behaviors are grouped into factors based on statistical methods. Many of the actual measurements are closely related and it's important to have the data for the original measurements evaluated. There was no positive control in this study, so authors should not use terms like 'demasculinization' and 'defeminization' to describe these effects. The CERHR review of this paper makes an important point about the interpretation of changes in play measurement that gets to the heart of the issue for many of the behavioral studies of this kind. It is not appropriate to use terms like masculinization and feminization simply because of change in direction of few measurements of uncertain biological significance especially when there are no data presented to indicate that these behaviors are clearly demonstrated to be sexually dimorphic nor that the direction of change is mediated by an organizational effect on the brain that is consistent with an estrogenic mode of action. The critical nature of the review made on this study needs to be applied to others (i.e. Adriani and Farabollini) claiming that changes in behavior are evidence of masculinization or feminization of the brain.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Razzoli (2005)	425 (p 271)	Mongolian gerbils	**Postnatal** Days 12 of pairing at approx 13 wks old	Oral (in oil); micropipette	0.002, 0.02;	13 week old F, days 1–21 of pairing (cohabitation of male and female) Pair behavior (n=12), free exploratory test (n=12)	N.A.	Y	BPA increased social investigation and reduced exploratory parameters.	Well conducted and suggests that 0.002 mg/kg bw/day shows greater effect on behavior than 0.02 mg/kg bw/day of BPA. Plasma analysis of BPA and potential metabolites was not performed. No strong evidence of dose response	Limited utility	Limited utility
Additional considerations: This paper evaluates the effect of BPA on adult behavior. The free exploratory test was well conducted, using a blinded observer scoring videotapes using pre-defined observations. It was not clear if the Pairs' behavior was conducted without knowledge of treatment level. The biological significance of the increased "social investigation" and "reduced exploratory parameters" is not clear.												
Ryan (2006)	375 (p 204)	C57/Bl-6 mice	**Prenatal, postnatal** GD3 – PND21	Oral; gavage	0.002 0.2;	14–16 F from an unknown # of litters Anxiety-related behaviors (elevated plus maze; n=14, age >PND 42), spatial memory (radial-arm maze, Barnes maze; n=16 for both mazes, age >PND 42) Other endpoints not relevant to neurological assessment are not included in this list. One female per litter was randomly selected for behavioral testing but there are different numbers of females for different behavioral tests and paper does not report total number of pregnant females tested	Y	Y	Developmental exposure to EE was found to masculinize behavior in all of the assays used.	Good selection of established measurements of sexually dimorphic behaviors. Behavioral evaluations conducted only on ovariectomized females. Data was then interpreted with respect to established dimorphic pattern rather than concurrent assessments of performance in males or intact females.	Useful	Adequate.
Additional considerations: The methods section did not state if the behavioral measurements were measured by observers unaware of treatment level, or if time of testing was balanced across different treatment levels. We disagree with CERHR that all measurements are established measurements of sexually dimorphic behaviors and the lack of baseline behavior of untreated males and females for every behavior limits ability to interpret results of this study in terms of organizational effects on the brain. This should not be stated as a strength.												
Shikimi (2004)	339 (p.172)	Fisher rats	**Postnatal** PND 6-9	Injection into cerebral spinal fluid daily for 4 days	50 or 500 μg/25 μL sesame oil (not converted to mg/kg)	M and F pups 6–9 days old Purkinje cell dendrite growth (n=4, age=PND 6-9)	N	Y	A high dose of BPA may induce dendritic growth in Purkinje cells.	Good use of 17β-estradiol as a positive control. Injection into cerebrospinal fluid and expression of dose as mg/day prevent comparison with other studies.	Little utility	Inadequate
Additional considerations: The route of exposure (cerebrospinal fluid injection) makes this study inadequate for risk assessment purposes. In addition, it is unclear what the functional consequence is for changes in dendritic process length, or whether this is a short-term or long term change.												

