Before the
Office of Management and Budget

OMB SHOULD NOT APPROVE EPA'S ICR FOR THE EDSP TIER 1 TEST BATTERY

COMMENTS ON EPA ICR NO. 2249.01

In the Matter of	)	
Agency Information Collection Activities	)	
Submission to OMB for Review and Approval	)	
Comment Request	)	
Tier 1 Screening of Certain Chemicals	)	
Under the Endocrine Disruptor	)	
Screening Program (EDSP)	)	Docket No. EPA-HQ-OPPT-2007-1081
EPA ICR 2249.01	)	FRL-8412-2
OMB Control No. 2070-New	)	

The Center for Regulatory Effectiveness
Suite 500
1601 Connecticut Avenue, NW
Washington, DC 20009
202.265.2383
www.TheCRE.com

May 15, 2009

OMB SHOULD NOT APPROVE EPA'S ICR FOR THE EDSP TIER 1 TEST BATTERY

The Center for Regulatory Effectiveness ("CRE") appreciates this opportunity to comment on an Information Collection Request ("ICR") that the Environmental Protection Agency has submitted for OMB review under the Paperwork Reduction Act ("PRA"). This ICR would authorize the collection of test results from all of EPA's proposed Tier 1 Battery of tests under the Agency's Endocrine Disruptor Screening Program ("EDSP").

EPA has proposed eleven of these Tests. EPA currently intends to order private parties who manufacture pesticides and other selected substances to perform some or all these tests and to report the results back to EPA. EPA will order additional Tier 2 endocrine effects testing based on the results of the Tier 1 Battery tests. According to its supporting statement for this ICR, EPA will also use the Tier 1 Battery data for risk assessments and other regulatory purposes.[1]

For the following and other reasons, we ask OMB to deny this ICR.

1) ***Approving this ICR would violate the PRA and OMB's practical utility rules.*** The PRA and OMB's rules require that EPA provide a record demonstrating that the EDSP Tier 1 tests have practical utility: *i.e.*, EPA has to show that these tests will generate accurate, valid, adequate and reliable information.[2] EPA has not provided this record for at least eight of the proposed Tier 1 tests. Even EPA admits that the ER binding test is not yet validated; consequently, EPA could not possibly demonstrate that the ER binding test is accurate, reliable, reproducible, unbiased, valid and complete. In addition, EPA asks OMB to approve seven tests in the Tier 1 Battery even though they have negative peer review reports and therefore have no practical utility.[3]

2) ***Approving this ICR would violate EPA and OMB's Information Quality Act Guidelines.*** Under OMB's and EPA's Information Quality Act ("IQA") guidelines, OMB cannot approve this ICR until and unless EPA demonstrates that the EDSP test data will be accurate, reliable, unbiased and complete.[4] EPA cannot make this demonstration for the ER binding test because even EPA admits that the ER binding test is not yet validated and ready to

---

[1] Page 31 of EPA's ICR Supporting Statement, available on line at www.regulations.gov, Document ID EPA-HQ-OPPT-2007-1081-0017

[2] 5 CFR §§1320.3(l); 1320.5(d)(1)(iii); 21 USC § 346a(p)(2); 5 CFR § 1320.5(e).

[3] These seven tests are the H295R, Uterotrophic, Fish Screening, Male and Female Pubertal Rat, AR binding, and Amphibian Metamorphosis tests. They are discussed *infra* in CRE's comments at page 14 *et seq*.

[4] *E.g.,* Page 12 of OMB IQA Guidance *available online at* http://www.whitehouse.gov/omb/inforeg/iqg_comments.pdf;; EPA IQA guidelines, page 15 and Section 6.5, page 28 , available online at http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf; OMB Government-wide IQA Guidelines, Section V.3, available online at http://www.whitehouse.gov/omb/assets/omb/fedreg/reproducible2.pdf.

use. EPA also can't make this demonstration for seven other EDSP tests because they have negative peer review reports.[5]

3) ***CRE and the rest of the public have not had an adequate opportunity to comment on this ICR.*** CRE and the rest of the public have a right to comment on this ICR. 5 CFR § 1320.10(a),(b)(public has 60 days to comment to EPA, and another 30 days to comment to OMB). Approving this ICR for the ER binding test would violate our right to comment. Even EPA admits that this test is not yet validated and therefore doesn't have a complete record for review and comment. Upon information and belief, EPA has a peer review report on the ER binding test but has not publicly disclosed it. We can't comment on a report that EPA won't disclose. OMB can't review a report that EPA won't disclose.

In addition, CRE filed a FOIA request seeking data and documents related to this ICR. In response, EPA sent CRE some documents, many of which were already publicly available. EPA also sent a 402-page single spaced document identifying all the EDSP documents that EPA refuses to disclose.[6] EPA's refusal to disclose these documents violates the Justice Department's new FOIA guidelines, prevents CRE and the rest of the public from having an adequate opportunity to comment on this ICR, and prevents OMB from having an adequate record to review this ICR.

