DRAFT FOR E.C. REVIEW ONLY – 6/15/99 - DO NOT QUOTE OR CITE

EPA-SAB-EC-99-0XX

Honorable Carol M. Browner
Administrator
U.S. Environmental Protection Agency
401 M Street, S.W.
Washington, DC 20460

Subject: Review of EPA's Proposed Endocrine Disruptor Screening Program.

Dear Ms. Browner:

In 1996, the passage of the Food Quality Protection Act (FQPA) and amendments to the Safe Drinking Water Act (SDWA) required EPA to develop a screening and testing strategy for endocrine disruptors. The Agency established the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) to provide advice on the screening and testing of pesticides and chemicals for their potential to disrupt the endocrine system. The EPA subsequently asked the Science Advisory Board (SAB) and the FIFRA Scientific Advisory Panel (SAP) to form a Joint Subcommittee to review a set of scientific issues being considered by the Agency concerning the development of the Agency's endocrine disruptor screening and testing program as required by the legislation noted above. A Joint Subcommittee (the Joint Environmental Disruptor Screening Program (EDSP) review Subcommittee met on March 30/April 1, 1999, in Arlington VA, and produced this report.

The Charge was broad and complex, posing 18 major questions within four broad areas: a) scope of the program; b) priority-setting; c) the high throughput pre-screening approach; and d) the proposed endocrine disruptor screening program (the complete Charge is provided in section 2.2 of this report).

At the outset, we wish to note that although our review identified several areas of concern, and the EDSP has provided recommendations to improve EPA's planned program, we wish to congratulate the Agency for dealing effectively with an extraordinarily complex set of issues, many of which are on the cutting edge of the relevant science. The EDSP's detailed response to each element of the Charge can be found in section 3 of the report, and our major issues and recommendations are summarized below:

a) **Evaluating the Program:** We find no provision for mid-course evaluation or optimization of the process. Although an approach may look fine on paper or in a small research setting, translating it into a volume-screening mode may be quite another thing. There was broad support among the Subcommittee for the concept that the Agency should convene a panel of independent scientists to review all the screening data for 50-100 compounds, with eye towards revising the process and eliminating those methods that don't work.

b) **Mixture Issues:** The Subcommittee agreed that the initial focus of the methods development effort must focus necessarily on single compounds and leave the question of testing of mixtures until accepted single-compound methods have been completed.

c) **Case Studies:** The Subcommittee strongly encourages the Agency to include more and better-detailed case studies in the evolution of the priority-setting scheme, enabling a realistic test of the plan, checking sensitivity of the system and it's practicality to prioritize properly chemicals for further testing.

d) **Sub-population Compartment:** The question of the need for a separate compartment to address sub-populations (e.g., developing children) was addressed to the EDPS. Our conclusions supported the use of sub-populations as a criterion within the existing compartments already identified, but not as a separate stand-alone compartment.

e) **Use of the Integrated Risk Information System (IRIS):** The priority testing scheme relies on the use of several data-bases summarizing the environmental fate and effects of chemicals. Several Members of the Committee expressed concern about problems with the validation of IRIS and other data-bases. Before placing heavy reliance on these computerized systems, users need to be aware of these validation problems and proceed with caution before incorporating these values unilaterally.

f) **Exposure:** The EDPS believes that consideration of the toxicological implications of exposure should include both dose *and* time of exposure. The current scheme does not adequately cover the time aspect of exposure and this needs to be remedied before broad-scale application of the approach.

g) **Use of Animals**: We are concerned about the large number of animals that would be needed by the EDSTAC program. The Subcommittee is cognizant of the essential role animals play in tests to detect endocrine disruption, and aware that there are no substitutes for tests currently available for the Tier 2 tests. This fact notwithstanding, the Agency has an obligation to conserve all resources in developing new testing protocols, and the use of animals in such tests poses both ethical and practical problems.

h) **Need for an Introductory Statement:** The previous EDSTAC meeting suggested that the final document needed, as a introductory section, a description of the problem or the scientific or health-based reason for the EDSTAC program. The EDPS urges the EPA's EDSTAC team to include a description of both the health and ecological problems associated with exposure to the endocrine disruptors and to show how the program is justified by these findings.

i) **Support for Decisions**: Decisions about which assays are selected, and which protocols are adopted for those assays, should be supported with data that are generally available.

j) **Exceptions**: Testing strategies will always have exceptions. Care should be taken to be aware of the imperfect nature of *any* future agreed strategy.

k) **Negative Control Agents**: There is a need to define and agree on some negative control agents for ED assay validation. Assay specificity will not be capable of assessment unless such agents can be made available for general study.

l) **Animal Tests and Routes of Exposure**: There is significant international concern on the proposed use of animals for screening. In this role of hazard assessment (as

We appreciate the opportunity to review these proposed revisions, and look forward to receiving your response to the issues raised.

Dr. Joan Daisey, Chair
Science Advisory Board, and
Co-chair, Endocrine Disruptor Screening
Program Review Subcommittee

Dr. Gene McConnell, Co-chair
Endocrine Disruptor Screening
Program Review Subcommittee

ENCLOSURE

REVIEW OF THE ENVIRONMENTAL ENDOCRINE DISRUPTOR SCREENING PROGRAM BY A JOINT SUBCOMMITTEE OF THE SAB/SAP

EXECUTIVE COMMITTEE REVIEW DRAFT

For Review only – Do not Quote or Cite

6/15/99

NOTICE

This report has been written as part of the activities of the Science Advisory Board (SAB) and the Scientific Advisory Panel (SAP), public advisory groups providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The SAB and SAP are structured to provide balanced, expert assessments of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

ABSTRACT

TO BE SUPPLIED

ROSTER

Joint SAB/SAP Subcommittee On Endocrine Disruptor Screening

March 30-April 1, 1999

CO-CHAIRS

Dr. Joan Daisey, Lawrence Berkeley Laboratory, Indoor Environmental Program, Berkeley CA
Dr. Ernest McConnell, ToxPath, Raleigh NC

MEMBERS AND CONSULTANTS

Dr. John Ashby, Zeneca Corporation, Cheshire, United Kingdom
Dr. Richard Bull, Molecular Biosciences, Battelle Pacific Northwest Laboratories, Richland, WA
Dr. Charles Capen, Department of Veterinary Biosciences, The Ohio State University, Columbus, OH
Dr. Kenneth Davis, Ecological Research Center, University of Memphis, Memphis, TN
Dr. John Doull, Department of Pharmacology, Toxicology and Therapeutics, The University of Kansas Medical Center, Kansas City, KS
Dr. Paul M.D. Foster, Chemical Industry Institute of Toxicology, Research Triangle Park, NC
Dr. James Gibson, Dow AgroSciences, Indianapolis, IN
Dr. Philippe Grandjean, Institute of Community Health, Odense University, Denmark
Dr. Diane Henshel, School of Public and Environmental Affairs, Indiana University, Bloomington, IN
Dr. Alan Maki, Exxon Company, USA, Houston, TX
Dr. Genevieve Matanoski, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, MD
Dr. Margaret McCarthy, Department of Physiology, School of Medicine, University of Maryland at Baltimore, Baltimore, MD
Dr. Michael McClain, University of Medicine and Dentistry of New Jersey, R.W. Johnson Medical School, Randolph, NJ
Dr. F.M. Anne McNabb, Department of Biology, Virginia Polytech. Institute & State University, Blacksburg, VA
Dr. Mary Anna Thrall, Department of Pathology, College of Veterinary Medicine & Biomedical Sciences, Colorado State University, Fort Collins, CO
Dr. John G. Vandenbergh, Department of Zoology, College of Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

Dr. Tim Zacharewski, Department of Biochemistry, Michigan State University, East Lansing, MI

FEDERAL EXPERTS

Dr. Robert Chapin, National Institute of Environmental Health Sciences, Research Triangle Park, NC

Dr. James Hanson, National Cancer Center, Bethesda, MD

CO-DESIGNATED FEDERAL OFFICIAL (SAP)

Mr. Larry Dorsey, FIFRA Scientific Advisory Panel, Office of Prevention, Pesticides and Toxic Substances, Environmental Protection Agency, Washington, DC

CO-DESIGNATED FEDERAL OFFICIAL (SAB)

Mr. Samuel Rondberg, Science Advisory Board (1400), Environmental Protection Agency, Washington, DC

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

In 1996, the passage of the Food Quality Protection Act (FQPA) and amendments to the Safe Drinking Water Act (SDWA) required EPA to develop a screening and testing strategy for endocrine disruptors within two years and implement the plan by August, 1999. EPA established the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) under the Federal Advisory Committee Act to advise the Agency on the screening and testing of pesticides and chemicals for their potential to disrupt the endocrine system. Consequently, the EPA asked the Science Advisory Board (SAB) and the FIFRA Scientific Advisory Panel (SAP) to form a Joint Subcommittee to review a set of scientific issues being considered by the Agency concerning the development of the Agency's endocrine disruptor screening and testing program as required by the legislation noted above. This Joint Subcommittee (the Joint Environmental Disruptor Screening Program (EDSP) review Subcommittee met on March 30/April 1, 1999, and produced this report.

The Charge was broad and complex, posing 18 major questions within four broad areas: a) scope of the program; b) priority-setting; c) the high throughput pre-screening approach; and d) the proposed endocrine disruptor screening program (the complete Charge is provided in section 2.2 of this report).

The EDSP's detailed response to each element of the Charge is found in section 3 of the report. The major issues and recommendations are:

a) **Evaluating the Program:** We find no provision for mid-course evaluation or optimization of the process. The Agency is mandated to assemble and evaluate this proposed panel of tests and then to implement them, but a correlate responsibility is to make sure that what's being done is the best that can be. Although something looks fine on paper or in a small research setting, translating it into volume-screening mode may be quite another thing. There was broad support among the Subcommittee for the concept that the Agency should convene a panel of independent scientists to review all the screening data for 50-100

compounds,

with eye towards revising the process and eliminating those methods that don't work.

b) **Mixture Issues:** The Subcommittee agreed that the initial focus of the methods development effort must focus necessarily on single compounds and leave the question of testing of mixtures until accepted single-compound methods have been completed. The Subcommittee concluded that very promising methods already exist in the field of ecotoxicology. These include the Whole Effluent Testing (WET) and Toxicity Identification Evaluation (TIE) procedures developed by the Agency in concert with the Society for Environmental Toxicology and Chemistry. Those methods have been developed to test effects of effluents and should have direct application to the Endocrine Disruptor Screening Program.

c) **Case Studies:** The Subcommittee strongly encourages the Agency to include more and better-detailed case studies in the evolution of the priority-setting scheme. Case studies will enable a realistic test of the scheme, checking sensitivity of the system and it's practicality to prioritize chemicals for further testing.

d) **Sub-population Compartment:** The question of the need for a separate compartment to address sub-populations (e.g., developing children) was addressed to the EDPS. Our conclusions supported the use of sub-populations as a criterion within the existing compartments already identified, but not as a separate stand-alone compartment.

e) **Use of the Integrated Risk Information System (IRIS)**: The priority testing scheme relies on the use of several data-bases summarizing the environmental fate and effects of chemicals. Several Members of the Committee expressed concern that there are numerous problems with the validation of IRIS and other data-bases.

