ECETOC Document No. 36

Comments on OECD Draft Detailed Review Paper: Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals

August 1997

Comments on OECD Draft Detailed Review Paper: Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals

CONTENTS

SUMMARY

The Draft Detailed Review Paper (DDRP): Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals is an OECD Environmental Health and Safety Publication which represents a monumental effort to review a large number of assays and forms the most comprehensive, well-written, and balanced overview of available test methods to date. Its focus is to list, describe and partly evaluate the potential methods that can be used to detect compounds which can act via oestrogenic/anti-oestrogenic and androgenic/anti-androgenic mechanisms. The DDRP clearly states that it does not cover other mechanisms of endocrine disruption (involving other hormonal systems). With a focus on these two endocrine mechanisms, there is an inherent potential risk of providing a false sense of security that enacting screening strategies based on these two endocrine classes will protect human health and the environment.

The DDRP endorses the definition of an endocrine disrupter[1] agreed at the European Workshop and correctly recognises that *in vivo* procedures are required to define potential hazard and ultimate risk. However, the DDRP does not address the fact that "endocrine disruption" is neither an endpoint nor an adverse effect *per se*, but one of many known mechanisms which may or may not result in an adverse effect. This has a major consequence since toxicological and ecotoxicological tests (especially OECD Guideline studies) have been designed to identify hazards and not the underlying mechanisms, e.g., endocrine disruption. Hazard identification through adequately designed *in vivo* studies is necessary for identification of inherent endocrine disruptive potential and along with dose-response data, is fundamental to meaningful risk assessment. Mechanistic data, on the other hand, may be important in defining appropriate experimental design (e.g. timing of dosage, period of observation of effects etc.) of the hazard identification studies.

A key factor in the interpretation of toxicological findings in the area of reproduction/developmental toxicity as well as in assessing endocrine disrupting potential is consideration of the role of general toxicity. It is so crucial for understanding and interpreting results, that it is difficult to think of a compound which would not cause endocrine disruption under some circumstances, namely at high doses in the range of overt (general) toxicity. Thus it would be advisable to include a special Section devoted to "Selection of Dose/Interpretation of Results", because of its importance to the process of obtaining valid and meaningful results.

[1] "An endocrine disrupter is an exogenous substance that causes adverse health effects in an intact organism, or its progeny, consequent to changes in endocrine function". This definition was agreed upon at the European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife, Weybridge, UK, Dec 1996 (MRC Institute for Environment and Health, 1996).

Despite the general adequacy of current Guideline studies to detect endocrine disrupting potential, it has been proposed in the DDRP that additional endpoints should be added to the existing protocols. The addition of endpoints should be carefully considered however, since current observations may provide adequate sensitivity which may not be enhanced by the added measurement thus increasing complexity and cost without the benefit of improving detection of potential hazards.

It is recommended that the DDRP addresses separately the specific protocols or tools which define mechanisms from those which assess hazards. *In vitro* assays may contribute to ultimate understanding of mechanism of action, but the potential application of *in vitro* assays needs to be put into context with the limitations noted in the DDRP and with the following: i) effects on reproductive organs are not always mediated by endocrine interaction; and ii) it is questionable if any interaction seen *in vitro* is a surrogate of an *in vivo* hazard. In this respect, *in vitro* assays do not seem suitable for hazard identification. Further, thorough validation of any assay is vital prior to its implementation as a Guideline.

The main focus of testing substances with endocrine-disruptive potential in wildlife organisms is the impairment of reproduction and development ultimately affecting the stability of populations. Therefore, the evaluation of testing procedures should address these effects, while the identification of mechanisms is of lesser importance. Existing, validated and standardised (regulatory) methods and potential enhancements should be evaluated for their suitability to address reproductive and developmental endpoints. Realistic ways of extending these methods to all stages of *in vivo* testing, short-term screening tests to long-term multi-generation studies should be a primary direction of the DDRP.

In conclusion:

- endocrine disruption is a mechanism (not an endpoint or an adverse effect, *per se*);
- the need for protocols or tools for mechanistic studies, in addition to those assessing hazards, warrants further discussion;
- *in vitro* assays are not suitable for hazard identification;
- at present only *in vivo* studies adequately reflect the great variety of subtle interactions and feed-back mechanisms of the endocrine system;
- introducing new procedures or new parameters in standard test protocols will need careful selection and a thorough validation process;
- addition of a special Section devoted to "Selection of Dose/Interpretation of Results", is strongly recommended because of high-dose phenomena and dose-response issues.

KEY TO DDRP PAGINATION

The specific page numbers of *Draft Detailed Review Paper: Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals* differ between the electronic version OECD web page and the OECD printed copy supplied to the member countries representatives and selected reviewers. The citation of a specific page number in these ECETOC comments on the OECD draft Detailed Review Paper are based on the printed distributed copy. The key is presented in the following Table.

Table: Key to DDRP Pagination

GENERAL COMMENTS

1. The Draft Detailed Review Paper (DDRP) evaluates the potential methods that can be used to detect oestrogenic/anti-oestrogenic and androgenic/anti-androgenic compounds. The review of *in vivo* and *in vitro* mammalian tests is clearly the most comprehensive, well-written, and balanced overview of available test methods to date for its stated goal. However, the DDRP does not cover other mechanisms of endocrine disruption (involving other hormonal systems), as clearly stated in the document. The issue of 'environmental oestrogens' originally focused on chemicals which mimic the action of the natural hormone oestrogen. However, the concern is now encompassing effects on the whole endocrine system and the collective term 'endocrine disrupter' is in general use (by scientists, journalists, environmental organisations, and the general public). Currently, the most favoured definition of an endocrine disrupter is "an exogenous substance that causes adverse health effects in an intact organism, or its progeny, consequent to changes in endocrine function". This definition was agreed upon at the European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife, Weybridge, U.K, Dec 1996 (MRC Institute for Environment and Health, 1996). Whilst the DDRP covers comprehensively the detection of oestrogenic/androgenic agonistic and antagonistic actions, it does not cover action mediated by other endocrine mechanisms.

   By focusing solely on these two endocrine mechanisms, there are at least two possible consequences. The first is provision of a false sense of security that enacting screening strategies on these two endocrine classes will protect human health and the environment from unintended exposure to endocrine active compounds (EACs). The second is that it will be necessary to expand the number of endpoints in the future to identify not only hazards, but also mechanisms of EACs. Hence, there is a potential for a proliferation of new/revised test methods. This could lead to a complex mixture of studies which cannot be conducted in a cost-effective manner. It should be borne in mind that toxicological tests (especially OECD protocols) have been designed to identify hazards (without the necessity to describe the underlying mechanisms such as, endocrine disruption); hence the emphasis (of the DDRP) should be concentrated on establishing tests which can detect biologically relevant effects caused by endocrine disruption. On the other hand, tests which describe mechanisms may be important in establishing the appropriate experimental design of the studies of effects (e.g. timing of dosage, period of observation of effects).

   It is strongly recommended that any testing strategies proposed should include reference to, and recognition of, these additional endocrine-disrupting mechanisms. An explanation of how these issues will be addressed and ultimately incorporated into the OECD process should be included in the DDRP.

There are many mechanisms by which a compound can act to enhance or attenuate hormone action and ultimately affect reproduction as well as numerous other functions. As indicated in the DDRP, *in vivo* tests are considered necessary to establish potential hazards for hormonally-active compounds using *in vitro* tests to prioritise and help determine mechanism of action. At present, several screening procedures have been proposed (Reel *et al*, 1996; Shelby *et al*, 1996; Carney *et al*, 1997). Another example being developed under the direction of Chemical Manufacturers' Association and Chlorine Chemistry Council is to validate a Tier I comprehensive testing scheme which looks at many types of mechanisms using 15 different types of endocrine-active control substances (O'Connor *et al*, 1996; Cook *et al*, 1997). It is strongly encouraged that the results and potential for implementation of this type of approach be one of the considerations before propagation of new/revised Guidelines.

2. Discussions are in progress on the best approach to detect substances having potential for endocrine activity and on how to assess potential hazards both for humans and wildlife. It is recognized that *in vitro* assays provide a valuable contribution to an ultimate understanding of mechanism of action. Further these assays require a limited amount of time and money. However, at present the application of *in vitro* assays needs to be put in context with the following:

- The currently known *in vitro* assays for oestrogenicity are capable of indicating interaction between the receptor and the chemical. Even though they are well established in pharmacological research for detection of hormonally-effective drugs, they serve, even in the pharmaceutical area, only as prioritisation tools. A more firmly based decision is made on the results of whole-animal experiments.

- As stated in the DDRP, such an interaction is only one (of many known) mechanism which can cause changes in the endocrine system. In addition effects on reproductive organs are not necessarily always mediated by endocrine interaction.

