ANNEX A: CLH PROPOSAL CONTAINS INCONSISTENT AND INCORRECT INFORMATION ON MALE AND FEMALE ENDPOINTS

The CLH proposal contains inconsistent and incorrect information on male and female endpoints. Based on the following reasons a classification of BPA as a Category 1B Reproductive Toxicant is unwarranted:

- The studies cited in the CLH proposal are largely non-guideline studies and their findings are not corroborated by the negative findings reported in more robust studies.
- Many of the studies employ non-validated, experimental methods with insufficient methodological information to examine the reported observations.
- A previous review of many studies cited in the CLH proposal, conducted e.g. by the National Institutes of Environmental Health Sciences' "Center for the Evaluation of Risks to Human Reproduction" (CERHR) in 2008, has found deficiencies in studies cited in the CLH proposal.
- The majority of the studies employed non-relevant, non-oral routes of exposure, e.g., subcutaneous, intravenous/intramuscular, slow release implants of BPA.
- Many of the studies fail to link observations to actual apical adverse outcomes that would impact reproduction and development.
- The CLH proposal contains inconsistent and incorrect information on individual studies.

More detailed information on experiments in male and female animals discussed in the CLH proposal and the conclusions drawn in the CLH proposal are discussed below:

Comments on male reproductive evidence in animals

Summary

The CLH proposal contains inconsistent and incorrect information. Overall, for the following reasons, a re-classification of BPA is unwarranted based on the available data on male reproductive endpoints in animal studies**:**

- The cited non-guideline studies relied upon on in the report are not corroborated by the negative findings reported in robust multigeneration guideline studies (e.g, Tyl et al. 2008 two-generation mouse study and the Tyl et al., 2002 three-generation rat study). Importantly, guideline studies are required and relied upon for registering substances in Europe under REACH and are also an important aspect for classification.
- Many of the studies employ non-validated, experimental methods with insufficient methodological information to examine the reported results.

- A previous expert panel review of many studies cited in the CLH proposal, conducted by the National Institutes of Environmental Health Sciences' "Center for the Evaluation of Risks to Human Reproduction" (CERHR) in 2008, has found deficiencies in these studies.
- Many of the studies fail to link observations, even if they are assumed to be valid observations, to actual apical adverse outcomes that would impact reproduction and development. According to GHS criteria for 1B Classification: *"Such data shall provide clear evidence of an adverse effect on sexual function and fertility or on development."* It should be noted that a large number of the studies cited in the CLH proposal provide no measure of functional impact on reproduction and development.

Finally, we note that the CLH proposal never concluded that the male reproductive effects were "Proven Effects" as was done for three of the female reproductive endpoints. Thus, there is some confusion as to whether or not the CLH proposal is really relying upon the male animal endpoints to support the more strict reproductive classification level.

1. Window of exposure and overall testing strategy

The CLH proposal presents data by window of exposure and type of effects, with exclusion from multigeneration studies from the sections on smaller windows of exposure. The scientific rationale for such a differentiation of intervals and exclusion of multigeneration studies from the assessment of defined intervals within the dosing scheme of the multigeneration studies is not given. In that context it is important to recognize that according to the logic of the OECD test guidelines as well as the globally used testing protocols and testing strategies it is assumed that a study result is relevant for the interval of dosing that is covers. This means that if a study with a large dosing interval reveals an adverse effect it might be worthwhile to further investigate the specific interval in which an adverse effect was induced, but where a study with a large dosing interval does not reveal an adverse effect there is no direct need to examine specific intervals or to exclude the large interval studies from the assessment of a specific interval.
Therefore the assessment of the respective endpoints or windows of exposure should include all available data, including the multigeneration reproductive toxicity studies conducted according OECD and EU test guidelines.

2. Inconsistency of reporting and quality of cited studies

In addition to the general remark on window of exposure the allocation of studies to the specific intervals is <u>not consistent</u> within the CLH proposal. Here an example:

On page 83 the CLH proposal indicates that "*The data are presented by window of exposure and by type of effect."* and continues: *"4.11.2.1.1 In utero and lactation exposure; Several researchers have studied BPA effects following an exposure during gestation and/or lactation period. These studies have been realized either in mice or rats with various routes of exposure. Six of these studies reported effects (Kabuto et al., 2004; Okada et Kai, 2008; Watanabe et al., 2003; Iida et al., 2002; Akingbemi et al., 2004; Timms at al., 2005), two reported limited effects (Kobayashi et al., 2002; Tinwell et al., 2002) and two no effects at all (Howdeshell et al., 2008, when control received phytoestrogens; LaRocca et al., 2011)."*

The studies mentioned cover premating, mating, gestation and lactation (e.g. Kabuto et al. 2004 or Okada and Kai 2008 covers the whole period from before mating to lactation). There are additional studies not

mentioned that also cover this exposure period, including e.g. Tyl et al. 2002, Tyl et al. 2008, Ema et al. 2001, Kobayashi et al. 2010, Stump et al. 2010. The statement that only 2 studies showed no effect is not correct.

The quality of some of the studies cited in the CLH proposal are not discussed throughout the report. As an example, no detailed evaluation was performed that acknowledged opinions from other expert assessments as to study validity and applicability for a comprehensive weight of evidence assessment, especially from the CERHR [NTP´s Center For The Evaluation of Risks To Human Reproduction] (Chapin et al 2008, CERHR 2008). Limitations of individual studies and inconsistencies within the CLH proposal concerning individual studies cited are discussed below:

2.1 In utero/lactation exposure (*CLH proposal chapter 4.11.2.1.1*)

Iida et al. 2002:

This study is inadequate for purposes of establishing an apical adverse outcome effect of reproductive toxicity. For example, the number of animals is not included in table 16 on page 99. As mentioned on page 83, offspring of 3 dams were examined in each group. The statistical unit for in utero experiments should be the litter. Therefore this study is insufficient according to the criteria defined in the ANSES interim report 2011 based on the criteria N ≤ 5.

CERHR evaluated this study in 2008 and concluded that this study is inadequate: *"Strengths/Weaknesses: The oral route of delivery is a strength of this study. The lack of information on details of husbandry, the small sample size (4–5 male mice from 3 litters per dose group) and the lack of adjustment for litter effects, inadequate methods for histopathological preservation and evaluation (i.e., use of paraformaldehyde for paraffin embedding) are weaknesses. Utility (Adequacy) for CERHR Evaluation Process: This study is inadequate for the evaluation process based on methodology."*

Overall, the study by Iida et al. 2002 is an inadequate study and no conclusion should be drawn on this study.

Howdeshell et al. 2008

The CLH proposal stated that the study did not demonstrate any changes in various reproductive parameters, but it is also stated on page 68 that *"it should be noted that in this study the rats' food contained some phytoestrogens on the contrary of most of the other cited studies."*

This statement is incorrect. According to the CERHR, the following information is available for the other studies that showed effects:

- Iida et al., 2002 – no information was provided about composition of feed, caging, or bedding
- Akinbemi et al., 2004 – rats were fed Purina chow, which contains soybean meal, and given drinking water in polycarbonate bottles
- Watanabe et al., 2003 – rats were fed standard laboratory feed. No information was provided about caging or bedding materials
- Kabuto et al., 2004 – ICR mice were fed standard laboratory chow containing 24% protein. No information was provided about caging or bedding materials

- Timms et al., 2005 – CD-1 mice were fed soy-based Purina 5008 chow, provided drinking water in glass bottles, and housed in polypropylene cages
- Okada and Kai 2008 – did not indicate the phytoestrogen content of the diet

The Howdeshell et al. study was conducted at a U.S.EPA Laboratory in Research Triangle Park, NC. Long Evans hooded rat dams were gavaged with corn oil vehicle, EE (0.05–50 µg/kg/day) or BPA (2, 20, and 200 µg/kg/day) during pregnancy through lactation from gestational day 7 to postnatal day (PND) 18. Male offspring were euthanized beginning at PND150, and sera and organs were collected for analyses. EE exposure ≥ 5 µg /kg/day resulted in developmental effects (for example, seminal vesicle and paired testes weights are reduced, epididymal sperm counts were also significantly decreased). These results demonstrate that developmental exposure to oral micromolar doses of EE can permanently disrupt the reproductive tract of the male rat. In contrast, treatment with 2, 20, or 200 µg BPA/kg/day did not significantly affect any male endpoint in the current study.

Overall, the statement that the diet used in this study contained phytoestrogens *"contrary to most of the other cited studies"* is incorrect.

Okada and Kai 2008:

Silicone tubes containing E2 (10 ng – 10 µg) or BPA (100 µg or 5 mg) were implanted into female ICR mice before mating. Tubes were kept into mice during mating, gestation and lactation, consequently the time of dosing is similar to the dosing in the generation studies and no rationale is given why this study is reported under *"in utero and lactation"* exposure while other studies that dose over the same time interval are not mentioned in this chapter.

Authors indicate *"During the period…BPA [was released] at 1.2 or 60 µg/day."* Since the body weight of female mice at the end of the experiment is given (approx. 26 g) the parental dose of BPA in this experiment might be approx. 50 µg/kg/day and 1500 µg/kg/day based on the assumptions reported by the authors. According to the authors the number of animals used in the experiment is 10 offspring from 4 litters in the low-dose group and 19 offspring from 5 litters in the high does group. Since the litter is the correct unit for statistics in this kind of experiment, this study is insufficient according to the criteria defined in the ANSES interim report 2011 based on the criteria N ≤ 5. In addition the authors give insufficient information to evaluate the validity of its methods and interpretation. Furthermore, the lack of raw data in the publication is problematic because they indicate in the statistical analysis that data were selectively removed from consideration: *"The data were expressed as mean +/- SEM after the elimination of some data among experimental groups using Smirnoff's elimination test."*

This non-guideline study cannot be relied upon for evaluating classification since an apical, adverse outcome, e.g., reduced sperm numbers culminating in reduced fertility, was not evaluated. The fact that the elongate spermatids were evaluated right around the onset time of puberty, introduces variability due to the natural variability in puberty onset. Therefore, the slight difference in the high BPA dose relative to controls (40.3% versus 53.8%) can be confounded by natural variability and not the BPA treatment. Further, the 40.3 versus 53.8% difference may not be statistically significant if evaluated on a litter basis. The same can be said for testosterone levels that will fluctuate as the male mice undergo puberty during this time. That is, the apparent differences in serum testosterone could reflect variability in Leydig cell function that is changing at the time of puberty and not as a function of BPA treatment. Finally, the non-relevant route of exposure eliminates this study as a meaningful study for assessing the reproductive classification category for BPA.

Overall, based on the incorrect statistics (unit should be litter), the low number of litter used (4 to 5) and the lack of transparency "*elimination of some data"* by the authors, Okada and Kai 2008 is an inadequate study and no conclusion should be drawn from this study.