Before placing heavy reliance on these computerized systems, users need to be aware of these validation problems and proceed with caution before incorporating these values unilaterally.

f) **Exposure:** The EDPS expressed concern that consideration of the toxicological implications of exposure should include both dose *and* time of exposure. The current scheme does not adequately cover the time aspect of exposure and this needs to be remedied before broad-scale application of the approach.

g) **Use of Animals**: We are concerned about the large number of animals that would be needed by the EDSTAC program. The Subcommittee is cognizant of the essential role animals play in tests to detect endocrine disruption. There are no substitutes, however, for tests currently available for the Tier 2 tests. This notwithstanding, the Agency has an obligation to conserve all resources in developing new testing protocols, and the use of animals in such tests poses both ethical and practical problems.

h) **Need for an Introductory Statement:** The previous EDSTAC meeting suggested that the final document needed, as a introductory section, a description of the problem or the scientific or health-based reason for the EDSTAC program. The EDPS urges the EPA's EDSTAC team to include a description of both the health and ecological problems associated with exposure to the endocrine disruptors and to show how the program is justified by these findings.

i) **Support for Decisions**: Decisions about which assays are selected, and which protocols are adopted for those assays, should be supported with data that are generally available.

j) **Exceptions**: Testing strategies will always have exceptions. Care should be taken to be aware of the imperfect nature of *any* future agreed strategy.

k) **Negative Control Agents**: There is a need to define and agree on some negative control agents for ED assay validation. Assay specificity will not be capable of assessment unless such agents can be made available for general study.

l) **Animal Tests and Routes of Exposure**: There is significant international concern on the proposed use of animals for screening. In this role of hazard assessment (as opposed to hazard definition) biologically relevant routes of exposure would be indicated (oral gavage, diet, water, inhalation, skin painting). At present, use of the subcutaneous injection or intraperitoneal injection routes are recommended in the frail quest of increasing assay sensitivity. In fact, irrespective of the outcome of this suggestion it should be noted that the current EPA synthesis of the EDSTAC recommendations is inconsistent on the matter of route of exposure.

Although the review identified several areas of concern, and the EDSP has provided recommendations to improve EPA's planned program, we wish to congratulate the Agency for dealing effectively with an extraordinarily complex set of issues, many of which are on the cutting edge of the relevant science.

2. INTRODUCTION

2.1 Background

Chemicals which may interfere with endocrine system functioning (endocrine disruptors) have concerned the U.S. Environmental Protection Agency (EPA) for some time. Such chemicals have the potential to impact human and wildlife populations. A variety of human health and ecological effects have been attributed to endocrine disruptors.

In 1996, the passage of the Food Quality Protection Act (FQPA) and amendments to the Safe Drinking Water Act (SDWA) required EPA to develop a screening and testing strategy for endocrine disruptors within two years and implement the plan by August, 1999. The legislation cites the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and the Toxic Substances Control Act (TSCA) as the two statutes under which EPA should implement an endocrine screening and testing strategy. EPA established the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) under the Federal Advisory Committee Act to advise the Agency on the screening and testing of pesticides and chemicals for their potential to disrupt the endocrine system.

Consequently, the Science Advisory Board (SAB) and the FIFRA Scientific Advisory Panel (SAP) were asked to form a Joint Subcommittee to review a set of scientific issues being considered by the Agency concerning the development of the Agency's endocrine disruptor screening and testing program as required by the legislation noted above.

2.2 Charge

The specific issues to be addressed by the Joint Subcommittee are:

a) Scope of the Program

1) The amendments to the Food Quality Protection Act (FQPA) and the Safe

Drinking Water Act (SDWA) mandate or support the development of a screening program that will determine whether pesticides and certain drinking water source contaminants "may have an effect in humans that is similar to an effect produced by a naturally-occurring estrogen, or other such endocrine effect as the Administrator may designate." Very early in its deliberations, EPA's Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) determined that there was both a strong scientific basis and feasibility, considering time and resource constraints, to expand the scope of the screening program to include the androgen- and thyroid-hormone systems, and to include evaluation of the potential impact on wildlife as well as on human health. EPA agrees and is developing a screening program which incorporates these modifications. **Does the Joint Subcommittee agree that this expanded scope is appropriate to serve as the starting point for the Endocrine Disruptor Screening Program (EDSP), given the understanding that the framework for the Program can support for further expansion at a later date?**

2) The FQPA and SDWA identify a universe of substances that should be evaluated in a an EDSP. EDSTAC noted that there exist many other substances in addition to pesticides and certain drinking water source contaminants that may exhibit endocrine-disrupting potential. They recommended that the "candidate pool" for the EDSP include substances on the Toxic Substances Control Act (TSCA) Inventory, certain complex environmental mixtures as well as non-pesticide food additives, cosmetics and nutritional supplements. EPA agrees that there are substances in addition to pesticides and certain drinking water source contaminants that warrant consideration for inclusion in the EDSP.. **Does the Joint Committee agree that this expanded universe of substances should be included in the EDSP process, at a minimum in the priority-setting phase, and continuing on if a potential for concern is identified?**

3) FQPA contains a provision which would exempt from the EDSP "any biological substance or other substance if the Administrator determines that the substance is anticipated not to produce any effect in humans similar to an effect produced by a

naturally-occurring estrogen" or, presumably, "such other endocrine effect as the Administrator may designate." EPA has identified some chemical categories that may be candidates for exemption. Examples include certain polymers with a number average molecular weight (NAMW) greater than 1000 daltons, certain List 4 pesticide inerts such as cookie crumbs, strong mineral acids and bases, which are most likely to interact with tissue at the portal of entry giving rise to localized lesions rather than systemic effects, certain biopesticides such as plant pesticides or microbials or non-chemical pesticides such as parasitic wasps. **Does the Joint Committee agree that there are categories of pesticides and other substances that should be exempt from the EDSP? In addition to the examples noted here, are there additional categories that should be considered for exemption?**

4) EDSTAC concluded, and EPA agrees, that there are important complex environmental mixtures that deserve inclusion in the EDSP. EDSTAC recommended that EPA include in the EDSP representative mixtures to which large or identifiable key segments of the population (e.g. children) are exposed. They suggested that high-priority mixture categories include: Chemicals in breast milk, phytoestrogens in soy-based infant formulas, mixtures commonly found at Superfund sites, common pesticide/fertilizer mixtures found in ground and surface water, disinfection byproducts, and gasoline. EPA proposes to screen and test (if appropriate) one representative mixture from each category, after it confirms that the screening and testing components of the EDSP are satisfactory for the handling of single substances.

A) **Is the proposal a reasonable way to address the practicality of screening and testing mixtures**

B) **Are the six categories of mixtures the most appropriate to address first?**

C) **Are there other mixture categories that should be included in addition to, or instead of those identified (Note: During the May Consultation, it was suggested that mixtures found in fish tissue, benthic sites and**

eggs of Great Lakes birds should replace gasoline as a priority mixture).

D) **Can/should standardized representative mixtures be developed? If so, how should the chemical combinations, ratios, and doses be selected for mixtures?**

E) **If a mixture is positive in the screening tier, should the whole mixture be tested in the testing tier or should only the active component(s) in the screen(s) be tested in the second tier?**

b) Priority-setting

1) EDSTAC recommended a compartment-based approach to priority-setting. EPA agrees that this is the appropriate framework. Under this approach, EPA will group chemicals into sets, based on the existence of factual information in a given area. Thus, priority ranking can be made fairly among substances, i.e., chemicals will compete for priority with others on the basis of comparable data and will not be assigned lower priority for lack of information. **Are these principles and the component-based approach to priority setting reasonable? Are there other approaches that would be more useful?**

2) EPA is developing a relational database to assist in developing priorities for screening. The relational database is intended to import existing data and allow its synthesis, as well as the estimation of certain parameters through modeling. The relational database was considered to have great value in helping to identify the specific compartments under the EDSTAC's compartment-based priority setting approach. The database will also be helpful in selecting chemicals for the first and subsequent rounds of screening. **Would the Joint Subcommittee comment on the approach and provide additional insights to improve the content of the relational database or its implementation?**

c) High Throughput Pre-screening Approach

1) EDSTAC recommended, and EPA proposes to implement a priority setting strategy that includes initial sorting based on an examination of existing information. This initial sorting strategy leads to four possible outcomes: I) polymers; ii) chemicals with sufficient data to proceed to testing; iii)chemicals with sufficient data to proceed to hazard assessment; and iv) chemicals with insufficient data, which presumably, would go into the screening phase. EPA anticipates that a large number of substances will end up in category iv-chemicals with insufficient data. To provide at least a minimum number of biological data to assist in the sorting process, EPA proposes to conduct High Throughput Prescreening on a significant number of substances (perhaps, as many as 15,000), using *in vitro* assay systems incorporating transcriptional activation or reporter gene systems for the estrogen-, androgen- and thyroid-hormone systems.

A) On the assumption that the technology can be shown to be applicable to the large number and wide range of chemical substances under consideration, and the limited relevant test data which are available for many industrial chemicals, **is this a reasonable approach and sorting strategy to support priority setting?**

B) EPA has been funding a pilot study, using about 80 chemicals, to determine the applicability of the high throughput technology in a prescreening component of the EDSP. **Based upon your review of the data developed to date, does the Joint Subcommittee believe that this technique can be used as a prescreening device? If not, what modifications/improvements must be made in order to assure its usefulness?**

d) **The Proposed Endocrine Disruptor Screening Program**

1) EPA is proposing to develop and implement an Endocrine Disruption Screening Program that consists of two phases. The first phase is screening, currently consisting of eight components. The second phase is testing, currently envisioned to have one to five components, depending upon the need for identifying effects in various sectors of the animal kingdom. **Is it reasonable and appropriate to develop and implement a two-phase program, the first phase focused on identifying a substance's potential to interact with one or more of the three hormone systems, the second phase to characterize the effects of concern that interaction with these hormone systems might elicit?**

2) EDSTAC recommended, and EPA proposes to implement, a screening battery consisting of eight assays, three *in vitro* and five *in vivo*, to address estrogen-, androgen- and thyroid-hormone system effects. At the time (a year ago or so) and continuing to today, based upon our knowledge of the state-of-the-science, the Agency believed that these eight assays, once validated and standardized, would detect all substances currently-known to interact with the three hormone systems to be covered in the Program. **Does the Joint Subcommittee agree with this assessment? If not, what changes should be made in the battery to assure the identification of substances of potential concern?**

3) Interaction with a receptor is the principal or key mechanism by which substances exert their effects on the estrogen- and androgen-hormone systems. This appears NOT to be the case for the thyroid-hormone system. **In light of this, does the Joint Subcommittee believe that there is adequate coverage of the thyroid provided in the proposed screening battery? If not, what modifications should/could be made?**

4) EPA would prefer to have a screening battery which included assays containing an *in utero* or *in ovo* exposure component, given its great and continuing concern about the potential for effects on the developing vertebrate organism. At the time

the proposed screening battery was being assembled, EPA was not aware of the existence of any such screens. **Is the Joint Subcommittee aware of any such assays that may exist or are under development that could supplant or complement one or more components of the proposed screening battery?**

5) EDSTAC recommended, and EPA would prefer, for efficiency and cost reasons given the numbers of substances that may be involved in the EDSP, to conduct each in vivo screening assay using only one dose, with the appropriate use of range finding studies and other information to inform dose selection. **Does the Joint Subcommittee agree with this approach, and if not, what suggestions would it have to modify the approach, keeping pace, volume, cost and efficiency in mind? What would be the public health consequences of these false negatives?** (Note: At the May consultation, some members raised concern about relying on a single dose and suggested that a minimum of two doses, and perhaps even three, be used to ensure that the screens do not yield false negative results.It has also been suggested, elsewhere, that this issue could/should be solved during validation/standardization.).