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Suzuki (2003)	372 (p.202)	ddY mice	**Prenatal, postnatal** Mating to weaning	Oral; diet	**estimated 0.4, 100 and 400 mg/kg/day** **CERHR incorrectly calculates 0.0004, 0.1 or 0.4 mg/kg bw/day**	6-10 MALES in pharmacology studies, unknown # in bio-chemical studies, from an unknown # of litters Conditioned place preference (n=6–10 mice/group each BPA group divided into saline and meth group, age not reported), methamphetamine-induced activity (n=9–10 mice 400 mg/k BPA group only, age not reported).	N	N	BPA potentiated effects of methamphetamine on hyperlocomotion and place preference	Inadequate description of what was done in study	Not useful	Inadequate

Additional considerations:

A major deficiency of this study is lack of food consumption data to allow calculation of chemical consumption of BPA. A second deficiency is that the authors overstate the conclusions that can be made from this data. A limitation is that only one low dose and two very high doses were used for the behavior, and only the highest dose level (estimated 400 mg BPA/kg/day) was used for the neuropharmacologic work. The molecular/biochemical data is very limited in terms of usefulness for risk assessment purposes. **PLEASE NOTE THAT CERHR INCORRECTLY CALCULATES DOSES ON PAGE 202. The error is that CERHR incorrectly reports Suzuki's feed concentrations in mg/kg instead of mg/gram food as stated on page 640 of Suzuki et al., 2003. This makes a major difference in estimation of dose level and relevance to risk assessment.**

This study does not appear in the neurodevelopmental sections of the CERHR draft report but is discussed in other developmental sections. Authors' conclusions about BPA's effects on psychological dependence on psychostimulants are a clear overstatement. CERHR should state that the study does not have sufficient data to support this conclusion. The weakness of the paper is that the authors only evaluated neuropharmacologic effects at just the highest dose level and overstated conclusions that could be drawn from this study about psychological dependence on psychostimulants. This study was not designed to study psychological dependence. We also agree with the weaknesses described in the CERHR review. However, there are some strengths that are not mentioned: a) route of exposure is dietary; b) the authors provide dose response data for methamphetamine place preference to support selection of probe dose; c) the authors used 3 doses of BPA to study effects of BPA on methamphetamine place preference; b) demonstrated that methamphetamine place preference can be antagonized by a D1 receptor antagonist; and e) provided preliminary data on BPA effects on D1 receptor function.

Takagi (2004)	305 (p.139)	Sprague Dawley rats	**Prenatal, postnatal** GD15 – PND10	Oral; diet	5–8, 50–80, 230–380;	5–8 Males and Females from 5–6 litters Body weight, litter parameters, endocrine-linked organs, size of PDN-POA (n=5 at PND 21, n=8 for PN week 11)	Y	Y	BPA at high doses retarded offspring growth but did not affect the endocrine/reproductive endpoints of offspring, whereas EE did.	Good range of endpoints measured, use of 17β-estradiol comparator group and complete statistical evaluation. Weaknesses include small sample size of dams and offspring for endpoints. Better than average range of endpoints and included gross assessment of volume of SDN-POA.	Barely adequate	Limited utility

Additional considerations:

Agree with the rating of limited utility. In addition to the strengths and weaknesses stated, use of the litter as the experimental units is a strength. However, the doses are all high, leading to a rating of this study as limited for risk assessment purposes.