- It is questionable if any interaction seen *in vitro* is a surrogate for an *in vivo* hazard. There are examples of substances showing low potency of receptor interaction but lacking adverse effects on endocrine-sensitive endpoints in appropriate *in vivo* assays. In this respect, *in vitro* assays are not necessarily suitable for hazard identification.

- The DDRP focuses on several further limitations of *in vitro* assays (Section 6.3, page 104 of the print version) including the following:

  - the interaction with the receptor construct or response elements may not mimic *in vivo* modes of action;

  - *in vitro* assays may be unable to distinguish agonists from antagonists;

- existing *in vitro* assays lack satisfactory metabolic activity (some yeast-based assays are not metabolically inert impeding interpretation of the results);
- some *in vitro* assays show limited chemical uptake;
- absorption, metabolism, excretion or bioaccumulation may play critical roles in determining activity.

In vivo assays, especially studies with repeated exposure, are not compromised by such limitations. At present, only such *in vivo* studies adequately reflect the subtle interactions and feed-back mechanisms of the endocrine system. Conducting repeat-exposure *in vivo* studies at different dose levels not only affords the opportunity to identify hazards but also to characterise the dose response. These are prerequisites for a meaningful risk assessment. Applying the current risk assessment paradigm on the results of such studies is the best known approach for adequate protection of humans and wildlife (Stevens *et al,* 1997a; Crisp *et al*, 1997).

NOTE: It should be recognised that currently there are questions regarding the availability of the technology for transfection assays in general. Specifically, Ligand Pharmaceuticals holds a patent on the technology which could make general implementation difficult. Prior to the availability of any Guidelines, these issues need to be resolved.

3. There are two large Sections of the DDRP which cover the review of the non-regulatory individual test methods - these are Chapter 5 "Critical Assessment of Non-Regulatory Test Methodologies" and Chapter 8 "Annexes". In Chapter 5, non-regulatory methods are reviewed, followed by a tabulated listing of methods and Sections in which certain methods are recommended. This structure is difficult to read and the basis for selecting and recommending certain methods or assays is not explained in sufficient detail. It is also confusing as some key references, which would influence the critique, are mentioned in the Annexes but not in the critical review. The tables list a number of procedures without any further qualification or comment on their usefulness. It would make it significantly clearer to draw conclusions from the vast amount of information available if these Chapters were combined. For example, for each test system, the details given in Chapter 8 could be followed by the critical review and recommendations (as given in Chapter 5). Further, the Annexes (Sections 8.1.5 and 8.2.6) are not comprehensive.

4. The terms "human relevant" and "wildlife relevant" (page 40, Section 5.1) are confusing and almost suggest that there should be Sections on "human non-relevant" etc. which of course is incorrect.

In addition, there is inconsistency in Section names which are termed "Systems of Relevance to Assessment of Human Toxic Potential" etc. later in the document (page 129, Section 8.1.1).

The authors' rationale for classifying a test method as relevant is not explained, and it would appear that none of the tests is relevant to anything but potential hazard assessment. In addition, we are not really dealing with "toxic potential" when referring to endocrine disruption. Identification and regulation of a genuine chemical hazard should depend on the observation of an adverse toxicological effect and not just on observation of mechanistic changes. For example, small perturbations in a hormone level in a study where no effects are observed in reproduction or chronic toxicity studies must be of questionable significance. Where possible the authors should clearly indicate the criteria upon which they base the statement relevant to human or wildlife.

5. It is suggested that a special Section be devoted to Selection of Dose/Interpretation of Results because of its importance to the process of obtaining valid and meaningful results (Foran *et al*, 1997) .

It is evident that *in vitro* and *in vivo* screening assays are being used without regard to appropriate or logical doses. There are examples in the literature of assays being used to measure the effects of oestradiol and xenobiotics at concentrations ranging over 8 or more orders of magnitude (Soto *et al,* 1995, Jobling and Sumpter, 1993). There is, however no effort to determine if the assay or the endpoint respond appropriately to high concentrations of chemicals or that the defined mechanism is functional at these high concentrations. Further, there is no effort to determine realistic concentrations over which the assay is valuable in identiying the potential hazard of a test material. In addition, Ren *et al* (1996) indicated that mRNA for vitellogenin in fish was induced by dimethyl formamide and ethylene glycol. It has been clearly demonstrated that ethylene glycol does not bind to the oestrogen receptor *in vitro*. The confounding effects produced by these solvents must be considered in selecting doses.

6. For this discussion, it needs to be stated that as more endpoints are added to standard tests, more random/spurious findings will be obtained. Interpretation and conclusions based on these findings must be treated with caution. It is likely that, except in cases of obvious and extreme effect, observation of related effects will be required to justify the conclusion of endocrine disruption. Therefore, these random/spurious findings will result in a large number of equivocal tests. Decisions on how these will be resolved are critical prior to embarking on a panoply of additions to standard test procedures.

7. A key factor in the interpretation of toxicological findings in the area of reproduction/ developmental toxicity is consideration of the role of general toxicity. The potential for this to be the explanation of effects (rather than a subtle or specific effect on the endocrine system) is real and should be brought out more in the DDRP. There is a mention of toxicity on page 99, paragraph 2, but its key role in attributing correct interpretation to findings should be stressed in other Sections that deal with the *in vivo* models.

8. The DDRP has avoided defining which *in vivo* endpoints definitely establish the sex-hormone disruption of a substance (Stevens *et al*, 1997b). These endpoints are essential for validating possible screening approaches. The DDRP presents numerous possible tests. However, it appears in the discussion that the establishment of sensitivity, relevance and relative importance of possible marker endpoints has been added as an afterthought. These features must be addressed with major emphasis, particularly after the unfortunate episode concerning synergy (McLachlan, 1997). The same lack of definition or identification applies to the need for a reference set of chemicals and dose-response profiles.

9. Parts of the environmental and wildlife discussions need to be considered for re-drafting as it is often difficult to follow what is being proposed by the authors. It is recommended that the draft be edited to improve significantly the ease with which it can be read. Also, the Executive Summary does not give sufficient attention to the wildlife issue given the scope of the full document and therefore should be amended to correct this imbalance. The definition given on page 15 should also be moved to the Executive Summary.

   There should be a stronger emphasis in the Executive Summary and Introduction on the ecological context of sex-hormone disrupting chemicals in wildlife. Although it is referenced on page 7 as an important consideration, this needs to be more prominent. It should be pointed out that the environmental hazard assessment for such hormone disrupters should be treated in the same way as any other agent that can cause adverse reproductive effects, regardless of the mechanism involved. As is true of other areas of wildlife (eco) toxicology, this Introduction should act as a signpost in ensuring that mechanistic studies are best targeted towards providing information that helps in ecosystem management, rather than such mechanistic data being seen as the final endpoint in their own right for environmental risk assessment.

   An additional concern is that much of the relevant ecotoxicology literature is not cited in the DDRP. There are many publications on reproductive toxicity in diverse wildlife species that should be considered in such a review. Specifically, a good deal of research work published by the US EPA over many years and also the standard methods for fish and invertebrates published by the US EPA and ASTM has not been included. Moreover scant attention has been given to

many important papers published on the comparative endocrinology of wildlife species used in the OECD testing scheme. For the papers that have been included, many appear as uncritical listings which do not make best use of the information. The authors are advised to refer to the report from the recent SETAC/OECD meeting on *Endocrine Modulators in Wildlife - Assessment and Testing* (EMWAT workshop), as well as the ECETOC Document 33 "Compendium of Test Methods (1996).

The authors of the DDRP should be wary of specifying tests using individual wildlife species that are not part of the OECD scheme, e.g., mosquito fish (Table 4 and 5). Such a recommendation does not reflect the current "state of the science" in either wildlife endocrine research or ecotoxicology. These tables should be supplemented with footnotes to indicate that, in principle, other species can be used for the same purposes.

10. In Chapter 5, non-regulatory methods are reviewed, followed by a tabulated listing of methods and a Section in which certain methods are recommended. This structure is very difficult to read and the basis for selecting and recommending certain methods or assays is not explained in sufficient detail. The tables list a number of procedures without any further qualification or comment on their usefulness. The different Sections of this Chapter should be consolidated.

Note: ECETOC together with EMSG has developed a research programme addressing existing data gaps related to specific testing approaches for wildlife organisms. ECETOC is happy to cooperate with OECD in providing further literature references and to update the Organisation about the on-going research activities being conducted by the chemical industry.

SPECIFIC COMMENTS

In this Section, detailed comments are listed under the chapter headings (italicised) employed in the DDRP.

1. EXECUTIVE SUMMARY

1.1 MODIFICATIONS TO EXISTING TEST METHODS

Page 5, Section 1.1: Although it is agreed that the most promising approach seems to be to make full use of the existing study Guidelines with inclusion of appropriate enhancements as necessary, it must be pointed out that the merit of adding some of the possible enhancements has yet to be established.

