Comments – Dec. 10-12 EDMVS meeting – JW Owens

This contains and summarizes my formal comments in several areas regarding subjects discussed at the EDMVS meeting on 10-12 Dec., 2003.

Phytoestrogens in Laboratory Diets

One of the vexing issues arising several times in the EDMVS history is the phytoestrogen content of laboratory rodent diets. To summarize comments that I have made in the meetings:

1. The degree of concern depends on the rodent species and the life stage. This is because a given dietary level of phytoestrogen is one determining factor. The other primary determining factors is the amount of diet consumed per unit body weight (grams of diet consumer per kg body weight). Together, these determine the dosage intake. The rat consumes less than the mouse, and the mature animal consumes less than the weanling. Intake may rise for the dam during pregnancy. On this basis, the "vulnerability" to phytoestrogens should occur in the following descending order: immature mouse > mature mouse > immature rat > mature rat. Hence, Thigpen's data with the immature mouse (where vaginal opening was accelerated with phytoestrogens) [Thigpen JE, Haseman JK, Saunders HE, Setchell KDR, Grant MF, Forsythe D. 2003. Dietary phytoestrogens accelerate the time of vaginal opening in immature CD-1 mice. *Comp. Med.* 53:477-485.] should not be immediately extrapolated at an equivalent level to older mice or to the rat.

2. The issue of sexual maturation will be complicated by caloric intake. This may be of equal or even greater importance than phytoestrogens. See:

Ashby J, Tinwell H. 1998. Oestrogenic activity of Burgen bread to female rats. *Hum. Exp. Toxicol.* **17:**598–599.

Ashby J, Tinwell H, Odum J. 2000. Uterotrophic activity of a "phytoestrogen-free" rat diet. *Environ. Health Perspect.* **108:**A12-13.

Ashby J, Tinwell H, Odum J, Kimber I, Brooks AN, et al. 2000. Diet and the aetiology of temporal advances in human and rodent sexual development. *J. Appl. Toxicol.* **20:**343-347.

Odum J, Tinwell H, Jones K, Van Miller JP, Joiner RL, Tobin RL, Kawasaki G, Deghenghi H, Ashby J. 2001. Effect of rodent diets on the sexual development of the rat. *Toxicol. Sci.* **61:**115-127.

Thigpen JE, Lockear J, Haseman J, Saunders HE, Caviness G, Grant MF, Forsythe DB. 2002. Dietary factors affecting uterine weights of immature CD-1 mice used in uterotrophic bioassays. *Cancer Detect. Prev.* 26:381-393.

3. The pertinent list of manuscripts, particularly for the uterotrophic, continues to grow. Overall, these data indicate no impact on the responsiveness of the uterotrophic bioassay when the response is measured with weak agonists such as BPA and NP. Kanno et al in contrast to others do indicate some influence on the baseline mean uterine weight. See:

Ashby J, Tinwell H. 1998. Oestrogenic activity of Burgen bread to female rats. *Hum. Exp. Toxicol.* **17:**598–599.

Degen GH, Janning P, Diel P, Bolt HM. 2002. Estrogenic isoflavones in rodent diets. *Toxicol. Lett.* 128:145-157.

Kanno J, Kato H, Iwata T, Inoue T. 2002. Phytoestrogen-low diet for endocrine disruptor studies. *J. Agric. Food Chem.* **50:**3883-3885.

Owens W, Ashby J, Odum J, Onyon L. 2003. The OECD program to validate the rat uterotrophic bioassay: Phase Two – Dietary phytoestrogen analyses. *Environ. Health Perspec.* 111:1559-1567.

Pocock VJ, Sales GD, Milligan SR. 2002. Comparison of the oestrogenic effects of infant milk formulae, oestradiol and the phytoestrogen coumestrol delivered continuously in the drinking water to ovariectomised mice. *Food Chem. Toxicol.* **40:**643–651.

Wade MG, Lee A, McMahon A, Cooke G, Curran I. 2003. The influence of dietary isoflavone on the uterotrophic response in juvenile rats. *Food Chem. Toxicol.* 41:1517-1525.

Yamasaki K, Sawaki M, Noda S, Wada T, Hara T, Takatsuki M. 2002. Immature uterotrophic assay of estrogenic compounds in rats given different phytoestrogen content diets and the ovarian changes in the immature rat uterotrophic of estrogenic compounds with ICI 182,780 or antide. *Arch. Toxicol.* **76:**613-620.

Conclusions

There current evidence does not indicate significant influences of laboratory dietary phytoestrogens in **rats**. Further, the uterotrophic and other data appear to be consistent with feeding studies using quantified amounts of genistein such as the work of Delclos, Wade, You, and others in various developmental and reproductive studies. Several preparatory actions could be warranted to follow-up on this issue: a) a review and possible comparison of the methods used to analyze dietary phytoestrogens and potentially some work with blinded samples (much depends on the comparisons made using different analytical techniques; this should be resolved first), and b) as the native phytoestrogens are glycosides, some experiments to confirm the expectation of no substantive difference between the glycoside and non-glycoside (these should be rapidly hydrolyzed and become bioactive; ensure that equivalent molar doses are used as the glycoside MW will be slightly higher).