Akingbemi et al*.* 2004:

Table 16 on page 100 mentions only the doses 0-2.4 µg/kg/day. This is not correct because the post natal study investigated 0, 2.4, 10 µg/kg and 100, 200 mg/kg/day BPA. Observations are only seen at the lowest dose investigated (2.4 µg/kg/day). No effects were reported in the other dose groups. This fact raises questions about the absence of a dose-response relationship for the findings and whether or not the reported, low-dose response, is BPA-related in a biologically plausible manner. This study is not reliable because experiment #1 is testing testicular endpoints in sexually immature rats (PND 36) rendering the results from this lifestage highly questionable and uninterpretable. In fact, the authors themselves note this limitation of testing BPA's effects in sexually immature stages: "*We used adult Leydig cells in these experiments because of the inconsistent sensitivity of immature Leydig cells to ER agonists in vitro*." The follow-up studies, where PND 90 rats were utilized, are all based on a single dose (2.4 ug/kg/day) given from GD21-PND21 or PND21-90. Thus, none of the findings from experiments using mature testes (Exp #2 and #3) are interpretable since only one dose was used. Experiments 4 and 5 are in-vitro studies in which only one BPA concentration (0.01 nM) apparently showed a suppression of T production in mature Leydig cell cultures, eg., alleged non-monotonic behavior. None of the higher concentrations showed increasing suppression of T production.

CERHR evaluated this study in 2008 and concluded that some experiments in this study are limited and one experiment is inappropriate: *"Weaknesses include an inadequate number of animals to obtain confidence about the hormonal changes (indeed, LH was decreased in the first experiment and increased in the third), the lack of histopathology evaluation, and lack of an estrogenic positive control. Utility (Adequacy) for CERHR Evaluation Process: Experiments 1 and 3 are adequate but of limited utility because of the mechanistic nature of the endpoints examined. Experiment 2 is inadequate for consideration due to inappropriate statistics that failed to account for litter effects."*

The CLH proposal did not include the study by Howdeshell et al*.* 2008 in table 16 on page 99 although the study used a very similar study design to Akingbemi et al. 2004. The Howdeshell et al. 2008 study employed the same rat strain (Long Evans) dams that were gavaged with corn oil vehicle, EE (0.05 to 50 µg/kg/day) or BPA (2, 20 and 200 µg/kg/day) during pregnancy and through lactation (from GD 7 to PND 18). Male offspring were investigated for multiple endpoints including all the endpoints measured by Akingbemi et al. (body weight, organ weights including testes, prostate, seminal vesicle as well as serum testosterone and LH level). Howdeshell et al. 2008 concluded that treatment with BPA *"did not significantly affect any male endpoint in the current study."*

Overall, the publication by Akingbemi et al. contains limited and inadequate experiments and should be evaluated in the context of a recent comprehensive study that could not reproduce the reported data with a very similar exposure design using the same strain of rats, substantially more animals and more endpoints.

Watanabe et al. 2003:

The study utilized SD rats. Importantly, results from the same rats evaluated in Watanabe et al. (2003) were reported in different publications (Kobayashi et al. 2002, Koybayashi et al. 2005, Watanabe et al. 2003). However, only the research by Watanabe et al. 2003 was included in the CLH proposal.

CERHR evaluated all three studies together and concluded: *"Sample sizes are too small to reliably judge postnatal endpoints. Utility (Adequacy) for CERHR Evaluation Process: This study is inadequate for the evaluation of bisphenol A effects on postnatal outcome."*

Overall, the conclusion drawn in the CLH proposal concerning statistically significant and insignificant hormone concentrations is not supported due to the statistical limitations on this study evaluated by CERHR. In addition the judgment of the strain sensitivity is inconsistent within the CLH dossier. Whereas in other sections of the report the SD strain is discussed as being insensitive sensitivity is not discussed in relation to this study.

Kabuto et al. 2004:

Mice were dosed from 1 week before mating through gestation and lactation; this dosing protocol is similar to the dosing regimen utilized in the multigeneration studies. No rationale is given why this study is reported under *"in utero and lactation"* exposure but other studies that dose over the same time interval are not mentioned in this chapter.

This study is irrelevant for establishing evidence of adverse effects on reproduction. The study only evaluated various endpoints of oxidative stress in mice pups exposed in-utero and during lactation with dams ingesting drinking water containing either 5 or 10 ug/ml. The oxidative stress endpoints included TBARS, superoxide dismustase, catalase and glutathione peroxidase activity, and GSH and LH concentrations in the brain, liver, kidney and testis of 28-day old mice. Thus, this study provided no direct evidence of an adverse apical outcome on reproduction.

CERHR evaluated this study in 2008: *"Weaknesses include small sample size of exposed dams (n = 6), inappropriate use of the pup as the experimental unit in statistics, and mechanistic data without functional correlates. Utility (Adequacy) for CERHR Evaluation Process: This study is inadequate for the evaluation process due to inappropriate statistical procedures and small sample size."*

Overall, in this study mice are dosed from before mating to lactation similar to the generation studies. Based on the small number of animals, the inappropriate statistics and lack of reproductive toxicity relevance, this study should be considered as inappropriate and no conclusion should be drawn on this study

Salinian et al. 2009:

This study is included in chapter *"Transgenerational Exposure"* 4.11.2.4.7 although animals were exposed from GD12 – PND21 only, consequently, this study should be considered to be *"in utero and lactation"* exposure.

Kobayashi et al. 2012:

The CLH proposal writes in Chapter 4.11.2 *"Effects on male reproductive tract": "However, in 5-week old females, there was a significant reduction in AGD and AGD index with exposure to 3.3 and 33 ppm of BPA."* In males AGD was reduced only at 33 ppm (control: 3.87+/- 0.27 mm; 33 ppm BPA group: 4.00 +/- 0.24 mm). There is no significant effect in males on AGD index (mm/ g body weight) in this study and the authors indicate that the AGD observation "*were not considered to be biologically significant*". Since a reduction was observed in males and females this interpretation by the authors seems to be the most likely. This study is not included in table 16.

Overall, no biologically significant on AGD or AGD-index was observed in this study in males or females.

DeCatanzaro et al. 2013:

The CLH proposal included information from the abstract of the publication but did not evaluate this publication. Three experiments are reported. Mice were orally dosed from GD10 to PND9 with different doses of BPA and different diets; Experiments 1 and 2: 0, 0.147, 1.75, or 17.5 µg BPA/g of peanut butter in combination with a high and low-phytoestrogen diet, respectively; experiment 3: 0, 17.5, 175, or 1750 µg/g peanut butter with a high phytoestrogen diet only. No effects on gestational length, litter size, sex and AGD are reported in any experiment. No effect on preputial gland, testes weight, or epididymides weights were reported in any experiment. Reduced vesicular-coagulating gland was observed only in experiment 1 in the high phytoestrogen diet group dosed with 17.5 µg BPA per gram of peanut butter. This effect was not reproduced in the second experiment at any dose (including the 17.5 µg BPA per gram of peanut butter; high phytoestrogen diet). Consequently, the initial findings on the vesicular-coagulating gland were not reproduced in the follow up study and should be considered as not reliable for the evaluation process. The authors report observations on "latency to insemination" in experiment 2 (parameter not measured in experiment 3) and number of "intromissions" (17.5 and 175 µg group; no observation at 1750µg) and "ejaculations" (17.5 µg group only, not 175 or 1750 µg) in experiment 3 (parameter not measured in experiment 2). Since these endpoints were measured only in one experiment and not investigated in the follow up / initial experiment, the data should be regarded as not reproduced.

Overall, initial findings by DeCatanzaro et al. 2013 on the vesicular-coagulating gland were not reproduced in the follow up study and should be considered as not reliable for the evaluation process. Other endpoints were measured only in one experiment and not investigated in the follow up / initial experiment; therefore, the data should be regarded as not reproduced.

2.2 Neonatal exposure *(CLH proposal chapter 4.11.2.1.2)*

Salian et al. 2009b:

The authors used an exploratory test system with a non-relevant route of exposure. They dosed rat offspring with high BPA doses (0.1 – 1.6 mg/kg/day) via s.c. injection from PND 1 to 5 and report only a limited number of parameters. Since no data on systemic effects e.g. body weight, kidney or liver toxicity, is given in the publication the data reported in this exploratory study concerning the endpoint reproduction cannot be interpreted. In addition the authors mention "*An additional control group that received sesame oil without ethanol was also incorporated into the present study*" but do not report any data on the additional control group.

Overall, based on the severe limitations in documentation the Salian et al. 2009b study should be regarded as invalid without further information.

Aikawa et al. 2004:

According to Figure 1 in the publication the N is 5-6 for incidence of abnormal sperm and 5-14 for sperm motility. No explanation if given for the discrepancy in the number of animals for which the sperm morphology and motility results should align.

CERHR evaluated all three studies together and concluded: *"This study provided follow-up information to that of Nakahashi et al. (2001). The use of 17b-estradiol as a positive control in the testis histology study is a strength. Weaknesses include subcutaneous route of administration, lack of clarity on exposure issues, small sample sizes, lack of husbandry and statistical information. Utility (Adequacy) for CERHR Evaluation Process: This study is inadequate and not useful based on small sample sizes and inadequate presentation of statistical methods of analysis."*

Overall, this study is inadequate considering the CERHR criteria and insufficient according to the criteria defined in the ANSES interim, report 2011 based on the criteria N ≤ 5.

Toyama and Yuasa 2004b:

CERHR evaluated all three studies together and concluded: *"…[It does not appear that any statistical analyses were conducted.] …Weaknesses include small sample size, unclear data analyses, and s.c. route of administration. Utility (Adequacy) for CERHR Evaluation Process: This study is inadequate and not useful due to critically small sample size, route of administration, lack of clarity of design, and inappropriate statistical procedures."*

Akingbemi et al. 2003:

This study is included in table 16 on page 100, but not included in chapter 4.11.2.1.2 Neonatal exposure. Table 16 does not mention that no effects are observed at 10 µg/kg, 100 mg/kg and 200 mg/kg. As mentioned above CERHR concluded on this publication "*Experiments 1 and 3 are adequate but of limited utility because of the mechanistic nature of the endpoints examined. Experiment 2 is inadequate for consideration due to inappropriate statistics that failed to account for litter effects."*

2.3 Prepubertal exposure *(CLH proposal chapter 4.11.2.1.3)*

The chapter starts with the statement "*Few studies investigated the effects of an exposure to BPA on male reproductive functions at doses comparable to human exposure*." This statement is not supported by data in this chapter, because all studies mentioned in this chapter are high dose studies in the mg/kg range (Akingbemi et al. 2004, Nakamura et al. 2010, Tan et al. 2003, Takahashi and Oishi 2003, Norazit 2012) or in the µg/kg range (Della Seta et al. 2006, Akingbemi et al. 2004) and used oral dosing, s.c or i.p. injection. That is, none of the studies in the CLH proposal investigated exposures relevant to human exposures.