The ICR record is inadequate for still other reasons. For example, EPA has not yet responded to the many long-standing comments on the EDSP tests which are the subject of this ICR. EPA's strategic decision to delay its comment response precludes adequate public comment on this ICR, and precludes adequate OMB review of the ICR.[7]

CRE's right to comment is also violated by EPA's heavy reliance on OECD validation and acceptance for five of the eleven Tier 1 tests.[8] CRE, most of the rest of the public, and OMB itself are not privy to the current OECD action on these tests, including but not limited to the Amphibian Metamorphosis test. The OECD peer review and validation process for one of

---

[5] These seven tests are the H295R, Uterotrophic, Fish Screening, Male and Female Pubertal Rat, AR binding,and Amphibian Metamorphosis tests. They are discussed *infra* in CRE's comments at page 14 *et seq*.

[6] This EPA document is available online at **http://thecre.com/pdf/EDSP%20list%20consolidated.pdf**

[7] EPA states that it will respond to comments on EDSP tests in a separate proceeding. *See, e.g.,* documents available online at http://edocket.access.gpo.gov/2009/E9-8709.htm,; http://www.epa.gov/QUALITY/informationguidelines/documents/08004-response.pdf , and http://www.epa.gov/fedrgstr/EPA-PEST/2007/July/Day-13/p13672.pdf. EPA has not responded to the long-standing public comments on these tests as of the time that CRE had to file comments on the ICR. This means that the Agency has not responded to most of the long-standing challenges to and criticism of the EDSP Tier 1 tests including, *e.g.*, the following comments to EPA available online at www.regulations.gov: EPA-HQ-OPP-2008-0012-0040.1 (American Chemistry Council comments); EPA-HQ-OPP-2008-0012-0001 (CropLife America comments) EPA-HQ-OPPT-2007-1080-0014 and EPA-HQ-OPPT-2007-1081-0008 (CRE comments); EPA-HQ-OPP-2008-0012-0021.1(PETA comments).

[8] These five tests are the Uterotrophic; Hershberger, Fish Screening, Amphibian Metamorphosis, and Human Cell Stably Transfected Transcriptional Activation Estrogen Receptor Binding tests.

EPA's Tier 1 tests--the Uterotrophic test--was condemned as unscientific and biased by the Science Advisory Committee of the European Centre for the Validation of Alternative Methods ("ECVAM"), and by other stakeholders in published articles.[9] ECVAM's unanswered critique of the OECD validation of the Uterotrophic test raises very serious questions about EPA's reliance on the OECD validation process both for the Uterotrophic and for several other Tier 1 tests.

4) ***A Science Advisory Panel ("SAP") needs to review EPA's new position that not all of the Agency's proposed tests are necessary for a useful Tier 1 screening batter, and a SAP needs to review EPA's determinations that all but one of the eleven tests in the Tier 1 Battery are currently validated.*** EPA is statutorily required to have a SAP review the Agency's EDSP screening program. 21 USC § 346a(p)(2). A SAP has only reviewed and approved an EDSP screening program based on a hypothetically complete Tier 1 Battery of tests. A SAP has never reviewed and approved a testing strategy that includes anything less than a complete battery of tests for each substance, which is what EPA now proposes. And no SAP has ever been allowed to review EPA's validation determinations for any EDSP test.

EPA apparently believes that once a test has been peer reviewed, it is validated, even if the peer review report is negative. Expert unbiased SAP review of the validation record for each proposed EDSP test is necessary to ensure that the statutory requirement of test validation is met.

5) ***EPA's ICR burden estimates are inaccurate, incomplete and developed contrary to the PRA's requirements.*** An analysis of EPA's burden estimates for this ICR yields three primary conclusions:

1. EPA has not included relevant cost components in their burden estimates;

2. The agency's burden estimation scheme places the "cart-before-the-horse" by using the Tier 1 test order itself as a *de facto* pilot project for the ICR that should be developed prior to issuing the Tier 1 test order; and

3. The burden on EPA of administering the Tier 1 test order is substantial and raises serious questions about the agency's priorities in their use of limited resources.

6) EPA has not demonstrated compliance with the 3 Rs. EPA's EDSP tests will kill and cause suffering in many, many animals. EPA itself admits that the first Tier 1 assays alone could kill 19,000 animals.[10] PETA claims that "The EDSP is by far the largest animal-testing program of all time, with the potential to kill tens of millions of animals."[11]

---

[9]See ECVAM documents quoted as pages 18-19 of CRE's ICR comments, *infra*.