6) EDSTAC recommended, and EPA is proposing, a testing phase in the EDSP which could have as many as five components (i.e. covering mammals, birds, fish, invertebrates and amphibians). Each test would be designed to delineate the dose-response relationships of effects of concern for chemicals which yielded positive results during the screening phase. The testing protocols to be used are either upgrades or modifications of existing guidelines, except for the amphibian. In this case, a protocol is being developed de novo. **Does the Joint Subcommittee believe that these test protocol designs will provide sufficient rigor to identify effects of concern and establish their dose responses for disruption in the estrogen-, androgen- and/or thyroid-hormone systems?**

7) There could be circumstances in which substances bypass the screening phase, and go directly into the testing phase. EPA is proposing for those cases that the chemical under evaluation be tested in all five tests. **Does the Join Subcommittee**

believe that the tests in the testing phase will be adequate to detect all known critical endpoints in the estrogen-, androgen- and thyroid hormone systems? If not, what modifications should be made?

8) If the results of any of the testing phase tests are negative, **what, if any, additional screening or testing should be conducted to assure that the chemical is not an endocrine disruptor in the estrogen-, androgen- or thyroid hormone systems of that sector of the animal kingdom?**

9) Testing phase tests will identify effects of concern that are the consequence of endocrine disruption. They may also identify effects of concern that are not the consequence of endocrine disruption. Thus, it may not be possible to determine if a substance is an endocrine disruptor if it has not been subjected to some or all components of the screening battery. **Is it important to be able to identify substances as endocrine disruptors from the standpoint of conducting a hazard assessment. If so, why? If not, why not?**

10) EPA is proposing a validation program in which the maximum validation effort will consist of conducting each assay in three laboratories. EPA believes that there currently is a wide variation in the state of validation of each of the proposed screens and tests, and that the validation efforts should be tailored for each assay/test accordingly. EPA plans to focus first on the validation of the mammalian assays as they are both better developed than the non-mammalian assays and are more directly relevant to meeting the FQPA and SDWA mandates for a screening program for potential human health impacts. EPA's preliminary assessment of the work needed is as follows:

> The uterotrophic assay requires the development of a standardized protocol but may need little or no additional laboratory/protocol development effort since the assay has been in extensive use for many years.

The Hershberger assay may require some, but not much, additional laboratory/protocol development in addition to standardization.

The pubertal male and pubertal female assays need some additional developmental work and will require the full regime of interlaboratory validation.

The mammalian two-generation reproduction test will require limited testing in one laboratory to validate the new endpoints since the basic protocol is already considered to be valid.

Both of the non-mammalian screens and some of the non-mammalian tests will require the full validation regime; some will require further pre-validation development (e.g. amphibian test).

The mammalian two-generation test will require limited testing in one laboratory to validate the new endpoints since the basic protocol is already considered to be valid. All of the non-mammalian assays will require the full validation regime and some will require further pre-validation development. **Does the Joint Committee agree with the Agency's assessment of the current status of the screens and tests? If not, what is the Joint Committee's own assessment of any screen or test which differs from EPA's, and what is the basis for your opinion?**

11) Does the Joint Subcommittee have any other suggestions or recommendations that would help EPA meet its charge?

3. Detailed Response to The Charge

3.1 Scope of the Program

The amendments (1996) to the Food Quality Protection Act (FQPA) and the Safe Drinking Water Act (SDWA) mandate or support the development of a screening program that will determine whether pesticides and certain drinking water source contaminants "may have an effect in humans that is similar to an effect produced by a naturally-occurring estrogen, or other such endocrine effect as the Administrator may designate." Very early in its deliberations, EPA's Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) determined that there was scientific basis for expanding the scope of the screening program to include the androgen- and thyroid-hormone systems, and to include evaluation of the potential impact on wildlife as well as on human health (EDSTAC, 1999). EPA agrees and is developing a screening program which incorporates these modifications. Sections 3.4.1 through 3.4.4 address significant issues in designing the screening program

3.1.i The proper Scope for the Endocrine Disruptor Screening Program (EDSP)

The initial element of the Charge (a) (1) for this review asks if the Joint Subcommittee agrees that this expanded scope is appropriate to serve as the starting point for the Endocrine Disruptor Screening Program (EDSP), given the understanding that the framework for the Program can support further expansion at a later date.

Expansion of the scope of the program under the FQPA and the SDWA from simply estrogen-like effects on human health to androgen and thyroid active compounds is reasonable. This expansion raises the sights above that demanded by the authorization language, but is clearly within the guidance provided.

The EDSTAC review of endocrine effects recognized that issues related to endocrine disruption are even broader than the three categories identified (EDSTAC, 1998). However, there will be significant technical difficulties in addressing estrogen, androgen and thyroid active

compounds. Further expansion at this stage in the development of the program would have created an unmanageable task. Nevertheless, it should be recognized that modification of the activity of other hormonal systems can be as important, or perhaps more important, than the systems identified. Moreover, the technology to begin integrating these systems is now becoming available. Overall, the proposed framework should enable the agency to integrate knowledge of these systems and assay techniques as they mature.

The expansion of the concerns to the broader environmental concerns over endocrine disruption is not only appropriate, but crucial. Modifications in reproductive and developmental processes in the environment have been related to endocrine disruption. Experience has shown that effects on wildlife and ecosystems are seen before there is any significant impact in humans, particularly for chemicals that bioaccumulate. It is unfortunate that effects on populations in actual ecosystems cannot be practically included in the program because in some cases these have been most sensitive measures for agents of this type. Consequently, there are substantive reasons for expanding the scope to non-human species. As with the expansion to the androgen and thyroid hormone systems,however, it is important to recognize that this further complicates implementation.

The Joint Subcommittee expressed considerable concern with respect to whether the program is focused on the critical endpoints. The focus on the endocrine system is very much a focus on mechanism(s) that do not necessarily relate in a holistic manner to the adverse health and environmental effects that are of ultimate concern. In general the major concern for endocrine disruptors are effects on normal reproduction and development that also extend to the induction of certain kinds of cancer. The endocrine disruptions, as characterized by interactions with these three hormone systems, are not the only way in which such effects can be produced. As a consequence, there is some inversion of the normal decision logic that makes it crucial that a concise working definition of endocrine disruption be developed. The EDSP is being developed on the apparent assumption that most compounds that affect the three endocrine systems identified are likely to be of toxicological concern. Secondary testing will be triggered based on broad-based screening. However, significantly more thought must be given to identifying quantitative "trigger points" as well as the question of whether the response is positive or not. These might include an exclusion based on excessive doses required to induce

the effect or a requirement for a minimum level of response in the test system. The most important issue is to identify the magnitude and perhaps the duration of a response. Most of this difficulty could be taken care of by crafting a more concise definition of what constitutes endocrine disruption. This could possibly be done by adding a phrase to the endocrine disruption definition that states "... reproducible effects on any endocrine sensitive system impairing successful reproduction and development ..." " Broader definitions could also raise issues relating to carcinogenesis or the ability to maintain homoeostatic balance in response to biotic, chemical and physical stresses. Extension of such definitions inevitably leads to complications in the application of the test scheme.

The Joint Subcommittee also identified a need to establish an on-going review of progress of the EDSP. There are both methodological and interpretative issues involved. The methodological issues are more straight-forward, involving a process for reviewing new screening and testing methods for incorporation or substitution for current methods. The interpretative problems have more to do with how the data are going to be applied to improving environmental protection. EPA put in a heroic effort in reviewing the available methodology and putting together a framework based on that methodology. They have addressed many of the interpretative problems that evolve from the current structure in the EDSTAC report (EDSTAC, 1998; Federal Register, 1998). It is not as clear how these processes are to be moved forward in a rational way. A regular plan to revisit both the methodological and interpretative issues can be used as a vehicle to stimulate that progress.

3.1.2 Use of the expanded set of Agents in the EDSP

The FQPA and SDWA identify a universe of substances that should be evaluated in an EDSP. The EDSTAC report (1998) noted that there are many other substances in addition to pesticides and certain drinking water source contaminants that may exhibit endocrine-disrupting potential. They recommended that the "candidate pool" for the EDSP include substances on the Toxic Substances Control Act (TSCA) Inventory, certain complex environmental mixtures as well as non-pesticide food additives, cosmetics and nutritional supplements. EPA agrees that there are substances in addition to pesticides and certain drinking water source contaminants that warrant consideration for inclusion in the EDSP. Charge element (a) (2) asked the Joint

Committee to

comment on the use of an expanded universe of substances in the EDSP process, both in the priority-setting phase, and continuing on to later phases if a potential for concern is identified.

Expansion of the scope of the mandated efforts under the FQPA and the SDWA makes sense if validated systems and clear decision criteria for the expansion are available. Under such circumstances there would little reason to exclude additional chemicals from consideration. At the present time, however, this particular expansion seems to add a level of complexity that may be counterproductive. The Subcommittee's concerns arise from considering what the underlying objective(s) of a screening and testing program are in the environmental programs administered by the EPA. The ultimate goal is to protect health and the environment from adverse effects. In one sense, the Agency is to be congratulated for attempting to focus on the endocrine disruptor issue, because it does move them in the direction of identifying and perhaps understanding more subtle environmental hazards. However, if the activity loses its connectivity to recognized or newly described forms of compromised health and ecological effects, it will difficult for the program to establish a solid rational basis. Expanded consideration of diverse types of candidate "endocrine disrupting" chemicals has the potential of getting ahead of our knowledge of the risks actually represented by screening and testing of large numbers of chemicals. The interest in how modified cellular function leads to adverse effects is a necessary step involved in improving hazard identification and risk assessment. These advancements must be built on careful development of the science that establishes clear causal associations between new testing tools and adverse impacts. The precedent that could be set by pursuing mechanisms without regard to effect can not only greatly increase the expense of testing of products, it can compromise confidence in the screening program that will eventually have to include all aspects of endocrine functions represented. The application of these tests to a wide variety of chemicals (as

many as 15,000) has the potential for building up a data base, but not necessarily increasing our knowledge about the significance of any effects that are observed.