Reference	CERHR Ref. # and page #	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	CERHR Strengths/Weaknesses	CERHR Ranking	Proposed Revised Ranking
Tohei (2001)	434 (p 279)	Wistar-Imam-ichi rats	**Adult**	s.c.	0.1, 1 mg/rat/day body weights between 300-350 g = approx 0.03, 0.3 mg/kg/day	Hormones (LH, FSH, PRL, testosterone, progesterone, inhibin) (n=5, 10^{-1} dose); testicular response to hCG (n=5-6, 10^{-2}, 10^{-1} dose); LH-RH challenge (n=5-6, 10^{-1} dose), sexual behavior (n=5-6, 10^{-1} dose)	N,A,	N	Decreases in testosterone, and testicular inhibin were seen, along with increases in prolactin and LH. Plasma inhibin remained the same. Testicular response to hCG was decreased. The pituitary response to LH-RH, and male sexual behavior, did not change. These results suggest that BPA directly inhibits testicular functions and the increased level of plasma LH is probably due to a reduction in the negative feedback regulation by testosterone. The testis is probably a more sensitive site for BPA action than the HPA.	RIAs appear competently conducted. SC not relevant route of exposure. Sample size limited. Blood collection via decapitation not appropriate. No mention of killing: may be serious confounding of data depending on order of kills and if animals were not in separate rooms. Rat plasma testosterone levels are normally highly variable, the low degree of variability in this study, given the sample small size, is remarkable (~±0.12 ng.mL). No functional consequence of alterations in hormone levels described.	Minimal utility	Inadequate
Additional Considerations: The study design issues identified by CERHR together with the subcutaneous routes of exposure make this study inadequate. In addition, data on male sexual behavior is not presented, and there is no statement indicated that these evaluations were conducted blind.												
Zsarnovszky (2005)	340 (p.172)	Sprague Dawley rats	**Postnatal** PND 4-9	Stereotaxic injection into cerebellar folia 6 & 7	10^{-12} to 10^{-6} M, inclusive of all orders of magnitude (as given by authors)	4-9 M and F from unknown number of litters. Immunostaining for ERK+ cells in the cerebellum (n=4–9, various ages from PND 4-16, and adult (age undefined)	N	Y	BPA can act as a highly potent E2 mimic and can also disrupt the actions of E2 at very low concentration (inverted 'U' shaped curve).	Good use of 17β-estradiol as a positive control. Dose was not completely clear.	Useful	Inadequate
Additional considerations: The intracebellar route of exposure does not reproduce the first pass metabolism of an oral dose, or the adult metabolism of the dam for BPA exposure through the dam's milk. Also, the ERK+ staining indicates the presence of protein but does not indicate functionality or any kind of an integrated tissue response as should be present for relevance to a human risk assessment. The lack of pharmacokinetic data to relate these levels to oral exposure makes this study inadequate for risk assessment purposes.												

[a] **Utility ratings for neurodevelopment/behavioral endpoints**

High Utility: Wide dose range, including low doses. Large number of relevant endpoints. Sufficient number of animals per dose group. Large sample size. Appropriate statistical analysis (e.g., litter should be the experimental unit after prenatal exposure). Oral route of exposure.

Adequate: Some relevant endpoints and doses (at least 2 low doses of different order of magnitude). Sufficient number of animals per dose group. Appropriate statistical analysis. Oral route or exposure. Studies claiming feminization or masculinization should support their interpretation by organizational effect.

Limited Utility: Small dose range (e.g., within single order of magnitude), or single dose. No dose-response trend for reported effects, or inconsistent patter of effects. Non-oral route of exposure. Some important data not reported. Limited statistical analyses. Inappropriate study design. Studies in which the biological significance of endpoints and correlation to human health risk assessment is not clear.

Inadequate: Direct injection into the brain. No relevant endpoints. Number of animals per dose group too low to draw conclusions. Much important, relevant data not provided. No or inappropriate statistical analyses. Animals with co-exposures to other compounds.

Table 8

Polycarbonate/BPA Global Group Comments on CERHR Draft Report of December 2006

February 2, 2007

This table lists studies that have some functional or morphologic endpoint and were not included in the CERHR draft report. Ishido, 2005 and McClusky, 2005 did not meet this criteria but were included because Shikimi, 2004, a study with related endpoint, was already included in the CERHR table. This review and the utility ratings focus primarily on neurodevelopmental results reported in the study. Some of the studies were not found in the CERHR neurodevelopmental sections, but were evaluated elsewhere in the CERHR review. The rating criteria used for this table, defined at the end of this table, are adapted from the criteria used to rank studies for reproductive and developmental endpoints other than neurodevelopmental endpoints.