Akingbemi et al. 2004:

As mentioned above CERHR concluded on this publication "*Experiments 1 and 3 are adequate but of limited utility because of the mechanistic nature of the endpoints examined. Experiment 2 is inadequate for consideration due to inappropriate statistics that failed to account for litter effects.*"

Nakamura et al. 2010:

The CLH proposal mentioned that body weight was reduced at 100 and 200 mg/kg. This is incorrect based on the data reported in Table 2 of the publication, the following body weight data are reported: control: 352+/-24.4 g, 20 mg/kg: 355+/-26.7 g, 100 mg: 325+/-20.8 g, 200 mg/kg: 309+/-11.2 g. Since reproductive organ weights are reduced only in the high dose group (and mainly absolute weights, and to a lesser extent relative weights are affected) it should be considered, that effects on reproductive organ weights are seen at BPA doses where reduction of body weight > 10 % occurred. In addition the number of animals investigated for hormone measurement is 5-9 according to Figure 1 and high variability might be anticipated according to the error bars indicated.

Overall, s.c. dosing of 200 mg/kg is a very high dose for this route of exposure and at this dose reduced body weight is observed. The reported observations on the reproductive organs are seen at doses clearly exceeding the maximal-tolerated-dose (MTD). The number of animals investigated for the hormone measurements is inadequate.

Tan et al. 2003:

The CLH proposal mentioned that no effect on body weight was observed, but the effects on systemic organs observed are not mentioned: kidney weight increased cortical thickness, hydronephrosis and multinucleated giant cells in seminiferous tubules; liver weight decrease. In addition increased thyroid weight is also reported in this study.

Overall, as already evaluated by EFSA (2006) observations on systemic and reproductive organs were reported in this study in which only one dose was investigated.

Takahashi and Oishi 2003:

EFSA 2006: evaluated this study and concluded that *"no overt signs of general or reproductive toxicity were observed"*.

CERHR evaluated this study and concluded: *"Strengths include multiple routes of exposure, use of two strains of mice and rats, and a comparison of the oral, i.p., and subcutaneous routes. Weaknesses include use of single high doses administered for different durations across groups using minimal sample sizes. Utility (Adequacy) for CERHR Evaluation Process: This study is adequate but of limited utility."*

Norazit et al. 2012:

The CLH proposal summarized the abstract of the paper but did not evaluate this paper. The authors report no observation in the BPA group on body weight, testes weight or estradiol serum level. There is only limited information on histopathology available. *"Both the soya extract treated and BPA treated groups showed cellular debris in the seminiferous tubules and sloughing of the germ cells, with the BPA*

treated group showing a higher degree of damage (Figure 1(b) and (c)). The tubules were also vacuolated with lipids. The presence of the vacuoles and the cellular debris resulted in the absence of a lumen. Both groups also showed absence of maturing spermatid. Only the soya extract treated group showed a disturbance in spermiogenesis. The soya extract group did not show Leydig cell hyperplasia as reported by previous studies (Sharpe et al. 2002)." There are no additional data (e.g. severity or incidence of these findings) given in the paper to judge the value of this observation. The observation on the reduced E2 level should be discussed in the context of the low number of animals used (N=6) and the high variability in the hormone level measurement based on the high SD reported by authors in Table 3 (control: 122.44 ± 90.37; BPA group: 12.35 ± 13.88). In addition the authors used only one dose level of BPA.

Overall, the study by Norazit et al. 2012 is limited based on the low number of animals used, the high variability in the data reported and the limited documentation which do not allow full evaluation of the reported observations.

2.4 Adult exposure *(CLH proposal chapter 4.11.2.1.4)*

The chapter starts with the statement: "*All the studies published since 2002 describing effect of BPA exposure during adulthood on male fertility are positives.*" This statement is incorrect the majority of studies published did not indicate positive effects on specific male fertility endpoints. E.g. Yamasaki et al 2002 (OECD TG 407); Tyl et al. 2002 and 2008 (F0 animals are exposed during adulthood).

Chitra et al. 2003:

CERHR evaluated this study and concluded: *"Strengths/Weaknesses: Strengths include the use of oral and low multiple doses and appropriate measures. A weakness includes the marginal sample size. Utility (Adequacy) for CERHR Evaluation Process: This study is adequate for inclusion but of limited utility based on small group size."*

Herath et al. 2004:

CERHR evaluated this study and concluded: *"Strengths/Weaknesses: This study appears to have been relatively well conducted. A major weakness of this study is the inconsistency in the hormone data (control data after 2 weeks were dramatically different than after 5 weeks even though both are from sexually mature rats). The subcutaneous route of administration with the use of DMSO as vehicle are weaknesses."Utility (Adequacy) for CERHR Evaluation Process: This study is inadequate and not useful for the evaluation process primarily due to the significant inconsistencies in the hormone data from control animals."*

Sakaue et al. 2001:

CERHR evaluated this study and concluded: *"A weakness is the brief exposure period. Variability in control daily sperm production between the first and second study is disturbing; given the small sample (5 or 8/group), this variability severely decreases confidence in the data. No histopathologic correlate was presented. Utility (Adequacy) for CERHR Evaluation Process: This study is adequate but of limited utility due to small sample and variable control values between experiments."*

Although ANSES included the Sakaue et al. 2001 study in this chapter, the CLH proposal fails to report a follow up study by Ashby et al. 2003. That study attempted to replicate earlier findings from Sakaue et al. 2001. Five independent experiments were conducted; some of the experiments used the same conditions as in the earlier study by Sakaue et al. In this comprehensive follow up study BPA failed to produce any evidence of changes in DSP as reported by Sakaue et al. in 2001.

CERHR evaluated this follow up study and concluded: *"Given the robustness and comprehensiveness of this study, it is highly useful. It strongly suggests that the NOAEL for potential bisphenol A-mediated effects on the adult rat reproductive system exceeds 200 mg/kg/ day. Absence of confirmation of the work of Sakaue et al. (2001) led to an extensive study of the potential variables (e.g., diet, housing, etc.) that might account for the discrepancies. These data suggests that subtle changes in study endpoints, especially daily sperm production and organ weights, may occur by random chance or genetic differences in the respective lab's supplier of rats may play a role. These data also strongly suggest bisphenol A administered orally has no effect on sperm production albeit following only 6 days of administration. Strengths/Weaknesses: This study reports a well conducted, comprehensive assessment of the potential effects of bisphenol A delivered by 6 daily doses on daily sperm production. The 6-day treatment period is a (understandable) weakness. Utility (Adequacy) for CERHR Evaluation Process: This study is adequate and useful for the evaluation process."*

Overall, the CLH proposal included the Sakaue et al. 2001 study but did not take into account a follow up study by Ashby et al. 2003 in which the observations could not be reproduced. Based on the evaluations of both studies CERHR concluded: "*These data also strongly suggest bisphenol A administered orally has no effect on sperm production albeit following only 6 days of administration*"

Toyama et al. 2004:

CERHR evaluated this study and concluded: *"Strengths/Weaknesses: Definite conclusions cannot be drawn from such a limited data set; the fertility assessment was not meaningful due to the sample size (2/group). The background incidence of the electron microscope findings was not discussed. Another weakness is the subcutaneous route with DMSO as a vehicle. Utility (Adequacy) for CERHR Evaluation Process: This study is inadequate and not useful in the evaluation."*

2.5 Multigeneration exposure *(CLH proposal chapter 4.11.2.1.6)*

The CLH proposal cited four studies for Multi-Generation Exposure (Ema et al. 2001, Tyl et al. 2008, Tyl et al. 2002, NTP 1985b). The studies by Tyl et al. 2002 and NTP 1985b were the basis for the decision on classification in 2002. BPA was classified as reproductive toxicity category 3 for fertility in 2002 (Directive 67/548/EEC; R 62; under CLP it translates to category (H361f)). For the overall assessment of these studies see Annex D.

Tyl et al. 2008

This 2-generation reproductive toxicity study in mice study was performed to resolve uncertainties as identified in the EU Risk Assessment in 2003. The study design and interpretation of the results were supervised by a Steering Group that was chaired by a representative of the European Chemicals Bureau and included experts from several EU Member States. This study confirmed that BPA is not a selective

reproductive toxicant. Neither the EU steering group on the 2-generation study nor the risk assessment report indicates that the new information has an impact on the classification previously defined in 2002.

The CLH proposal cites a critique from Myers et al. 2009 on the Tyl et al. 2008 study. The information as given is not complete and inconsistent with the overall assessment of that study.

For example, the EU Risk Assessment Report from 2008 reads: *"The effects of BPA on fertility and reproductive performance in mice have been investigated in a two-generation study, conducted in compliance with GLP (Tyl et al, 2007). The study design and interpretation of the results were supervised by a Steering Group, that was chaired by a representative of the European Chemicals Bureau and included experts from several EU Member States. The overall design of this study was based on OECD Test Guideline 416, enhanced by incorporation of a second vehicle control group, a positive control group, a total of 6 exposure levels of BPA, the retention of additional F1 male offspring for organ weight and other assessments, and extending histopathological examinations to all treatment groups. This study was conducted in response to an ESR risk assessment conclusion that further research is needed to resolve the uncertainties surrounding the potential for BPA to produce adverse effects on development of the male reproductive tract at low doses (0.002- 0.05 mg/kg/day) in mice.…As we consider this investigation by Tyl et al. (2007) as the gold-standard, definitive study of the reproductive toxicity of BPA (for the endpoints examined),"*

The CLH proposal did not mention on page 94 or table 16 that systemic effects were observed at 50 mg/kg (liver histopathology) and 600 mg/kg (significant effects on body weight, kidney and liver). Therefore, BPA should not be considered a selective reproductive or developmental toxicant based on reproductive effects observed at the highest dose 600 mg/kg (Tyl et al. 2008).

On page 95 the CLH proposal contends that this generation study is limited because E2 was used as positive control and *"abnormally high prostate weights"* and cites Myers *et al.* 2009. The CLH proposal did not mention or evaluate the responses by Tyl to this criticism (Tyl 2009 and Tyl 2010).

Tyl 2009 response to Myers et al. (2009) claiming that the positive control group used a very high E2 dosage. *"The dietary positive control (0.5 ppm E2) was selected based on one-generation (Tyl et al. 2008a) and two-generation (Tyl et al. 2008b) CD-1 mouse studies using E2 doses of 0.001–50 ppm (8 mg/kg/day), with the most complete spectrum of effects seen at 0.5 ppm [estrogenic effects were observed at 0.05 ppm (increased weanling uterine weight) with a NOAEL of ~ 1 µg/kg/day]. Because E2 was dietary (animals were fed ad libitum, with ADME occurring as they fed), it took higher feed doses versus gavage or other routes to achieve the same test chemical response. Because almost all human BPA exposure is oral (during episodic eating/drinking), dosed feed is the most relevant exposure route.*

No studies mentioned by Myers et al. (2009) investigated E2 after oral administration: Gupta (2000), Timms et al. (2005), and vom Saal et al. (1997) used DES; Putz et al. (2001a, 2001b) used oral EE; Timms et al. (2005) used estradiol benzoate by subcutaneous injection in rats, and Leranth et al. (2008) [see Leranth et al. 2008a and b and Mathews 2008] *used it by subcutaneous capsules in primates; vom Saal et al. (1997) used subcutaneously implanted Silastic E2 capsules; and Richter et al. (2007) used E2 (and BPA) in cell culture. We rejected EE (synthetic and greater oral bioavailability) and DES (multiple modes of action, no reproducible data that it is orally active at low doses) as positive controls.*

In Table 9 of our paper (Tyl et al. 2008a), we summarized key findings in rats and mice at the highest dietary concentrations producing viable offspring (2.5 ppm in rats; 0.5 ppm in mice), at doses of similar order of magnitude: mice, ~ 0.08–0.12 mg/kg/day; rats, ~ 0.17–0.2 mg/kg/day. Although there are differences in some end points, effects of exogenous estrogen in rats and mice are similar at similar doses. The concerns about our positive control and dietary concentration are unfounded."

