Inside EPA's
# Risk Policy Report

September 19, 1997 • Volume 4, Number 9

## INSIDE THIS ISSUE



## Administration Offers Mixed Reviews On Senate Risk Bill

OMB's Katzen lauds bipartisan Senate bill's intent, but raises seven major areas of concern

— *page 3*

## EPA Enforcement Office Will Set Priorities Using Risk-Based Tools

Enforcement program reviewing data-rich, multi-media risk models

— *page 7*

## Point-Counterpoint: Can Today's Risk Tools Handle Endocrine Mimics?

Experts argue the merits of traditional risk assessment tools for assessing endocrine disruptor risk

— *page 30*

**Edited exclusively for scientists interested in environmental policymaking and policymakers interested in science**

*NOTE: In recent* Risk Policy Report *articles, researchers and activists have raised concerns about the adequacy of traditional risk assessment approaches for coming to grips with the unique hazards that may be posed by endocrine disrupting chemicals (*Risk Policy Report, *July 18, p10; May 23, p12). Expert proponents on two different sides of the debate offer their views on this important issue in this special section featuring in-depth essays and rebuttals that touch on emerging data related to "U" and "inverted-U" shaped dose-response curves, effects (adverse and otherwise) and longstanding assumptions of toxicological research.*

# CAN TODAY'S RISK ASSESSMENT PARADIGMS DEAL WITH ENDOCRINE ACTIVE CHEMICALS?

## By James C. Lamb

Is the current risk assessment process adequate for managing the risks of chemicals that might act through the endocrine system? Are toxicology and current study designs able to identify and predict the hazard of chemicals that act through the endocrine system? These are interesting and provocative questions raised by new data on hormonally active chemicals.

Recently, risk assessment and toxicological methods have come under attack as part of the debate about endocrine active chemicals (EACs). The EAC debate is complicated, but generally comes down to three mechanisms, including when chemicals cause adverse effects by acting like natural hormones at the receptor, when they cause adverse effects by interfering with natural hormone action at the receptor, or when they cause adverse effects by affecting the endocrine system through a non-receptor event.

Every cell in the body is affected by some part or parts of the endocrine system. Therefore, it does not take much effort to infer an endocrine mechanism for nearly any outcome.

Decreases in sperm count, changes in reproductive organ structure, increases in various types of cancer, and behavioral changes are among the outcomes reported in humans and animals. It has been proposed that those changes could be caused by EACs. The scientific debate has been intense, centering on critical issues such as:

Is there a threshold for the effects of EACs?
Is the dose response curve fundamentally different from other toxicants? and
Do environmental agents with weak hormonal activity cause adverse effects at relevant exposures?

### Threshold

Is there a threshold? It is an emotionally charged question that may not be answerable. It may be unnecessary to prove or disprove that a threshold exists to be able to manage risks, but the dose response models that are used by risk assessors are guided by this issue. It has been said that, because the hormone system is already "turned on" by endogenous hormones, "some degree of disruption will occur at any dose, although a particular measurement may not reveal a change, since there can be no threshold for a response. . . already operating above the threshold" (vom Saal et al., In press, *Toxicol. Envtl. Health*). Not everyone agrees.

Lucier et al. (1993) have presented a strong case that supports this position, in part. First, a very simple system with no other input or output, adding a like molecule should add proportionately to the response. However, Lucier et al., note that occupancy of a single receptor, in a system of tens of thousands of receptors per cell, is unlikely to cause a response that can be detected. Thus, even if there is theoretically or actually no threshold for the biochemical response under those circumstances, the practical consequences are that a threshold exists for measuring a response in the whole animal.

The simple case of adding the natural ligand to a closed system is not representative of the typical EAC. Most synthetic estrogenic chemicals are weaker agonists and can act as antagonists in the presence of a potent endogenous ligand, like estradiol. Biological responses to endogenous and exogenous estrogens are diverse; thus, one cannot assume that they will act exactly like the endogenous ligand. As noted by Lucier et al., the diversity of response is probably due to tissue- and cell-specific factors that affect the receptor ligand interaction. One could even have different responses because of different receptors that may not interact proportionately with ligands.

The simple case assumes that the body does not monitor the levels of ligand, and does not adjust accordingly; then any amount of hormone added, from any source, would add to the concentrations in the target cell. However, the intact organism has feedback loops for many hormone systems and increased concentrations are met with decreased synthesis or release, or

*continued on page 32*

*James Lamb is Vice President of Toxicology and Environmental Sciences at the consulting firm Jellinek, Schwartz & Connolly, Inc., in Arlington, Virginia.*

# LOW DOSE EFFECTS OF ENDOCRINE DISRUPTORS- A CHALLENGE FOR RISK ASSESSMENT

By Daniel M. Sheehan and Frederick S. vom Saal

**The views expressed here are those of the authors and do not represent the opinions or policies of the Food & Drug Administration or the University of Missouri.*

## Introduction

We believe that current approaches to risk assessment may not accurately predict risk based on recent research on environmentally relevant exposures to endocrine disrupting chemicals.

Toxicologists have long studied adverse effects of hormonally active chemicals which disrupt the functioning of the endocrine system. Most utilize accepted study designs and data analysis, which include high dose testing to predict low dose outcomes and the assumption of a threshold dose and a monotonic dose-response curve. Recently, chemicals found globally in the environment have been associated with adverse effects in wildlife at concentrations lower than doses typically examined in toxicological studies (Colborn and Clement, 1992). Such findings conflict with toxicology studies which conclude that these concentrations are safe. At the same time, significant advances in understanding hormone action have occurred. Here, we will explore some challenges to traditional toxicology which have explanatory power to help resolve the conflict among the several disciplines concerned with endocrine disruptors.