Reference	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	Rationale for Ratings and Recommended Strengths/Weaknesses	Recommended Rating[a]
Ishido (2005) (Current Topics in Pharm)	NB-1 Human neuroblast-oma tumor cell line	Immortalized cells in culture	In-vitro cell culture 24hrs post-plating treated with cadmium or bisphenol A. (does not give vehicle or volume)	See endpoints.	For Neurite growth: Cadmium: 0, 0.5, 1.0, 5.0 µM for 48 hours. Bisphenol A: 0, 1.0, 5.0, 50.0 µM for 48 hours. Replicate numbers of cultures not given. For Microarray: sham, cAMP 1mM, 17β-estradiol 5 µM, cadmium 5µM. Methylmurcury 0.5µM, and dioxin (no dose given) Replicates not given.	N/A	Y cAMP for one study	Exposure to cadmium and Bisphenol A facilitated neurite growth.	This study is inadequate for risk assessment purposes not only because the *in vitro* exposure cannot be related to the oral route of exposure, but because of the inadequate methods and analyses. Flaws in the study design and analysis, and the scant Methods section that omits important experimental details are critical weaknesses. The use of a cancer cell line rather than a primary culture makes the choice of NB-1 cells a poor model to extrapolate to normal cells, especially when access to primary culture techniques and support solutions are widely available. This paper was intended to be a demonstration of the use of specific techniques to address the question of biological effects caused by environmentally available chemicals, not an in depth study of any compound in particular.	Inadequate
Ishido (2005) (Reg Pep) * see below	Wistar rats	Postnatal PND5	Intracisternal injection to pup	2 (approx),	17 M for motor activity, unknown number for genetic assays, from 10 litters Motor activity (n=17, 4-5 weeks), DNA macroarrays (n=?, 4-8 weeks), TH+ immunohistochemistry (n=?, age=8 weeks)	N	N	BPA caused hyperactivity at 4-5 weeks of age, and decreased the gene expression of the dopamine transporter by 0.5-fold at 8 weeks of age.	This study is inadequate for risk assessment purposes because intracisternal injections were used. *see below for further evaluation for a series of studies of which this was a part. Strengths include good use of motor activity system, including 24-hour observations. Pharmacokinetic data is needed to determine relevance of this finding for oral dosing. This result should be evaluated against other studies evaluating motor activity following oral exposures. Weaknesses include a lack of information on control of confounders for motor activity such as randomization of testing across treatments, and no information on litter as an experimental unit. Also, the molecular data is not well described in the methods, and is over interpreted for the data derived. The use of intracisternal injection as the route of exposure makes comparison to oral exposure difficult because of the lack of information for dose extrapolation and the lack of metabolism of the parent BPA as occurs by the oral route of exposure.	Inadequate