Tyl 2009 response to Myers et al. (2009) claiming the "large" prostate weights suggest poor dissection technique, and that use of dissection, weight, and histopathology of the seminal vesicles plus coagulating glands (SVCG) together was "inappropriate:"*Prostate weights (lobe or whole gland) in all mammals evaluated to date have increased with age as males mature sexually (Sinowatz et al. 1996, for mice), as observed in our CD-1 mice. Mean control mouse prostate weights by lobe and age in our studies are presented in Table 1. Table 2 shows mean mouse whole-prostate weights of mice of various ages reported by others.*

My laboratory has extensive experience in weighing rat and mouse prostate glands at various ages. In the Hershberger interlaboratory validation study, my laboratory had some of the most precise prostate weight data and robust treatment-related effects of the 17 participating laboratories (according to one senior OECD/U.S. EPA reviewer; personal communication). We performed power calculations on our rat and mouse prostate data versus data of Nagel et al. (1997). Because of larger sample sizes and smaller coefficients of variation, our data have greater power to detect small effects than the Nagel data (unpublished data). Examination of paraffin block faces and slides of the rat and mouse prostates in the studies (Tyl et al. 2002, 2008c) indicated no evidence of extraneous tissue/fat or excessive inflammation.

vom Saal's CF-1 offspring mice are routinely group housed by sex for several months beginning at weaning, and then singly housed for 1 month before necropsy. It is highly unlikely that 1 month of single housing compensated for months of group housing and its effect on male mouse sexual development; the dominant cage male develops large androgen-dependent accessory sex organs (ASOs), and subservient cage males have smaller ASOs (Bartos and Brain 1993). vom Saal's CF-1 mice did not exhibit increased prostate weights with age, likely because of the postwean caging regimen utilized in the study. Adult control prostate weights in our reproductive toxicity studies with E2 (Tyl et al. 2008a, 2008b) and BPA (Tyl et al. 2008c) (Table 1) are well within the weight range of other published studies (Table 2), and, as expected, our mouse prostate weights increased with increasing age.

We dissected, weighed, and examined the SVCG together to prevent tissue damage from necropsy separation of these intimately associated organs for histopathology, especially in mice."

Overall, the EU RAR 2008 considers the *"investigation by Tyl* et al. *(2007) as the gold-standard, definitive study of the reproductive toxicity of BPA (for the endpoints examined)";* the statement in the CLH proposal concerning the limitations of the Tyl 2008 study based on Myers et al. 2009 is inappropriate taking into account responses by Tyl 2009 and 2010 and the fact that the study design and interpretation of the results were supervised by a Steering Group that was chaired by a representative of the European Chemicals Bureau and included experts from several EU Member States.

3. Conclusion *(CLH proposal chapter page 97-98)*

The conclusion of the CLH proposal mentions several observations by indicating positive studies in this chapter and summarizes some data in table 16. Studies without significant effect are also mentioned in the CLH proposal: "In contrast to these previous studies, several authors found only limited effect (Tinwell et al., 2002 and Kobayashi et al., 2010 and 2012) and other failed to *demonstrate significant effects of BPA exposure on the male reproductive tract, especially at low doses (Howdeshell et al., 2008; LaRocca et al., 2011)."* In table 16 it is indicated that only Howdeshell showed no significant effects. But no information on Kobayashi et al. 2001 and 2012

or LaRocca et al. 2011 are included in table 16. Endpoints measured by Tinwell et al. 2002 without effect are also not mentioned in table 16.

Overall, only "positive" effects are included in table 16 and no weight of evidence taking into account positive and negative data is performed.

Therefore, a weight-of-evidence evaluation taking into account positive and negative data was not performed. The CLH proposal did not make conclusions on individual studies based on scientific criteria (number of animals, appropriate statistics, consistency of data, documentation and reporting of data,…).

No weight of evidence is performed for any endpoint

Two examples are given to illustrate that not all available data were included in the CLH proposal.

Sperm production and quality:

On page 97 the CLH proposal reads: **"***In the animals treated in utero and/or lactation,* ***most of the studies performed in mice or rats found effects on sperm production or quality*** *(Tinwell et al. (2002); Salian et al. (2009c))…".* This statement is not correct because there are many additional studies mentioned in the CLH proposal which report no effect on sperm. For example page 85 of the CLH proposal reads: *"In contrast to these previous studies, several authors found only limited effect (Tinwell et al., 2002 and Kobayashi et al., 2010 and 2012) and other failed to demonstrate significant effects of BPA exposure on the male reproductive tract, especially at low doses (Howdeshell et al., 2008; LaRocca et al., 2011).";* Kobayashi et al. 2012 (mentioned in the CLH proposal on page 86: "*It is difficult to find any specific reason explaining those contradictory results with the state of actual knowledge.")* and Kato et al. 2006 *(*mentioned in the CLH proposal on page 88 *"All BPA groups in which analysis were performed at PND10, 35 and 150 showed normal reproductive parameters (for instance preputial separation, sperm analisys, serum testosterone levels, copulatory and fertility rate, sexual organ weight…"*).

In addition, the CLH proposal failed to note several additional studies reported in the literature which investigated sperm parameter. The CLH proposal provides no rationale for the inclusion/non inclusion of studies on this topic. Based on published weight of evidence evaluations from Gray *et al.* 2004, Goodman *et al.* 2006, Goodman *et al* 2009 and a list of publications from an additional literature search that was done to bridge the time between the last published weight of evidence assessment until December 2012 for example the following studies were identified: Aikawa et al. 2004, Toyama et al, Kubo et al 2003, Wistuba et al 2003, Peknicova et al 2002, Park et al 2005b, , Kubo et al 2001, Cagen et al 1999, vom Saal et al 1998, Ashby et al 1999, Talsness et al 2000, Ema et al 2001 and more recently Kendig et al. 2012, Xie et al. 2010, Zhang et al. 2012b. (Note: In our comments concerning the CLH proposal we have not evaluated these additional studies for quality, but include them as examples for apparently available data that the CLH proposal failed to collect or analyze).

Testes weight and histology:

There are several studies reported in the literature which investigated testes weight and/or histology. On page 98 the CLH proposal indicates: *"...a decrease of the testis weight and an increase of the ventral prostate weight have been observed. (Chitra et al., 2003; Herath et al., 2004). …In 2 independent studies (Tan et al., 2003 and Toyama et al., 2004) histology of the testes was precisely investigated and morphological changes and abnormalities were reported."* The CLH proposal provides no rationale for the

inclusion/non inclusion of studies on this topic. Based on published weight of evidence evaluations from Gray et al. 2004, Goodman et al. 2006, Goodman et al 2009 and a list of publications from an additional literature search that was done to bridge the time between the last published weight of evidence assessment until December 2012 for example the following studies were identified: Ichihara et al. 2003, Kim et al. 2003, Kobayashi et al. 2002, Kubo et al. 2003, Negishi et al. 2003, Ashby et al 2003, Kim et al. 2002, Wistuba et al. 2003, Saito et al. 2003, Takahashi and Oishi 2003 (no effect on testes weight), Toyama and Yuasa 2004, Kawai et al. 2003, Peknicova et al., 2002, Al-Hiyasat et al., 2002a and b, Takao et al., 2003, Park et al 2005b, Park et al., 2004, Negishi et al. 2004, Kwon et al. 2000, Howdeshell et al 2008, Talsness et al., 2000, Takao et al 1999, Pacchierotti et al., 2008, Noda et al., 2005, Gui and Lu 2007, Okada and Kai 2008, Naciff et al. 2005, Song et al 2008, Nanjappa et al., 2012, Dobrzynska and Radzikowska 2013, Kendig et al. 2012, Kobayashi et al 2010, Xu et al, 2011b, Lopez-Casas et al 2012, Xie et al 2010, Cardoso et al 2011, Nanjappa et al. 2012, Wu et al 2011, Horstman et al, 2012, Hutanu 2011 (Note: In our comments concerning the CLH proposal we have not evaluated these additional studies for quality, but include them as examples of available data that the CLH proposal failed to collect or analyze).

Overall, the CLH proposal discussed only a limited and selected number of studies for each endpoint and did not evaluate all positive and negative results in a single weight of evidence determination.

Comments on female reproductive evidence in animals

Summary

The CLH proposal alleges increased ovarian cysts, reduced luteinizing hormone/luteinizing hormone surge, and decreased fertility as "Proven Effects" from BPA administration in female rats and mice. The CLH proposal also cited and discussed studies that attributed BPA treatment with changes in estrous cyclicity, pubertal onset (vaginal opening) and oocyte aneuploidy. These effects (e.g., estrous cycle and puberty onset changes), however, were not identified as "Proven Effects" in the report. Therefore, the following comments focus on the "Proven Effects" evidence and how ANSES interpreted this evidence. Overall, based on the following reasons a re-classification of BPA is unwarranted**:**

- The cited studies that purportedly establish these "Proven Effects" are non-guideline studies.
- They are not corroborated by the negative findings reported in robust guideline studies (e.g, Tyl et al. 2008 two-generation mouse study and the Tyl et al. 2002 3-generation rat study that evaluated the ovaries and fertility and did not find any effects in the absence of maternal toxicity). Importantly, guideline studies are required and relied upon for registering substances in Europe under REACH and are also an important aspect for classification.
- Many of the studies employ non-validated, experimental methods with insufficient methodological information to examine the reported results.
- Many of the studies fail to link observations, even if they are assumed to be valid observations, to actual apical adverse outcomes that would impact reproduction and development. According to GHS criteria for 1B Classification: *"Such data shall provide clear evidence of an adverse effect on sexual function and fertility or on development."* It should be noted that a large number of the studies cited in the CLH proposal provide no measure of functional impact on reproduction and development.
  The majority of the studies employed non-relevant, non-oral routes of exposure, e.g., subcutaneous, intravenous/intramuscular, slow release implants of BPA.

Each of the "Proven Effects" for female reproductive endpoints reported in the animal studies is discussed in greater detail below.

1. Increased Ovarian Cysts (Proven Effect #1 for Female Reproduction)

The studies reporting ovarian cysts are briefly summarized in Table 1.

These studies are of limited value for hazard assessment due to their non-relevant route of exposure utilizing excessive exposures – all involve subcutaneous administration. It is well-known that oral 1st pass clearance of BPA, via intestinal and hepatic glucuronidation, severely limits oral bioavailability and results in ineffective blood levels of BPA aglycone (Volkel et al. 2002; Teeguarden et al. 2011). Therefore, the route and magnitude of exposure is highly relevant for hazard classification since a hazard cannot exist unless humans are administered BPA via a non-oral route .