There is consensus that the primary concern about endocrine disruptors is exposure during critical periods in development when organs are differentiating (Colborn and Clement, 1992). This concern has been ignored in some recent reviews. In addition, the phrase "endocrine disruptor" has been replaced by some authors with the term "endocrine modulators," which conveys an entirely different meaning. Specifically, endocrine modulation applies to adulthood, when effects are typically transient (referred to as activational effects) and not adverse. In marked contrast, chemicals that irreversibly disrupt normal endocrine functioning and thus organogenesis due to exposure during critical developmental periods or that induce adverse effects in adults are "endocrine disruptors." The greatest threat posed by such chemicals is developmental exposure, because adult functions, and particularly reproductive capability, are irreversibly disrupted.

We will begin by establishing the lowest concentration of estradiol that leads to responses in vitro and the use of this and other information to predict low doses that induce in vivo responses. Results demonstrate that developmental sensitivity to estrogenic endocrine disruptors occurs at much lower doses than doses commonly used in toxicological studies because the latter are selected based on the assumptions of a threshold and a monotonic dose-response curve. We will then examine the implications of ligand and tissue-specific effects of estrogens, summarize results showing that the assumptions can be incorrect, and conclude with a summary of the characteristics described for endocrine disrupting chemicals (EDCs) that are the basis for the challenge before toxicologists.

## Low Dose Activity of Estrogens in Vitro

There is sometimes confusion with regard to the concentration of a hormone which is biologically active. MCF-7 cells proliferate in response to concentrations of free (unbound) estradiol in the low picomolar range. A useful biochemical parameter (the dissociation constant) Kd — the concentration of estradiol required to occupy 50% of receptors — is sometimes erroneously assumed to be the lowest biologically active concentration. As the Kd for estradiol is about two orders of magnitude higher than the concentration of estradiol which stimulates cell proliferation, this low dose proliferation response might appear inconsistent. However, studies have shown that only 1-5% of the receptors need to be occupied to induce a half-maximal proliferation response. While xenoestrogens (environmental estrogens) of lower potency require correspondingly higher concentrations of free chemical to occupy the small fraction of the receptor population necessary to drive a half-maximal response, recent studies have shown that a similar small percent of receptors is typically required for this to occur. In addition, very low doses of xenoestrogens have been shown to be active in vivo (Nagel et al., 1997; vom Saal et al., 1997).

## Predicting Activity of Very Low Doses of Xenoestrogens

Toxicologists often focus on "low doses" in the part per

*continued on page 35*

*Daniel M. Sheehan is Research Biologist, The National Center for Toxicological Research, USFDA, Jefferson, AR.*
*Frederick S. vom Saal is Professor of Biology, Division of Biological Sciences, University of Missouri, Columbia, MO.*

## Point - Counterpoint . . . continued from page 31

increased excretion, of the endogenous ligand. Such feedback loops can certainly be overwhelmed, but only above certain "threshold" dose levels.

The endocrine system is not just a ligand and a receptor. The body has many levels and modes of control. Receptors are different for different cells. Hormone response elements, which influence cell response, are also diverse. There are controls at the tissue, cellular, and molecular levels. Even receptor levels are not constant in a cell. The body has certain barriers to absorption and transport, and the mode of administration affects the amount of exogenous material that is available. Adding ligand to an in vitro system bypasses important processes. For example, if a ligand is not absorbed through the skin, gut, or lungs, the dose, on a mg/kg body weight basis, will not add proportionately to the concentrations at the target site. Disposition, metabolism, and protein-binding can dramatically affect the target dose. The models, therefore, are more complicated than a simple addition. Some processes will decrease and others may increase the target dose and response.

### Shape of the Dose Response Curve

One controversial point is whether the shape of the EAC dose response curve can invalidate current toxicological study designs. The most heavily cited examples are the inverted-U or U dose response curves. It has been incorrectly presumed that toxicologists assume increased dose always results in increased response, to a maximal response. The dose response relationship is often observed within the confines of particular studies as a sigmoidal curve. There are three phases to such a curve. The lower end of the curve shows little or no measurable response, the middle range shows a response increasing proportionally to increased dose, and the plateau of the curve reaches a limit of no additional response regardless of dose.

The inverted-U is nothing more than a second response limiting and even decreasing the initial maximal response. For example, developmental toxicity studies often show increases, followed by decreases, in developmental abnormalities with increasing dose. This occurs because fetal mortality increases and masks the developmental effect. Another example is where body weight increases for some offspring because they have greater access to nutrients as others perish due to the effects of a developmental toxicant. As the dose increases further, the toxicity can decrease the weight for the survivors as well. Inverted-U dose response curves are common with in vitro systems. For example, low doses of estrogens may cause increased cell proliferation, but even higher doses often decrease that proliferation, and this may be the result of cytotoxicity or differentiation.

A U-shaped dose response curve can be another type of interaction of two dose response curves. This is not new to toxicology. For example, *Casarett and Doull's Toxicology: The Basic Science of Poisons* (Klaassen et al., 1996) illustrates a U-shaped dose response to a nutrient. The text shows a curve that starts with a high incidence of death at insufficient levels of nutrients. As the nutrient levels increase, survival improves and death falls into the normal range (or region of no effect), followed by an increase in death as the dose of the nutrient itself becomes toxic.

Regulatory toxicology studies usually include a top dose level that is high enough to definitely cause some toxicity, an intermediate dose level that is somewhat lower and which may be toxic, and the lowest dose level, which is supposed to approximate a no observed adverse effect level (NOAEL). The concern has been raised that the lowest level tested may be too high to observe outcomes at environmentally relevant levels. Changes will be missed, because the NOAEL for the very high dose effect is still above the effect level for the lower end of the inverted-U, but the effect is masked by the start of the high dose effect.

For example, what if prostate weight were increased by very low levels of a chemical, but higher doses of that chemical actually decreased prostate weight? Will outcomes be missed by a toxicology study design by current standards? How will such outcomes be handled in risk assessment? Are the changes important? The clear answer is, it depends.