As the sensitivities of these endpoints need to be established, it would appear inappropriate to add these endpoints in test systems without first establishing (relative) sensitivity and then proposing possible extension of existing test-Guidelines.

Further, the modification to existing test methods is presented as if applicable for all (or most) reviewed OECD tests. It is very doubtful that it is desirable to include these modifications in all tests.

Since the added value of some of the modifications have not been clearly established perhaps the modification should only be made to a limited number of the most relevant OECD tests and then only on a trial basis. Because a base set of information is required in Europe, the best approach is to have the proposed extra parameters included in only one of the required tests, most logically a 28-day type study such as a 'modified OECD 407' Guideline.

1.2 NON-REGULATORY TEST MODELS PROPOSED FOR FURTHER DEVELOPMENT/ ADOPTION

Page 6, Section 1.2, third bullet: As written, the castrated rat model appears to be focused on detecting androgens. It should clearly be designed to detect androgen-receptor antagonists. If the castrated model is to be implanted with testosterone implants O'Connor *et al* (1997) recommend using silastic tubing rather than pellets.

Page 6, Section 1.2: Strongly agree that at present it is not possible to recommend adoption of any of the *in vitro* assays as a regulatory acceptable model because of the various limitations and difficulties inherent in current designs.

In addition to the list of limitations and difficulties with *in vitro* assays, it should also be mentioned that lack of well-validated assays and proven reproducibility in different labs is also a limitation of many *in vitro* assays (Ashby *et al*,1997a).

The recommendation of the MCF-7 cell line assay for further development is not supported for a number of reasons. In particular there are reported differences in cell proliferation due to cell-strain variability, culture conditions, receptor-level differences, clone heterogeneity, loss of responsiveness to oestradiol and proliferation in response to compounds acting *via* non-oestrogenic mechanisms, e.g. EGF (Herman and Katzenellenbogen, 1984; Karey and Sirbasku, 1988; Godden *et al*, 1992; Welshons *et al*, 1992; Villalobos *et al*, 1995; Zacharewski, 1997).

1.3 REQUIREMENTS FOR BASIC RESEARCH

Page 7, Section 1.3, bullet 3: Strongly agree with the concept "elucidate the dose-response profiles for endocrine disruptive mechanisms and apply to dosage selection during testing". An understanding of appropriate dose selection and interpretation is vital to the overall development of screening assays.

2. INTRODUCTION

2.1 GENERAL BACKGROUND

Page 8, Section 2.1: The report from the European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife has now been issued by the EC (MRC Institute for Environment and Health,1996; Crisp *et al*, 1997).

2.2 SEXUAL DETERMINATION AND REPRODUCTIVE CONTROL SYSTEMS

No specific comments on this Section.

2.3 DEFINITION OF SEX-HORMONE DISRUPTERS

Page 15, Section 2.3: There is strong agreement with the definitions and the emphasis on *in vivo* rather than *in vitro* responses for characterising this class of compounds. In the US, there is a growing move toward using the term "endocrine active compounds", (EACs) among scientists involved in risk assessment since it does not evoke value judgements associated with terms such as endocrine disruption.

The proceedings of European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife, Weybridge, U.K, Dec 1996 (MRC Institute for Environment and Health, 1996) have been published and should be referenced accordingly.

2.4 MECHANISMS OF SEX-HORMONE DISRUPTION

Pages 15-16, Section 2.4: Strongly agree that an endocrine disrupter can only be adequately defined using an *in vivo* test model (using the accepted definition of an endocrine disrupter). Also agree that *in vitro* models may be useful in setting priorities in a testing strategy for determining the potential of a chemical to cause endocrine disruption, however, the proper application of *in vitro* assays needs further validation.

2.5 REVIEW OBJECTIVES

Page 17, Section 2.5: The focus of the DDRP is to evaluate the potential methods that can be used to detect oestrogenic/anti-oestrogenic and androgenic/anti-androgenic compounds. Whilst the review document covers the detection of sex hormones and antagonistic actions very comprehensively, it does not cover actions mediated by other endocrine mechanisms.

3. INVENTORY OF INTERNATIONAL DATA REQUIREMENTS FOR SEX-HORMONE DISRUPTING CHEMICALS

No specific comments on this Section.

4. OVERVIEW OF EXISTING REGULATORY TEST METHODS

4.1 REGULATION BACKGROUND

No specific comments on this Section.

4.2 OECD ACTIVITIES

No specific comments on this Section.

4.3 REVIEW OF OECD GUIDELINES

Page 32, Section 4.3.1: In the discussion on "Mammalian acute toxicity" the statement is made "The inclusion of any additional endpoints should, however, be approached with caution for these study types since interpretation of data may be difficult given the high dose levels employed ...".

Single, high doses used in acute studies will probably be of little value in determining potential endocrine-disrupting effects. Although the focus of the DDRP is only on sex-hormone disruption, the assessment of oestrus cycle and organ weights after a single toxic dosage provides questionable interpretable information, especially when it concerns moribund animals. Clearly stress induced by toxic doses of a chemical can influence the biological response, such as the oestrus cycle. It is important to recognise, as stated in the OECD document (Section 4.3) that chemicals with overt endocrine toxicity, regardless of mechanism, have been and will be identified in standard procedures. The primary issue being addressed in this arena is whether exposure to low doses (e.g. environmental exposure) can result in a significant hazard. Therefore, it is arguable that doses even in typical subchronic, chronic, developmental toxicity and reproduction studies will typically be higher than is useful to answer the ultimate questions. Hence, addition of endpoints as suggested in this Section under "Mammalian subchronic toxicity" must also be put in context of dose - that is, how will some effect on a single endpoint be interpreted if the test is conducted at very high doses? Does this represent an endocrine toxicity? At some dose, it is probable that many materials will affect the endocrine system but in the vast majority of these cases, the effects will be secondary to other toxicity - e.g. CCl_4 produces liver toxicity that increases uterine weight in intact females due to decreased excretion of estradiol via conjugation.

Page 35, Section 4.3.2 (Non-mammalian study designs): The review states that Guidelines 203 and 204 are unlikely to be useful. Guideline 203 is indeed focused on an unsuitable exposure time and measurement endpoints. On the other hand, it may be possible to develop Guideline 204 as a screening tool; for example adult or sub-adult fish may be used. However, the ability to measure the endpoints listed below in small juvenile fish may be explored in order to be able to discount the need to know the sex of the fish. The measurements listed below, require that certain quantities of blood plasma are obtained. In small fish this can pose technical problems, although it is also possible to use whole body homogenates.

The test duration would be expected to be approximately 7 to 21 days, (depending on the time required to detect a measurable response, compared with controls) but protocols should be optimised to as short an exposure period as possible.

The following endpoints should be validated against reproductive and developmental endpoints and the assays fully standardised.

- sex steroid levels
- vitellogenin (VTG) induction
- gonadosomatic index (GSI) (although time period may be too short)

Page 36, Section 4.3.2 (Fish early life stage (OECD 210)): It is stated in the review that it is 'unlikely that additional endpoints could be included without a significant extension to the duration and nature of the test...'. However, this test could actually be of considerable value with some modifications.

Potential endpoints include:

- reproductive success
- resulting sex ratio of individuals treated during sexual differentiation
- sex steroid concentrations
- VTG induction
- time to production of first eggs
- fecundity (number of eggs produced per female)
- assessment of sperm production
- gonad histopathology (reproductive tract and gonadal tissue).

The exposure phase can be followed by a period during which the fish are held in water without further test substance treatment. This procedure reduces the experimental effort.

4.4 POTENTIAL ENHANCEMENTS TO EXISTING OECD GUIDELINES

Page 36, Section 4.4, paragraph 1: There is total agreement about the necessity to conduct cost-benefit analysis on potential enhancements to the OECD Guidelines as well as for any new testing strategy.

Page 37, Section 4.4, paragraph c): The measurement of hormone levels was also proposed as a potential enhancement to existing OECD Guidelines. Again, a cost-benefit review of these measurements is warranted prior to implementation. There is clear evidence indicating that

measurement of serum hormone levels in the cycling female rat is often confounding and of much less value and utility than measurement of changes in vaginal cytology (Eldridge *et al*, 1994; Eldridge *et al*, 1995; Eldridge *et al*, 1996; Stevens, 1994; Wetzel *et al,* 1994).

Page 38, Section 4.4, paragraph f), first sentence: It is difficult to envisage the quantitation of mating behaviour and its ultimate sensitivity.