Hazard classification must be based on validated and reliable methodologies, preferably, guideline studies suitable for REACH registration. In contrast, the studies cited in the CLH proposal represent often non-validated, non-guideline studies that suffer from a lack of methodological validation for reporting ovarian cysts (see Table 1 below). For example, careful step-sectioning of the ovaries beginning in the middle of the ovary, and stepping out, is necessary to quantify cysts or other follicular structures within the ovary (Step sectioning is noted in the 3-generation rat study of Tyl et al. 2002). While a few studies suggest that step sectioning was utilized, the information was not sufficient to fully understand if step sectioning was adequately done to evaluate the entire ovary. For example, Adewale et al (2009) only evaluated two slides per animal demonstrating an inadequate step-sectioning of ovarian tissue for quantifying cysts. Secondly, with step-sectioning, great care would be required to avoid the double counting of larger cysts that penetrate multiple cross sections. It is not explained how this potential double counting was avoided in methods were additional sectioning was employed. Consequently, the challenge of quantifying ovarian structures may be the reason why none of the studies reported quantitative morphometry of ovarian cysts. Instead, all of the studies largely relied on a qualitative visual reporting of ovarian cysts. While reporting ovarian cysts are relatively straight-forward from a histological perspective, there are differences in how they were approached. For example, Fernandez et al. (2010) described their cyst as atretic follicles whereas Adewale et al (2009) described their cysts as antral-like follicles. Even within the same lab there are inconsistencies. Newbold et al. 2007 and 2009 provide inconsistent findings between their two studies for increases in ovarian cysts, neither of which show a biologically plausible dose-response relationship (Table 2). What is needed is a validated methodology for the quantitative morphometric assessment of ovarian cysts. Only a validated quantitative methodology can provide for the necessary understanding of the background and treatment variability in ovarian cysts and the histological criteria used to define cysts as a function of species, age, and estrous cycle. Consequently, these studies present a number of problems with validating ovarian cysts as an apical, adverse outcome, or, a "Proven Effect" according to the CLH proposal. Furthermore, these studies provide no linkage to an apical outcome (reproductive/developmental impairment) so they cannot satisfy criteria for classification under CLP.

The CLH proposal cites studies beyond those summarized in Table 1 in attempts to support their ovarian cyst conclusion. For instance, they cited studies reporting aneuploidy and meiotic changes following BPA administration in rodent oocytes but they concluded this evidence provides" too few studies on this subject are available and it is thus difficult to draw conclusions (citing Hunt et al. 2003; Susiarjo et al. 2007 and Rodriguez et al. 2010). The CLH proposal seems to cite three other studies as further support of the ovarian cyst claim (Markey et al. 2005; Mendoza-Rodriquez et al. 2011 and Schönfelder et al. 2004). Markey et al. 2005 did not look specifically at the ovary making this paper irrelevant to the

question of ovarian cysts. Mendoza-Rodriguez (2011) histologically examined the ovaries for estrous in the controls and BPA-treated animals but never commented on abnormal histology in the ovaries. Schönfelder et al. 2004 is irrelevant since they only looked at the uterus response in pups born to mother gavaged with either 0.1 or 1.0 mg/kg/day given on GD6-GD2. No ovarian data were provided in this study.

The CLH proposal cites Hunt et al. 2003 and Susiarjo et al. 2007. Not included in the CLH proposal are two subsequent studies (Pacchierotti et al. 2008, Eichenlaub-Ritter et al. 2008) that could not reproduce these findings. This inconsistent picture was noted in the 2008 EU Risk Assessment. Also not mentioned in the CLH proposal is a further study published by the Hunt group, Muhlhauser et al. (2009), is a study in which the authors could not replicate their own initial findings. In conclusion, the initial observations reported by the Hunt laboratory were not reproduced in the same laboratory or in other independent laboratories. This important information is not included in the CLH proposal.

Table 1: Studies Reporting Ovarian Cysts.					
Study	Route	Dose	Methods[A]	Results	Comments
Fernandez et al., 2010	sc	250 to 62,600 µg/kg/day PND1-PND10	H&E Stains of ovarian sections to count PFs, CL, AF, AtrF using light microscopy[1] Female Sprague Dawley Rats	1. ↓ CL at 25-62.6 mg/kg[B] 2. ↓ AF at 2.5-6.25 and 25-62.5 mg/kg[B] 3. ↑ AtrF at 2.5-6.25 and 25-62.5 mg/kg[B] 4. No change in preovulatory follicles	• No information on step sectioning required to correctly conduct histology examination. • Non-relevant route of exposure with high doses for that route of exposure-Atretic follicles counted as cysts with no explanation • No change in pre-ovulatory follicles (primordial follicles) which would be expected if a true estrogenic response was being induced. • Observation not confirmed in 3-generation study in the same strain (Sprague-Dawley rats)
Adewale et al., 2009	sc	50 µg/kg/day or 50 mg/kg/day PND0-PND3	2 slides/animal examined for abnormal histology, including corpora lutea. Long Evans Rats	1. *"Ovaries from both BPA-treated groups also contained a number of large antral-like follicles that were similar in appearance to those seen in the PPT-treated animals and were structurally consistent with ovarian cysts."*	• Histological methods lack capability of accurately assessing quantitative nature of ovarian cysts. • Irrelevant route of exposure using high dosages for that route of exposure
Newbold et al., 2007	sc	10, 100 and 1000 µg/kg/day PND1-PND5	6 µm sections were evaluated. If a lesion was observed, further serial sections were evaluated to complete-ly assess the lesions. CD-1 Mice	1. No statistically significant decline in CL 2. ↑ in ovarian cysts only at 100 µg/kg/day	• Irrelevant s.c. route at high exposures for that route of exposure • Inadequate histological methodological information to assess quantitative relationship with BPA • No dose-response relationship (only increased at one dose) • Inconsistent with the Newbold et al. 2009 findings.
Newbold et al., 2009	sc	0.1, 1., 10, 100, and 1000 µg/kg/day on GD9-GD16	Nine ovarian sections/animal with additional serial sections made to fully account for lesions CD-1 Mice	1. No statistically significant decline in CL 2. ↑ in ovarian cysts only at 1 µg/kg/day	• No dose-response relationship (only increased at one dose) • Inconsistent with the Newbold et al., 2007 findings.
Signorile et al., 2010	sc	100 or 1000 µg/kg/day GD1-PND7	Four sections every 100 µm with additional serial sections to fully assess lesions	1. ↑ in ovarian cysts at both dosages 2. No difference in CL reported.	• Irrelevant route of exposure at high exposures for that route of exposure. • Inadequate histological information to validate results • No dose response since the incidence of ovarian cysts was similar in the two dose groups (45% and 50% compared to 10% in the controls)

1. No initial mention of how cysts were to be quantified. The investigators finally described the counting of atretic follicles "..as anovulatory follicles with a granulose layer (two or three layers) and non-detectable theca" as the population representing ovarian cysts.

A PF (primordial follicles); CL (Corpora Lutea); AF (Antral Follicle); AtrF (Atretic Follicles).

B. No results, or mention, of ovarian histology was provided for the 0.25 to 0.625 mg/kg/day dose groups with no explanation for the missing data in this dose group.

Table 2: Comparison between Newbold et al., 2007 and Newbold et al., 2009

	% Incidence of Ovarian Cysts	
Dose (µg/kg/d)	**Newbold et al., 2007** (exposure PND1-PND5), Determined at 18 months of age	**Newbold et al., 2009** (exposure GD9-GD16) Determined at 16-18 months of age
Control	39%	25%
0.1	Not Done	29%
1.0	Not Done	**67% ($p < 0.05$)**
10	35%	14%
100	**70% ($p < 0.05$)**	36%
1000	38%	38%

An important aspect of the alleged ovarian cyst increase is its biological plausibility. That is, can the CLH proposal establish internal dosimetry consistent with the toxicokinetics of the substance and its proposed mode-of-action that are consistent with the relevant exposure route for humans?

CLP Regulation EC1272/2008 indicates:

> *1.1.1.5 For the purpose of classification for health hazards (Part 3), route of exposure, mechanistic information and metabolism studies are pertinent to determining the relevance of an effect in humans. When such information, as far as there is reassurance about the robustness and quality of the data, raises doubt about relevance in humans, a lower classification may be warranted. When there is scientific evidence that the mechanism or mode of action is not relevant to humans, the substance or mixture should not be classified*
>
> *3.7.2.3.2 „if it is conclusively demonstrated that the clearly identified MOA has no relevance for humans or when TK differences are so marked that it is certain that the hazardous property will not be expressed in humans then a substance which produces an adverse effect on reproduction in experimental animals should not be classified"*

Mode-of-action is an important aspect in the assessment of a substance and to determine if it is e.g. even biologically plausible for a weak estrogenic substance to cause ovarian cysts. For example, is estrogen involved in polycystic ovary development in laboratory animals? Had the CLH proposal explored this question it would have acknowledge that it is testosterone agonist administration (e.g., testosterone, adrostenedione, dihydrotesterone) that is associated with producing ovarian cysts and a metabolic condition in rodents that mimics PCOS in humans (Walters et al. 2012). Thus, BPA is an unlikely biologically plausible candidate for PCOS.

Finally, additional criticism towards the CLH proposal emerges from how the alleged ovarian cyst claim was linked to evidence in humans. To these points, ANSES cited two epidemiological publications on the topic of BPA exposure and PCOS (Kandaraki et al. 2011; Takeuchi et al. 2004). Takeuchi et al. 2004 in which serum BPA levels, measured by an unreliable ELISA method, claimed to show a relationship between BPA and PCOS. ANSES then went on to criticize this study:

> *The results remain difficult to interpret because of the imprecision of the sampling, the very small size of the sampling, the lack of information on inclusion criteria, and the lack of accounting in the results of disorders in the controls. Then, the results from this study could not be considered as totally proven.* (pg. 79 of the CLH proposal).

ANSES also criticized the Kandaraki et al. (2011) study because of its use of a non-specific ELISA method that does not distinguish between free-BPA and BPA's conjugates. The apparent 1.05 ng/ml BPA reported in serum of the PCOS subjects could be the inactive BPA glucuronide. According to the recent publication by Patterson et al. 2013 the likelihood that this 1.05 ng/ml represents the weak estrogenic form of BPA aglycone is not credible.

> *"….aglylcone BPA levels reported in human blood in the ng/ml range from routine sample collections are implausible and could represent sample contamination artifacts."* (Patterson et al., 2013)

To summarize, none of the studies regarding ovarian cysts are guideline studies. None use a relevant route of exposure (all are subcutaneous). None of these rat and mouse studies report dose-response trends in the increase in ovarian cysts. The studies either don't define how they identified ovarian cysts, or, they used different histological definitions, e.g. antral follicle-like versus atretic follicles. The endpoints assessed in these studies did not appear to be obtained in a blinded manner and validation for the true range of variability in the control groups, using objective means to measure ovarian and ovarian cyst pathology, was not provided. Standardized, reliable, and valid methods to accurately quantify changes in ovarian follicles related to chemical exposure are not available. Ragan et al. 2005 discussed histological methods for qualitatively and quantitatively assessing ovarian follicles noting limitations:

> *"The utility of follicle counts to detect female reproductive toxicity in rats that cannot be detected by other means has not been demonstrated adequately. …Data generated from follicle counts can be highly variable with large standard deviations both between animals and between groups, making interpretation difficult. In part, this variability can be caused by sampling methods that are inadequate for consistent quantitative evaluation…..studies are needed to determine and confirm a reliable standard method to prepare rat ovarian sections for quantitative assessment and to determine the value of follicle counts in rats."*

Altogether, these studies claiming to show a BPA-dependent increase in cystic ovaries are unreliable and of low quality for purposes of supporting a weight-of-evidence basis for a change in the reproduction classification.