Organ weight is a nonspecific endpoint that can be increased by normal (e.g., growth) or abnormal (e.g., swelling) mechanisms. However, even routine toxicology studies used for risk assessment purposes do not rely on organ weight alone, but use histopathology or various clinical chemistry endpoints, even when organ weight is in the normal range.

A low dose increase in organ weight might be missed, if, at even higher levels, the organ weight decreases. At some point on the dose response curve, the initial elevation will drop and pass through the normal range. At that point, the weight changes could be missed and mistaken for a NOAEL. But if other endpoints are taken into consideration one could determine whether an outcome is missed, or even important.

### Adverse Effect

Many factors determine whether a change in a study parameter is an adverse effect: What is the nature of the study; are the group sizes adequate; has the study been reproduced in that laboratory or other laboratories; is the route or mode of administration an appropriate model; are the endpoints relevant, reliable, and reproducible; are the changes outside the normal range or the range for historical controls; and do the changes reflect adverse effects on structure or function?

Is a relatively small increase in mouse prostate weight an adverse effect? It has been proposed that a shift in weight predisposes the prostate for disease (vom Saal et al., 1997). Prostate weight is influenced by many natural factors. Location in utero has been reported to affect mouse prostate weight. A male between two females (2F) is said to have a larger prostate than a male between two males

(two males because of endogenous estrogen exposure). This shift in the prostate weight can be mimicked by diethylstilbestrol or estradiol. Investigators have hypothesized that this shift to a higher weight will predispose these males for disease (vom Saal et al., 1997).

Humans have a relatively high incidence of prostate disease, both as benign prostatic hypertrophy (BPH) and prostate cancer. However, data do not yet show a relationship between prenatal diethylstilbestrol (DES) exposure and prostate cancer (Santti et al., 1994); this should be closely watched. It has been proposed that increases in mouse prostate weight from estradiol or DES will serve as a model for such changes in humans (vom Saal et al., 1997). However, the mouse may not be the best model for human prostate disease for at least two reasons:

Rather than concentrating on creating simplistic assumptions about threshold and dose response, researchers should focus on developing additional information for biologically based dose response assessment of EACs.

First, the background incidence of prostate cancer in mice is very low. For example, out of 2,343 control mice from the National Toxicology Program (NTP) historical database, no cases of prostate cancer were reported (Haseman et al., 1984). Of course, these controls randomly included 2M, 2F, and 1MF males. The absence of prostate cancer in any group, including 2F, makes it impossible to conclude that one sub-population is predisposed to disease over another.

Second, the mouse prostate does not appear to be sensitive to chemical toxicity. The NTP has an extensive database of long-term studies. A search of the abstracts of the NTP bioassay reports (on the NTP Web page from over 400 long-term studies in rats and/or mice on diverse types of chemicals) indicates that the mouse prostate has rarely been seen as a target. In fact, no increases were reported in mouse or rat prostate cancer after chemical exposure. Only two cases of non-neoplastic effects were found in mice. No BPH increases were reported. Phenylephrine hydrochloride and furosemide caused inflammation of the mouse prostate. The mouse prostate does not appear to be a sensitive target for adverse effects.

## Conclusions

Current risk assessment and toxicology methods can be improved with information on mechanism of toxicity. Rather than concentrating on creating simplistic assumptions about threshold and dose response, researchers should focus on developing additional information for biologically based dose response assessment of EACs. Improved information about dosimetry and mode-of-action should improve the risk assessment process.

EACs act through natural signaling systems; therefore, they have the potential to stimulate responses through those systems. Such responses would not necessarily be adverse. In fact, some of the reported effects of EACs, like the phytoestrogens, are beneficial (e.g., breast cancer decreases or blood lipid changes) having the potential to inhibit natural responses if they compete with the endogenous ligand. However, they may also cause responses through mechanisms that do not involve the natural receptor system. Thus, one cannot generalize that the response is necessarily adverse.

Before advocating additional routine testing at dose levels below the current NOAEL, because of concern about the possible inverted-U dose response, the phenomenon needs to be better understood. Even the responses at the lower end of the inverted-U have no effect levels. If the effects are adverse, then more sensitive endpoints may have to be assessed in certain cases and the risk assessment can be conducted using the new, more sensitive endpoint. However, there may be changes that are judged not to be adverse. The magnitude of the response, or nature of the response, may or may not lead to risk management control. Like cancer and other toxic effects, response to EACs can be a multistage process with various manifestations, such as changes in biochemical or cellular parameters that precede toxicity. The key is to determine the most effective way to identify adverse responses, without expending huge amounts of resources and time chasing effects that may not be there or may not be adverse.

## Bibliography

Klaassen, C.D., et al. (1996) The Basic Science of Poisons. *Casarett and Doull's Toxicology,* 5:18-25.

Haseman, J.K. et al. (1984) Use of Historical Control. Data in Carginogenicity Studies in Rodents*. *Toxicologic Pathology,* 12:126-135.

Lucier, G.W., et al. (1993) Receptor Mechanisms and Dose—Response Models for the Effects of Dioxins. *Environmental Health Perspectives,* 101:36-44.

Santti, R., et al. (1994) Developmental Estrogenization and Prostatic Neoplasia, 67-78.

vom Saal, F.S., et al. (1997) A Physiological and Ethotoxicological Approach to the Study of Estrogenic Endocrine Disruptors on Reproductive Organs and Behavior, *Toxicology and Industrial Health,* In press.

vom Saal, F.S., et al. (1997) "Prostate Enlargement in Mice Due to Fetal Exposure to Low Doses of Estradiol or Diethylstilbestrol and Opposite Effects at High Doses." *The National Academy of Sciences,* 94:1-6.