[10] EPA Comment Response Document for Endocrine Disruptor Chemical Selection/PrioritySetting (Nov. 2004), Page D-3, *available online at* http://www.regulations.gov as document EPA-HQ-OPPT-2004-0109-0003

[11] http://www.peta.org/feat/greenwash/nrdc.html

Reduction in animal suffering/death is an essential part of the test validation process. As EPA explained, the Tier 1 Battery validation program must include "adequate animal welfare considerations (3Rs)": *i.e.*, will the Tier 1 Battery **reduce, refine** to make less stressful, and **replace** animal tests? [12]

In Professor Cass Sunstein's words, "At the very least, I suggest that suffering and harm to animals should count, and that any measures that impose suffering and harm should be convincingly justified."[13]

EPA has not "convincingly justified" its EDSP Tier 1 Battery. There is no basis in the ICR record for concluding that most of these tests will even generate accurate and reliable data, in violation of the very statutes that EPA purports to be working under. *See, e.g.*, 21 USC § 346a(p)(1)(EPA has to validate Tier 1 tests before using them); 42 U.S.C. § 281-4(c)(same).

There follows additional discussion of these and other reasons for denying this ICR.

[12] Who's Who in the Validation of Assays for the EDSP, Briefing for new EDMVAC Members, March 2, 2005, *available online at* http://www.epa.gov/oscpmont/oscpendo/pubs/edmvac/validation_briefing_edmvac_030205.pdf.
[13] The Rights of Animals: A Very Short Primer, *Cass R. Sunstein*, *available online at* http://www.law.uchicago.edu/academics/publiclaw/resources/30.crs.animals.pdf

BEFORE OMB APPROVES AN ICR FOR AN EDSP TEST, EPA MUST PRODUCE A PUBLIC RECORD DEMONSTRATING THE TEST IS ACCURATE, RELIABLE, UNBIASED, COMPLETE, ADEQUATE, AND VALID

EPA asks OMB to approve an ICR for all eleven tests that EPA is thinking about using in the EDSP tier 1 Battery. In order for OMB to approve this ICR, EPA must demonstrate that all these tests will generate information which meets the IQA quality standards of accuracy, reliability, no bias, and completeness. EPA cannot make this demonstration for tests that are not validated: *i.e.*, that have not been demonstrated to generate accurate, reliable, unbiased and complete information.

OMB's IQA guidance is unambiguous and unequivocal on this requirement:

> "...we note that each agency is already required to demonstrate the 'practical utility' of a proposed collection of information in its PRA submission, i.e., for draft information collections designed to gather information that the agency plans to disseminate. Thus, we think it important that each agency should declare in its guidelines that it will demonstrate in its PRA clearance packages that each such draft information collection will result in information that will be collected, maintained, and used in a way consistent with the OMB and agency information quality standards. It is important that we make use of the PRA clearance process to help improve the quality of information that agencies collect and disseminate. Thus, OMB will approve only those information collections that are likely to obtain data that will comply with the OMB and agency information quality guidelines."[14]

EPA's own IQA guidelines require EPA to demonstrate to OMB and the public that the EDSP ICR will generate information that complies with the IQA quality standards:

> "For all proposed collections of information that will be disseminated to the public, EPA intends to demonstrate in our Paperwork Reduction Act clearance submissions that the proposed collection of information will result in information that will be collected, maintained and used in ways consistent with the OMB [IQA] guidelines and these EPA [IQA] Guidelines." [15]

---

[14] Page 12 of OMB IQA Guidance *available online at* http://www.whitehouse.gov/omb/inforeg/iqg_comments.pdf

[15] *E.g.*, EPA IQA guidelines, Section 6.5, page 28, available online at http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf; OMB Government-wide IQA Guidelines, Section V.3, available online at http://www.whitehouse.gov/omb/assets/omb/fedreg/reproducible2.pdf

The OMB and EPA IQA guidelines both require accurate, reliable, complete and unbiased information. [16] Consequently, before OMB can approve this ICR for any EDSP test, EPA has to demonstrate that that test will generate accurate, reliable, complete and unbiased information.

Independent of the IQA/PRA interface, OMB's ICR rules under the PRA require that EPA demonstrate that these tests will generate accurate, valid, adequate and reliable information. OMB cannot approve an EDSP ICR until and unless EPA makes this demonstration, and EPA has not yet made this PRA-required demonstration.

OMB's ICR rules define the term practical utility as "the actual, not merely the theoretical or potential, usefulness of information to or for an agency, taking into account its **accuracy, validity, adequacy, and reliability**...." 5 CFR §1320.3(l)(emphasis added).

With regard to EPA's duties, the ICR rules state that "[t]o obtain OMB approval of a collection of information, an agency shall demonstrate that it has taken every reasonable step to ensure that the proposed collection of information...has **practical utility**." 5 CFR 1320.5(d)(1)(iii)(emphasis added).