Page 38, Section 4.4, next to last paragraph: The DDRP states "The scope of the current OECD test protocols for non-mammalian, wildlife species is limited." Despite this fact, there are protocols which have been established and evaluated for determining the effects of chemicals on non-mammalian wildlife species under other regulatory jurisdictions, i.e. the US Environmental Protection Agency. The fish partial life-cycle test (PLC) of the US EPA Guideline could be adapted for the purpose of testing endocrine modulation in sub-chronic exposures. Fish should be exposed over a period of 2 to 4 weeks, during which they are held in mating pairs or groups. Endpoints which could be evaluated include:

- mating behaviour
- fish colouration (secondary sexual characteristics)
- time to first hatch
- fecundity (number of eggs produced per female)
- assessment of sperm production
- sex steroid concentrations
- VTG induction

Conduct of ELS or PLC testing will depend on determination of the most critical life stages and sensitivity of endpoints.

Additional tests includes the fish full life-cycle/multi-generation tests. A testing Guideline for fish life-cycle studies has been standardised by the US EPA (US EPA, 1986). The existing procedures will need adapting to include the analysis of endocrine-specific effects on reproduction and development.

Such technically-challenging and costly testing of chemicals for their endocrine effects in such a complex regime is unlikely to be viable on a routine basis. Shorter-term tests (as discussed above) should be developed and validated to be predictive of/ correlative with full life-cycle effects.

Page 38, Section 4.4: last paragraph on page 38: Strongly agree that a careful analysis of the pros/cons of adding additional endpoints to existing Guidelines versus new stand-alone tests needs to be conducted.

5. CRITICAL ASSESSMENT OF NON-REGULATORY TEST METHODOLOGIES

5.1 INTRODUCTION

Page 40, Section 5.1: The introduction to the Section on assessment of individual methods should be expanded to include definitions of criteria for tests being classified as "relevant" and "non-specific" (see Section 5.2.1.1).

5.2 ASSESSMENT OF INDIVIDUAL METHODS

Page 48, Section 5.2.1.2.7: Too much emphasis has been placed on behavioural modification. It is not clear that behavioural change in rodents is a sensitive endpoint. Indeed, the majority of work would suggest that such measurements would be unlikely to be observed without other toxic effects. The cost-benefit of addition of these measurements to routine testing should be a major point of discussion and the scientific value verified prior to consideration.

Page 50, Section 5.2.1.3.4, fourth line from the bottom of the page: The statement is made in regard to the rat prostate models that "It must, however, be recognised that weight increases in these organs may arise as a result of other mechanisms, e.g., tissue oedema or inflammation." This statement describes the interpretation concerns for virtually all current and modified procedures. If doses of chemicals are used which can result in significant tissue effects, then "apparent" endocrine-related toxicity is likely. This situation can lead to the conclusion that such chemicals are endocrine disrupters at low doses when, indeed, no such mechanism or potential toxicity would be manifest.

Page 51, Section 5.2.1.3.4, paragraph 2: An important point from the paper by Cook *et al* (1993) appears to have been overlooked. In this paper, the sensitivity of measuring the accessory sex gland unit weight (prostate, seminal vesicles with fluid, and coagulating glands) versus individual organ weights was compared. The authors found that the unit weight was more sensitive and had more statistical power than measuring the individual weights. In addition, the levator ani muscle weight was not a sensitive marker.

Page 54, Section 5.2.2: The classification of test methods into "non-specific" or "oestrogen-related" is somewhat confusing based on the headings which the authors have used for the test methods. For example the heading T-47D cell lines appears in both Sections for "non-specific" and "oestrogen-related"

test systems for endocrine-disrupting chemicals. All the cell lines, T-47D, MCF-7 and ZR-75-1, possess oestrogen receptors and can be used in an "oestrogen-related assay". Hence it is not the cell lines themselves which are non- specific, but the assays in which they are used which are either "oestrogen-related" (e.g. proliferation in response to oestrogenic compounds) or "non-oestrogen related" (e.g. proliferation in response to a non-oestrogenic compound). It is suggested that the headings for the test systems should be more explicit in order to avoid this confusion.

Page 57, Section 5.2.2.2.2.1. (MCF-7 cell line systems): The authors are correct in their statement that the MCF-7 cell line has been the subject of considerable scientific research into its ability to detect oestrogenic agonistic/antagonistic activity. However, the specificity of the cell line to discriminate between cellular proliferative responses from oestrogenic mechanisms and those from non-oestrogenic mechanisms should be discussed in this Section. For example, for this assay to be taken as a measure of the direct interaction with the oestrogen receptor and equated with oestrogenicity, we must be certain that the cells do not respond to non-oestrogenic substances and do not proliferate in response to binding to other receptors (e.g. growth factors). In the case of the MCF-7 cells used by Soto *et al* (1995), it has been reported that the cells were unresponsive to non-oestrogenic compounds or growth factors. Other workers, however, have reported that MCF-7 cells do respond, for example, to epidermal growth factor (EGF) by proliferation (Herman and Katzenellenbogen, 1984; Karey & Sirbasku, 1988; Godden *et al,*1992; Welshons *et al,*1992). It is also known that the behaviour and response to 17β-oestradiol of MCF-7 cells grown in different laboratories varies (Osborne *et al,* 1987; Villalobos *et al,* 1995). Furthermore it has been demonstrated that proliferative responses of MCF-7 cells can be manipulated by varying culture conditions (Jones *et al,* 1997; Berthois *et al,* 1986; Katzenellenbogen *et al,* 1987; Ruedl *et al,* 1990). Differences in the strains of MCF-7 cells and the protocols used in different laboratories may lead to significant inter-laboratory variability for this assay.

Page 64, Section 5.2.2.2.5. (Oestrogen related - yeast screens): This Section on oestrogen-related yeast screens should be expanded considerably to include several key references which have been published in the last six to nine months (since the document was written) as the use of yeast screens for oestrogenicity is expanding extremely rapidly. Routledge & Sumpter (1996) have used the *S. cerevisiae* yeast strain expressing the human oestrogen receptor, linked to the β-galactosidase reporter gene to assess the oestrogenic potential of a wide range of chemicals, particularly focusing on alkylphenols. Gaido *et al* (1997) have examined a range of yeast transformants expressing the human oestrogen, androgen or progesterone receptor (again linked to the β-galactosidase reporter gene) for their ability to detect different classes of endocrine-disrupting compounds. In addition, Ramamoorthy *et al* (1997a) demonstrated a lack of synergy with selected compounds using a yeast strain transformed with the human oestrogen receptor and linked to a β-galactosidase reporter gene. Reference should also be made to the excellent review of *in vitro* bioassays by Zacharewski (1997) which discusses the types of yeast screens which have been used to date and the advantages/disadvantages of yeast screens

compared to other *in vitro* methods. Finally for this Section, reference should also be made to recent publications suggesting the use of oestrogen receptor-dependent transcriptional-expression assays in *in vivo/in vitro* testing strategies for oestrogenic compounds (Reel *et al,* 1996; Klotz *et al,* 1996; Ashby *et al,* 1997a).

Page 67, Section 5.2.2.3.2 (Androgen related - yeast screens): The authors should include the recent paper by Gaido *et al* (1997) which examined the ability of a yeast strain transformant with human androgen-receptor expression to distinguish between endocrine disrupters acting via different mechanisms.

In general, it would appear (from the literature and also from discussions among scientists working in this area) that yeast-based assays are becoming more widely accepted and favoured as potential *in vitro* screens. They are becoming more widely recommended than cell-line based assays because of specificity, ease and speed of use, reliability and reproducibility between different laboratories.

Page 67, Section 5.2.2.4.1 (Receptor-binding studies): Strongly agree with the authors comments made in this Section regarding the utilisation of receptor-binding assays for detection of endocrine-disrupting potential *in vitro.*

Page 70, Section 5.2.3 (Wildlife Relevant - *In vivo*): The review concludes that none of the non-regulatory models could be considered as suitable for direct regulatory adoption, but that current OECD fish-study Guidelines could be reviewed for possible incorporation of additional endpoints. Comments on issues relevant to this conclusion (the critical life stages, range-finding tests, test species and testing strategy) are given below:

Critical life stages

The review states that the timing of exposure is critical in determining toxicity (ED effects). Research is needed in order to define this critical period; for example, tests utilising mature and sexually undifferentiated stages of development (e.g. Early Life Stage and Partial Life Cycle) should be compared.

Range-finding tests

The review states most non-regulatory studies employed preliminary (range-finding) studies to define treatment levels. This is recommended as the best design for such tests although treatment ranges may also be defined on the basis of known/expected exposure levels in the environment.