## Daniel Sheehan and Fred vom Saal Respond to James Lamb's Agruments

Dr. Lamb identifies three critical issues that he contends are the subject of debate among toxicologists. For the first two issues, the existence of thresholds and shape of the dose response curve, both theory and data reveal that there can be no threshold dose and non-monotonic curves. The answer to the third issue "do environmental agents with weak hormonal activity cause adverse effects at relevant exposures?" is that our findings are consistent with and offer an explanation for the findings in wildlife.

### Threshold

The threshold issue is a scientific, not emotional, question. A practical threshold is a function of statistical resolving power in the determination of a NOAEL and has no relevance for a curve that regresses to background incidence at zero dose. How can low levels of risks be managed under the assumption of a practical threshold used to derive a safe dose that is not tested? Additionally, for non-monotonic curves, this procedure can give entirely wrong answers. A hidden assumption is that there is a level of risk in some circumstances and it is at an acceptable level although we do not estimate it. It is well recognized that a NOAEL is a statistically limited value while a threshold is a concept; the concept of a threshold is used to justify the use of a NOAEL. The circular reasoning here is apparent. We agree that diverse mechanisms underlie different responses, but it is unclear how this implies that a threshold will be observed for these responses. In fact, diverse mechanisms increase the likelihood that one or more responses will fail to have a threshold dose. In vitro systems were not used to argue against a threshold, but to select a dose range for in vivo studies. For estradiol, the dose of target was determined as suggested by Lamb; this integrates all the events that could modify proportionality of administered dose to target dose. Finally, to identify other low dose effects we need to test for them. If the data analysis in specific situations shows simple additivity of endogenous and exogenous doses or extrapolates to the control response at zero added dose, why should we generically argue to the contrary? Let the data speak, and discard the blindly formulatic approach. Lamb proposes that organisms have feedback systems that can only be overwhelmed above threshold dose levels. While this may apply to individual adults, they simply do not apply during development when components of the endocrine system are changing continuously. This reality is fundamentally incompatible with the concept of homestatic regulation. For example, when serum estradiol in male mouse fetuses was increased 8-fold above controls, serum testosterone levels increased, an outcome opposite from that found in adult males (vom Saal et al., 1997). Feedback mechanisms governing placental endocrine function are not well understood for any mammal. Fetuses and children do not function as little adults.

### Shape of the dose-response curve

The assumption in risk assessment is that there is no change of sign for the slope across the dose range. Because application of a safety factor can provide a low dose that carries increased risk above the NOAEL, denial of this assumption invalidates conclusions regarding safe doses for chemicals inducing such responses. Lamb uses high dose examples of inverted-U of U dose-response curves which are well known and irrelevant to the low dose issue. The low dose example of nutrients is similar to our examples of vitamin A and thyroid hormone. What is astonishing is that toxicologists consider them to be idiosyncratic and rare, and have not developed a study design to detect low dose effects. For this reason, we simply don't know how many chemicals on the market display these curve shapes. Toxicology studies use a few high doses (3 doses are typical), and do not produce information regarding the potential for adverse effects by the much lower doses of chemicals consumed by people and wildlife. Only 2% of toxicological studies examine effects of doses below the NOAEL (Calabrease, E.J. and Baldwin, L.A. The dose determines the stimulation (and poison): Development of a chemical hormesis database. Int. J. Tox, in press, 1997.) Lamb suggests that only organ weight was examined. However, low-dose endpoints other than prostate weight were presented. Finally, for low dose estrogen sex reversal in turtles, organ weight is irrelevant, since ovaries develop instead of testes. Lamb also raises the question of how such outcomes will be handled in a risk assessment. The answer in our paper is to use a predetermined acceptable risk, as opposed to assuming safety at some low dose which is not tested. Finally, it is asserted that other endpoints could determine whether an outcome is missed when an effect level is mistaken for a NOAEL. But if the other endpoints are less sensitive than the original, all will be missed. It is certainly more logical to examine doses below the NOAEL where effects will be seen even though they exist at lower doses. However, additional endpoints could be useful in the dose region below the NOAEL.

### Adverse effect

Lamb asks "Is a relatively small increase in mouse prostate weight an adverse effect?" but then inappropriately focuses on cancer, high dose effects of DES, and the species-specific nature of effects. The major assumption is that effects in rodents predict those in other species, and that such predictive effects may not even be of the same type. For example, finding that fetal exposure to a very low dose of estrogen can lead to hyperplasia and an increase in hormone responsiveness in the mouse prostate could be relevant for prostate disease in humans, such as benign prostate hyperplasia.

### Conclusions

Lamb implies that we have created simplistic assumptions about threshold and dose response. In stark contrast, we assert that the simplistic assumptions on these issues were made about five decades ago and have not been carefully re-examined since. In any such examination, experimental models must be defined that allow clear testing of the traditional assumptions as hypotheses. There must be evidence that endogenous chemicals are causing a response and that dosing with the same chemical increases the response in a manner incompatible with a threshold dose and/or that as dose is lowered below the NOAEL, there is a change in the sign of the slope. As this has been accomplished in the two examples presented here, such issues as the complexities of the response system and the general applicability of these principles can now be examined. Such studies could be improved by using, where possible, biologically based dose-response assessments as the prostate studies have demonstrated. Breast cancer is another example raised by Lamb. While the proposal that phytoestrogens may lower breast cancer risk is attractive, equally credible evidence shows that phytoestrogens are developmental toxicants. Additionally, because endogenous estrogens are causally associated with human breast cancer risk, exogenous estrogen agonists will increase population risk, even if homeostasis provides protection to some individuals. This is because some members of the population are already above the threshold that can result from homeostatic mechanisms. In general, however, we agree that the responses should not be assumed to be adverse but be determined for each measured response. Routine testing at doses be-

low the NOAEL is an expedient route to examine the low dose phenomena we have described, and in particular to identify adverse responses not seen in the inadequate high dose testing procedure. We have spent about 50 years not spending money to do so because, in part, we repeatedly verified the assumptions by failing to design adequate procedures to test these assumptions, analogous to the drunk looking for the lost key under the lamp post. This is implicitly acknowledged in Lamb's last paragraph: "... more sensitive endpoints may have to be assessed in certain cases and the risk assessment can be conducted ..." using these endpoints.