The PRA itself requires that

> "As part of the agency submission to OMB of a proposed collection of information, the agency (through the head of the agency, the Senior Official, or their designee) shall certify **(and provide a record supporting such certification)** that the proposed collection of information--
> (a) Is necessary for the proper performance of the functions of the agency, **including that the information to be collected will have practical utility...."**

21 USC § 346a(p)(2). emphasis added).

With regard to OMB's duties, the ICR rules require that

> "OMB shall determine whether the collection of information, as submitted by the agency, is necessary for the proper performance of the agency's functions. In making this determination, OMB ...will consider whether the burden of the collection of information is justified **by its practical utility."**

5 CFR § 1320.5(e)(emphasis added).

---

[16] *E.g.*, EPA IQA Guidelines , page 15, available online at http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf; OMB Industry wide IQA Guidelines, Section V.3, available online at http://www.whitehouse.gov/omb/fedreg/final_information_quality_guidelines.html

In other words, OMB has an independent, mandatory duty under its own PRA Information Collection rules to determine whether EPA has produced a public record demonstrating that each EDSP test covered by this ICR will generate valid, accurate, reliable and adequate information. OMB cannot approve an ICR for an EDSP test until and unless EPA has produced such a record,

EPA has not produced an adequate public ICR record for each of its proposed eleven EDSP tests. In fact, as discussed below, EPA incorrectly claims it doesn't have to produce such a record at all.

EPA NOW INCORRECTLY CLAIMS THAT IT DOESN'T HAVE TO MEET THE IQA AND PRA REQUIREMENTS

EPA asserts in its response to ICR comments, "Completion of the validation process for all of the assays in the final tier 1 battery is not a prerequisite for compliance with the PRA." [17]

EPA's statement is equivalent to saying that EPA does not have to provide a record demonstrating IQA and practical-utility compliance for an EDSP test in order for OMB to approve an ICR for that test. This statement is both incorrect and inconsistent with EPA's prior representations.

EPA cannot possibly demonstrate IQA guideline compliance and PRA practical utility for a test that has not successfully completed the validation process. A record showing successful completion of the validation process is necessary to comply with the IQA and PRA requirements that each Tier 1 test be accurate, reliable, etc....

EPA's current statement is also inconsistent with EPA's prior representations. Before it became convenient to argue otherwise, EPA claimed that the entire Tier 1 Battery of tests is necessary for any of these tests to have practical utility. For example, EPA told its EDSP Science Advisory Panel that the entire Tier 1 Battery of tests is necessary to screen for endocrine effects:

> "In interpreting the battery... using the weight of evidence includes professional judgment, you know, some end points more diagnostic specific than others and, and really **it's the weight of various effects seen in multiple endpoints and across multiple assays that carry the most weight**. We're looking for that confirmation of corroboration across the assays and two possible interpretational outcomes, either the potential for the activity on estrogen action hormones, that would require some further analysis to the Tier 2 family of the patient or we can interpret that there's low and no potential for EAT activity so that the compound can be, you know, pushed aside instead of somewhat harping on it.
>
> In summary, your multiple assays are required to comprehensively screen estrogens and androgens, the thyroid hormone systems…. **The complete battery is needed to support a weight of the evidence finding** something lower, low potential for EHE activities."[18]

---

[17] Page 4 of Response to Comments on the Public Review Draft of the Information Collection Request (ICR) entitled: "Tier 1 Screening of Certain Chemicals Under the Endocrine Disruptor Screening Program (EDSP)," EPA ICR No. 2249.01, OMB Control No. 2070–new (72 FR 70839, December 13, 2007), April 10, 2009, Document ID EPA-HQ-OPPT-2007-1081-0017.10, available online at http://www.regulations.gov/fdmspublic/component/main?main=DocumentDetail&o=0900006480954383

[18] SAP Meeting Transcript. Pages 61-62(emphasis added), available online at http://www.epa.gov/scipoly/sap/meetings/2008/march/transcript2008-03-25.pdf

The purpose of this SAP was to determine whether, as a whole, EPA's proposed Tier 1 Battery of tests was adequate to screen for endocrine effects. The SAP's first charge question was:

> "Please comment on the ability of the proposed Tier 1 Screening Battery to provide sufficient information to determine whether or not a substance potentially interacts with the estrogen, androgen, and thyroid hormonal systems based on the modes of action covered within the battery."[19]

No SAP has ever opined as to whether a Tier 1 Battery would have practical utility as a screen if the Battery contained less than the 11 tests reviewed by 2008 SAP. SAP review of this issue is necessary for OMB's review of EPA's sudden change of mind, if EPA continues to maintain this abrupt change of position. See 21 USC § 346a(p)(2); 21 USC § 346a(p)(2).

It is important to note that, even aside from the test validation/non-validation issue, EPA has had a remarkable change of heart and decided that they may not need all of the Tier 1 battery of tests. The agency also raises the possibility of their not needing all of the tests for all of the chemicals. Moreover, the agency explains that they have not yet even defined what tests constitute the Tier 1 screening battery.