Androgen Receptor (AR) Assay

I have some very real concerns with the direction of this program. I'll try to summarize in bullet form:

- There is a tendency to use more complex systems than are necessary to obtain what is effectively simple and preliminary information: does a test substance bind to the AR and with what approximate affinity does it bind? A bacterial

recombinant, even of just the ligand binding region of the AR, is sufficient to get this basic information.

- Using animals to get this information (prostate AR source) does not appear to be justified.
- The prostate system has drawbacks that haven't been vetted: a) the AR receptor concentration is not high in the prostate (compared to the ER in the uterus, at least), b) there are several competing/confounding receptors (GR, PR and so on), c) there are proteases which may produce an AR ligand region in this system unless significant precautions are taken, and d) to address the competing receptors, an inhibitor is introduced – more on this in the next point.
- The data presented do not show a perfect concordance between the bacterial recombinant and the prostate. In fact, they hint at a slightly greater sensitivity for the bacterial system. The plausibility of this is supported by the complexity of the prostate system and the addition of an inhibitor. An important note: the greatest impact (negative) would occur for low affinity ligands – aren't these a primary area of interest to the EPA?? Therefore, I would strongly recommend a careful look at this sensitivity and detection issue.
- The downside of the bacterial recombinant appears to be its stability, which may influence its reproducibility. Again, this area should be experimentally attacked.

Conclusions
The Agency needs to stand back and look at the offsetting plusses and minuses between the two systems. Long-term, the simplicity, possible sensitivity, and the lack of animal use of the bacterial recombinant has to be recognized. Investigate both the sensitivity issue and the stability issue.

Pubertal Assays
The comments on the pubertal assays are broken into several subject areas. The premise of my comments is that the pubertal assays are effectively the cornerstone of the USEPA's 'screening' strategy.

Specificity
The foremost concerns are that: a) the apical endpoints used in the pubertal assays will not provide clear mode of action or mechanistic information and b) the current data base has not yet addressed the overall "specificity" of the assay.

Again, the Agency is faced with the offsetting plusses and minuses to deal with and rationalize. A great deal of regulatory effort hinges on a limited number of apical endpoints for a "screen." Implicit to this commenter in this effort is a desire by the Agency to cover a very wide range of mechanisms, well beyond estrogens, androgens, and thyroid (EAT). The data to date indicate that the uterotrophic, Hershberger, and enhanced TG407, respectively, could adequately cover these three EAT mechanisms. EAT coverage could be done then in a more internationally harmonized fashion than the pubertals; the TG407 is required in most regulatory jurisdictions. Thus, the pubertals are in-effect a large addition to the numbers of animals to be used internationally. If the implicit criteria of the Agency is the desire to cover more mechanisms than EAT, then the

Agency should explicitly say so and enumerate the actual mechanisms. This would make the rationale for the pubertal assay transparent and make the validation objectives as to mechanisms that need to be included in the validation program clear.

The Agency also needs to wrestle with mechanistic concerns in two other aspects. The adult male assay appears to be more mechanistically specific due to its employment of the hormonal endpoints. Again, the offsets to be dealt with: the prepubertal animal has the apical endpoints to its advantage and the adult animal has the hormonal endpoints to its advantage. Once more, the Agency needs a clear rationale to justify its selection of one or the other. As part of this, if the pubertal is not mechanistically specific and the regulatory mandate from Congress is mechanistic screens, how will the Agency justify its selection and defend it against any legal challenges?

Negative Test Substances

The most outstanding gap as discussed at the meetings is the data with endocrine-negative compounds. This gets to the overall specificity of the data, and is an essential undertaking for the validation. As noted by several parties, there are a host of classical hepatotoxicants, nephrotoxicants, neurotoxicants, immunotoxiants, anti-hematopoietic agents and so on to choose from. This also confronts the Agency lack of clear definitions on what constitutes a potential endocrine disruptor and an actual endocrine disruptor due to the potential area of dispute between primary and secondary modes of action.

Some of the same dispute has arisen in the OECD's Peer Review of the uterotrophic validation. One panel member and one panel observer have been very critical of the use of only a single negative chemical in that program (dibutylphthalate). The Agency should expect similar, even expanded, questions in regards to the pubertals based on the multiple mechanisms that the Agency apparently wishes to address and the apical endpoints lack of mechanistic specificity. Despite the costs, approximately five negative compounds, acting by other general toxicity mechanisms, should be considered. If rationales can be developed not to use as many as five negatives, then the technical, logistical, and cost points constituting the rationale should be made explicit.