Test species

The review indicates that zebra fish have been used in the majority of reviewed non-regulatory studies. However, there is no comparison of the suitability or sensitivity of the potential range of fish species that could be used. Determination of choice of species may be assessed when the relative sensitivities of the test species to endocrine-disrupting chemicals are established. Their reproductive physiology and endocrinology should be understood, and they must be suitable for laboratory handling. In addition, baselines for endpoints should be established and variability assessed. For example, freshwater species recommended by participants in the EMWAT workshop for use in screening and/or higher tier testing included:

- Rainbow trout (*Oncorhynchus mykiss*)
- Goldfish (*Carassius auratus*)
- Carp (*Cyprinus carpio*)
- Channel catfish (*Ictalurus punctatus*)
- Japanese medaka (*Oryzias latipes*)
- Zebra fish (*Danio rerio*)
- Fathead minnow (*Pimephales promelas*)

However, it was recommended that the study of reproductive effects should be performed with small, warm-water species with shorter life-cycles like the medaka, zebra fish and fathead minnow in order to minimise duration of the test.

Testing strategy

A tiered-testing strategy would be most useful for determining effects of endocrine-disrupting chemicals. The first tier of screening and prioritisation tests could be rapid *in vivo* assays for endocrine endpoints (e.g. hormone levels/VTG - research is needed to validate such measures against reproductive/developmental endpoints). The second tier of definitive tests could be sub-chronic or short-term assays including developmental and reproductive endpoints (research is needed to determine critical periods and durations of exposure, and most useful endpoints). The third tier of definitive tests would be chronic or long-term assays, but with similar endpoints to the sub-chronic tests (such tests are technically challenging and costly and are therefore not viable on a routine basis).

Page 71, Section 5.2.3.2 (Sex-reversal assays):

a) *Chicken.* Elbrecht and Smith (1992) examined the effects of aromatase inhibition on gonadal differentiation of early chicken embryos. Phenotypic sex was determined by vent sexing (dimorphism in cloaca) whilst genotypic sex was determined by the use of a sex-linked feather colour mutation. Treatment with aromatase inhibitors (AI) produced hatchlings which were all of the male phenotype. Inhibition of aromatase in embryos results in a testosterone increase and a decrease in oestradiol leading to testicular development. Genotypic females treated with the AI had low serum oestradiol concentrations and high testosterone concentrations similar to those of males. Gonads of females exposed to AI resembled testes. Administration of testosterone or tamoxifen did not reverse the female sexual phenotype. Co-administration of oestradiol with AI abolished the effects of AI. Administration of oestradiol on its own feminises males. The results appear specific for AI and not other steroids or anti-oestrogens. The authors did not expose embryos to a non-aromatisable androgen.

   The authors specifically looked at AI but this assay appears to have the possibility of detecting oestrogens and possibly androgens as well although the latter would require further investigation.

b) *Alligator.* Lance and Bogart (1992) examined sex differentiation in alligator embryos. Embryos were exposed (at either female- or male-producing temperatures; up to 75 days) to anti-oestrogens, anti-androgen, oestradiol, dihydrotestosterone (DHT) and AI. Oestradiol at male-producing temperature produced 100% females. Anti-androgen, DHT and anti-oestrogen did not have any effects at either temperature. The aromatase inhibitors disrupted ovarian development but gonads were not masculinised. Inhibition of oestrogen synthesis or exposure to androgen is insufficient to masculinise an ovary and testicular differentiation is dependent on other factors.

   Oestrogens could be identified in this assay but the complete sex determination process is not fully understood. Further research is required before it can be determined if other endocrine disrupters could be detected.

c) *Tadpole.* Yu *et al* (1993) exposed female tadpoles (*Rana catesbeiana*; 9 month old; 3-month exposure) to an AI. About 80% of ovaries were transformed to testes; oestradiol secretion in AI-treated ovaries was diminished whilst testosterone levels increased compared to control tadpoles.

   This assay can detect AIs but it is unclear how relevant it could be for the detection of endocrine disrupters such as environmental oestrogens.

d) *Turtle.* Wibbels and Crews (1995) investigated the effects of various steroids on sex determination using incubation temperature that produced mixed sex ratios in the turtle, *Trachemys scripta*. Oestradiol and the aromatizable androgen (testosterone) induced female sex determination whilst the non-aromatizable androgen (DHT) induced male sex determination. Much higher concentrations of DHT were required to induce male sex determination compared to the concentrations of oestradiol necessary to cause feminisation.

Richard-Mercier *et al* (1995) observed sex-reversal of gonads by AI in the turtle, *Emys orbicularis*. Exposure at a 100% female-producing temperature induced sex-reversal of gonads. Mullerian ducts regressed and ovarian aromatase was reduced and in 50% of the embryos was similar or somewhat higher than that seen in control testes of embryos incubated at male-producing temperatures.

Bergeron *et al* (1994) exposed the turtle, *T. scripta,* to polychlorinated biphenyls (PCB) and some of their metabolites at all-male-producing temperatures. Two hydroxylated metabolites and oestradiol significantly feminised turtle hatchlings as compared to controls. The authors claim synergy was observed (about 10-fold) when the PCB metabolites were administered together.

Turtle assays appear able to detect the effects of endocrine disruptors such as oestrogens, androgens and AI. Further work appears necessary for the effects of anti-oestrogens as Wibbels & Crews (1995) observed feminisation in *T. scripta* after exposure to tamoxifen whilst in *E. orbicularis* this substance was found to cause masculinisation (Richard-Mercier *et al*, 1995 and references therein).

Page 74, Section 5.2.3.5 (*Daphnia):* Recent investigations have examined the physiological and biochemical changes in *Daphnia magna* exposed to oestrogens such as diethylstilbestrol (DES) and endosulfan (Zou & Fingerman, 1997; Baldwin *et al*, 1995) and physiological changes in *Daphnia galeata mendotae* exposed to nonylphenol (Shurin & Dodson, 1997). Chronic exposure to DES (0.5 mg/l) reduced moulting frequency (first-generation juveniles only) and decreased fecundity of second generation daphnids (Baldwin *et al*, 1995). These authors also found that steroid metabolism in DES-exposed animals was altered. Zou and Fingerman (1997) observed no difference in the proportion of male broods in *Daphnia* exposed to DES or endosulfan and control animals (both control/exposed animals kept under conditions of food limitation, crowding and short daily photoperiod). These authors did observe a reduced moulting frequency in exposed Daphnids compared to the controls. Shurin and Dodson (1997) found numbers of male offspring to be unaffected by nonylphenol exposed *Daphnia galeata mendotae* kept under crowded conditions whilst production of females showed an inverse dose-response.

Under favourable conditions *Daphnia* populations mostly consist of parthenogenic females but alterations in food availability, photoperiod, temperature and crowding will cause parthenogenic females to produces male and females which can undergo sexual reproduction. The factors controlling parthenogenesis and sexual reproduction are poorly understood. Sexual differentiation, vitellogenesis as well as moulting in Crustacea are controlled by a number of different hormones including peptides, steroids and terpenoids (Quackenbush, 1989). Fecundity and moulting in *Daphnia* appear to be affected by oestrogenic chemicals although concentrations that cause significant effects are very high in comparison with those inducing vitellogenesis in fish. Zou and Fingerman (1997) speculated that moulting inhibition could be due to oestrogens like DES or endocrine-disrupting chemicals such as endosulfan and endogenous ecdysteroids but also considered it possible that this inhibitory effect was a general response to stressors.

There is a lack of knowledge of the endocrine system of crustacean species such as *Daphnia magna* and, more specifically, a poor understanding of how fecundity and moulting is influenced by receptor-mediated interactions. Further work should be considered in both parthenogenic and sexually-reproducing crustacea. Therefore, at present, *Daphnia* cannot be recommended for laboratory testing of endocrine-disrupting chemicals as a representative example of invertebrates.

5.3 OVERVIEW OF NON-REGULATORY TEST METHODS

Pages 81, 83; Table 4 (Summary of duration and complexity of non-regulatory test methods): This table suggests that the test duration for the MCF-7 cell proliferation assay and transfected yeast strains is the same, but this is generally not the case - results from yeast strains can be obtained far quicker than from MCF-7 cell proliferation assays, which require longer incubation times.

Page 82, Section 5.3, Table 4, Leydig cells: Isolated Leydig cells from rats can also be used to detect steroid biosynthesis inhibitors. Detailed methodology has been published (Chapin and Heindel,1993).

Page 88, Section 5.4.3 (Wildlife *in vivo*): The proposal for testing many different species for sex-hormone disrupting activity appears somewhat contradictory with the notion of phylogenetic conservation of receptor systems.

Page 89, Section 5.4.3 near the end of the Chapter (and page 9, Section 1.2 bullet 4): As an interim measure, investigate the use of models using changes in secondary sexual morphology of fish in response to sex-hormone disrupter exposure: Such a morphological change is considered to be a toxic endpoint (adverse health effect) for endocrine disruption, and can therefore be used to validate the significance of a screen for identification of potential endocrine disrupters. It is a bit strange to call it a screening for possible endocrine disruption by itself.

Page 90, Section 5.3, Table 5: An additional use of the Sf9 (Baculovirus) system is the ability to produce large quantities of proteins. Specifically, the steroid receptors of interest can be produced in mg quantities which has two advantages: (1) eliminate sacrificing animals to obtain tissues for preparing cytosolic receptor preparations (i.e. prostate, uterus), and (2) the high concentrations of receptor improve the sensitivity of receptor-competition studies.