This point may be illustrated by considering some multifaceted problems that could

evolve from the EDSP in the form of questions some of which are actually discussed in the report. Is the intent to focus entirely on direct effects mediated through recognized estrogen, androgen and thyroid receptors? How will indirect effects on the endocrine receptor be identified? Are indirectly mediated effects on the estrogenic, androgenic and thyroid systems to be handled differently from a risk assessment standpoint? The High Throughput Pre-screening System (HTPS) is unlikely to detect indirect effects. However, this may provide the only data that is available on most of those chemicals. Are the ones that are "negative" then neglected? Their ability to harm health and the environment has not been evaluated. Reporter systems can only dependably detect those interactions that are mediated at the level of the receptor. Many results from higher level tests are likely to arise from these indirect effects. Does the inability of seeing the effect on a cellular system containing a reporter system linked to the hormone response element provide a rationale for dismissing endocrine disruption as a mechanism that is likely to be active at low doses? What if the steroid hormone response is mimicked by a membrane hormone effect, as has been demonstrated with Epidermal Growth Factor? Where does mechanism of action fit into this process and how does it modify perceived risks at low dose?

On the other hand, there is the clear long-term advantage of beginning to focus on the association of adverse health impacts with modifications in hormonal control mechanisms. If an adverse effect of a chemical can be clearly associated with an endocrine effect (or any other biochemical/molecular response that can be clearly related to adverse effects), the dose-response relationships can be explored across test systems and into the impacted species. Thus, the impacts of low dose exposure to environmental agents can be explored in greater detail and with greater understanding. This is true, however, only if the health impacts that are associated with changes are understood in fairly explicit ways. Developing massive amounts of screening information on a large universe of chemicals does not necessarily expedite the development of the appropriate scientific underpinning that the Agency needs to broaden this effort. Consequently, the Subcommittee recommends that EPA should not expand the set of agents until validated systems and clear decision criteria are available.

3.1.3 Exemptions from the EDSP

The Food Quality Protection Act (FQPA, 1996) contains a provision which would exempt from the EDSP "...any biological substance or other substance if the Administrator determines that the substance is anticipated not to produce any effect in humans similar to an effect produced by a naturally-occurring estrogen" or, presumably, "...such other endocrine effect as the Administrator may designate." EPA has identified some chemical categories that may be candidates for exemption. Examples include certain polymers with a number average molecular weight (NAMW) greater than 1000 daltons, certain List 4 pesticide inert substances (such as cookie crumbs, strong mineral acids and bases), which are most likely to interact with tissue at the portal of entry giving rise to localized lesions rather than systemic effects, certain bio-pesticides such as plant pesticides or microbials or non-chemical pesticides such as parasitic wasps. In Charge element (a) (3), EPA asked if the Joint Subcommittee agreed with the Agency's position that there are categories of pesticides and other substances that should be exempt from the EDSP. The Subcommittee was also asked to identify any additional categories that should be considered for exemption.

The Subcommittee believes that there are clearly categories of chemicals that should be exempt. However, the boundaries between those compounds that would be exempted and those that would not, must be carefully considered. The selection of 1000 daltons as a cutoff for polymers based on nominal molecular weight appears to have precedent under TSCA, but the scientific justification for this limit was not provided. Many compounds with molecular weight approaching 500 are known to have biological activity. Presumably this precedent will allow polymers with as much as 10% of their total mass to be 500 daltons or less. A more concise statement of the scientific reason for taking the specific action on polymers would have been useful. Clearly, there would be rationales for dismissing other types of chemicals (e.g. amino acids, fatty acids, sugars that are part of normal diets) from the EDSP.

The Joint Subcommittee did not respond to the second issue of this Charge element. We did not think it appropriate for the Subcommittee to identify additional classes of chemicals for exemption. The Subcommittee suggests that the Agency consider handling of exemptions through a rule making process that is transparent and open to public comment.

3.1.4 Mixtures

EPA recognizes that there are important complex environmental mixtures that deserve inclusion in the EDSP. Consequently, EPA plans to include in the EDSP representative mixtures to which large or identifiable key segments of the population are exposed. Initial choices for these high-priority mixture categories include: chemicals in breast milk; phytoestrogens in soy-based infant formulas; mixtures commonly found at Superfund sites; common pesticide/fertilizer mixtures found in ground and surface water; disinfection byproducts; and gasoline. EPA proposes to screen and test (if appropriate) one representative mixture from each category, after it confirms that the screening and testing components of the EDSP are satisfactory for the handling of single substances.

For this review, EPA asked (in Charge elements (a) (4) (A-E)) if the proposal is a reasonable way to address the practicality of screening and testing mixtures; if the chosen six categories are the right ones to address first; if there other mixture categories that should be included in addition to, or instead of those identified; and if standardized representative mixtures be developed. The Agency also sought advice on dealing with those mixtures found to be positive in the screening tier, asking if the whole mixture should be tested in the testing tier or only the active component(s) identified in the screens(s)

The recommendations for handling mixtures are outlined in Section VII of Chapter 4 of the EDSTAC Final Report (EDSTAC, 1998) and were discussed by EPA staff and were the subject of several comments from the public during the Subcommittee's public meeting. Although there was general recognition of the key importance of mixtures as a part of the overall EDSP, there were concerns about the selection process, the experimental design for testing the mixtures, and the ability of the Agency to evaluate and interpret the results of the studies and to effectively communicate this information to the risk managers and the public. There was a consensus by the Subcommittee that the mixtures section of the EDSTAC document needed to be re-worked and there were several public comments recommending that mixtures not be included in the program.

The Subcommittee suggests a compromise proposal: delay starting the mixtures testing program until most of the single agent testing was completed. This would have two advantages; first it would provide a more extensive data base to use in selecting mixtures for testing; and

second it would enable the Agency to benefit from some of the current efforts underway with pesticides (within EPA and by outside research groups) to improve our ability to define and test mixtures. The Subcommittee recognizes that the Agency is currently testing some mixtures (wastewater, cholinergic pesticides etc.) and that relatively little effort would be required to incorporate these into EDSP. We are also aware of studies that compare the effects of mixtures having independent actions with those having identical actions as a way to characterize the risk assessment of mixtures (Yang *et al.*, 1998; Feron *et. al.*, 1995; NAS/NRC/COT, 1989). Similar approaches could be used to standardize or characterize mixtures for testing in the EDSTAC program and would provide more interpretable results than those proposed in this report.

The EPA's final question in the mixtures section addressed phase 2 testing of mixtures and/or the components. The Subcommittee believes that would be prudent to test both the mixture and its components.

3.2 Priority Setting

3.2.1 The Compartment-based Approach to Priority-setting

The EDSTAC report (1998) recommended a compartment-based approach to priority-setting. EPA agrees that this is the appropriate framework, and plans to group chemicals into sets, based on the existence of factual information in a given area. Thus, comparisons can be made between like substances (i.e., chemicals will compete for priority with others on the basis of comparable data and will not be assigned lower priority for lack of information). In Charge element (b) (1), EPA asked the Joint Subcommittee to comment on the principles of the component-based approach to priority setting, and to suggest any other approaches that would be more useful.

The Subcommittee finds that the compartment-based approach[1] is supportable when

[1]This approach first assigns environmental toxicants into four categories (based on available data): a) Specially-targeted priorities; b) Exposure-related information; c) Effects-related information; and d) Integrated Effects and Exposure. Compartments (or sets) are defined within each category, into which agents are assigned on the basis of exposure and/or effects information. The individual agents are then ranked within each compartment

ranking is based on *both effect and exposure* data following guidance in NRC and EPA risk assessment literature (NRC, 1994, 1983; EPA 1997, 1992, 1986). The greatest weight should be given to chemicals for which we have data that indicates actual human or environmental exposure and effects. Lower weight should be given to agents for which the data are indicative of *probable* exposure (in food or drinking water) or *probable* effects (from well-conducted animal studies). The lowest weight and priority ranking should be given to chemicals for which the data are indicative of *possible* exposure (based on Toxics Release Inventory data or known high production volume) or *possible* effects from (*in vitro* research or from Structure Activity Research). The Subcommittee supports the nomination concept (i.e., the process of identifying ("nominating") probable/possible exposure or probable/possible effects as noted above by citizens who are disproportionately exposed because of the group or community to which they belong, or because an ecosystem is disproportionately exposed (EPA, 1999)) but advises the Agency that the process needs further definition and that no unsubstantiated claims be allowed.

The Subcommittee is concerned that the prioritization process is not as "transparent" as it needs to be for public understanding. Also, concern was expressed at the public meeting that health care professionals (both personal health clinicians and public health) may not understand EPA's process, intent, and implementation methods, especially in the context of their own work. Communication at this level needs to improve, or the results will be valueless. Additional concern was expressed that the process appears to have undue emphasis on chemistry and toxicology, with less clear emphasis on health effects. In summary, prioritization should be based on a sound scientific basis.

Other than the comments provided above, the Subcommittee has no suggestions for alternative approaches.

3.2.2 The Relational Database and Priority Setting

EPA is developing a relational database to assist in developing priorities for screening. This database is intended to import and synthesize existing data, allowing EPA to estimate

---

from highest concern to lowest concern to set priorities for eventual Tier one screening.

certain parameters through modeling. It is expected to have great value in helping to identify the specific compartments under the EDSTAC's compartment-based priority setting approach. The database will also be helpful in selecting chemicals for the first and subsequent rounds of screening. The EPA asked the Joint Subcommittee to comment on this approach, and to provide advice to improve the content of the relational database or its implementation (Charge element (b) (2)).

The Subcommittee believes that the proposed relational database shows strong promise of being a useful tool, as long as it does prove to be truly relational, i.e. provides a means to relate environmental exposure data with toxicological effects information. This step is the very core of the risk assessment process and will ultimately provide the most valuable guidance in the priority setting approach.

When designing a relational database, it is important to consider relationships that accomplish more than simple indexing of component data, so that the database is not just a resource to access information on a specific agent. Although this capability is valuable in of itself, one needs to consider biologic relationships in modeling the database, so that it can become a more active investigational tool. For instance, the data need to be collected and organized in a way that can respond to our growing knowledge in gene sequence at specific loci and implications for health and disease. Likewise, the data need to be organized so that developmental gene networks and other biological hierarchies can be reflected appropriately in the database. We are moving beyond single major risk factors for particular outcomes and into the complex gene-gene and gene-environment relationships which characterize common diseases (cancer, heart disease, behavioral disorders, aging etc.) This is a very difficult challenge for genetic epidemiologists and has enormous design and analytical ramifications. Since the type of information provided by the database will affect greatly the interpretation of available data, its design will inevitably affect prioritization.

The database needs to reflect knowledge throughout the specified organism's life cycle, and should be able both to examine longitudinal developmental changes within a system, and to make cross-sectional comparisons across the organism. The goal is to facilitate creation of a biologically-plausible chain of causal inference. The database also needs to be prepared to deal

with a rapidly growing genetic database on variation in endocrine system-related genes.