In their April 15, 2009 ICR Support Statements, EPA states: "the Agency has identified the universe of assays that might be included in the final Tier 1 battery. The Agency may, however, not include all of these assays in the final Tier 1 battery, or may otherwise decide to only include a subset of these assays in an individual Tier 1 Order that is issued under the EDSP for a particular chemical and respondent." Thus, EPA has not yet defined the Tier 1 screening battery.

EPA does not explain how they will decide which Tier 1 tests to use. On what basis will EPA decide whether all or some of the tests included in this ICR are needed for all chemicals? How will the agency decide which tests are needed for only certain chemicals? When does the public have the opportunity to comment on these decisions? When does an SAP have an opportunity to comment on these decisions? What assurance does the public have that EPA test decisions will not be made on an arbitrary and capricious basis?

One of the certifications that EPA is required, by law, to make to OMB is that the information sought is "is necessary for the proper performance of the functions of the agency…." Since EPA has stated in their Supporting Statement that some of the tests they seek authority to require may not be needed, such certification of necessity is, by EPA's own admission, false. On this basis alone, OMB should disapprove the ICR and require EPA to make a determination as to which tests constitute the Tier 1 screening battery before submitting a revised ICR.

---

[19] SAP Meeting Transcript, pages 61-62, available online at http://www.epa.gov/scipoly/sap/meetings/2008/march/transcript2008-03-25.pdf

OMB SHOULD NOT APPROVE AN ICR FOR THE ER BINDING TEST BECAUSE THERE IS NO ER BINDING TEST

EPA admits there is no validated ER binding test. EPA cannot demonstrate IQA compliance and practical utility for a test that even EPA admits is not yet validated. Therefore, OMB cannot approve an ICR for the ER binding test.

There are other questions that EPA needs to answer about the ER binding assay.

On or after March 13, 2009, EPA sent an ER binding test out for peer review for use as a EDSP Tier 1 test. According to EPA's public statements, peer review and validations should be completed soon. See
http://www.epa.gov/endo/pubs/test guidelinevalidation/estrogen.html

EPA's public statements conflict with an EPA internal document dated "11/14/08," which provides the following "Interlaboratory Validation Schedule" for the ER binding test:

> "•Data analysis and completion: Fall 2009
> •Final report: winter 2010
> •Peer review: Spring 2010?"

This schedule pertains to both the FWA and CERI "Human Recombinant Estrogen Receptor Binding Assay Validation." [20]

Has EPA prematurely rushed the ER binding assay out to peer review in order to speed up its Tier 1 Battery implementation?

The answer to this question is yes if you believe EPA's Integrated Summary Report for the ER binding assay. There, EPA explained that

> "While these results were disappointing, the study continued with these laboratories. Given the experience in finding these laboratories, finding other laboratories would have delayed the study -- and the Screening Program -- significantly, and the deviations were judged to be marginally acceptable in this context."[21]

The Integrated Summary Report further explains that

---

[20] EPA document available online at http://thecre.com/pdf/VMG%20Slides%20scan.pdf

[21] Integrated Summary Report for Validation of an Estrogen Receptor Binding Assay using Rat Uterine Cytosol as Source of Receptor as a Potential Screen in the Endocrine Disruptor Screening Program Tier 1 Battery (PDF), page 49, available online at http://www.epa.gov/endo/pubs/assayvalidation/er_ruc_isr.pdf.

1) EPA couldn't find a negative control for the ER binding assay, even though the test's performance criteria require negative controls. There were recurring problems finding both positive and negative controls which is "a serious deficiency of the assay."

2)"Intralaboratory variability was disappointingly high."

3) The "original performance criteria were not followed by participating labs."[22]

[22] *Id.*, at pages 50, 56, 64-65, 67.

APPROVING THIS ICR WOULD VIOLATE THE PUBLIC NOTICE AND COMMENT REQUIREMENT IN OMB'S PRA RULES

Approving an ICR for the ER binding test would violate the public notice and comment requirements in OMB's PRA rules. These rules require that the public be given a 60-day comment period to EPA, and a following 30-day period to comment to OMB, on this ICR and on its supporting documentation. 5 CFR § 1320.10(a),(b). The ER binding test lacks a peer review report and lacks an EPA response to peer review. The public cannot comment on a test whose record is incomplete.

There is no prior opportunity to comment that mitigates this violation of the PRA rules. Test protocols and records that aren't publicly available now, weren't publicly available then. In fact, the public has never had access to the entire record for any of the Tier 1 tests because EPA is withholding data.

In order to submit complete and informed comments on this ICR, CRE filed a Freedom of Information Act request with EPA to uncover the record for all the EDSP tests. EPA's response was to produce some documents, most of which were already available, **and to send CRE a 402-page single-spaced list of all the documents that EPA is withholding from disclosure.**[23]

Upon information and belief, many of the documents on this list are essential to determining whether EPA's Tier 1 Battery will generate accurate, reliable, unbiased, adequate, complete and valid information.