## Point - Counterpoint . . . continued from page 31

million range. However, evidence is emerging that the endocrine system operates at dramatically lower concentrations of these hormones. With regard to predicting effects at these levels, toxicologists need to begin to explore doses far below this range.

We know that xenoestrogens bind with varying affinity to estrogen receptors, and that this factor partially determines potency; therefore a physiologically-based method that takes this factor into account should be useful to predict the dose at the target tissue required to stimulate estrogenic responses in the tissue in vivo. Information concerning metabolism and affinity for serum binding proteins (albumin and specific estrogen-binding glycoproteins, as well as other serum factors that influence the bioavailable fraction of circulating xenoestrogen are similarly useful for predicting the proportion of administered dose that reaches specific target tissues. In contrast, in the absence of any information regarding the mechanism of action, metabolism or plasma transport of a chemical (the general situation in toxicology), this is not possible. A recent study serves as an example of this new approach.

One example of the dramatic sensitivity of fetuses to natural estrogen is recent research with fetal mice indicating that only 0.2%, or 0.2 pg/ml (about 0.7 picomolar) of endogenous serum estradiol-17ß is free (this is the bioactive fraction of total circulating estradiol for many tissues) (vom Saal et al., 1997), similar to findings in rats (Montano et al., 1995). This concentration is in the dose range which also stimulates proliferation in MCF-7 cells. Evidence shows that fetuses with different endogenous estradiol concentrations in this range show different behaviors and reproductive organ weights as adults (vom Saal et al., 1997). It is also this concentration of estradiol with which xenoestrogens are presumed to act additively (that is, without a threshold dose; Gaylor, et al., 1988) in increasing the total estrogenic bioactivity leading to changes in responses. Increasing free serum estradiol experimentally by 50% in fetal male mice from 0.2 to 0.3 pg/ml significantly increased the fetal formation of prostate glands, permanently enlarged the adult prostate (by 30%) due to hyperplasia, and induced a 2-fold increase in adult prostatic androgen receptors per cell (vom Saal et al., 1997). This concentration increase in estradiol stands in contrast to typical doses used in toxicological studies. It also serves as the reference increase in estradiol which significantly disrupts normal fetal development, just as does the MCF-7 cell growth response to estradiol at less than 0.1 pg/ml culture medium (0.1 ppt) (Welshons and Jordan, 1987).

A permanent increase in prostate size (vom Saal et al., 1997) and in territorial marking behavior (vom Saal et al., 1995) in adult male mice resulted from maternal administration of 20, 200 and 2,000 parts per trillion (ppt) per day of DES during days 11-16 of pregnancy. These low dose effects of DES were predicted based on findings from prior DES bioavailability studies as well as the increase (above endogenous levels) in serum estradiol previously shown to produce a significant change in development of the reproductive system in male mice (vom Saal et al., 1997).

The actual in vivo potency, relative to estradiol, of administered doses of DES and other chemicals with limited binding to plasma steroid binding proteins will be underestimated if serum binding is not taken into account (Sheehan and Young, 1979; Brenham, et al., 1987; vom Saal, et al., 1995; Nagel et al., 1997). This is particularly important during development when dramatic increases in plasma binding proteins over adult levels keep the concentration of free, bioavailable estradiol very low (Montano et al., 1995; vom Saal et al., 1997). Specifically, whereas 0.2-0.3% of total circulating estradiol is free in fetuses, 4% is free in adults (Montano et al., 1995). Thus, during development, the potency of DES, relative to estradiol, can be as much as 100-fold greater than in adults (Branham et al., 1988). Taken together, all of the above findings and calculations suggested that the 20 ppt dose of DES would have an effect on the developing prostate in male mice similar to the previously described 0.1 pg/ml increase in free serum estradiol, and this prediction was confirmed (vom Saal et al., 1997).

### Non-monotonic, Inverted-U Dose-Response Curves

The issue of the importance of the selection of doses in toxicological studies has recently received renewed attention. The importance of dose in toxicological experiments was emphasized in a study in which a high (200 ppb/day) dose of DES was fed to pregnant mice and was found to inhibit normal development of the prostate in male offspring. In contrast, a 10,000-time lower dose of DES (20 ppt/per day based on body weight) led to the opposite response, a permament increase in prostate size in male offspring; the dose-response curve based on a 5-log range of doses of DES formed an inverted U (vom Saal et al., 1997). This finding is of considerable importance with regard to the traditional use of high doses of chemicals in toxicological studies, and extrapolation to predict effects of

lower doses based on the assumption of a monotonic dose-response curve for systemic toxicants.

A non-monotonic, inverted-U dose-response curve is unlikely to occur for all responses to estrogens or other hormones or hormone-mimicking chemicals. Indeed, no single dose-response function is likely to be applicable to all endocrine disrupting chemicals. However, inverted-U functions are not uncommon in endocrinological studies in which responses are mediated by receptors for hormones and other intercellular signalling molecules (vom Saal et al., 1997). The conclusion from these studies is that responses to endocrine disruptors cannot be assumed to be monotonic across a wide dose range, and unique outcomes may occur in response to environmentally relevant doses of endocrine disruptors that may be much lower than doses used in toxicological research. We use environmentally relevant here to refer to doses of chemicals that are entering the bodies of humans and wildlife via contact dermal with food, water and air. What follows is a discussion of the methodological and testing implications of findings that extrapolation from high dose studies to predict low dose effects of endocrine disrupting chemicals is sometimes problematic. We propose that this process can lead to false conclusions concerning safety.