The database should be designed so it can be readily interfaced with human health surveillance data on disorders such as birth defects and cancer. The National Institutes of Health, the Center for Disease Control, and other agencies are working with states to strengthen these surveillance systems, and in some states (e.g. Iowa) these systems have been constructed together with environmental quality databases. These have been used for aggregate (ecologic) studies of health outcome risk factors. It is expected that such capabilities will be substantially expanded in the next few years with corresponding implications for priority setting.

Priority setting should also address those persons or organisms found to be "most-susceptible", but not be limited to this subpopulation alone. There needs also to be focus on population disease burden. Individual rare genes may be major risk factors for a few persons, but may contribute less to the burden of a disease in a population than do "minor risk factor genes" which are common in the population.

Finally, there is one important problem which must be considered in using the relational database as proposed. The Subcommittee expressed concern that the lack of effects data on the universe of chemicals currently in commercial use will lead to a relational data base that only identifies known problem chemicals that are already well-studied. The Subcommittee encouraged the development and use of new techniques including quantitative structural activity relationships, molecular modeling, and androgen binding, in addition to solubility (K_{ow}) and other measures to help identify the bio-available, potentially active compounds for further testing in the EDSP.

3.3 High Throughput Pre-Screening Approach

Based on recommendations from the EDSTAC (1998), EPA proposes to implement a priority setting strategy that includes an initial sorting, based on an examination of existing information. This initial sorting strategy leads to four possible classifications: i) polymers; ii) chemicals with sufficient data to proceed to testing; iii) chemicals with sufficient data to proceed to hazard assessment; and iv) chemicals with insufficient data, which presumably, would go into

the screening phase. EPA anticipates that a large number of substances will end up in category iv. To provide biological data to assist in the sorting process, EPA proposes to conduct High Throughput Pre-Screening (HTPS) on a significant number of substances (perhaps as many as 15,000), using *in vitro* assay systems incorporating transcriptional activation or reporter gene systems for the estrogen-, androgen- and thyroid-hormone systems.

In Charge element (c) (1) EPA asks two questions about this approach: first, is this a reasonable approach and sorting strategy to support priority setting?; and second, based upon the data developed through the pilot study to date, can this technique can be used as a prescreening device, and what modifications/improvements must be made? (The latter two questions are addressed in section 3.3.2, below.)

3.3.1 High Throughput Technology As A Tool for Priority Setting

EDSTAC has recommended the use of HTPS in order to address the problem that most chemical substances on the TSCA Inventory have little or no data regarding their potential to interact/modulate/disrupt the endocrine system. HTPS is designed to: a) provide priority setting information for chemicals to be examined in Tier one Screening (T1S); b) provide a prospective on the effectiveness of HTPS relative to other methodologies such as QSARs; and c) satisfy the receptor binding/in vitro gene expression T1S requirement for those chemicals that go through HTPS.

EPA does not intend to use HTPS data to establish the endocrine disrupting status of a chemical. Nevertheless, there is considerable concern that results from HTPS will be the first available data, and will thus be (inappropriately) used, resulting in a certain stigma or in product de-selection. This is a concern and appropriate measures should be spelled out and taken in order to ensure that the data from HTPS is not misused.

Unfortunately, the EPA-funded demonstration project with OSI Pharmaceuticals failed to sufficiently demonstrate the utility of their HTPS system for the purpose of chemical sorting and priority setting of estrogen, androgen and thyroid active chemicals. The Joint Subcommittee

raised several concerns regarding the responsiveness and selectivity of the assays developed to date. We believe that the currently available data obtained from the OSI assays would not be of assistance in chemical sorting and priority setting. However, it was acknowledged that this was a work in progress and that, in general, the HTPS approach had merit but required further development prior to implementation. Therefore, the Joint Subcommittee agrees that in conjunction with other priority setting data, results from estrogen and androgen receptor HTPS assays could contribute to chemical priority setting provided the assays are validated and standardized. The Subcommittee also questioned the utility of the thyroid receptor HTPS assay, since there are no known examples of endocrine disruption that occur as a result of chemical interaction with this receptor.

The Subcommittee had one additional concern. EPA's plan for increasing the quality of the assay is appropriate, but there appears to be no contingency plan in the event that it is eventually discovered that the assay is not working. Also, the plan says nothing about a time frame for making adjustments to the assay, nor at what point it would be prudent to discontinue it and seek other approaches.

3.3.2 High Throughput Technology As A Pre-screening Device

Eight transcriptional activation assays have been recommended by EPA. These assays include the estrogen receptor (ER) alpha and beta, the androgen receptor (AR), and the thyroid receptor (TR) in the absence and presence of metabolic activation/detoxification system. The OSI Pharmaceuticals Corporation (under an EPA contract as noted above) initiated a study to determine the feasibility of using AR and TR transcriptional activation assays to pre-screen chemicals in the presence and absence of a metabolic system. Sixty-one chemicals were examined including known ER or AR agonists and antagonists. The known ER and AR active chemicals were selected in order to span a wide range of potencies. As of March 5, 1999, stably transfected ER and TR transcriptional activation assays in the absence of metabolic systems have been used to assess the 61 selected chemicals. A stable AR cell line has been selected and was used in an initial pilot screen of 16 chemicals.

Following a review of the data and an up to date presentation by OSI at the public

meeting, the Joint Subcommittee believes that the OSI HTPS assays were not ready to be used as a pre-screen device. The following modifications/improvements are suggested in order to ensure its usefulness:

a) improve responsiveness and selectivity of assays

b) conduct a thorough statistical analysis of the results to identify significant chemical effects on gene expression

c) validate and standardize all HTPS assays using a training set of agents known to be either positive or negative with regard to endocrine disruption. Use of this set should identify the error rate, i.e., the percentage of false positive and false negative findings.

d) verify the results by comparing to other bench gene expression assays

e) develop assays that would be capable of distinguishing interactions between estrogen receptors alpha and beta.

f) establish/define criteria for positive, negative and equivocal results

The Joint Subcommittee also made the following suggestions regarding the use of HTPS assays for the purposes of priority setting:

a) re-open the bidding process to include other assays (e.g., receptor binding) and to identify additional analytical resources

b) consult with intramural EPA scientists and extramural scientists with expertise in receptor binding/gene expression assays to evaluate responses to any Agency request for proposals

c) investigate the development and utility of other HTPS assays such as gene chip/cDNA array assays and computer modeling of receptor ligand binding domain-chemical interactions

3.4 The Proposed Endocrine Disruptor Screening Program.

EPA is proposing to develop and implement an Endocrine Disruption Screening Program (EDSP) that consists of two phases: screening, currently consisting of eight components, and testing, currently envisioned to have one to five components, depending upon the need for identifying effects in various sectors of the animal kingdom. EPA posed a number of questions to the Subcommittee concerning the proposed EDSP, comprising Charge elements (d)(1-10). The response to these questions are provided in the following report sections (3.4.1-3.4.10).

3.4.1 The Two-phase Sorting Strategy

In Charge element (d) (1), the Agency asked if it is reasonable and appropriate to develop and implement a two-phase program, the first phase focused on identifying a substance's potential to interact with one or more of the three hormone systems, the second phase to characterize the effects of concern that interaction with these hormone systems might elicit.

The Subcommittee supports the proposal to develop a two-phase program for endocrine disruptor screening and testing (EDST). Further, a formal reevaluation of the screening and testing process at regular intervals should be part of the program. The purposes of this reevaluation process would be to evaluate the effectiveness of the protocols initially adopted for screening and testing and to adopt new protocols in cases where none currently exist for identifying endocrine alterations or the effects of those alterations. Adoption of new screens and tests should also mean the elimination of previous, less useful ones.

The suggestion was made that non-mammalian systems might serve as developmental

screening tests. Amphibians, birds and fish have all been used for developmental screening to provide an integrative assessment system. The fish and the bird assays seem to be the most sensitive. Of the three, the basic mechanisms underlying development are best understood in the bird to date, but some fish (especially zebra fish and medaka species) are rapidly catching up. Of the amphibians, the frog embryo teratogenesis assay Xenopus(FETAX) may be adaptable to a fully integrative screening assay (Fort, 1995, 1996).

As multiple laboratories are likely to be running the prescribed assays, it is important to establish procedures for standardization among laboratories and for training of the technicians and scientists who will run the screens and tests. Significant consideration and planning needs to be conducted on how to ensure inter-laboratory standardization.

The Joint Subcommittee is concerned that the planned process of testing for endocrine disruption is "**ahead of the science.**" Without a clear statement of the goals of identifying endocrine disruption, it is difficult to prepare a valid testing program for this particular effect. The absence of an anticipated report from the National Research Council (NRC) study committee on endocrine disruptors made the task of this review Subcommittee more difficult from a procedural perspective. We may need to reevaluate our recommendations on the underlying science for screening and testing once conclusions drawn by the NRC study committee are available.

3.4.2 Adequacy of the Screening Battery

The EDSTAC (1998) recommended, and EPA proposes to implement, a screening battery consisting of eight assays, three *in vitro* and five *in vivo*, to address estrogen, androgen, and thyroid systems (EAT) effects. The Agency believes that these eight assays would detect all substances currently known to interact with the EAT, and asked the Subcommittee (in Charge element (d)(2)) if it agrees with this position. Also, EPA sought advice on making changes in the battery to assure the identification of substances of potential concern.

It is difficult to evaluate the proposals provided to the Subcommittee without a clear determination of the scope and nature of the problem for both humans and wildlife. The

Subcommittee agreed that, as a minimum, the Agency must develop an acceptable methods and standardization and validation program for all proposed testing methods. The program proposed is clearly screening more for mechanism than adverse responses. There are other potential mechanisms in the EAT systems of which we are not fully aware. Thus, the information on di-n-butyl phthalate (Gray, 1999a) presented by EPA staff at the public meeting clearly show major adverse effects produced by alterations in the androgen system during development. These effects might not be found using the screens currently employed in Tier 1. This indicates that there are potential critical events that would not be detected with the current screening battery. An *in utero* (or *in ovo*) screen that is recognized as the most sensitive exposure window for endocrine disruptor event (see section 3.4.) should be utilized. The Joint Subcommittee firmly agreed that an *in utero* assay should be developed by the EPA and that it should be considered as a substitution, not an addition to the proposed battery. However, it is imperative that it be validated before becoming a required assay.

Thus, the screens in whole animals would provide access to more potential mechanisms than receptor based screens since these animals would have intact hypothalamic/pituitary/gonadal or thyroid axes and also have multiple end points in the same animals related to endocrine disturbances. Moreover, such screens would also provide positive information on reproductive (and developmental) toxicants that act, for example, directly on the gonads via non-endocrine primary mechanisms, such as methoxyethanol (Foster *et al.*, 1986), but would also affect endocrine end points subsequent to gonadal damage over several days.

The Subcommittee also suggests that developmental nervous system endpoints should be incorporated into the screening assays. This could be done at the level of a "to be developed" integrative screening assay.

The Subcommittee was also of the view that due regard should be given to the dose route employed in the *in vivo* screens. Some flexibility should be employed but it was considered that the most appropriate route of exposure (that which mimics the typical route of exposure in humans and/or wildlife) be chosen with the oral route being the default exposure route. The use of the intra peritoneal route, especially for the uterotrophic assay, was considered to be inappropriate.