Withholding these documents violates the Justice Department's new FOIA guidelines. The new guidelines state that DOJ will defend non-disclosure only if an agency "reasonably foresees that disclosure would harm an interest protected by one of the statutory exemptions," such as national security and law enforcement interests, or if disclosure is prohibited by law. [24] EPA has not made this showing with regard to any of the withheld documents.

Withholding these documents also means that none of the EDSP tests are validated. EPA's own validation principles require the Agency to "[m]ake publicly available all data supporting the assessment of the validity of the assays including the full data set collected during the validation studies and publish results in independent, peer-reviewed scientific journals." [25] EPA is

---

[23] This EPA document is available online at **http://thecre.com/pdf/EDSP%20list%20consolidated.pdf**

[24] U.S. Attorney General's March 19, 2009 Memorandum for Heads of Executive Departments, available online at http://www.dhs.gov/xlibrary/assets/foia/ag_foia_memo2009-03-19.pdf.

[25] EPA adopts these validation principles at , e.*g.*, Pages 10-11 of EPA's document ENDOCRINE DISRUPTOR SCREENING PROGRAM (EDSP): PROPOSED TIER 1 SCREENING BATTERY (March 7, 2008), available online at http://www.epa.gov/scipoly/sap/meetings/2008/march/technical_review.pdf.

concealing much of the validation data, and EPA has not published the validation results for these tests in peer reviewed scientific journals.

There follows a detailed test-by-test explanation the problems with most of the other tests in the EPA's proposed Tier 1 Battery. This explanation is based primarily on peer review results.

Even a cursory review of the peer review reports shows that these Tier 1 tests are not accurate, reliable, reproducible, complete, valid, adequate and/or unbiased. Consequently, they do not meet the EPA/OMB IQA guidelines, and they do not meet the PRA practical utility requirement. EPA has not made its required showing, and has not produced the required record; therefore, OMB should not approve an ICR for these tests.

---

and Page 7 of EPA's Validation Paper, *Validation of Screening and Testing Assays Proposed for the EDSP* (October 23, 2006), available online at http://www.epa.gov/scipoly/sap/meetings/2007/february/edsp-validation-paper.pdf;

OMB SHOULD NOT APPROVE AN ICR FOR THE H295R TEST BECAUSE THIS TEST FAILED PEER REVIEW. EVEN EPA ADMITS THAT THIS TEST IS INCOMPLETE

EPA told OMB:

> "In addition, the H295R assay peer review has since been completed (June 2008). Although EPA did not ask for a consensus peer review panel opinion, the comments of one panel member are typical: "The H295R steroidogenesis assay is biologically and toxicologically relevant to the stated purpose. The assay would fit perfectly in the Tier 1battery of assays to screen for endocrine disruptors. The assay has a series of strengths that would make it an excellent screening tool for endocrine disruptors of sex steroid hormone synthesis."[26]

EPA's statement about the H295R peer review is misleading.

EPA publicly disclosed the Peer Review Report for H295R only after repeated inquiries by CRE representatives and others. Upon information and belief, EPA did not make the H295R peer review report publicly available until March 26, 2009, even though it was completed in June 2008. There's a reason why EPA didn't want to disclose this document. Actually, there are many reasons why.

For example, the Peer Review Report includes the following comment, "Overall, the test guideline has the potential to be a screening tool for steroidogenesis but requires further testing and refinement."[27]

EPA's only response to this comment is "No response needed."

EPA responds to 10 other separate and significant peer review criticisms of H295R by stating, "This is being investigated and the protocol modified, if appropriate." To the best of CRE's knowledge, EPA is still investigating these peer review criticisms of H295R. EPA's responses to the many peer review criticisms of H295R acknowledge that the test needs work and is not final.

This needs-work conclusion is consistent with the following EPA response to a peer review criticism:

---

[26] EDSP-EPA response to PCRM-2008-10-20.doc, Docket ID No. EPA-HQ-OPPT-2007-1080 Additional Materials Provided to OMB During EO 12866 Review of the Draft Federal Register document entitled: "Endocrine Disruptor Screening Program (EDSP); Policies and Procedures for Initial Screening; Notice" (08/11/2008).

[27] *Disposition to Peer Review Panel Comments on the H295R Steroidogenesis Assay*, finally available online at http://www.epa.gov/endo/pubs/assayvalidation/h295r_pr.htm .

> "There will be a complete and separate H295R protocol after this assay undergoes peer review both by the US-EPA and OECD which will combine all of these aspects."

To the best of CRE's knowledge, H295R has not been additionally peer reviewed by EPA and the OECD. Consequently, by EPA's own admission, there is no complete H295R test protocol.