Unique outcomes may occur in response to environmentally relevant doses of endocrine disruptors that may be much lower than doses used in toxicological research.

**Ligand and Tissue Specific Effects of Xenoestrogens**

For a single estrogen, potency may vary greatly for responses in the same (or different) tissues or species, despite the same level of receptor occupancy. This suggests that events downstream from receptor binding provide further control of response intensity. Examples include inhibition of gland genesis (Branham, et al., 1985) or estrogen receptor concentration (Medlock, et al., 1991) in the rat uterus which are up to an order of magnitude more dose sensitive than uterine weight gain. Across species and tissues, the effect level of bisphenol A is 5-500 mg/kg in uterotrophic assays in rats, while it is 2 ug/kg in the developing mouse prostate, a difference of 2,500-250,000-fold. Thus, use of a single endpoint is inadequate to characterize estrogens; rather, the most sensitive and relevant responses and adverse effects across species need to be defined. Additionally, the pharmacological nature of a hormone also shows cell type- and species-specificity. The realization that not all estrogens are created equal was greatly accelerated by studies of tamoxifen, which is in widespread use as an antagonist to reduce the recurrence of estrogen receptor positive breast cancer. However, the agonist actions of tamoxifen may be responsible for the increase in endometrial cancer seen in the treated women. Additionally, tamoxifen is an agonist in the mouse uterus, a partial agonist/antagonist in the rat uterus and an antagonist in the chicken oviduct. Furthermore, a new estrogen receptor (ER ß) was recently discovered which has a different tissue distribution and ligand selectivity from the well described ER a. There are numerous potential explanations for the biological observations other than the type and mixture of estrogen receptors in tissues, including differences in transcription factors and estrogen response elements. These examples illustrate some of the complexities we must take into account to help explain the diversity in dose-response curve shapes and low dose sensitivity.

**Testing the Low Dose Hypothesis with Xenoestrogens In Vivo**

The physiologically-based approach to testing for effects of environmental chemicals that test positive for estrogenic activity has also been used in a study that compared the effects of two estrogenic chemicals: bisphenol A (the monomer used to make polycarbonate plastic, the resin lining in food cans, and many other products) and octylphenol (an alkylphenol also used in many products). A physiologically relevant dose range for these chemicals was determined prior to administering them to pregnant mice to test for estrogenic effects on the reproductive organs in male offspring. The doses administered were based on determining in prior experiments the potency relative to estradiol-17ß of these chemicals using an in vitro assay that also takes into account the binding of xenoestrogens in plasma (Nagel et al., 1997). As above, the administered doses were referenced to the increase in serum estradiol previously shown to produce a significant change in development of the reproductive system in male mice (vom Saal et al., 1997). The additional important information about metabolism, which was available for DES, is unavailable for bisphenol A, octylphenol, and most other xenoestrogens, and the doses administered were thus based on the assumption of zero metabolism, which would thus be expected to overestimate the effect of a particular dose.

Estrogenic effects of bisphenol A have been seen in human breast cancer MCF-7 cells at $10_{-8}$ M or 2.3 ppb (Olea et al., 1996), and in rat prolactin secreting GH3 pituitary cells at 1 nM or 0.23 ppb (Steinmetz et al., 1997). However, the in vivo potency of bisphenol A, but not octylphenol, was hypothesized to be higher than would have been predicted based on in vitro assays that did not take into account the inhibitory effect of plasma proteins on uptake of estradiol. The basis for this hypothesis was that bisphenol

A was inhibited from entering cells by serum to a lesser degree than was estradiol. In contrast, serum dramatically inhibited the uptake of octylphenol into cells relative to inhibitory effects on estradiol. Taking all of this information into account, Nagel et al. (1997) fed pregnant female mice a 2 or 20 ng/g body weight/day dose of bisphenol A and octylphenol. The prediction was that the 20 ng/g dose of bisphenol A, but not octylphenol, would result in a similar increase in prostate size in male offspring as did the 0.02 ng/g dose of DES and the 0.1 ppt increase in free serum estradiol (vom Saal et al., 1997). The results showed a significant increase in prostate weight with both the 2 and 20 ng/g doses of bisphenol A, but not octylphenol. Bisphenol A was thus approximately 100-times less potent than DES when fed to pregnant mice. In another study, estradiol and bisphenol A were administered and measurements of serum prolactin indicated that bisphenol A was within 100-fold of the potency of estradiol which is consistent with the finding comparing bisphenol A and DES in fetal mice (Steinmetz et al., 1997).

Where mechanistic information is not available, use of environmentally relevant doses will ensure that extrapolation from studies with high doses will not dramatically underestimate risk.

The findings described above for bisphenol A were not predicted by a prior study using rats and mice in which high doses of bisphenol A (160 - 1,250 mg/kg or 0.16 - 1.25 parts per thousand) were administered (Morrissey et al., 1987). The 2 ppb dose that altered development of the fetal prostate is 625,000-times lower than the highest dose used in this prior study. The 2 ppb dose is also 25,000-times lower than the 50 µg/g body weight/day (50 ppm) dose that was previously reported to be the NOAEL for bisphenol A (Society of the Plastics Industry, 1996). Taken together, these findings suggest the possibility of a huge error in estimating safety of endocrine disrupting chemicals using data only from high dose studies. If current risk assessment methods were applied to the findings for bisphenol A as described by Nagel et al. (1997) where a 2 ppb maternal dose produced an adverse effect on differentiation of the prostate, the acceptable daily intake (ADI) would be calculated to be 2 ppt. In contrast, a NOAEL of 50 ppm, as reported by the Society of the Plastics Industry (1996), would lead to calculation of an ADI which was 25,000-times higher or 500 ppb (Welshons et al., 1997). This discrepancy illustrates the differences in interpretation of the risks posed by bisphenol A based on environmentally relevant doses versus on doses typically examined in toxicology. Where mechanistic information is not available, use of environmentally relevant doses, in addition to determining the maximum tolerated dose in toxicological studies, will ensure that extrapolation from studies with high doses will not dramatically underestimate risk.