Further, addressing technique, it was not clear why the fish protocol presently being refined and tested by EPA's Duluth laboratory specifically says not to use the organ weight corrected for body weight (organ/somatic index). This practice is generically used and recommended to account for any changes in overall body weight induced by the chemical treatment. The protocols should be consistent with each other. Since exposure to the chemicals may well induce changes in body weight as a separate phenomenon from any gonad or other endocrine effects, it would seem that the protocol should be to calculate and report both the raw organ weights and the body-weight-corrected somatic indices.

The Subcommittee was also aware that there is a huge gulf in terms of effort, complexity and cost between Tier 1 screening and Tier 2 testing. The EPA may wish to consider if an intermediate tier would be warranted that would provide valuable information without the expense of multi-hundred thousand-dollar efforts.

3.4.3 Adequacy of Thyroid Coverage

Interaction with a receptor is the principal or key mechanism by which substances exert their effects on the EAT. There is an exception, however, in that this appears not to be the case for the thyroid-hormone system. Consequently, the Agency asked the Joint Subcommittee to comment on the adequacy of coverage of the thyroid provided in the proposed screening battery, and suggest modifications if needed (Charge element (d)(3)).

The Subcommittee believes that the proposed screening battery should detect alterations in thyroid function. However, the screens proposed are more general and less robust than those designed to detect alterations in estrogens and androgens. It would be prudent to have thyroid-hormone-sensitive tests in the screen. Most known thyrotoxicants produce changes in thyroid-related hormones and/or clearance and/or thyroid histology. The proposed EDSTAC screening process for thyroid hormone appears to address these requirements. Measuring hormone levels and thyroid histopathology in rats, and amphibian tail resorption, should effectively capture the strongest thyrotoxicants. The Subcommittee supports the inclusion of Thyroid Stimulating Hormone (TSH) and T3 in addition to the measurement of T4, histopathology and amphibian tail

resorption.[2] Only the proposed amphibian tail resorption test specifically evaluates an effect of thyroid hormones on target tissues.

No data were offered by EPA to support the inclusion of additional tests, other than the fact that T3 is the biologically active form of the hormone, and that an elevation in TSH would confirm a physiologically-relevant reduction in T4 or T3 levels. The data from Cook and O'Connor (in press) showed that for every thyrotoxicant that reduced T3, there were also changes in T4, which offers direct support for the EPA proposal. If the EPA wishes to have the extra confirmation of a thyroid effect (or lack thereof), the Subcommittee would support the inclusion of TSH and T3 in addition to the measurement of T4, histopathology and amphibian tail resorption.

Because few chemicals that alter thyroid function do so by binding to thyroid receptors, binding assays and gene reporter screens for thyrotoxicants have been omitted from Tier 1. This omission is appropriate because it would have generated false negatives for thyroid effects. However, if these receptor assays are to play a large part in priority setting for testing in Tier 1, it needs to be recognized that less information about potential thyroid alterations will be available from the Tier 1 screens than is the case for E and A.

There is broad agreement that most *known* thyrotoxicants produce changes in thyroid-related hormones and/or clearance and/or thyroid histology. And the Subcommittee agrees that it is prudent to have some degree of overlap and complementarity in the screening tests. Consequently, the same function should be evaluated by more than one test. Finally, because hormone signals are both amplitude and frequency modulated signals, and a single-time-point measurement may not capture or identify an exposure-related change when one is primarily measuring a hormone, it is desirable to also measure some downstream functional result of that hormone. The proposed EDSTAC screening process for thyroid appears to address these requirements by measuring hormone levels and thyroid histopath in rats, and tail resorption in

[2] T3 and T4 are two forms of thyroid hormone, the digits "3" and "4" indicating the number of iodine molecules in its atomic structure.

amphibians. These measures should effectively capture the strongest thyrotoxicants.

3.4.4 *In utero* and *In Ovo* Screens and Single Dose Screening

EPA would prefer to have a screening battery which included assays containing an *in utero* or *in ovo* exposure component, given the concern about the potential for effects on the developing vertebrate organism. At the time the proposed screening battery was being assembled, EPA was not aware of the existence of any such screens. The question posed to the Subcommittee in Charge element (d) (4) asks for comment on any such assays that may exist or are under development, and that could supplant or complement one or more components of the proposed screening battery? Charge element (d) (5) addressed EPA preference to conduct each *in vivo* screening assay using only one dose, selected through the use of range finding studies and other information. The Agency asked if the Subcommittee agreed with the single dose approach, and what suggestions it had to modify it. The EPA also sought advice on the possible public health consequences of these false negatives. The Subcommittee decided that, since the issues were inter-related, it would be best to address both elements of the Charge in a single response, which follows below.

The Subcommittee prefers those tests which bundle several endpoints into a single "test unit." The Subcommittee consequently supports the use of gene reporter and binding assays as part of Tier 1. Problems may be encountered because of differences in the specificity of different cell systems and because of patent control of some assay components. A screen, using animals exposed *in utero* and possibly during lactation, is appealing. The Subcommittee strongly encourages the continued development and evaluation of such a protocol. It could replace several individual assays. No protocol for such a test has been evaluated or validated as a screen to date. However, such a test is easily developed by taking pieces of existing protocols (see the discussion in section 3.4.1). The development of such a protocol would significantly improve screening effectiveness, reduce the numbers of animals used, and could improve overall efficiency of screening.

The consensus of the Subcommittee regarding dose levels for *in vivo* screens was focused around two issues. First, for those relatively non-toxic agents, the employment of a single limit

dose (as specified in the *Federal Register* (1998) document, e.g., 1 g/kg/d oral) was considered to be appropriate. Second, in other cases where non-specific toxicity could be possible, the highest dose level tested should elicit some, but not overt, systemic toxicity in line with the establishment of a maximum tolerated dose (MTD). A second dose level should then be employed at one quarter of the MTD. The Subcommittee felt that the application of a multiple of exposure as the highest dose level tested would not be appropriate since exposure data would only be available infrequently. The Subcommittee also found that false negatives from high dose exposure were unlikely, since even where multiple mechanisms may be operating at different parts of the dose response curve, one would not expect to see effects only at low dose levels. A second lower dose level would also resolve some of these questions.

The potential for low dose effects of putative endocrine disruptors was discussed. In view of the preliminary nature of these potentially important findings (see Nagel, *et al.*, !997), the Subcommittee recommends that EPA continue to remain alert to new information on low dose effects. The Subcommittee was pleased to learn that EPA is sponsoring a workshop on this issue in May, 1999.

It should be noted that the current EPA synthesis of the EDSTAC recommendations is inconsistent on the matter of route of exposure. The uterotrophic assay uses subcutaneous or intra peritoneal injection, the Hershberger assay oral gavage, the multi-generation assay uses diet/oral/inhalation, and no route is identified for the pubertal male and female assays. Consistency is preferred unless evidence requires otherwise.

An integrative developmental assay using the chicken was proposed (Henshel, 1998, 1996; Henshel *et al.*, 1997). The assay integrates both a rapid 5-day screening component with a more complete developmental assessment. Many chemicals that are developmental toxicants interact with the embryo during organogenesis. Therefore, using modifications of established procedures, and modeling the system after the mammalian embryo culture systems, the avian embryo may provide a useful assay.

3.4.5 Rigor of The Five Compartment Test Protocol Design

EDSTAC recommended, and EPA is proposing, a testing phase in the EDSP which could have as many as five components (i.e., mammals, birds, fish, invertebrates and amphibians). Each test would be designed to delineate the dose-response relationships of effects of concern for chemicals which yielded positive results during the screening phase. The testing protocols to be used are either upgrades or modifications of existing guidelines, except for the amphibian. In this case, a protocol is being developed *de novo*. The Subcommittee was asked if the planned test protocol designs would provide sufficient rigor to identify the effects of concern and establish their dose-response relationship for disruption in the EAT systems (Charge element (d) (6)).

The Subcommittee concluded that many of the proposed tests were valid assays of endocrine disruptors. They also concluded that methods must be standardized and validated, based on accepted criteria for validation and regulatory acceptance of toxicological test methods. Other tests, however, met with considerable criticism:

a) A more comprehensive, *in-utero* test battery should be assembled to replace several tests in Tier 1 (see sections 3.4.1 and 3.4.4)

b) The Daphnia developmental assay should be considered as a replacement for the mysid assay because there is a better understanding of the endocrine mechanisms in Daphnia (Baldwin *et al.*, 1998; Baldwin and LeBlanc, 1994).

c) The fish assay for endocrine disruption should include the measurement of vitellogenin in male fishes. Vitellogenin is a yolk precursor protein made by the liver in response to estrogen in female but normally not in male oviparous animals. Its detection in male fish is a highly sensitive assay for estrogenic activity. Many laboratories have the ability to detect vitellogenin by radio immunoassay in a variety of species. There are no known barriers to the development of such a vitellogenin test, although it would still have to be standardized and validated.

d) The fish reproduction assay should include some measure of the reproductive fecundity of the selected compounds. Egg production and developmental success

will detect effects which may not be obviously toxic to the organism but might have detrimental effects at the population level. None of the proposed tests with wildlife detect breeding success. Further, exposure to a variety of compounds can alter the sex ratios to favor one sex or the other. The effect of the test chemicals should include an evaluation of the sex ratios of eggs (or other stages of development) treated with the chemicals.

e) The Subcommittee recommends that the EPA examine the use of the Japanese quail to substitute for the proposed avian tests on Bob-white quail and mallard ducks. Japanese quail have the advantage of short generation time and provide a model with a great deal of background information.

f) Although the Tier 2 tests designed to indicate thyroid alterations should "identify effects of concern" they will not effectively determine whether those hormone alterations have adverse effects on the development or function of the target tissues for thyroid hormones. Thus, the proposed tests may be adequate for detecting hormone perturbations but they don't give information about the effects of those perturbations (see additional information in section 3.4.6).

3.4.6 Adequacy of Detection of Critical Endpoints in The EAT Systems

There could be circumstances in which EPA decides to bypass the screening phase for a particular agent and go directly into the testing phase. EPA is proposing for those cases that the chemical under evaluation be tested in all five tests. The Agency is asking the Subcommittee if the tests in the testing phase will be adequate to detect all known critical endpoints in the EAT hormone systems, and what modifications the Subcommittee might recommend (Charge element (d) (7)).

A central point in the Subcommittee's discussion was that the tests employed in Tier 2 will be the ones used in risk assessment. The number of tests employed would be dependent on the use and potential exposure for each chemical. Clearly no single test, or group of tests, has the ability to cover all critical endpoints for the EAT systems; they should, however, cover most

endpoints. Many research techniques, especially with regard to the thyroid system, are not at a stage where they can be ready for application in a regulatory testing scenario. The Subcommittee recommends that the EPA should remain alert for new techniques and end points to improve testing protocols when these become robust and applicable for routine testing. The Subcommittee also suggests that specific consideration be given for the use of Japanese quail in the avian reproduction study and the use of *Daphnia* spp. as a useful alternate species for invertebrates. It was also unclear during the discussions if the proposed *Mysid* species did indeed have a functioning EAT system. Since these were the specific endocrine systems laid out in the EDSP, it would be inappropriate to propose a species in which estrogen, androgen and thyroid hormones did not have a physiological role.