Reference	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	Rationale for Ratings and Recommended Strengths/Weaknesses	Recommended Rating[a]
MacLusky (2005)	Sprague-Dawley rats	adult	s.c.	0.04, 0.12, 0.4	N=3 for each treatment Density of pyramidal cell dendritic synapses, uterine weight	N	Y	BPA may interfere with the development and expression of normal sex differences in cognitive function, via inhibition of estrogen-dependent hippocampal synapse formation	This study is of limited utility for risk assessment because of the subcutaneous route of exposure and the use of ovariectomized animals. This paper clearly describes the methods and results. The experiments are well conducted with attention to controlling for potential confounders. The sample size is small (n=3). Although the authors address this statistically using power calculations, this study would need to be repeated before firm conclusions can be made. Two preliminary conclusions can be made from this paper: (1) acute subcutaneous doses of BPA (40-400 μg/kg) reduced the number of hippocampal synaptic densities induced in 30 minutes by acute subcutaneous doses of estradiol (45-60 μg/kg) in ovariectomized young adult females; and (2) acute subcutaneous dose of 300 ug/kg BPA reduced hippocampal synaptic densities in similar animals untreated with estradiol. The relevance of these effects is unclear since the number of synapses in hormonally intact female rats varies by as much as 30% over a 5 day estrus cycle (Nimchinsky EA, Sabatini BL, Svoboda K. Structure and function of dendritic spines. *Annu Rev Physiol* 2002; 64:313-53). The authors examine only female rats, so their hypothesis that these observations indicate that BPA may interfere with normal sex differences is unsubstantiated. Also, these effects are seen under dynamic physiological conditions in which the 30 minute response of a hormonally deficient animal is evaluated after administration of the estrogen dose that showed the maximal response. This experimental paradigm is significantly different from hormonally intact humans. It should also be noted that BPA's *antagonistic* effects to estradiol in this paper appear to be inconsistent with the underlying hypothesis of their cited neurobehavioral studies that BPA disrupts sexual differentiation of the brain during development through it's estrogen *agonist* effects Note: This MacLuskey paper did not meet our criteria for inclusion into this list of additional studies because it did not have a morphologic or functional endpoint. However, this research is very similar to that of Shikimi et al (2004) which was included in the CERHR review. Shikimi et al (2004) injected BPA into cerebral spinal fluid in pups and examined Purkinje cell dendritic growth.	Limited utility
Masuo (2004a) (Reg Pep) *see below	Wistar rats	Postnatal PND5	Intracisternal injection to pups	87 nmol/10 μl	6 males for motor activity from unknown # of litters; 6-16 used for catecholamine analysis, unknown number used for macroarray. 2 rats each group pooled for striata, and one for midbrain per group for macroarray analysis. 3 experiments with 3 different groups. Motor activity (4-5 weeks), brain catecholamines (10 weeks), cDNA macroarray (8 weeks)	N	N	Gene-expression profiles showed variation after treatment with endocrine disruptors, and suggest an animal model of attention-deficit hyperactivity disorder.	This study is inadequate for risk assessment because of intracisternal route of exposure and weak molecular data. *see below for further evaluation for a series of studies of which this was a part. . Strengths include good use of motor activity system, including 24 hour observations. Weaknesses include a lack of information on control of confounders for motor activity such as randomization of testing across treatments, and no information on litter as an experimental unit. Also, the molecular data is not well described in the methods, and is over interpreted for the data derived. The use of intracisternal injection as the route of exposure makes comparison to oral exposure difficult because of the lack of information for dose extrapolation and the lack of metabolism of the parent BPA as occurs with the oral route of exposure. The observation that BPA may increase motor activity under admittedly non-physiological dose conditions has value, but needs to be compared with other oral studies evaluating motor activity	Inadequate
Masuo (2004b) (Neuro Plasticity) * see below	Wistar rats	Postnatal PND5	Intracisternal injection to pups	0.087, 0.87, 8.7, 87 nmol/10μl	6 males for motor activity, unknown # for genetic assays, from unknown # of litters Motor activity (n=6, 4-5 weeks)), catecholamine analysis (n=16 treated, 6 control, 10 weeks), cDNA macroarray 2 rats pooled for striata, I for mid brain, 3 experiments, 3 groups, 4-5 weeks)	N	N	Rats given BPA or 6-OHDA both displayed clinical signs of hyperactivity but changes in gene expression caused by BPA differed from 6-OHDA , indicating that hyperactivity is caused by different mechanisms.	This study is inadequate for risk assessment because of intracisternal route of exposure and weak molecular data.. *see below for further evaluation for a series of studies of which this was a part. . Strengths include good use of motor activity system, including 24-hour observations. Weaknesses include a lack of information on control of confounders for motor activity such as randomization of testing across treatments, and no information on litter as an experimental unit. Also, the molecular data is not well described in the methods, and is over interpreted for the data derived. The use of intracisternal injection as the route of exposure makes comparison to oral exposure difficult because of the lack of information for dose extrapolation and the lack of metabolism of the parent BPA as occurs with the oral route of exposure. The observation that BPA may increase motor activity under admittedly non-physiological conditions has value, but needs to be compared with other oral studies evaluating motor activity	Inadequate