In addition there are additional studies that investigate ovaries morphologically are not mentioned in the CLH proposal. The CLH proposal provides no rationale for the inclusion/non inclusion of studies on this topic. Based on published weight of evidence evaluations from Gray et al. 2004, Goodman et al. 2006, Goodman et al 2009 and a list of publications from an additional literature search that was done to bridge the time between the last published weight of evidence assessment until December 2012 for example the following studies were identified: Mendoza-Rodriguez et al. 2011, Rodriguez et al. 2011, Karavan and Pepling 2012, Kwon et al. 2000, Noda et al. 2005. (Note: In our comments concerning the CLH proposal we have not evaluated these additional studies for quality but include them as examples for apparently available data that the CLH proposal failed to collect or analyze).

Ovaries are also investigated in the recent NCTR study. In this study BPA was administered to rat dams by oral gavage from gestation day 6 until parturition and then directly to pups from postnatal day 1 until termination at postnatal day 90 at doses up to 300 mg/kg bw/day. In this robust study, which features a comprehensive range of dose levels and wide array of health endpoints, BPA exhibited Increased cystic follicles, depleted corpora lutea and antral follicles at the highest BPA dose tested. A dose where the authors indicate *"the interpretation of the high dose BPA effects are confounded by systemic toxicity"*. Thus, the study supports earlier research and European Risk Assessments concluding that BPA produces adverse effects only at high doses associated with systemic toxicity.

Overall the data in this in this comprehensive study do not support the premise that BPA is a selective reproductive toxicant *(see Annex C for further details).*

2. **Reduced LH Levels and LH Surge - Amplitude and Frequency ("*Proven Effect*" #2 for Female Reproduction)**

Four studies in sheep and one in rats were cited as evidence of a "Proven Effect" that BPA causes reduced LH concentrations and a reduction in the LH surge.

1. Collet et al. 2010 (sheep)
2. Evans et al. 2004 (sheep)
3. Savabieasfahani et al. 2006 (sheep)
4. Mahoney and Padmanaban 2010 (Sheep)
5. Rubin et al. 2001 (Sprague-Dawley rats)

The CLH proposal provided the following commentary regarding this alleged "Proven Effect."

> *However, a decrease in LH concentration and LH pulse amplitidue, suggesting a negative feedback exerts (sic) by BPA, was found in several studies in sheeps (sic) and rats after a gestational, perinatal postnal and prepubertal exposure (Savabieasfahani et al., 2006; Rubin et al., 2001; Evans et al., 2004; Colet et al. 2010) and can thus be considered as a proven effect. (page 65).*

While these studies are interesting for their use of non-oral routes of administration and attention to BPA blood levels associated with changes in LH physiology, they cannot be used to support classification because of issues of route, toxicokinetics, and metabolism factors that must be accounted for.

To provide additional information as to why these LH/LH surge data are not conclusive for a "proven effect", the data from three of the sheep studies cited in the CLH proposal are briefly summarized in Table 3. The short paragraphs that follow briefly touch on each study and why the study does not support the classification upgrade request. The Rubin et al. and Mahoney and Padmanabhan 2010 studies are briefly discussed separately below since one is a rat study (Rubin et al. 2001) and the other did not evaluate LH serum levels or the LH surge following BPA exposure (Mahoney and Padmanabhan, 2010).

Table 3: Sheep Studies Evaluating Non-Oral, High Dose BPA Administration's Effects on Luteinizing Hormone

Study	Dose, Route, Period	Serum LH*	LH Pulsatile Nature	Assessment by CERHR 2008
Savabieasfahani et al., 2006	5 mg/kg/day, **subcutaneous administration** on GD30-GD90	Elevated in lambs born to BPA injected ewes that were primed with $PGF_{2\alpha}$ at 40 weeks of age**	Reduced by BPA	*"Weaknesses are the use of a single dose level and the relatively small sample size. The single time point for bisphenol A plasma determination at an unknown time relative to s.c. injection is a weakness.* *Utility (Adequacy) for CERHR Evaluation Process: This study is adequate though of limited utility."*

Collet et al., 2010	0.5 to 80 mg/kg **intravenous** infusion for 54 hours A second experiment where lambs were **intramuscularly** injected, twice a week, with 3.5 mg/kg for 8 weeks	Reduced serum conc. At intravenous dosages exceeding 1 mg/kg	Reduction in LH pulsatile fluctuations	-
Evans et al., 2004	3.5 mg/kg **intramuscularly** every week x 7 weeks with ovariectomy conducted at 5 weeks of treatment. Pre-ovariectomy and post-ovariectomy LH determinations were made	BPA had no effect on LH levels relative to controls pre-ovariectomy. Post-ovariectomy, LH pulse frequency and amplitude was reduced in the BPA-treated animals. Also noted was an increase in uteri size.	Reduced amplitude and frequency of LH pulse determined every 15 minutes over a 6 hour period 2-weeks post ovariectomy.	*"The high-dose level via i.m. injection is a weakness as are small sample sizes (n = 6). The statistical tests for LH trends did not appear to take into account the repeated nature of the sampling leading to over stating the significance of trend effects. Utility (Adequacy) for CERHR Evaluation Process: This study is adequate for inclusion but of limited utility for the evaluation process."*

** Elevated relative to controls during "early" postnatal period in ewes born to mothers injected with BPA

Savabieasfahani et al. 2006 reported blood levels of BPA aglycone following daily subcutaneous injection of BPA at dosages of 5 mg/kg/day to pregnant Ewes from gestational day 30 to 90. Lambs born to these Ewes were later evaluated, every 2 hours for changes in plasma LH over a 120 hour period at 40 weeks after birth. The lambs were primed with injection of $PGF_{2\alpha}$. BPA prenatal treatment reduced the LH surge in some of the lambs. No differences in FSH were seen between groups. Progesterone secretion pattern showed no difference between groups, despite perturbations in LH patterns. Although there was no apparent effect on progesterone production, the authors suggested that the changes induced by prenatal exposure of females could interfere with fertility.

Notably, the maternal BPA blood levels at 50, 70 and 90 days during gestation ranged from approximately 35 to 55 ng/ml (0.15 to 0.24 µM). Collet et al. 2010, on the other hand administered BPA intravenously over a 54 hour period at daily dosage that ranged from 0.5 to 80 mg/kg/day. These intravenous dosages are hundreds to thousands of times in excess of background exposures in the population (EFSA, 2013). What Collet e al. 2010 demonstrated is that dosages ranging from 0.5 to 1 mg/kg/day given intravenously, result in steady state BPA aglycone blood levels that exceed approximately 30 ng/ml, but had no effect on LH serum levels in ovariectomized sheep. Orally exposed humans could never achieve blood levels of 30 ng/ml (Patterson et al. 2013).

Evans et al. 2004 administered weekly intramuscular dosages of BPA (3.5 mg/kg) for 7 weeks in prepubertal lambs. BPA blood levels were not determined in this study. At five weeks the lambs were ovariectomized thus greatly sensitizing the lambs to any subsequent estrogenic stimulus. At 7 weeks of treatment (2-weeks post ovarectomy), LH pulses were sampled for 6 hours every 15 minutes (with 6 lambs in the controls and 6 lambs in the BPA-treated group). In these estrogen-sensitized animals, the authors reported: "Uteri were visually larger in size in the DES- and BPA-treated animals

indicating a positive uterotrophic effect of treatment. Prior to the ovariectomy, and an important observation, the 3.5 mg/kg weekly intramuscular dosages of BPA did not affect serum LH levels relative to controls. It was only after ovariectomy that sensitized the sheep to BPA's weak estrogenic properties responded. In the ovariectomized and estrogen-sensitized animals, BPA produced a decrease in both the amplitude and frequency of the LH pulse. As to this study's relevance to classification, it too, like the Savabieashfani et al. 2006 and Collet et al. 2010 studies cannot be relied upon due to its non-relevant route and high dose exposures that do not take into account BPA's important toxicokinetic and metabolism barriers that are route-specific.

Mahoney and Padmanabhan 2010 differs from the other three sheep studies since they did not measure LH levels or the LH surge. They injected pregnant sheep subcutaneously with 5 mg/kg/day of BPA on GD30-GD90 and it appears that they were lambs used in the Savabieasfahani et al. 2006 study. In the previous study (Savabieasfahani et al. 2006), free BPA blood levels of around 40 ng/ml (0.175 µM) were achieved with this systemic administration of BPA. The mature F_1 females were then primed with prostaglandin $F_{2\alpha}$ twice, 11 days apart. Then, according to unspecified information obtained from the controls, the ewes were sacked at their onset of their LH preovulatory surge. The hypothalamus was removed and stained with ^{35}S- mRNA probes for GnRH, ERα and ERβ mRNA. Based on just a few cells from the hypothalamus, e.g., 7-15 cells to count staining for mRNA for GnRH, the authors reported the differences in mRNA probe staining (5 Controls versus 7 BPA-treated F_1 females) for the three endpoints. They reported that GnRH, and ERβ mRNA staining was reduced in the BPA group whereas ERα staining was increased. The results were presented with their standard error of the mean. Overall, this study reports findings from a high dose, non-relevant route of administration reported to produce free BPA blood levels potentially within the range of where BPA can exert biological activity. The results reflect a non-validated morphological quantification of mRNA staining in whole tissue using a very small number of animals per treatment group, e.g, to account for normal variability in the timing of the LH pre-ovulatory surge and in the normal variation from animal to animal.

Rubin et al. 2001 claimed that drinking water exposure in Sprague Dawley rats between GD6 and lactation produced female offspring that experienced a reduction in serum LH levels at 4 and 6 months of age. This result and conclusion, however, is unreliable. For one, Rubin et al. 2001 did not indicate how they controlled for the large changes in serum LH as a result of LH pulsatile frequency. For example, Matt et al. 1998 demonstrated that LH concentrations can vary over a 6 hour period with a normal range that includes LH concentrations of 14.97 ng/ml (Control values for Rubin et al., 2001) and approximately 12 ng/ml (two BPA-treated groups in Rubin et al). Thus, the control and two BPA-group LH values reflect normal mid-range amplitude variability. What Rubin et al. 2001 failed to control for is the LH pulsatile amplitude and frequency nature of LH. This biological phenomenon requires frequent sampling techniques (every 2-3 minutes for a blood sample) in order to accurately capture the rapidly changing blood LH concentrations that result in the normal pulsatile frequency activity. Rubin et al.'s 2001 single blood sample, taken at some unspecified time and not controlling for pulsatile frequency and amplitude with frequent blood sampling, does not allow one to make any interpretation out of the normal LH concentration data reported in their study. As seen in the three sheep studies, the investigators employed frequent blood sampling over a period of hours to capture the rapidly fluctuating changes in the pulsatile nature of LH production.