The immediate focus of many of the proposed tests is on mechanisms. The ultimate goal, however is the capability to detect adverse effects on reproduction and development in a variety of species. Thus, all chemicals interfering with reproduction and development should be detected in these test systems, including those whose primary mechanism is not via a disturbance in the endocrine or EAT systems. Although the risk assessment for any adverse effect and the dose-response data for that adverse effect will be provided by these tests, it will be unfortunate if all reproductive and developmental toxicants are labeled as "endocrine disruptors." This issue further raises the need for a clear definition of an endocrine disruptor – if it is to receive special consideration - as opposed to being treated as any normal reproductive or developmental toxicant.

The advent of new test end points (especially for incorporation into the mammalian two-generation reproduction study) has raised questions about the adversity of specific responses and the normal range for these end points (e.g. anogenital distance, preputial separation, vaginal opening). Guidance from the EPA would be especially welcome in these specific areas of testing.

The Tier 2 tests include few endpoints that will detect critical target tissue effects of thyroid hormone alterations. Such tests are needed to provide suitable information about whether alterations in thyroid hormones (which should be detected by the proposed screening and testing) will affect other developmental, morphological or physiological endpoints in target

tissues. Currently such tests are not available for quick adoption. However, the research information about these effects is available and could be used as the basis for development of such tests at later stages of this program. With respect to the evaluation of thyroid function, the proposed Tier 2 tests seem marginally adequate for providing information to the final program stages of hazard evaluation and risk assessment. Thus, although the addition of more tests to Tier 2 should not be done lightly, there are serious questions about the adequacy of the thyroid tests for assessing whether there are adverse effects of thyroid alterations.

The proposed Tier 2 tests include some endpoints affected by alterations in thyroid function - e.g. growth. However, the proposed measurements are not very sensitive and most are ones that involve the interactions of several hormone systems.

3.4.7 Additional Screening for Agents Initially Found to Be Negative

EPA wished to know what, if any, additional screening or testing would be required to assure that an agent is not an EAT disruptor, if the results of any of the testing phase tests are negative (Charge element (d) (8)) .

The Subcommittee agreed that, if an agent is found to be inactive in the Tier 2 tests, it would be regarded as being inactive as an endocrine disruptor. This is axiomatic, as the Tier 2 tests were selected to define the endocrine toxicity of agents found to be potentially active in the Tier 1 tests. So the answer to the question posed is that no further testing would be required.

The Members also noted that an agent found to be active in Tier 1 tests, but inactive in Tier 2 tests, should be considered to be inactive as an endocrine disruptor. In particular, the positive Tier 1 data should not assume" a life of its own" after the Tier 2 tests are found to be negative.

3.4.8 Endocrine Disruptors and Hazard Assessment

Testing (as opposed to screening) phase tests will identify effects of concern that are the

consequence of endocrine disruption. They may also identify effects of concern that are not the consequence of endocrine disruption. Thus, it may not be possible to determine if a substance is an endocrine disruptor if it has not been subjected to some or all components of the screening battery. Because of this, EPA has asked the Subcommittee if it is important to be able to identify substances as endocrine disruptors from the standpoint of conducting a hazard assessment, and if so, why (Charge element (d) (9).

It is important to be able to identify substances as endocrine disruptors from the standpoint of conducting a hazard assessment. If a compound causes toxicity, it should be treated like all other toxicants. On the other hand, knowing that a compound is more toxic to developing hormonal systems means that the particularly vulnerable populations are more likely to be protected. The perception among the Subcommittee Members is that hazard assessment will not likely be impacted, but that risk assessment will be improved.

3.4.9 Validation of The Proposed Screens and Tests

EPA is proposing a validation program in which the maximum validation effort will consist of conducting each assay in three laboratories. EPA believes that there currently is a wide variation in the state of validation of each of the proposed screens and tests, and that the validation efforts should be tailored for each assay/test accordingly. EPA plans to focus first on the validation of the mammalian assays, as they are both better developed than the non-mammalian assays and are more directly relevant to meeting the FQPA and SDWA mandates for a screening program for potential human health impacts. EPA's preliminary determination of the areas needing development are: a) the uterotrophic assay; b) the Hershberger assay; c) the pubertal male and pubertal female assays; d) the mammalian two-generation reproduction test; and e) the non-mammalian screens and some of the non-mammalian tests.

EPA asked if the Joint Subcommittee agrees with the Agency's assessment of the current status of the screens and tests, and, if it reached differing conclusions, to provide the background and rationale for its findings (Charge element (d) (10)).

It was agreed that the new mammalian multi-generation assay protocol would require

validation of its *practicality*. It cannot be validated *per se* because it is an apical Tier 2 test. The Subcommittee recommended that the validation should proceed sequentially. One laboratory should establish practicality, and that result should then be confirmed in one or two additional laboratories. An objective appraisal of the result of the first run could well indicate that the protocol is practical, and the second phase of validation may be canceled. This point is important, given the time taken to conduct the assay, and the present need for the assay as an apical (Tier 2) test.

The Subcommittee also agreed that the non mammalian Tier 2 tests, as well as the mammalian tests, would require formal validation as to their practicality *and* sensitivity/specificity.

The purpose of the Hershberger assay is to quantify the effects of potential anti-androgenic and androgenic compounds on the hormone-dependant tissues in the immature male rat (Hershberger, *et al.* 1953). Castrated immature male rats, reared under standardized housing conditions, are treated with a potential xenobiotic or the vehicle daily for seven to ten days via oral gavage from 28-37 days of age. The animals are then euthanized and the relevant target tissues are fixed and stained and examined for histopathology. Serum thyroxin (T4) and TSH is measured,. as well as measurements of serum Luteinizing Hormone and androgen measurements are optional. The data are then analyzed for statistical significance of any differences found between the treated animals and the controls.

The Subcommittee was concerned that the existing animal assays in Tier I may not be sensitive to events occurring uniquely in the foetus or in the developing neonate/weanling. The development of a limited *in utero* assay is currently under study, and several laboratories are evaluating the effects of a range of endocrine active chemicals on sexual development of perinatal rats and mice (Gray *et al.*,, 1999a; 1999b). Uterotrophic effects in the female weanling rat are already incorporated as an alternative assay in the EDSTAC proposals, and work is being done on the male weanling at present (Ashby, and Lefevre, 1997; Gray *et al.*, 1997).

We recommend that these assays be kept under close review, with attention focused on

the results obtained when testing the activity of the same agents with the different types of assay. It may be that data will eventually indicate that one or other of these classes of assay can replace the existing rodent assays in Tier I, but a well constructed, robust, database will be needed before such a decision can be made. The Subcommittee endorsed strongly the continuing evaluation of endocrine disruption assays that cover the periods of gestation and sexual development.

3.4.10 Subcommittee Recommendations to Help EPA Meet its Charge

The final element of the Charge (d) (11) asked the Joint Subcommittee for any other suggestions or recommendations that would help EPA meet its charge.

The body of this report provides specific recommendations concerning the screening and testing of endocrine disruptors, as posed by the Charge. The following section of this report contains a summary of our major findings and recommendations.

4. MAJOR FINDINGS AND RECOMMENDATIONS

This section highlights a variety of recommendations and concerns discussed at the public meeting, or generated during the preparation of this report. These findings are:

a) **Evaluating the Program:** We wish to reinforce the comments concerning the lack of on-going program evaluation in section 3.1.1 of this report. We find no provision for mid-course evaluation or optimization of the process. The Agency is mandated to assemble and evaluate this proposed panel of tests and then to implement them, but a correlate responsibility is to make sure that what's being done is the best that can be. Edmund Burke's "You can never plan the future by the past," and Robert Burns' "The best-laid plans of mice and men oft gang agley." both apply here. For example, evaluation of minced testis and minced-ovary assays finds them to be only 50% effective in identifying compounds that inhibited steroid biosynthesis (Powlin *et al.*, 1998). Although something looks fine on paper or in a small research setting, translating it into volume-screening mode may be quite another thing. There was broad support among the Subcommittee for the concept that the Agency should convene a panel of independent scientists to review all the screening data for 50-100 compounds, with eye towards revising the process and eliminating those methods that don't work.

A second reason for doing this is that the more removed a screen is from a whole model, the more wrong the answers are likely to be. For example, the phthalates don't show up as anti-androgens in the Hershberger assay, although they clearly have this activity in intact rodents Gray, 1999a).(.

Finally, we believe that the regulated community and the public interest groups would be more willing to participate if they knew that the system was going to be optimized as it proceeded. The Agency should have one or more evaluations of the process as we proceed with this, and the Subcommittee strongly encourages this.

b) **Mixture Issues:** Discussions at the public meeting focused on whether to include mixtures in the listing of materials to be screened and tested. The Subcommittee agreed that the initial focus of the methods development effort must necessarily focus on single compounds and leave the question of testing of mixtures until accepted single-compound methods have been completed. However, Agency representatives underscored the need to apply the methods to testing of effluents and source waters which are obviously complex mixtures. The Subcommittee concluded that very promising methods already exist in the field of ecotoxicology. These include the Whole Effluent Testing (WET) and Toxicity Identification Evaluation (TIE) procedures developed by the Agency in concert with the Society for Environmental Toxicology and Chemistry. Those methods have been developed to test effects of effluents and would have direct application to the Endocrine Disruptor Screening Program.

c) **Case Studies:** The Subcommittee strongly encourages the Agency to include more and better-detailed case studies in the evolution of the priority-setting scheme. Case studies will enable a realistic test of the scheme, checking sensitivity of the system and it's working practicality to actually prioritize chemicals for further testing.

d) **Sub-population Compartment:** The question of the need for a separate compartment to address sub-populations (i.e., human babies) was addressed to the Subcommittee. Our conclusions supported the use of sub-populations as a criterion within the existing compartments already identified, but not as a separate stand-alone compartment.

e) **Use of IRIS:** The priority testing scheme relies on the use of several data-bases summarizing the environmental fate and effects of chemicals. Caution was expressed by several members of the Committee that there are numerous problems with the validation of IRIS and other data-bases. Before placing heavy reliance on these computerized systems, users need to be aware of these validation

problems and proceed with caution before incorporating these values unilaterally.

f) **Exposure:** The Subcommittee expressed concern that consideration of the toxicological implications of exposure should include both dose *and* time of exposure. The current scheme does not adequately cover the time aspect of exposure and this needs to be remedied before broad-scale application of the approach.