**A Challenge to the Default Assumption of a Threshold for Systemic Toxicants**

A major source of uncertainty in risk assessments is the default assumption of a threshold dose. If the assumption is incorrect in some cases then greater risks may be present. Nonetheless, it is defended with great vigor by many toxicologists despite emerging data to the contrary. If there is a detectable incidence of an adverse outcome in an unexposed population due to an endogenously produced chemical, then the threshold for the outcome is already exceeded. If an administered chemical acts via the same mechanism, no threshold dose will be observed experimentally. Note that this does not deny that a threshold exists (although this assumption can be questioned) but only that a threshold dose does not exist. To test this important hypothesis the hormone must be present and must create an adverse effect. The hypothesis is verified by a demonstration that the dose-response curve for the endocrine disruptor is inconsistent with a threshold assumption.

Results from recent studies argue that endogenous estrogen production is responsible for female sex determination and that a threshold exists for this effect. Temperature sex determination (TSD) is found in some turtles and lizards and all crocodilians. These animals possess no sexually dimorphic chromosomes; rather, the incubation temperature of the egg in which the embryo and fetus are developing determines sex. In the red-eared slider turtle, low temperatures produce males and higher temperatures females. At intermediate temperatures a mixture of males and females is found, with the female proportion increasing with temperature. Estrogen applied to the eggshell produces females at male-determining temperatures; lower doses are required as temperature increases within the male-determining range and into that which generates a mixture. Inhibitors of aromatase, which converts testosterone to estradiol, decrease the female fraction at female-determining temperatures. These findings (Crews, et al., 1994) argue that endogenous estrogen production is responsible for female sex determination and that there is a threshold for this effect. Incubation of eggs at temperatures that produce a minority of females meets the criteria for testing the threshold dose hypothesis. Examination of published estradiol dose-response data reveals no threshold dose: additionally, the data sets are well-fitted by a simple Michaelis-Menten model which has no threshold term. Estriol is about 25-fold more potent than estradiol while in rats estradiol is the more potent. Sex reversal is

induced by very low doses of estradiol (0.4 ug/kg).

In women, numerous studies implicate endogenous estrogens as causal agents in the development of breast cancer. Likewise, in experimental animals, ovarian estrogens greatly increase the incidence of breast cancer in a variety of experimental models; this effect is not seen in ovariectomized animals. This substantial and concordant literature suggests that women, as a population, are above the threshold dose for estrogen causation of this disease, and that exposure to estrogens in the environment could increase the risk for breast cancer.

**Conclusions**

If several major assumptions in risk assessment are flawed for examining endocrine disruptors, risk may be dramatically underestimated. We have examined this possibility for hormones — specifically natural and manmade estrogens — effects on development, and their relevance to other toxicants and adverse effects is apparent. There are several important reasons that the effects of estrogens on development are a good starting point to investigate these issues. First, estradiol has a very high affinity for the ER and produces responses at very low doses, so low dose effects are also expected even for chemicals that are 1,000-10,000-times less potent than estradiol. Second, the observation that a low occupancy level of the ER can induce some responses explains another significant factor in low dose sensitivity. Third, failure to bind to serum binding proteins increases the potency relative to estradiol. Finally, the greatest sensitivity can be expected when a xenoestrogen is adding to the effects of the endogenous estrogen because a dose sufficient to exceed the threshold does not have to be attained before responses can be elicited.

The non-monotonic dose-response curves represent mechanisms to regulate responses within the physiological range of hormone concentrations. Traditionally, toxicology has not been concerned with physiological doses but rather with higher doses which are presumed to overwhelm various defense and repair mechanisms to induce adverse outcomes. However, when such outcomes occur at some measurable rate as a consequence of normal hormone levels, the physiological and toxicological dose ranges overlap or are identical. Thus, no threshold is expected and non-monotonic curves may be found. While estrogens are active at many life stages, they increase greatly during both rodent and human pregnancies and, moreover, are critically important for the proper development of many organs, particularly the reproductive tract.

These features help explain the experimental observations and thereby place them into a biological context with explanatory power. High dose experiments cannot predict low dose outcomes if the dose-response curve is non-monotonic. Additionally, rather than the threshold assumption — which asserts that there are safe doses of chemicals — we assert that such curves must be analyzed using Occam's Razor. If the data fit a curve inconsistent with a threshold dose and this is supported by underlying biological knowledge, then an acceptable risk must be agreed upon and the dose at that acceptable risk level becomes the appropriate exposure value. Note that the finding of a NOAEL under these circumstances is not evidence of a threshold but simply a measure of the statistical resolving power of the experiment.

The results of the findings presented here demonstrate that a paradigm-shift is in progress.

The results of the findings presented here demonstrate that a paradigm-shift is in progress. We need to reconsider some of the important assumptions used in toxicological risk assessments and redesign the approaches we use to test chemicals. Such approaches should include use of low doses rather than — or in addition to — the high doses currently tested and should utilize low dose sensitive models. Finally, the data and explanations presented here have explanatory power in accounting for findings from studies of wildlife species demonstrating a variety of adverse effects in environments contaminated with low concentrations of chemicals, some of which are known to be toxicants in experimental animals at high doses.

**REFERENCES**

Branham, W.S., Sheehan, D. M., Zehr, D. R., Ridlon, E. and Nelson, C. J. The postnatal ontogeny of rat uterine glands and age-related effects of 17B-estradiol Endocrinology, 117(5): 2229-2237, 1985.

Colborn, T. and Clement, C. (1992). Chemically-induced alterations in sexual and functional development: The wildlife/human connection. In: Advances in Modern Environmental Toxicology (M. A. Mehlman, eds.). pp. 403. Princeton Scientific Publishing C., Inc., Princeton.