CERHR assessment of the study: *"Strengths/Weaknesses: This study incorporates a range of basic developmental and gross functional reproductive endpoints, but the sample sizes are small (6 dams/group) and the statistical approach does not appear to use litter as the unit. Actual exposures are poorly defined, particularly postnatally. The plausibility of the estrous cycle changes is a strength. Utility (Adequacy) for CERHR Evaluation Process: This study is inadequate for the evaluation process, based on a lack of adequate control for litter effects"*

To summarize, the ability of BPA to influence LH levels in these sheep studies depended on a non-relevant route of administration with high dosages In order for these sheep studies to be relevant with respect to toxicokinetics, metabolism and mode-of-action, they would require relevant exposure route BPA treatments with confirmation that free BPA (the aglycone with weak estrogenic activity) blood levels are sufficient to explain the observed changes in LH consistent along the lines of its proposed estrogenic mode-of-action.

3. Reduced Fertility (Proven Effect #3 For Female Reproduction)

The CLH proposal identifies reduced fertility as a "Proven Effect" following BPA administration in rats and mice. This opinion is primarily based on the evidence published in the Cabaton et al., 2010; Fernandez et al. 2010; Al Hiyasat e al. 2004 and the three publications by Berger et al. 2007, 2008 and 2010). Two other publications appear to play a more limited role in support of the report: Tachibana et al. 2007 and Varyoud et al., 2011. All of these studies are non-guideline studies. A number of studies utilize non-relevant, non-oral routes of administration. For example, Varayoud et al. 2011 administered 50 or 20,000 µg/kg/day dose subcutaneously to rat pups on PND 1, 3, 5, and 7. As observed in the sheep studies, non-oral dosages in this range can produce effects that do not take into account important hazard classification factors including toxicokinetics, metabolism and route. While the CLH proposal notes a number of larger, more robust guideline studies that do not corroborate these studies, ANSES provides numerous reasons for why these negative guideline studies are not persuasive, e.g., done in so called non-sensitive Sprague Dawley rats, failure to use a continuous breeding method. As discussed in the following comments, however, the non-guideline studies are not sufficient to support a change of the existing classification nor do they refute the negative findings from much better, robust guideline studies (Tyl et al., 2002, 2008, Ema et al. 2001, Stump et al. 2010; see also Annex D).

Cabaton et al. 2010 implanted CD-1 mice at GD8 with alzet pumps administering approximately 25 and 250 ng/kg/day, and/or 25 µg/kg/day. The resulting F_1 females were then placed with a male from an unexposed F_0 dam for 32 weeks beginning at 8 weeks. Each resulting litter produced by the F1 female was euthanized and the F1 female returned to its original male cage mate (no ability to control for male partner effects on the alleged changes in reproductive capacity). The Cumulative # of pups/dose, b) time to 1st litter, the number of deliveries/dam, the number of neonates/litter and the number of mice having litters were recorded. Besides using an irrelevant route of exposure (implanted alzet pumps), the apparent BPA-related impact on fertility endpoints was not dose related. This is shown by the cumulative number of pups per treatment group born over the 32 week period. The following relative dose-response relationship for a reduction in the cumulative number of pups was observed: Controls = 250 ng/kg/day > 25 ng/kg/day = 25 µg/kg/day (Figure 2 in the paper). Thus, an expected dose-related reduction in cumulative pups, if BPA was truly acting to suppress continuous fertility, as exhibited by a monotonic decrease in fertility as a function of dose, was not observed. The same lack of a dose-response relationship between BPA and other fertility endpoints (number of litters per dam) and fecundity (# of pups born at each delivery). Therefore, the same non dose-dependent obseravtions were carried through in the other endpoints suggesting nothing more than random variation in this strain (something the authors did not address) or the impact of the male partner (that was not addressed). These discrepancies in dose-response, route of administration, and validation of the method used by Cabaton et al. with consideration of historical

Fernandez et al. 2010 is not relevant for classification purposes because of its non-relevant route of exposure (s.c.) and the very high dosages (2.5 to 6.5 mg/kg/day and/or 25 to 62 mg/kg/day). Interestingly, the positive response in Sprague Dawley rats in this specific case was not dismissed for being a non-responsive strain (in contrast to other sections in the report).

The CERHR (NTP´s Center For The Evaluation of Risks To Human Reproduction) concluded that the 2004 Al-Hiyasat et al. 2004 study was not useful due to its small sample size (Chapin et al., 2008, CERHR 2008). Confounding by the presences of other glicidyl ether substances in the dental extract exposure study, and the absence of dose-related declines in fertility in the BPA-treated groups are two other limitations in this study. In the face of much larger, robust guidelines studies that do not corroborate the Al-Hiyasat et al. 2004)findings, this study of limited usefulness for classification.

The Berger et al. studies (2007 and 2008) are not relevant due to their excessively high subcutaneous dosages administered on GD1-GD4. The dosages were estimated to range from approximately 0.01 to over 300 mg/kg/day (milligrams per kilogram) (Goodman et al. 2009) and effects were observed only in the very high dose range.
Berger et al. 2007 was evaluated by CERHR: *"Weaknesses include the limited/unequal number of mated mice in each dose group, absence of maternal data to ascertain the potential impact of maternal toxicity on pregnancy, methodological deficiencies regarding fertility assessment, and the use of a diet that contains phytoestrogens. Utility (Adequacy) for CERHR Evaluation Process: Due to the limited number of mated mice per dose level (n 5 5–15), methodological concerns, absence of key statistical information as well as maternal information, this study is inadequate for the CERHR evaluation process."*

To summarize, the cited studies on female reproductive chages in laboratory rodentes are not, and cannot be compelling reasons to support a change in classification. The studies employed large dosages via a non-relevant route. The studies used limited study designs including low number of animals, limited parameter investigated and/or limited number of doses investigated. The studies are contradicted by more comprehensive guideline studies required under REACH registration requirements (see Annex D).

References (if not cited in CLH proposal)

Al-Hiyasat, A. H., Darmani, H., and Elbetieha, A. M. 2002a. Effects of bisphenol A on adult male mouse fertility. European Journal of Oral Sciences. 110:163-167.

Al-Hiyasat AS, Darmani H, Elbetieha AM. 2002b. Effects of bisphenol A on adult male mouse fertility. Eur J Oral Sci.; 110: 163–167. Erratum

Ashby, J., Owens, W., Odum, J., and Tinwell, H. 2003. The intact immature rodent uterotrophic bioassay: Possible effects on assay sensitivity of vomeronasal signals from male rodents and strain differences. Environmental Health Perspectives. 111:1568-1570.

Cardoso, N., Pandolfi, M., Lavalle, J., Carbone, S., Ponzo, O., Scacchi, P., and Reynoso, R. 2011. Probable gamma-aminobutyric acid involvement in bisphenol A effect at the hypothalamic level in adult male rats. Journal of Physiology and Biochemistry. 67(4):559-567.

CERHR: NTP-CERHR monograph on the potential human reproductive and developmental effects of bisphenol A. Shelby MD., NIEHS, Research Triangle Park, NC 27709, USA. NTP CERHR MON. 2008 Sep;(22)

Chapin et al.: NTP-CERHR expert panel report on the reproductive and developmental toxicity of bisphenol A. Birth Defects Res B Dev Reprod Toxicol. 83, 157-395, 2008

CLP Regulation EC1272/2008. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:353:0001:1355:EN:PDF

Dobrzynska, M. M., Gajowik, A., and Radzikowska, J. 2012. Male-mediated developmental toxicity in the offspring of adult male mice exposed to bisphenol A alone or in combination with X-rays. Mutagenesis. 27(1):123-124.

Eichenlaub-Ritter, U., Vogt, E., Cukurcam, S., Sun, F., Pacchierotti, F., and Parry, J. 2008. Exposure of mouse oocytes to bisphenol A causes meiotic arrest but not aneuploidy. Mutation Research/Genetic Toxicology and Environmental Mutagenesis. 651(1-2):82-92.

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. 2006. An updated weight of the evidence evaluation of reproductive and developmental effects of low doses of bisphenol A. Critical Reviews in Toxicology. 36(5):387-457.

Goodman et al. 2009. Weight-of-evidence evaluation of reproductive and developmental effects of low doses of bisphenol A. Crit Rev Toxicol. 39, 1-75,

Gray, G. M., Cohen, J. T., Cunha, G., Hughes, C., McConnell, E. E., Rhomberg, L., Sipes, I. G., and Mattison, D. 2004. Weight of the evidence evaluation of low-dose reproductive and developmental effects of bisphenol A. Human and Ecological Risk Assessment. 10(5):875-921. Supplemental information in Eval document.

Gui, J. and Lu, F. 2005. Effects of bisphenol A on reproductive system of puberty make rats. Zhongguo Gonggong Weisheng. 21(6):696-698.

Horstman, K. A., Naciff, J. M., Overmann, G. J., Foertsch, L. M., Richardson, B. D., and Daston, G. P. 2012. Effects of transplacental 17-α-ethynyl estradiol or bisphenol A on the developmental profile of steroidogenic acute regulatory protein in the rat testis. Birth Defects Research Part B: Developmental and Reproductive Toxicology. 95(4):318-325.

Hutanu, D. 2011. Experimental investigations regarding the effects of bisphenol A in adult mice spermatogenesis. Annals of RSCB. 16(2):74-78.

Ichihara, T., Yoshino, H., Imai, N., Tsutsumi, T., Kawabe, M., Tamaoa, S., and Inaguma, S., S. Suzuki, and Shirai, T. 2003. Lack of carcinogenic risk in the prostate with transplacental and lactational exposure to bisphenol A in rats. Journal of Toxicological Sciences. 28(3):165-171.

Kawai, K., Nozaki, T., Nishikata, H., Aou, S., Takii, M., and Kubo, C. 2003. Aggressive behavior and serum testosterone concentration during the maturation process of male mice: the effects of fetal exposure to bisphenol A. Environmental Health Perspectives. 111(2):175-178.

Kendig, E. L., Buesing, D. R., Christie, S. M., Cookman, C. J., Gear, R. B., Hugo, E. R., Kasper, S. N., Kendziorski, J. A., Ungi, K. R., Williams, K., and Belcher, S. M. 2012. Estrogen-like disruptive effects of dietary exposure to bisphenol A or 17α-ethinyl estradiol in CD1 mice. International Journal of Toxicology. 31(6):537-550.

Kim, M. J., Choi, B. S., Park, J., Hong, Y. 2002. Male reproductive toxicity of subchronic bisphenol A exposure in F344 rats. Chung-Ang Journal of Medicine. 27(3-4):111-120.

Kim, P., Lee, N., and Hwang, S. 2003. The bisphenol A: A modulator of pregnancy in rats. Korean Journal of Envoirnmental Health Society. 29(4):27-43.