g) **Use of Animals:** During the public meeting, concern was expressed about the large number of animals that would be needed in the EDSTAC program. The Subcommittee was asked whether alternatives and approaches to minimize animal use had been appropriately considered in developing the protocols. The Subcommittee pointed out the essential role animals play in tests to detect endocrine disruption to reveal adverse effects on humans. There are no substitutes for tests currently available for the Tier 2 tests using animals. Because of the complexity of the biological systems involved in endocrine disruptor detection, animals will remain a necessary model for the foreseeable future. Additional comments by Subcommittee Members and others described protocol or method modifications which would be less expensive, faster and use fewer animals. The Agency has an obligation to conserve all resources in developing new testing protocols, and the use of animals in such tests poses both ethical and practical problems

h) **Need for an Introductory Statement:** The previous EDSTAC meeting suggested that the final document needed, as a introductory section, a description of the problem or the scientific or health-based reason for the EDSTAC program (1998). Although the anticipated NAS/NRC report is expected to address this issue, the Joint Subcommittee now urges the EDSTAC team to include a description of both the health and ecological problems associated with exposure to the endocrine disruptors and to show how the program is justified by these findings.

i) **Support for Decisions**: Decisions about which assays are selected, and which

protocols are adopted for those assays, should be supported with data that are generally available.

j) **Exceptions**: Testing strategies will always have exceptions. Care should be taken to be aware of the imperfect nature of *any* future agreed strategy. In particular, there is the present danger that the two chemicals dibutyl phthalate (DBP) and methoxychlor (MC) will have an undue influence on the future. MC stands as the only recognised precedent for using animals in a screening mode. However, even this is not secure as this chemical has been published as being active as an estrogen *in vitro* (Gray *et al.*, 1999c; Bulger *et al.*,1978).

k) **Negative Control Agents**: There is a need to define and agree on some negative control agents for ED assay validation. It has been suggested that the only valid one at present is diethyl phthalate. (Foster, 1980). This position is supported by the fact that it gave negative results in a full and updated rodent multi- generation study at the National Institute of Environmental Health Sciences (Chapin, 1997). Assay specificity will not be capable of assessment unless further such agents can be found and made available for general study.

l) **Animal Tests and Routes of Exposure**: As noted above, MC stands alone as the only precedent for why animals should be used in the screening mode (Tier 1). There is significant international concern on the proposed use of animals for screening. Ashby and Lefevre (1997) proposed that short term animal studies should be recognised as an intermediate Tier, much as happens now with the anticipation of animal carcinogens [screen *in vitro,* assess in short term *in vivo* assays, and then define in lifetime bioassays]. In this role of hazard assessment (as opposed to hazard definition) biologically relevant routes of exposure would be indicated (oral gavage, diet, water, inhalation, skin painting). At present, use of the subcutaneous injection or intraperitoneal injection routes are recommended in the frail quest of increasing assay sensitivity. In fact, irrespective of the outcome of this suggestion it should be noted that the current EPA synthesis of the EDSTAC recommendations is inconsistent on the matter of route of exposure -

the uterotrophic assay uses subcutaneous or intra peritoneal injection, the Hershberger assay oral gavage, the multi generation assay uses diet/oral/inhalation, and no route is identified for the pubertal male and female assays.

The main contributors to differences in test outcome between assays conducted in vitro and assays in rodents will be delivered dose, pharmacodynamics, and pharmacokinetics. Route of exposure will dominate these factors. The ultimate role to be adopted for animal studies, and the route of animal exposures, will have the greatest impact on the successful implementation of the EPA initiatives in the area of endocrine disruption.

REFERENCES

Ashby, J., and P.A. Lefevre. 1997. The weanling male rat as an assay for endocrine disruption: preliminary observations. Regul. Toxicol. Pharmacol 26(3): 330-337.

Baldwin, W.S., Graham, S.E., Shea, D., and G.A, LeBlanc. 1998. Altered metabolic elimination of testosterone and associated toxicity following exposure of Daphnia magna to nonylphenol polyethoxylate. *Ecotoxicol. Environ. Saf.*.39 (2): 104-111.

Baldwin, W.S., and G.A. LeBlanc. 1994. Identification of multiple steroid hydroxylases in Daphnia magna and their modulation by xenobiotics. *Environ. Toxicol. Chem.*(13) (7): 1013-1021

Bulger, W. H., *et al.* 1978. Interactions of methoxychlor, methoxychlor base-soluble contaminant, and 2,2-bis(p-hydroxyphenyl)-1,1,1-trichloroethane with rat uterine estrogen receptor. *J. Toxicol Environ Health.* 1978 4(5-6):881-93.

Chapin, R. 1997. Reproductive assessment by continuous breeding: evolving study design and summaries of 88 studies. *Env. Health Perspectives* 105 (1):199-395.

Cook. J., and J. C. O'Connor (In Press).

EDSTAC (Endocrine Disruptor Screening and Testing program Advisory Committee). 1998. EDSTAC Final report (Vols 1 and 2). U.S. Environmental Protection Agency, Washington DC.

EPA (Environmental Protection Agency). 1999. "Strawman" priority setting process for endocrine disruptor screening and testing. Washington, DC.

EPA (Environmental Protection Agency). 1997. Guiding principles for Monte Carlo analysis. EPA/630/R-97/001.

EPA (Environmental Protection Agency). 1992. Guidelines for exposure assessment. Federal Register 57 (104), 22888-22938.

EPA (Environmental Protection Agency). 1986. Guidelines for the health risk assessment of chemical mixtures. Federal Register 51 (185), 34014-34025.

FR (Federal Register). 1998. Endocrine Disruptor Screening Program: Statement of Policy. United States Government Printing Office, Volume 63, Number 248:71541-71568. Washington DC

Feron, V.J., Groton, J.P., vanZorge, J.A., Cassee. F.R., Jonker, D., and P.J. vanBladerem. 1995. Toxicity studies in rats of simple mixtures of chemicals with the same or different target organs. *Toxcology Lett.* 82/83: 505-512

FQPA, 1996

Foster, P.M., Blackburn, D.M., Moore, R.B., and S.C. Lloyd. 1986. Testicular toxicity og 2-methoxyacetaldehyde, a possible metabolite of glycol monomethyl ether, in the rat. *Toxicol Appl Pharmacol* 32(1-2):73-80.

Foster, P.M., Thomas L., Cook, M.W., and S.D. Gangoli. 1980. Study of the testicular effects and changes in zinc excretion produced by some n-alkyl phthalates in the rat. Toxicol. Appl. Pharmacol. 54(3):392-398.

Fort, D. J., Stover, E. L., and J.A. Bantle. 1996. Integrated ecological hazard assessment of waste site soil extracts using fetax and short-term fathead minnow teratogenesis assay. ASTM Spec Tech Publ. (1262): 93-109.

Fort, D,J., Stover, E.L., and D. Norton. 1995. Ecological hazard assessment of aqueous soil extracts using FETAX. *J. Appl. Toxicol.* 15, (3):183-191.

Gray, L.E., Wolf, C., Lamhright, C., Mann, P., Price, M., Cooper, R.L.,and J. Ostby. 1999a. Administration of potentially antiandrogenic pesticides (procymidone, linuron, iprodione, chlozolinate, p,p'-DDE, and ketoconazole) and toxic substances (dibutyl- and diethylhexyl phthalate, PCB 169, and ethane dimethane sulphonate) during sexual differentiation produces diverse profiles of reproductive malformations in the male rat. Toxicology And Industrial Health(15)1/2: 94-118.

Gray, L.E., Ostby, J., Cooper, R.L., and W.R. Kelce. 1999c. The estrogenic and antiandrogenic pesticide methoxychlor methoxychlor alters the reproductive tract and behavior without affecting pituitary size or LH and prolactin secretion in male rats. *Toxicology and Industrial Health* (15) (1/2): 37-47.

Gray, L.E., Ostby, J., Monosson, E., and W.R, Kelce. 1999b. Environmental antiandrogens: Low doses of the fungicide vinclozolin alter sexual differentiation of the male rat. *Toxicology and Industrial Health* 15 (1/2): 48-64.

Gray, L.E., Jr., Kelce, W.R., Wiese, T., Tyl, R., Gaido, K., Cook, J., Klinefelter, G., Desaulniers, D., Wilson, E., Zacharewski, T., Waller, C., Foster, P., Laskey, J., Reel, J., Giesy, J., Laws, S., McLachlan, J., Breslin, W., Cooper, R., Di Giulio, R., Johnson, R., Purdy, R.,

Mihaich, E., Safe, S., and T. Colborn. 1997. Endocrine screening methods workshop report: detection of estrogenic and androgenic hormonal and anti-hormonal activity for chemicals that act via receptor or steroidogenic enzyme mechanisms. *Reprod. Toxicol.* 11(5):719-750.

Henshel, D.S. 1998 Developmental and neurotoxic effects of dioxin and dioxin-like compounds on domestic and wild avian species. *Environ. Toxicol Chem.* 17(1):88 - 98.

Henshel, D.S., Sparks, D.W., Allen, C.A., Benson, K., Fox, C., Lam, Y., Sobiech, S.A., Wagey, R.. 1997. Preliminary results using the early embryo teratogenesis assay: a comparison of early embryo abnormalities with late embryo and hatchling teratogenic changes. Environmental Toxicology and Risk Assessment: Modelling and Risk Assessment (6th Volume),ASTM STP 1317, F. James Dwyer, Thomas R. Doane, and Mark L. Hinman, Eds., American Society for Testing and Materials, 391-401.

Henshel, D.S. 1996. An argument for the chicken embryo as a model for the developmental toxicological effects of the polyhalogenated aromatic hydrocarbons (PHAHs). Environmental Toxicology and Risk Assessment: Biomarkers and Risk Assessment (5th Volume),ATSM STP 1306, David A. Bengston and Diane S. Henshel, Eds., American Society for Testing and Materials, Philadelphia, pp. 219 - 229.

NAS/NRC/COT (National Academy of Science/National Research Council/Committee on Technology). 1989. Safe Drinking Water Committee, Vol. 9, Drinking Water and Health, Part II Mixtures. National Academy Press, Washington DC.

NRC, National Research Council. Risk Assessment in the Federal Government: Managing the Process. Washington, D.C., National Academy Press (1983).

NRC, National Research Council. Science and Judgment in Risk Assessment. Washington, D.C., National Academy Press (1994).

Nature, 385,494 1997).

Hershberger, *et al.* 1953) Proc. Soc. Exp.Biol. Med. 83:175)

Nagel, S.C., von Saal, F.S., Thayer,K.A., Dhar, M.G., Boecher, M., and Welshons, WV. 1997. R elative binding affinity-serum modified access (RBA-SMA) assay predicts the relative in vivo bioactivity of the xenoestrogens bisphenol A and octylphenol. *Enviiron. Health Perspect.* 105:70. (Nature, 385,494 1997).

Powlin, S.S., Cook, J.C., Novak, S., and J.C. O'Connor. 1998. Ex vivo and in vitro testis and ovary explants: utility for identifying steroid bisosysthesis inhibitors and comparison to a tier 1 screening battery. *Toxicological Sciences* 46:51-74..

Yang, R.S.H., Thomas, R.S., Gustafson, D. L., Campain, J. A., Benjamin, S.A., Verhaar, H.J.M., and M.M. Mumtaz. 1998. Approaches to developing alternative and predictive toxicology based on PBPK/PD and QSAR modeling. *Envir. Health Perspectives*: 106 (sup) 1385-1393.