Other Test Substances

In the estrogen area, I would nominate nonylphenol (NP). It is now very well characterized with two published multi-gens and a third just completed (RTI). NP has been used in the uterotrophic (positive) and 407 programs (positive at high doses at or approaching the MTD); Japan has generated Hershberger data (negative); and a recent adult male assay has been done (positive). Further, NP is intermediate in potency between methoxychlor (positive with several endpoints in the pubertals) and BPA (negative in the pubertals); so this will help establish the response boundaries of the pubertal assays.

Similarly, a review of moderate potency substances for other mechanisms should be done. Again, this applies to the response boundaries or degree of sensitivity the pubertal will have with various mechanisms.

Thyroid and SOPs (Standard Operating Procedures)
The phenobarbital results are a concern for application of the pubertals to thyroid toxicants [a more robust reponse was expected, see McClain RM, Levin AA, Posch R, Downing JC. (1989). The effect of phenobarbital on the metabolism and excretion of thyroxine in rats. *Toxicol. Appl. Pharm.* **99:**216-228 plus several publications from Klassen's lab with pehnbobarbital]. In particular, questions are raised in regards to both the histopathology and the T_4 and TSH assays. I have forward background data from the OECD enhanced TG407 studies to Drs. Cooper and Stoker. In those studies, the histopathology was clearly more consistent, picking up PTU at very low doses and also the effects of thyroxine, methyl testosterone and *p,p'*-DDE on the thyroid in both of the duplicate studies with each compound. In contrast, the thyroid hormone analyses were somewhat erratic. Further, TSH CVs were high, indicating limited power. As with the phenobarbital studies, there was precious little description of the SOPs in the TG 407 protocols in regards to the thyroid hormones – including how the animals were handled to reduce stress and circadian issues with thyroid hormones, blood sampling, and the thyroid hormone analyses themselves (reproducibility, LOD/LOQ, and so on). Note that in a sidebar with Dr. O'Connor of Dupont, that he indicated their SOPs included transporting the animals early into the necropsy room and allowing a minimum of one hour's rest and conducting the necropsy in a short two hour time window to address circadian concerns. This is consistent with reviews on issues with thyroid hormones:

Döhler KD, Wong CC, von zur Mühlen A. 1979. The rat as model for the study of drug effects on thyroid function: Consideration of methodological problems. *Pharmac. Ther.* **5:**305-318.

Davies, DT. 1993. Assessment of rodent thyroid endocrinology: Advantages and pitfalls. *Comp. Haematology Int.* **3:**142-152.

These same issues and concerns were noted in an EPA sponsored workshop a few years ago:

DeVito M, Biegel L, Brouwer A, Brown S, Brucker-Davis F, Cheek AO, Christensen R, Colborn T, Cooke P, Crissman J, Crofton K, Doerge D, Gray E, Hauser P, Hurley P, Kohn M, Lazar J, McMaster S, McClain M, McConnell E, Meier C, Miller R, Tietge J, Tyl R. 1999. Screening methods for thyroid hormone disruptors. *Environ Health Perspect* **107:**407–415.

My recommendation is to carefully review the SOPs, histopathology, and analytical techniques used in the EPA studies. Additional studies may then be clearly justified or, alternatively, rationales for the lack of further work can be made explicit.

International Harmonization
Given that most other regulatory jurisdictions appear to be leaning towards the uterotrophic, Hershberger, and enhanced TG 407 assays for the endocrine screening programs, how will the EPA use and apply data from these? If data are available indicating the compound is not estrogenic, (anti)androgenic, or thyroid toxicant from these assays, what is the position of the EPA in using these available data and in the need

for applying the pubertal assays (now redundant in several major aspects) to a compound? The question is raised as the likelihood is growing that there will different, non-harmonized regulatory approaches as EPA pursues its different path. This also raises the issues of animal welfare and resource use that the Agency may need to address.

Other Issues
Many of the endpoints used by the pubertal assays have background data indicating that the dam (or also called the litter) is a variable. What precautions will be necessary or not in the randomization of the animals into treatment groups??

OECD fish
Note is taken of the following points as related in the report on the OECD fish work:

1. I concur that 'pruning' of the endpoints list is warranted. My previous comments had, for example, questioned the use of the gonadal somatic index based on EPA work with the fathead.

2. I fully support the effort to conduct a professional review of the histological findings. The use of the fish gonads is challenged by the lack of guidance and the lack of thorough baseline and toxicological understanding. The limited reproducibility, the difficulties in the OECD work and in the USEPA work with the ovaries, and the potential differences in pathologist interpretation all indicate full investigation of this area is necessary before regulatory applications are undertaken.

3. Consistent with my previous comments, I support the OECD work groups recommendation concerning voluntary nature of the fecundity measures. Again, fecundity is a fundamental apical measure of reproductive function (that is, it is a Tier 2 test endpoint and not a Tier 1 endocrine screen endpoint), and it is non-specific as to any mechanism (endocrine or otherwise). The EPA should also make this point voluntary or at least acknowledge that it is not to be used as a screening trigger. The best overall approach is to use the fecundity in a Tier 1.5 setting where it can be applied to risk assessment for compounds in its proper context.