Kobayashi, K., Miyagawa, M., Wang, R. S., Suda, M., Sekiguchi, S., and Honma, T. 2005. Effects of in utero and lactational exposure to bisphenol A on thyroid status in F1 rat offspring. Industrial Health. 43(4):685-690.

Kubo, K.,, Arai, O., Ogata, R., Omura, M., Hori, T., and Aou, S. 2001. Exposure to bisphenol A during the fetal and suckling periods disrupts sexual differentiation of the locus coeruleus and of behavior in the rat. Neuroscience Letters. 304. 73-76.

Kubo, K., Arai, O., Omura, M., Watanabe, R., Ogata, R., and Aou, S. 2003. Low dose effects of bisphenol A on sexual differentiation of the brain and behavior in rats. Neuroscience Research. 45(3):345-356.

Leranth, C., Hajszan, T., Szigeti-Buck, K., Bober, J., and MacLusky, N. J. 2008a. Bisphenol A prevents the synaptogenic response to estradiol in hippocampus and prefrontal cortex of ovariectomized nonhuman primates. Proceedings of the National Academy of Sciences. 105(37):14187-14191.

Leranth, C., Hajszan, T., MacLusky, N., and vom Saal, F. 2008b. Reply to Mathews: Misplaced focus on experimental detail. Proceedings of the National Academy of Sciencesof the United States of America (PNAS). 105(4):E98.

López-Casas, P. P., Mizrak, S. C., López-Fernández, L. A., Paz, M., de Rooij, D. G., and del Mazo, J. 2012. The effects of different endocrine disruptors defining compound-specific alterations of gene expression profiles in the developing testis. Reproductive Toxicology. 33(1):106-115.

Mathews, R. A. 2008. Publication incorrectly characterizes dose of bisphenol A. Proceedings of the National Academy of Sciences. 105(48):E97

Matt, D.W., Gilson, M.R., Sales, T.E., Krieg, R.J., Kerbeshian, M.C., Veldhuis, J.D. and Evans, W.S. 1998 Characterization of attenuated proestrous luteinizing hormone surges in middle-aged rats by deconvolution analysis. Biol Reprod 59: 1477-1482.

Muhlhauser, A., Susiarjo, M., Rubio, C., Griswold, J., Gorence, G., Hassold, T., and Hunt, P. 2009. Bisphenol A effects on the growing mouse oocyte are influenced by diet. Biology of Reproduction. 80(5):1066-1071.

Naciff, J. M., Hess, K. A., Overmann, G. J., Torontali, S. M., Carr, G. J., Tiesman, J. P., Foertsch, L. M., Richardson, B. D., Martinez, J. E., and Daston, G. P. 2005. Gene expression changes induced in the testis by transplacental exposure to high and low doses of 17alpha-ethynyl estradiol, genistein or bisphenol A. Toxicological Sciences. 86(2):396-416.

Nagel, S. C., vom Saal, F. S., Thayer, K. A., Dhar, M. G., Boechler, M., and Welshons, W. V. 1997. Relative binding affinity-serum modified access (RBA-SMA) assay predicts the relative in vivo bioactivity of the xenoestrogens bisphenol A and octylphenol. Environmental Health Perspectives. 105(1):70-76.

Nanjappa, M. K., Simon, L., and Akingbemi, B. T. 2012. The industrial chemical bisphenol A (BPA) interferes with proliferative activity and development of steroidogenic capacity in rat Leydig cells. Biology of Reproduction. 86(5):135, 1-12.

Negishi, T., Kawasaki, K., Ishii, Y., Kyuwa, S., Kuroda, Y., and Yoshikawa, Y. 2003. Perinatal exposure to bisphenol A and nonylphenol in F344 rats causes behavioral alterations of offspring. The 6th Annual Meeting of Japan Society of Endocrine Disrupters Research, December 2-3. Abstract PD-37.

Negishi, T., Kawasaki, K., Suzaki, S., Maeda, H., Ishii, Y., Kyuwa, S., Kuroda, Y., and Yoshikawa, Y. 2004. Behavioral alterations in response to fear-provoking stimuli and tranylcypromine induced perinatal exposure to bisphenol A and nonylphenol in male rats. Environmental Health Perspectives. 112(11):1159-1164.

Noda, S., Muroi, T., Mitoma, H., Takakura, S., Sakamoto, S., Minobe, A., and Yamasaki, K. 2005. Reproductive toxicity study of bisphenol A, nonylphenol, and genistein in neonatally exposed rats. Journal of Toxicologic Pathology. 18(4):203-207.

NTP (1985b). Bisphenol-A: Reproductive and Fertility Assessment in CD-1 Mice when Administered in the Feed. National Toxicology Program. Report NTP-85-192; Order No. PB86-103207 (NTIS) 1-346.

Pacchierotti, F., Ranaldi, R., Eichenlaub-Ritter, U., Attia, S., and Adler, I.-D. 2008. Evaluation of aneugenic effects of bisphenol A in somatic and germ cells of the mouse. Mutation Research/Genetic Toxicology and Environmental Mutagenesis. 651(1-2):64-70.

Park, D. H., Jang, H. Y., Park, C. K., Cheong, H. T., Kim, C. I., and Yang, B. K. 2004. Effect of bisphenol A administration on reproductive characteristics and blood metabolites in mice. Journal of Animal Science and Technology. 46(6):957-966.

Park, D.-H., Jang, H.-Y., Kim, C.-I., Cheong, H.-T., Park, C.-K., and Yang, B.-K. 2005. Effect of bisphenol A administration on reproductive toxicant of dam and sex ratio of pups in pregnant mice. Journal of Toxicology and Public Health. 21(2):161-165.

Patterson, T. A., Twaddle, N. C., Roegge, C. S., Callicott, R. J., Fisher, J. W., and Doerge, D. R. 2013. Concurrent determination of bisphenol A pharmacokinetics in maternal and fetal rhesus monkeys. Toxicology and Applied Pharmacology. 267(1):41-48.

Peknicova, J., Kyselova, V., Buckiova, D., and Boubelik, M. 2002. Effect of an endocrine disruptor on mammalian fertility. Application of monoclonal antibodies against sperm proteins as markers for testing sperm damage. American Journal of Reproductive Immunology. 47:311-318.

Putz O, Schwartz CB, Kim S, LeBlanc GA, Cooper RL, Prins GS. 2001a. Neonatal low- and high-dose exposure to estradiol benzoate in the male rat. I. Effects on the prostate gland. Biol Reprod 65(5):1496–1505.

Putz O, Schwartz CB, LeBlanc GA, Cooper RL, Prins GS. 2001b. Neonatal low- and high-dose exposure to estradiol benzoate in the male rat. II. Effects on male puberty and the reproductive tract. Biol Reprod 65(5):1506–1517.

Ragan et al. 2005. STP Position Paper: Ovarian Follicular Counting in the Assessment of Rodent Reproductive Toxicity. Toxicologic Pathology. 33: 409-412.

Rodríguez, H. A., Santambrosio, N., Santamaría, C. G., Muñoz-de-Toro, M., and Luque, E. H. 2010. Neonatal exposure to bisphenol A reduces the pool of primordial follicles in the rat ovary. Reproductive Toxicology. 30(4):550-557.

Saito, D., Minamida, G., Tani-Ishii, N., Izukuri, K., Ozono, S., Koshika, S., and Teranaka, T. 2003. Effect of prenatal exposure to dental composite resin monomers on testosterone production in the rat testis. Environmental Sciences. 10(6):327-336.

Song, Q., Hao, W., Shang, L., Wei, X., Jiang, J., Xing, L., Wang, H., Wu, S., Zhao, Y., Qiao, Y., and Hao, C. 2008. The effects of bisphenol A on reproduction and development of pubertal male rats. Aibian, Jibian, Tubian. 20(4):267-271.

Stump et al.,: Developmental neurotoxicity study of dietary bisphenol A in Sprague-Dawley rats. Toxicological Sciences 115, 167-182, 2010

Takao, T., Nanamiya, W., Nagano, I., Asaba, K., Kawabata, K., and Hashimoto, K. 1999. Exposure with the environmental estrogen bisphenol A disrupts the male reproductive tract in young mice. Life Sciences. 65(22):2351-2357.

Takao, T., Nanamiya, W., Nazarloo, H.P., Matsumoto, R., Asaba, K., and Hashimoto, K. 2003. Exposure to the environmental estrogen bisphenol A differentially modulated estrogen receptor and immunoreactivity and mRNA in male mouse testis. Life Sciences. 72:1159-1169.

Talsness, C., Fialkowski, O., Gericke, C., Merker, H.-J., and Chahoud, I. 2000. The effects of low and high doses of bisphenol A on the reproductive system of female and male rat offspring. Congenital Anomalies. 40:S94-S107.

Toyama, Y. and Yuasa, S. 2004. Effects of neonatal administration of 17beta-estradiol, beta-estradiol-3-benzoate, or bisphenol A on mouse and rat spermatogenesis. Reproductive Toxicology. 19(2):181-188.
Tyl RW, Myers CB, Marr MC, Castillo NP, Veselica MM, Joiner RL, et al. 2008a. One-generation reproductive toxicity study of dietary 17β-estradiol (E2; CAS No. 50-28-2) in CD-1® (Swiss) mice. Reproductive Toxicology. 25(2):144-160.

Tyl, R. W., Myers, C. B., Marr, M. C., Sloan, C. S., Castillo, N. P., Veselica, M. M., Seely, J. C., Dimond, S. S., Van Miller, J. P., Shiotsuka, R. S., Stropp, G. D., Waechter, J. M., and Hentges, S. G. 2008b. Two-generation reproductive toxicity evaluation of dietary 17ß-estradiol (E2; CAS No. 50-28-2) in CD-1® (Swiss) mice. Toxicological Sciences. 102(2):392-412.

Tyl, R. W. 2009. Basic exploratory research versus guideline-compliant studies used for hazard evaluation and risk assessment: bisphenol A as a case study. Environmental Health Perspectives. 117(11):1644-1651.
Walters, K.A., Allan, C.M. and Handelsman, D.J. 2012 Rodent models for human polycystic ovary syndrome. Biol Reprod 86: 1-12

Wistuba, J., Brinkworth, M.H., Schlatt, S., Chahoud, I., and Nieschlag, E. 2003. Intrauterine bisphenol A exposure leads to stimulatory effects on Sertoli cell number in rats. Environmental Research. 91(2):95-103.

Wu, J.-H., Jiang, X.-R., Liu, G.-M., Liu, X.-Y., He, G.-L., and Sun, Z.-Y. 2011 Oral exposure to low-dose bisphenol A aggravates testosterone-induced benign hyperplasia prostate in rats. Toxicology and Industrial Health. 27(9):810-819.

Xie, M.-N., Li, F.-J., Ding, B.-Q., Zhao, H. 2010. Effects of Bisphenol A exposed to lactating mice on the testis development of their male offspring. Chinese Journal of Zoology. 45(4):31-38.

Xu, X., Tian, D., Hong, X., Chen, L., and Xie, L. 2011. Sex-specific influence of exposure to bisphenol-A between adolescence and young adulthood on mouse behaviors. Neuropharmacology. 61(4):565-573.