5.4 RECOMMENDED NON-REGULATORY MODELS

Page 94, Section 5.4.1, line 3 from top: Regarding the point that this screen is too expensive, the purpose of this paper and the ongoing work is to evaluate which of the selected endpoints are most predictive and dose-related. After the validation phase it should be possible to reduce the number of endpoints. The balance to be paid is having enough redundancy of endpoints to ensure a low false-positive and false-negative response. To be cost-effective, it is critical that tiered-testing strategies be integrated so that as many types of EACs as possible are identified.

Page 94, Section 5.4.1, paragraph 2, line 4: An important point was not captured in the summary of the paper by Cook *et al* (1993). In this paper, the sensitivity of measuring the accessory sex gland unit weight (prostate, seminal vesicles with fluid, and coagulating glands) versus individual organ weights was compared. The authors found that the unit weight was more sensitive and had more statistical power than measuring the individual weights. In addition, the levator ani muscle weight was not a sensitive marker.

Page 96, Section 5.4.2.1: Four important papers have been published since that of Arnold *et al* (1996). These are the papers of Gaido *et al* (1997) who evaluated the ability of three transformed yeast strains (which contain either androgen, oestrogen, or progesterone receptors) for their ability to detect several different types of EACs and both Ramamoorthy *et al* (1997a,b) and Ashby *et al* (1997b) who demonstrated the absence of synergy by several different methodologies. In addition, this inability to observe synergism for these compounds is not due to differences in receptor number. Lastly it should be noted that McLachlan (1997) has retracted the original paper in *Science* that indicated significant synergism in yeast assays.

Page 96, Section 5.4.2.2: A more-comprehensive evaluation of yeast transformed with the androgen receptor has been conducted since the DDRP was written. Specifically, Gaido *et al* (1997) evaluated several compounds for their ability to be detected using a transformed yeast strain that contains the androgen receptor.

Page 97, Section 5.4.2.4: Strongly agree with the stringent criteria required for *in vitro* assay development and the need for a set of defined reference chemicals for inter-laboratory validation of any methods recommended as part of a testing strategy.

It is noteworthy that issues exist surrounding the general availability of transformation procedures owing to patent rights. These issues need to be resolved prior to proposing transformation assays as a general Guideline procedure.

6. PROPOSED ENHANCEMENTS TO CHEMICAL TESTING PROCEDURES

6.1 CHEMICAL TESTING REQUIREMENTS AND STRATEGIES

Page 99, Section 6.1, paragraph 3, last sentence: This is an excellent point. Clearly, we need to improve our knowledge of comparative differences, not only from mammalian to wildlife, but also to understand the comparative endocrinology differences between rodent models and humans. This knowledge will greatly improve the accuracy of predictions in assessing the potential risks to humans and wildlife of endocrine disrupting chemicals.

Page 100, Section 6.1 (last paragraph before diagram 1 on tiered testing): The approach delineated for risk assessment of a potential endocrine disrupter is supported.

6.2 EXTENSIONS TO EXISITING REGULATORY STUDY DESIGNS

Applying the current risk assessment paradigm to the results of such studies is the best known approach to protect human and wildlife reproductive health from endocrine disrupting chemicals (Crisp *et al*, 1997; Stevens *et al,* 1997a). However, based mainly on theoretical grounds the suitability of the current Guideline studies to detect endocrine-disrupting potential has been questioned. It is worth noting that selection of certain endpoints as an indicator of effect on others is a well-established concept to make best use of limited resources. For example, oestrogens may affect sperm production but most frequently also alter testis morphology. Thus, the use of routinely-employed histopathology of the testis can serve to indicate adverse effects on sperm production. Such an approach is considered to be adequate if the parameter investigated is the more-sensitive endpoint. In this respect, the experience gained by the ICH (International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, Smith, 1996) is of special interest. The detection of substance-related effects on male reproduction was an issue of the ICH. To put the recommendation on a firm basis, the issue was investigated both by experiments (Takayama *et al*, 1995a) and by a literature survey (Ulbrich and Palmer, 1995).

Experiments were carried out on 16 substances with different pharmaceutical activity (2 hormones, 4 anticancer drugs, 2 psychotropic drugs, 2 nootropic drugs, 2 vitamins, 1 antihypertensive agent, 1 diuretic drug, 2 general chemicals) by Japanese pharmaceutical companies and by the National Institute of Health Sciences (Japan). Comparing two treatment periods (4 and 9 weeks) and various parameters (organ weights, spermatogenic endpoints, mating behavior, caesarean-section findings and histopathology), it was concluded that treatment for 4 weeks before mating is sufficient to detect adverse effects on male fertility with histopathology of the testis being the most-sensitive index for the substances investigated. From the data presented, it appears that genital organ-weight determination is at least as sensitive as the spermatogenic endpoints. Tests of reproductive activity were generally found to be insensitive (Takayama *et al*, 1995b)

A literature survey (Ulbrich and Palmer, 1995) covering 117 substances or substance classes also supports the assumption that histopathology and organ-weight analysis provide the best general-purpose means of detecting substances with the potential to affect male fertility, and examinations after a treatment period of 4 weeks appear to be as effective as examinations conducted at later times.

In summary, it appears scientifically sound to assume that substances affecting male fertility by various mechanism (and not only oestrogenicity) will be sufficiently detected by parameters (histopathology and organ weight) and after exposure regimes (treatment period of 4 weeks or longer) already employed in current Guideline studies. Such studies are frequently conducted at early stages of chemical development or are required in the first steps of chemical notification. The multiple dosing *in vivo* studies are capable of detecting mechanisms of endocrine disruption for which, at present, *in vitro* tests are not available. Thus it seems prudent (both scientifically and on a cost-effective basis) to evaluate the suitability of sub-chronic tests in rodents as screening tools for detection of endocrine disruption.

Page 102, Section 6.2, first sentence, last paragraph: It is agreed that there is an outstanding need to establish the sensitivity of the various endpoints and rank their importance as markers of toxic hazards. The effort to validate the reliability and sensitivity of various endpoints and to establish and standardise these measures of endocrine disruption has been (Eldridge *et al*, 1995; Eldridge *et al*, 1996; O'Connor *et al*, 1996; Reel *et al*, 1996; Shelby *et al*, 1996, MRC Institute for Environment and Health, 1996) and continues to be (Cook *et al*, 1997; Stevens *et al*, 1997b, the DDRP, 1997) the focus of such work. It is important to ensure that this process of development continues and is not prematurely abbreviated in favour of a hurried solution.

6.3 NON-REGULATORY TEST METHODOLOGIES SUITABLE FOR FURTHER DEVELOPMENT

Page 103, Section 6.3, paragraph 3: As written, the castrated-rat model appears to be focused on detecting androgens. It should clearly be designed to detect androgen-receptor antagonists. If the castrated model is to be implanted with testosterone implants, O'Connor *et al* (1997), recommend using silastic tubing rather than pellets .

An important point appears to have been overlooked in the summary of the paper by Cook *et al* (1993). In this paper, the sensitivity of measuring the accessory sex gland unit weight (prostate, seminal vesicles with fluid, and coagulating glands) versus individual organ weights was compared. The authors found that the unit weight was more sensitive and had more statistical power than measuring the individual weights. In addition, the *levator ani* muscle weight was not a sensitive marker.

6.4 OUTSTANDING RESEARCH REQUIREMENTS

Page 107, Section 6.4, paragraph 1: It is agreed that there is an urgent need to define a set of reference chemicals to use in validation of proposed test methods by EDSTAC and OECD.

7. REFERENCES

Additional references supporting the comments made in this Document are included in the Bibliography page 28.

8. ANNEXES

8.1 DETAILED REVIEW OF NON-REGULATORY TEST METHODS IDENTIFIED AS SUITABLE FOR FURTHER DEVELOPMENT

No specific comments on this Section.

8.2 DETAILED REVIEW OF OTHER NON-REGULATORY TEST METHODS

No specific comments on this Section.

BIBLIOGRAPHY

References used in ECETOC Comments

Ashby, J., Houthoff, E., Kennedy, S.J., Stevens, J., Bars, R., Van Miller, J., Carpanini, F.M. and Randall, G. 1997a. The challenge posed by endocrine disrupting chemicals. *Environ. Health Perspect.* 105, 1-11.

Ashby, J., Harris, C.A., I, P.A., Odum, J., Routledge, E.J. and Sumpter, J.P 1997b. Failure to confirm oestrogenic synergism between dieldrin and endosulfan. *Nature.* 385, 494.