Some of the many peer review criticisms of H295R are included in the Appendix to CRE's ICR comments. They are too lengthy and detailed to be included in the text of our comments.

Under these circumstances, OMB cannot approve an ICR for H295R, because EPA cannot demonstrate that H295R will generate accurate, reliable, unbiased and complete information.

The Aromatase Test Peer Review Report emphasizes the Importance of the H295R test to the EDSP Tier 1 Test Battery. There, in response to criticism of the aromatase test, EPA stated:

> "EPA has recognized this limitation and is currently validating the H295R cell-based assay to identify inhibitors of aromatase activity in addition to chemicals that induce aromatase (CYP 19) mRNA that results in increased aromatase activity in addition to chemicals that induce aromatase (CYP 19) mRNA that results in increased aromatase activity" [28]

The aromatase test peer reviewers again commented on this test's limitations, and EPA again stated that a validated H295R would solve this problem:

> "As noted before, an assay that detects both inhibition and induction would be more useful. This assay may become outdated in the near term."
>
> ***
>
> "EPA agrees and is validating the H295R assay."[29]

In sum, there is no H295Rtest yet, and this test is essential to an effective EDSP Tier 1 Test Battery.

---

[28] Page 36 of EPA Response to Peer Review Comments on Aromatase Assay, available online at **http://epa.gov/endo/pubs/aromatase_peer_review_response.pdf**

[29] *Id.* at page 50

OMB SHOULD NOT APPROVE AN ICR FOR THE UTEROTROPHIC TEST BECAUSE ITS PEER REVIEW DID NOT MEET THE OMB PEER REVIEW GUIDELINES, AND BECAUSE IT FAILED PEER REVIEW ANYWAY

EPA stated that validation of this test was completed in June 2003, that a draft test was approved by OECD coordinators in March 2007, and that "[t]his assay is ready for use." [30]

EPA further stated that "the uterotrophic assay has been peer reviewed by OECD. EPA will not conduct a peer review of this assay." [31]

The two international bodies in charge of validating new tests have both stated that the uterotrophic assay is not yet valid. One of them challenged the integrity and objectivity of the OECD peer review process itself, and the other hasn't disagreed.

ICCVAM rejected the Uterotrophic test and stated that it is not yet validated. As noted in ICCVAM's letter rejecting the Uterotrophic test, EPA is the only agency on ICCVAM's Endocrine Disruptor Working Group which believes that the Uterotrophic test is validated and ready for use. EPA's disagreement with every other agency in the Group suggests bias on EPA's part:

> "On behalf of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), we are pleased to provide the enclosed comments on the most recent revised version of the proposed draft OECD Test Guideline (TG) 'The Uterotrophic Bioassay in Rodents: a short-term screening test for oestrogenic properties.' These comments are in response to your January 29, 2007 notification that the OECD has requested a round of comments on a second revised version of the original draft TG. The previous sets of comments relating to this test method are attached and it is requested that they be addressed before proceeding with further consideration of this draft TG.
>
> As stated in previous comments submitted by ICCVAM to the U.S. National Coordinator on this test method and the proposed test guideline, ICCVAM and its Endocrine Disruptor Working Group have concluded, with only one agency in disagreement, that the uterotrophic bioassay has **not** been adequately validated for its intended purpose. Thus, ICCVAM's recommendation is that this material be placed in an OECD Guidance Document, which could then be used as the basis for further studies that could lead to an adequate demonstration of validation. This is identical with the approach proposed by the OECD for the uterotrophic bioassay to detect estrogen antagonists.

[30] EPA's EDSP December 2008 Status Filing, available online at http://thecre.com/pdf/EPA%20Wall%20letter.pdf.

[31] EPA document available online at http://www.epa.gov/endo/pubs/assayvalidation/peerreview_process.htm

> In addition, the OECD needs to fully recognize the importance of having TGs based on adequately considered and evaluated draft TGs and that providing a revised TG with only a few days to consider it is entirely inappropriate. It is critical to the success of the TG program and the acceptance of data under MAD that adequate time for a careful review be provided."[32] [Emphasis added]

EPA should defer to ICCVAM's opinion on this issue. ICCVAM is the U. S. Interagency Coordinating Committee for the Validation of Analytical Methods. The federal statute creating ICCVAM states that

> "The purposes of the ICCVAM shall be to...
>
> (1) increase the efficiency and effectiveness of Federal agency test method review;
>
> (2) eliminate unnecessary duplicative efforts and share experiences between Federal regulatory agencies;
>
> (3) optimize utilization of scientific expertise outside the Federal Government;
>
> (**4) ensure that new and revised test methods are validated to meet the needs of Federal agencies**; and
>
> (5) reduce, refine, or replace the use of animals in testing, where feasible."