Reference	Species/ Strain	Lifestage during dosing	Exposure Route	Dose Levels (mg/kg-day)	Sample Size, Endpoint(s)	Litter as Experimental Unit?	Positive Control?	Author's Conclusions	Rationale for Ratings and Recommended Strengths/Weaknesses	Recommended Rating[a]
Rubin (2006)	CD-1 mice	Prenatal, postnatal GD8-PND16	s.c. osmotic pump	0.000025, 0.00025;	14-17 dams were dosed but assignment of animals to different tests not clearly stated TH+ staining AVPV nucleus (n=7-8/sex; 2 pups/sex/litter, PND 22-24) - So litter is not the experimental unit. TH+ staining arcuate nucleus (n=4/sex; presumably 1 pup/sex/litter) open field (n=10-12/sex; 1 pup/sex/litter, 4 weeks of age; n=14-17; 1 pup/litter, 6-9 weeks of age) There is no explanation of how pups were selected from subset of litters for various endpoints.	Mixed	N	Control offspring had significant sex-related differences (TH+ neurons in the APV and in open field behavior) but differences were not observed in offspring of BPA-treated dams.	This study is inadequate because of the subcutaneous route of exposure and significant limitations in study design and methods. This study uses osmotic pumps to deliver a low continual dose subcutaneously rather than episodic oral dosing from daily feeding activity of the dams. A limitation is that there is no data presented that allows extrapolation from this route of exposure to expected human oral exposures. This study would have been rated limited utility except for substantial weaknesses in study design and methods. The major weakness of this study is that the methods are inadequately reported. The authors do not clearly state the number of dams tested and method of assignment of subsets used to test different endpoints. Based on information from figure legends and methods section, there were 14-17 litters/dose group tested for motor activity at 6-9 weeks of age, 10-12 litters/dose group at 4 weeks of age); and 4 litters/dose group for TH+ staining (2 pups/litter through AVPV; 1 pup/litter through arcuate nucleus). It appears that the prepubertal motor activity test was planned after the start of the study (page 16 of paper), which accounts for fewer animals tested on motor activity (the authors state that the earlier time point was decided later in the study which presumably means only later replicates were tested). Confidence in the behavioral tests is decreased by (a) inability of observers to measure a primary behavior they planned to measure – distance traveled - because animals were too active (this should have been determined in control animals prior to study rather than working out methods during the conduct of the study); (b) lack of information on balancing time of testing across observers and replicates, and (c) lack of information on whether observers were unaware of treatment level. The TH+ cell counts were conducted by 3 observers unaware of treatment level, but it is not clearly stated if the counting of cells for each dose group was balanced across the different observers. The sample size for TH+ staining is low, there was no counterstaining of neurons (i.e. Nissl stain) to confirm decrease in number of neurons; the number of sections through area is a very crude method of evaluating size of a nuclei. This study provides preliminary evidence that requires follow up with more rigorous morphometric or unbiased stereological methods and counterstaining of TH+cells with nissl based stains. It was somewhat surprising to see differences in activity measurements in males and females prior to puberty. There was no positive control included in this study. Dose-response seen for decrease in ratio of female to male TH+ neurons in the AVPV, and open field behaviors are different in treated vs. non-treated animals. This study lacks a positive control, although the authors cite references indicating that sex hormones drive the dimorphism. Rubin et al state that estrogen receptor α is critical in the decline of sex ratio. If BPA is acting through a non-estrogenic mechanism both sexes should be affected.	Inadequate
Sashihara (2001)	Julia type chicks	Postnatal PND 4	Intracerebral injection to chicks	0.1, 0.2	Male chicks Body weight (n=13 control, 12 low, 12 high, age=4, 5, 6 days), organ weight (n=10 control, 8 low, 9 high, age=20 days), jumping, distress vocalization, motor activity, duration of crouching (n=7 chicks, age=12 days: 8 days after injection)	NA	N	BPA may somewhat induce a behavioral change in a stressful situation.	This study is inadequate for risk assessment purposes because the endpoints are not clearly defined and difficult to relate to humans, the route of exposure is via direct injections into the brain, and the chicken is not a validated animal model for extrapolation to humans for such endpoints. The authors inject BPA directly into the brain, leaving the conversion to human oral intake questionable. The methods section presented very few details, although some information could be extracted from the results and discussion section. There was no clear operational definition for “distress vocalization”. Observations were not conducted blind to treatment level. There was no positive control data.	Inadequate
Sato (2001)	Jcl-ICR mice	Prenatal GD11-19	s.c. injection	100	12–62 pups from 7 litters, gender not specified Sensorimotor development (n=38 control, 17 EE, 51 BPA, age=as appropriate for development endpoint), open field (n=56 control, 37 EE, 62 BPA, age=PND 60), diameter of tractus mamillothalamicus (n=PND 40: 30 control, 19 EE, 12 BPA; PND 60: 51 control, 68 EE, 52 BPA, age=PND 40 and 60)	N	Y	Prenatal BPA and estradiol exposure impairs litter size, openfield behavior, and spermatogenesis.	Usefulness of the study is limited by the single high subcutaneous dose used, which is well beyond the doses encountered in the human population. Strengths are having a concurrent positive control, and a large group size in the open field test that allows detection of subtle differences (n=56-62). Anatomic measurements also have good group sizes (n=12-68). Weaknesses include s.c. dosing which lacks the first pass metabolism of oral dosing, and lack of controlling for litter as the experimental unit.	Limited utility