Baldwin, W.S., Milam, D.L. and Le Blanc, G.A. 1995. Physiological and biochemical perturbations in *Daphnia magna* following exposure to the model environmental oestrogen diethylstilbestrol. *Environ. Toxicol. and Chem.* 14, 945-952.

Bergeron, J.M., Crews, D. and McLachlan, J.A. 1994. PCBs as Environmental Oestrogens: Turtle sex determination as a biomarker of environmental contamination. *Environ. Health Perspect.* 102(9), 780-781.

Berthois, Y., Katzenellenbogen, J.A. and Katzenellenbogen, B.S. 1986. Phenol red in tissue culture medium is a weak oestrogen: implications concerning the study of oestrogen-responsive cells in culture. *Proc. Natl. Acad. Sci. USA* 83, 2496-2500.

Carney, E.W., Hoberman, A.M, Farmer, D.R., Kapp, Jr., R.W., Nikiforov, A.I., Bernstein, M., Hurtl., M.E., Breslin, W.J., Cagen, S.Z. and Daston, G.P. 1997. Oestrogen modulation: Tiered testing for hazard evaluation. *Repro. Toxicol.* (In press)

Chapin, R.E. and Heindel, J.J., 1993. Male Reproductive Toxicology. Volume 3. Part A. In *Methods in Toxicology.* C.A. Tyson and H.P. Witschi, eds., Academic Press, Inc., San Diego.

Cook , J.C., Mullin, S.L., Frame, R.S. and Biegel, B.L. 1993. Investigation of mechanisms for Leydig cell tumorigenesis by Linuron in rats. *Toxicol. Appl. Pharmacol.*, 15, 297-312.

Cook, J.C., Kaplan, A.M., Davis, L.G. and O'Connor, J.C. 1997. Development of a Tier I screening battery for detecting endocrine active compounds (EACs). *Regul. Toxicol. Pharmacol.* (in press).

Crisp, T.M., Clegg, E.D., Cooper, R.L., Anderson, D.G., Baetcke, K.P., Hoffmann, J.L., Morrow, M.S., Rodier, D.J., Schaeffer, J.E., Touart, L.W., Zeeman, M.G., Patel, Y.M. and Wood, W.P. 1997. *Special report on environmental endocrine disruption: An effects assessment and analysis.* EPA Office of Research and Development Report for the Risk Assessment Forum. Washington, 1-111.

ECETOC 1996. Environmental Oestrogens : Compendium of Test Methods. ECETOC Document No. 33. ECETOC, Brussels, 27 p.

Elbrecht, A. and Smith, R.G. 1992. Aromatase enzyme activity and sex determination in chickens. *Science*, 225, 467-470.

Eldridge, J. C., Fleenor-Heyser, D. G., Estrom, P. C., Wetzel, L. T., Breckenridge, C. B., Gillis, J.H., Luempert III, L. G. and Stevens, J. T. 1994. Short-term effects of chlorotriazines on estrus in female Sprague-Dawley and Fisher 344 Rats. *J. Toxicol. Environ. Health.* 43(2), 169-182.

Eldridge, J.C., Wetzel, L.T., Tisdell, M.O. and Stevens, J.T. 1995. Pattern of Reproductive Aging in Female Rats can affect Mammary Tumor Incidence. Ed. J. J. Li, In *Hormonal Carcinogenesis. Volume 2.* Springer-Verlag, New York, 23-27.

Eldridge, J.C., Stevens, J.T., Wetzel, L.T., Tisdel, M.O., Breckenridge, C.B., McConnell, R.F. and Simpkins, J.W. 1996. Atrazine: Mechanisms of Hormonal Imbalance in Female SD rats. *Fund. Appl. Toxicol.* 24(12), 2-5.

Foran, J. and the ILSI Risk Science Working Group on Dose Selection (*Ashby, J., and Stevens, J., members*). 1997. Principles for the selection of doses in chronic rodent bioassays. *Environ. Health Perspect.* 105(1), 18-20.

Gaido, K.W., Leonard, L.S., Lovell, S., Gould, J.C., Babai, D., Portier, C.J. and McDonnell, D.P. 1997. Evaluation of chemicals with endocrine modulating activity in a yeast-based steroid hormone receptor gene transcription assay. *Toxicol. Appl. Pharmacol.*143, 205-212.

Godden, J., Leake, R. and Kerr, D.J. 1992. The response of breast cancer cells to steroid and peptide growth factors. *Anticancer Res.* 12, 1683-1688.

Herman, M.E. and Katzenellenbogen, J.S. 1984. Alternations in transforming growth factors -α and -β production and cell responsiveness during the progression of MCF-7 human breast cancer cells to oestrogen-autonomous growth. *Cancer Res.* 54, 5867-5874.

Jobling, S. and Sumpter, J.P. 1993. Detergent components in sewage effluent are weakly oestrogenic to fish: An *in vitro* study using rainbow trout (*Oncorhynchus mykiss*) hepatocytes. *Aquat. Toxicol.* 27, 361-372.

Jones, P.A., Baker, V.A., Irwin, A. and Earl, L.K. 1997. Modulation of MCF-7 cell proliferative responses by manipulation of assay conditions. *Toxicology in vitro* (in press).

Karey K.P. and Sirbasku D.A. 1988. Differential responsiveness of human breast cancer cell lines, MCF-7 and T47D to growth factors and 17β-estradiol. *Cancer Res.* 48, 4083-4092.

Katzenellenbogen, B.S., Kendra, K.L., Norman, M.J. and Berthois, Y. 1987. Proliferation, hormonal responsiveness and oestrogen receptor content of MCF-7 human breast cancer cells grown in the short-term and long-term absence of oestrogens. *Cancer Res.* 47, 4355-4360.

Klotz, D.M., Beckman, B.S., Hill, S.M., McLachlan, J.A., Walters, M.R. and Arnold, S.F. 1996. Identification of environmental chemicals with oestrogenic activity using a combination of *in vitro* assays. *Environ. Health Perspect.* 104, 1084-1089.

Lance, V.A. and Bogart, M.H. 1992. Disruption of ovarian development in alligator embryos treated with an aromatase inhibitor. *Gen. Comp. Endocrinol.* 86, 59-71.

McLachlan, J.A. 1997. Synergistic effect of environmental oestrogens: Report withdrawn. *Science.* 277(5345), 462-463.

MRC Institute for Environment and Health. 1996. *European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife. Report of Proceedings.* Weybridge, U.K. EUR 17549, 1-125.

O'Connor, J.C., Cook, J.C, Craven, S.C., Van Pelt, C.S. and Obourn, J.D. 1996. An *in vivo* battery for identifying endocrine modulators that are oestrogens or dopamine regulators. *Fund. Appl. Toxicol.* 33, 182-193.

O'Connor, J.C., Cook, J.C., Van Pelt, C.S., Arnold, S.F. and Obourn, J.D. 1997. Weak oestrogenic activity from continuous-release pellets. *Reproductive Toxicol.* 11, 101-106

Osborne, C.K., Hobbs, K. and Trent, J.H. 1987. Biological differences among MCF-7 human breast cancer cell lines from different laboratories. *Breast Cancer Res. Treat.* 9, 111-121.

Quackenbush, L.S.. 1989. Vitellogenesis in the shrimp *Panaeus vannamei in vitro. Comp. Biochem. Physiol. B. Comp. Biochem.* 94(2), 253-262.

Ramamoorthy, K., Wang, F., Chen, I-C., Safe, S., Norris, J.D., McDonnell, D.P., Gaido, K.W., Bocchinfuso, W.P. and Korach, K.S. 1997a. Potency of combined oestrogenic pesticides. *Science* 275, 405.

Ramamoorthy, K., Wang, F., Chen, I-C., Norris, J.D., McDonnell, D.P., Gaido, K.W., Bocchinfuso, W.P., Korach, K.S. and Safe, S. 1997b. Oestrogenic activity of dieldrin / toxaphene mixture in the mouse uterus, MCF-7 cells, and yeast-based receptor assays: No apparent synergism. *Endocrinology* 138, *1520-1527.*

Reel, J.R., Lamb, J.C. IV and Neal, B.H. 1996. Mammalian oestrogen biological assays: current state of the art. *Fund. Appl. Toxicol.* 34, 288-305.

Ren, L., Lattier, D. and Lech, J.J. 1996. Oestrogenic activity in rainbow trout determined with a new cDNA probe for vitelligen pSG5Vgl.1. *Bull. Environ. Contam. Toxicol.* 56, 287-294.

Richard-Mercier, N., Dorizzi, M., Desvages, G. Girondot, M. and Pieau, C. 1995. Endocrine sex reversal of gonads by the aromatase inhibitor Letrozole (CGS 20267) in *Emys orbicularis*, a turtle with temperature-dependent sex determination. *Gen. Comp. Endocrinol.* 100, 314-326.