Colborn, T., vom Saal, F.S. and Soto, A.M. (1993). Developmental effects of endoocrine disrupting chemicals in wildlife and humans. Environ. Health Perspect. 101:378-384.

Morrissey, R.E., George, J.D., Price, C.J., Tyl, R.W., Marr, M.C. and Kimmel, C.A. (1987). The developmental toxicity of bisphenol A in rats and mice. Fundam. Appl. Toxicol. 8:571-582.

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. Environ. Health Perspect. 105:70-76.

Olea, N., Pulgar, R., Prez, P., Olea-Serrano, F., Rivas, A., Novillo-Fertrell, A., Pedraza, V., Soto, A.M. and Sonnenschein, C. (1996). Estrogenicity of resin-based composites and sealants used in dentistry. Environ. Health Perspect. 104:298-305.

Shah, H.C. and McLachlan, J.A. (1976). iThe fate of diethylstilbestrol in the pregnant mouse. J Pharmacol Exp Ther. 197:687-696.

Sheehan, D. and Young, M. (1979). Diethylstilbestrol and estradiol binding to serum albumin and pregnancy plasma of rat and human. Endocrinol. 104:1442-1446.

Society of the Plastics Industry (1996). "Report on the potential exposures to bisphenol A from epoxy can coatings". Society of the Plastics Industry, Washington, D.C.

Steinmetz, R., Brown, N.G., Allen, D.L., Bigsby, R.M. and Ben-Jonathan, N. (1997). "The environmental estrogen bisphenol A stimulates prolactin release in vitro and in vivo." Endocrinol. 138:1780-1786.

Vom Saal, F.S., Nagel, S.C., Palanza, P., Boechler, M., Parmigiani, S. and Welshons, W.V. (1995). Estrogenic pesticides: Binding relative to estradiol in MCF-7 cells and effects of exposure during fetal life on subsequent territorial behavior in male mice. Toxicol. Lett. 77:343-350.

vom Saal, F.S., Timms, B.G., Montano, M.M., Palanza, P., Thayer, K.A., Nagel, S.C., Dhar, M.D., Ganjam, V.K., Parmigiani, S. and Welshons, W.V. (1997). Prostate enlargement in mice due to fetal exposure to low doses of estradiol or diethylstilbestrol and opposite effects at high doses. Proc. Natl. Acad. Sci. 94:2056-2061.

Welshons, W.V. and Jordan, V.C. (1987). iAdaptation of estrogen-dependent MCF-7 cells to low estrogen (phenol red-free) culture. Euro J Cancer Clin Oncol. 23:1935-1939.

## *James Lamb Responds to Daniel Sheehan's and Fred vom Saal's Arguments*

Rebuttal to Drs. Sheehan and vom Saal

I would like to amplify and respond to a few of the points raised by Drs. Sheehan and vom Saal regarding the challenges posed to toxicology by hormonally active agents.

First, I want to address the use of the term *endocrine disruptor* which they have defined as hormonally active agents "... that irreversibly disrupt normal endocrine functioning and thus organogenesis due to exposure during critical developmantal periods . ..." This definition implies an adverse effect through a developmental mechanism. However, one of the authors (Dr. vom Saal) has published a list of more than 40 endocrine disruptors that includes many chemicals that have been shown neither to cause adverse effects, nor to act through a developmental mechanism.[1] many chemicals were listed because of in vitro studies only, which would not provide the basis for labeling a chemical as an endocrine disruptor under their definition. When ssuch lists are based on hazard identification only, they serve to confuse and misrepresent critical issues and short-circuit the inclusion of critical information such as dose response, exposure, and risk.

Second, the authors present a "challenge to the defult assumption of a threshold" that I do not find persuasive. Although we agree that elevated endogenous estrogens, from early menarche or delayed menopuse, correlate with increased breast cancer, the addition of exogenous estrogens has been not shown to cause breast cancer in women. In fact, just the opposite is true. Most environmental estrogens are weak estrogens, such as those used in hormone replacement therapy, do not increase breast cancer risk according to the results of an American Cancer Society study. Thus, one molecule added to an already turned on system does not necessarily exceed a threshold.

Third, a great deal has been made of testing low or "environmentally revevant" levels of chemicals. Toxicologists should carefully consider their dose levels, and in some cases, levels below the no observed adverse effect level (NOAEL) might be useful. Those occasions do not, however, appear to be universal. Toxicologists should also consider group size, dosing method, enpoint reliability, sensitivity, and detail (e.g., histopathology rather than just weight), so the studies are as reliable and valid as possible. Moreover, some definition of *low* or *environmentally relevant* should be given before sweeping changes are proposed. Because such studies may raise important, controversial issues, these investigations must be conducted with sufficient sizes, with analytical chemistry to confirm the dosing solutions, and with carefully selected endpoints.

Finally, I am in complete agreement with the authors that "all estrogens are not created equal." Even within a single system, the responses can be different for each tissue and sell type. One of the real dangers in this field may result from the search for quick and easy endocrine disruptor screens, which will oversimplify or over-extrapolate from simple tests leading to faulty conclusions about complicated adverse effects. Even Dr. Sheehan's and vom Saal's decision to use ppb or ppt, rather than mg/kg body weight, implies a simple system is being assessed, such as would be used for a homogeneous concentration in a test tube. It is more complicated than that. However, the body has many redundant systems to protect from insults, or to repair damage. It also has more targets of insultthan we could ever hope to put into a simple screen. We have to consider the trade-offs inherent in screening many chemicals quickly or testing fewer chemicals, but testing them well. Even if financial resources were unlimited for this exercise, our intellectual resources are not.

Note:[1] Conlborn, T., vom Saal, F. Soto, A. (1993) " Developmental Effects of Endocrine-disrupting Chemicals in Wildlife and Humans." *Environ. Health Perspec.*, 101: 378-384