*(Ishido, Masuo) This group of researchers has published a series of studies in which they inject BPA directly into the brains of 5 day old rats, then measure motor activity in an infrared sensor box over 24 hours at 4-5 weeks of age. The route of exposure limits the applicability of these studies for human risk assessment since a mg/kg dose is difficult to approximate, and there is a complete lack of the first pass metabolism that occurs in oral dosing. Line graphs of the motor activity indicate that BPA-treated rats have greater activity over 24 hours, especially at the start of the 24 hour period, which is the beginning of their awake period (dark phase) combined with a novel environment. The strength of these studies is the consistency in which this result is seen over studies. However, the group size is small for activity testing, and researchers do not control for litter, so it is unknown how many litters are represented in each study or over the series of studies. Also, the researchers do not state whether testing of animals is randomized across treatment groups. The motor activity observation is combined with molecular studies on components of the dopaminergic system, under the hypothesis that movement disorders could be due to disruptions in that neurochemical system. On the basis of changes seen in DNA macroarrays, such as changes of 0.77 fold in one of the dopamine receptors (Ishido 2005), or of 0.33 fold in one galanin receptor (Masuo 2004), claims are made for deficits in the entire system. This is extrapolation beyond the data since macroarray data is notoriously variable and only semi-quantitative, and expression of genetic data should be correlated with protein levels for the observation to be physiologically meaningful. Also, if the changes seen were physiologically meaningful, changes in other proteins within the same and related neurochemical pathways should be seen and are not. Nonetheless, the overall conclusion made by the authors of this series of papers is that BPA may be a risk factor for attention-deficit hyperactivity disorder based on the two data points of increased activity and isolated alterations in macroarray findings for the dopamergic system after BPA injection into the brain. This conclusion must be treated only as a hypothesis because the connection between hyperactivity and variation in macroarray data seen after intracisternal injection requires proof of many intermediate steps before a connection to ADHD could be seen.