Routedge, J.E. and Sumpter, P.J. 1996. Oestrogenic activity of surfactants and some of their degradation products assessed using a recombinant yeast screen. *Environ. Toxicol. Chem.* 15, 241-248.

Ruedl, C., Cappelletti, V., Coradini, D., Granata, G. and Di Fronzo, G. 1990. Influence of culture conditions on the oestrogenic cell growth stimulation of human breast cancer cells. *J. Steroid Biochem. Molec. Biol* . 37, 195-200.

Shelby, M.D., Newbold, R.R., Tully, T.B., Chae, K. and Davis, V.L. 1996. Assessing environmental chemicals for oestrogenicity using a combination of *in vitro* and *in vivo* assays. *Environ. Health. Perspect.* 104, 1296-1300.

Shurin, J.B. and Dodson, S.I. 1997. Sublethal toxic effects of cyanobacteria and nonylphenol on environmental sex determination and development in *Daphnia*. *Environ. Toxicol. Chem.* 16, 1269-1276.

Smith, P. F. 1996. Third international conference on harmonization of technical requirements for the registration of pharmaceuticals for human use - A toxicologist's perspective. *Toxicol. Pathol.* 24, 519-528.

Soto, A.M., Sonnenschein, C., Chung, K.L., Fernandez, M.F., Olea, N. and Serrano, F.O. 1995. The E-screen assay as a tool to identify oestrogens: An update on oestrogenic environmental pollutants. *Environ. Health. Perspect.* 103 (Suppl. 7), 113-122.

Stevens, J.T. 1994. Relevance of Hormonal Imbalance to Dose Selection. *The Toxicology Forum.* 1994 Annual Summer Meeting, 259-278.

Stevens, J.T., Tobia, A., Lamb IV, J.C., Tellone, C. and O'Neal, F. 1997a. FIFRA Subdivision F testing Guidelines: Are these tests adequate to detect potential hormonal activity for crop protection chemicals? *J. Toxicol. Environ. Health.* 50, 415-425.

Stevens, J.T., Gfeller, W., Machemer, L. and Leist, K.H. 1997b. Adequacy of required hazard testing for the detection of potential hormonal activity of crop protection chemicals. *J. Toxicol. Environ. Health* (in press).

Takayama, S., Akaike, M., Kawashima, K., Takahashi, M. and Kurokawa, Y.A . 1995a. Collaborative study in Japan on optimal treatment period and parameters for detection of male fertility disorders induced by drugs in rats. *J. Am. Coll. Toxicol.* 14 (4), 266-292.

Takayama, S., Akaike, M., Kawashima, K., Takahashi, M. and Kurokawa, Y. 1995b. Studies on the optimal treatment period and parameters for detection on male fertility disorder in rats - Introductory summary. *J. Toxicol. Sc.* 20, 173-182.

Ulbrich, B. and Palmer, A.K. 1995. Detection of effects on male reproduction - A Literature Survey. *J. Am. Coll. Toxicol.* 14 (4), 293-327.

US EPA.1986. Hazard Evaluation Division. Standard Evaluation Procedure - Fish Life-cycle Toxicity Test. EPA 540/9, 86-137.

Villalobos, M., Olea, N., Brotons, J.A., Olea-Serrano, M.F., Ruiz de Almodovar, J.M. and Pedraza, V. 1995. The E-screen assay: A comparison of different MCF-7 cell stocks. *Environ. Health. Perspect.* 103, 844-850.

Welshons, W.V., Grady, L.H., Engler, K.S. and Judy, B.M. 1992. Control of proliferation of MCF-7 breast cancer cells in a commercial preparation of charcoal-stripped adult bovine serum. *Breast Cancer Res. Treat.* 23, 97-104.

Wetzel, L.T., Luempert III, L.C., Breckenridge, C.B., Tisdel, M.O., Stevens, J.T., Thakur, A.K., Extrom, P.J. and Eldridge, J.C. 1994. Chronic effects of atrazine on estrus and mammary tumor formation in female Sprague-Dawley and Fischer 344. *J. Toxicol. Environ. Health.* 43(2), 182-196.

Wibbels, T. and Crews, D. 1995. Steroid-induced sex determination at incubation temperatures producing mixed sex ratios in a turtle with TSD. *Gen. Comp. Endocrinol.* 100, 53-60.

Yu, N.W., Hsu, C.Y., Chang, L.T., and Liu, H.W. 1993. Gonadal differentiation and secretions of estradiol and testosterone of the ovaries of *Rana catesbeiana* tadpoles treated with 4 - hydroxyandrostenedione. *J. Exp. Zool.* 265, 252-257.

Zacharewski, T. 1997. *In vitro* bioassays for assessing oestrogenic chemicals. *Environ. Sci. Technol.* 31, 106-116.

Zou, E. and Fingerman, M. 1997. Synthetic oestrogenic agents do not interfere with sex differentiation but do inhibit moulting of the cladoceran *Daphnia magna*. *Bull. Environ. Contam. Toxicol.* 58, 596-602.

AUTHORS OF THE REVIEW DOCUMENT

J. ASHBY	ZENECA UK-Macclesfield
V. BAKER	UNILEVER UK-Sharnbrook
R. BARS	RHÔNE-POULENC AGRO F-Sophia Antipolis
J. COOK	DUPONT USA-Newark
E. HOUTHOFF	AKZO NOBLE CHEMICALS NL-Amersfoort
F.W. JEKAT	BAYER D-Wuppertal
R. LÄNGE	SCHERING D-Berlin
J. STEVENS	NOVARTIS USA-Greensboro
J. VAN MILLER	UNION CARBIDE USA-Danbury

MEMBERS OF THE ECETOC ENVIRONMENTAL OESTROGENS TASK FORCE

G. RANDALL (Chairman)	ZENECA UK-Brixham
J. ASHBY	ZENECA UK-Macclesfield
V. BAKER	UNILEVER UK-Sharnbrook
R. BARS	RHÔNE-POULENC AGRO F-Sophia Antipolis
P. CAMPBELL	PROCTER & GAMBLE B-Brussels
J. COOK	DUPONT USA-Newark
C. DE ROOIJ	SOLVAY B-Brussels
S. EINARSON	NORSK HYDRO N-Porsgrunn
P. FOSTER	CIIT USA-Research Triangle Park
C. HEGELE-HARTUNG	SCHERING D-Berlin
E. HOUTHOFF	AKZO NOBLE CHEMICALS NL-Amersfoort
J. JACKSON	MONSANTO B-Brussels
F. JEKAT	BAYER D-Wuppertal
R. LÄNGE	SCHERING D-Berlin
M. MARTENS	MONSANTO B-Brussels
P. McCAHON	RHÔNE-POULENC F-Sophia Antipolis

R. McKEE	EXXON B-Brussels
N. MOORE	BP CHEMICALS UK-London
D. PALLAPIES	BASF D-Ludwigshafen
A. POOLE	DOW EUROPE CH-Horgen
N. SCHOLZ	HÜLS D-Marl
J. STEVENS	NOVARTIS USA-Greensboro
L. TATTERSFIELD	SHELL INTERNATIONAL UK-Chester
J. TOMENSON	ICI UK-Northwich
R. VAN EGMOND	UNILEVER UK-Wirral
J. VAN MILLER	UNION CARBIDE USA-Danbury
F. CARPANINI	ECETOC B-Brussels

MEMBERS OF THE SCIENTIFIC COMMITTEE

(Peer Review Committee)

W.F. TORDOIR (Chairman), Group Adviser, Environmental Health and Human Toxicology	SHELL INTERNATIONAL NL- Den Haag
H. VERSCHUUREN (Vice-Chairman), Head, Toxicology Department	DOW EUROPE CH - Horgen
O.C. BØCKMAN, Scientific Adviser	NORSK HYDRO N - Porgrunn
N.G. CARMICHAEL, Toxicology Director Worldwide	RHÔNE-POULENC F - Lyon
H. DE HENAU, Research Fellow,Europe BVBA Professional and Regulatory Services	PROCTER AND GAMBLE B - Strombeek-Bever
C. D'HONDT, Head, Ecology Department	NOVARTIS CH-Basel
B. HILDEBRAND, Director, Experimental Toxicology	BASF AG D - Ludwigshafen
J.R. JACKSON, Director, Medicine and Health Science	MONSANTO EUROPE B - Louvain-La-Neuve
E. LÖSER, Head, Institute of Industrial Toxicology	BAYER D - Wuppertal
R. MILLISCHER, Head, Industrial Toxicology Department	ELF ATOCHEM F - Paris
G.L.P. RANDALL, Director, Environmental Laboratory	ZENECA UK - Brixham
A. SARRIF, Associate Director, Toxicology Affairs	DUPONT D-Bad Homburg
J. SOLBE, Head, Ecotoxicology	UNILEVER UK-Wirral
H. WIEGAND, Head, Product Safety Department	HÜLS AG D - Marl