[a] **Utility ratings for neurodevelopment/behavioral endpoints**

High Utility: Wide dose range, including low doses. Large number of relevant endpoints. Sufficient number of animals per dose group. Large sample size. Appropriate statistical analysis (e.g., litter should be the experimental unit after prenatal exposure). Oral route of exposure.

Adequate: Some relevant endpoints and doses (at least 2 low doses of different order of magnitude). Sufficient number of animals per dose group. Appropriate statistical analysis. Oral route or exposure. Studies claiming feminization or masculinization should support their interpretation by organizational effect.

Limited Utility: Small dose range (e.g., within single order of magnitude), or single dose. No dose-response trend for reported effects, or inconsistent patter of effects. Non-oral route of exposure. Some important data not reported. Limited statistical analyses. Inappropriate study design. Studies in which the biological significance of endpoints and correlation to human health risk assessment is not clear.

Inadequate: Direct injection into the brain. No relevant endpoints. Number of animals per dose group too low to draw conclusions. Much important, relevant data not provided. No or inappropriate statistical analyses. Animals with co-exposures to other compounds.

Table 9

Polycarbonate/BPA Global Group Comments on
CERHR Draft Report of December 2006
Section 3 and 4 (Sub-mammalian Studies)

February 2, 2007

Page	Lines	Comment
216	24	A weakness of the Roepke et al. study is the lack of analytical confirmation of BPA concentration over the course of a 96 hour study in which BPA can substantially biodegrade.
216	45-46	A weakness of the Andersen et al. study is the method of exposure, which involves use of algae that had been exposed to BPA as a food source for the test species. No analytical confirmation of the amount of BPA sorbed by the algae was conducted and the authors have no idea of the actual exposures.
217	8	Concentrations of BPA in spiked sediment were NOT confirmed in this study. Only spiking solution concentrations were verified.
217	23-24	A weakness of the Watts et al. study is that no analyses were made of the spiked sediment exposure system to confirm concentrations of BPA, which is readily biodegradable. Consequently, the authors had no idea of the actual exposures.
218	34-35	The Iwamuro et al. study is of no utility for the CERHR evaluation since effects were reported only at high concentrations in the ppm range. These levels are not relevant environmentally or for human health.
219	17	The Oka et al. study is of no utility for the CERHR evaluation since effects were reported only at high concentrations in the ppm range. These levels are not relevant environmentally or for human health.

220	7	The Sone et al. study is of no utility for the CERHR evaluation since effects were reported only at high concentrations in the ppm range. These levels are not relevant environmentally or for human health.
221	9-12	A weakness of the Levy et al. study is that the study is statistically flawed due to the pooling of two replicates into one, giving only a single replicate. The study is of no utility for the CERHR evaluation.
222-223	50-51 2-3	The histology endpoint with a reported LOEL of 10 µg/L was not dose-related. The dose-related NOEL for this study based on growth and sex ratio, which are population relevant parameters, is 200 µg/L (120 µg/L mean percent nominal concentration). These levels are not relevant environmentally or for human health and the study is of no utility for the CERHR evaluation.
223	22-25	The histology endpoint with a reported LOEC of 13 µg/L was not dose-related. The dose-related NOEC for this study based on growth and sex ratio, which are population relevant parameters, is 355 µg/L. This level is not relevant environmentally or for human health and the study is of no utility for the CERHR evaluation.
272	42-45	Fundamental weaknesses of the Oehlmann et al. study are the lack of replication and the improper statistical evaluation. It is not possible to determine statistically significant effects with only a single replicate.
294	19-21	Fundamental weaknesses of the Oehlmann et al. study are the lack of replication and the improper statistical evaluation. It is not possible to determine statistically significant effects with only a single replicate.