42 USC 285*l*-3 (emphasis added).

EPA itself has broadly described the "ICCVAM Role" in EPA's EDSP validation process:

> "•determination of readiness for validation in consultation with ICCVAM
>
> •Validation
>
> •Peer review by SAB/SAP and review by ICCVAM."[33]

---

[32] ICCVAM letter available online at http://iccvam.niehs.nih.gov/methods/endocrine/OECDdocs/Uterotrophic/ICCVAM_UB_Ltr_Smrchek.pdf

[33] http://74.125.93.104/search?q=cache:G_7WqpdJ_SEJ:epa.gov/endo/presentations/stdvalgt.ppt+edsp+validation+process+IccvAMA+role&cd=2&hl=en&ct=clnk&gl=us

ECVAM agrees with ICCVAM that the Uterotrophic test has not been validated and is not supported by peer review. ECVAM is the EU equivalent of ICCVAM. Its duties include "coordinating the validation of alternative test methods at the European Union level." [34]

With regard to the Uterotrophic test, ECVAM disputed the integrity of the peer review process. ECVAM's letter is available online at http://ecvAM.jrc.it/index.htm.
This letter follows in its entirety:

"Ispra 15/2/2005

ECVAM-European Centre for the Validation of Alternative Methods

ESAC Statement on the status of validation of the uterotrophic assay

At the 22nd meeting of the ECVAM Scientific Advisory Committee (the ESAC) held 14 - 15 February, 2005, the outcome of the recent OECD Endocrine Disruptor Testing and Assessment (EDTA) committee meeting regarding the uterotrophic assay was discussed.

The ESAC understands that, at the meeting of the EDTA, it was concluded that the uterotrophic test has been validated for the purposes intended for the assay. This decision was taken, despite the fact that a final report of the peer review panel has yet to be produced.

The peer review was organised and managed by the OECD and involved members, some of whom were part of the validation study itself. As a consequence, the ESAC cannot accept that the validation study has been endorsed properly, by an independent peer review process.

The ESAC is particularly concerned that the decision by the EDTA contravenes the processes laid down for the peer review of validation studies, by the OECD in its draft guidance document 34, and as recommended by ICCVAM and ECVAM.

One of the roles of the ESAC is to make formal recommendations on the scientific validity of alternative and advanced methods and to endorse validation studies. Such independent advice is provided to the European Commission. From the information available, the ESAC believes that a robust and defensible peer review process for the validation study of the uterotrophic assay has not demonstrably been conducted, and that this sets an unacceptable precedent for future validation studies and their peer reviews.

[34] http://ecvAMA.jrc.it/ .

The ESAC considers that there should be a proper peer review, which should be organised and coordinated completely independently of the OECD, and which should be fully transparent. The peer review panel should comprise individuals who have no conflicts of interest, but who are experienced in the science involved in the assay and in the process of validation, in order to achieve an objective evaluation of the validation study.

Thomas Hartung
Head of Unit, ECVAM
Ispra 15/2/2005

Jan van der Valk
Moderator of ESAC
Ispra 15/2/200

Non-Commission Members of the ESAC present at the 22nd ESAC Meeting:

Dr Nathalie Alépée (EFPIA)
Dr Sonja Beken (Belgium)
Dr Argelia Castaño (Spain)
Professor Robert Combes (ERGATT)
Dr Maija DAMbrova (Latvia)
Professor André Guillouzo (France)
Dr Julia Fentem (ECETOC)
Dr Katalin Horvath (Hungary)
Professor Elisabeth Knudsen (Denmark)
Dr Manfred Liebsch (Germany)
Dr Mykolas Maurica (Lithuania)
Dr Efstathios Nikolaidis (Greece)
Professor Milan Pogačnik (Slovenia)
Dr Jon Richmond (UK)
Professor Michael Ryan (Ireland)
Dr Odile de Silva (COLIPA)
Dr Dariusz Sladowski (Poland)
Dr Annalaura StAMmati (Italy)
Professor Eric Tschirhart (Luxembourg)"

These problems and flaws in the OECD peer review/validation process are the subject of published aticles. For example,

The involvement of the OECD in managing the validation of the rat uterotrophic assay for endocrine disruptors, and in organising the peer review of the results of this study, has been assessed and compared with the many conclusions and recommendations in several published reports of international workshops on validation, and information in guidance documents, produced by the European Centre for the Validation of Alternative Methods (ECVAM), the US Interagency Coordinating Committee on the Validation of Alternative

Methods (ICCVAM) and the OECD itself. It is concluded that the OECD has not followed the recommendations for full transparency and independence of the peer-review `process. This is based on the fact that it has published a draft guidance document that differs from the report of a recent OECD workshop on validation, in such a way as to give the OECD the flexibility to fully control the peer-review process and, in so doing, to avoid full transparency. Comparison of the timing of the organisation of workshops by the OECD and the progression of the validation study, together with the fact that a draft test guideline for the assay was written before completion of the peer review, suggest that the OECD has given a higher priority to the expedition of the validation and regulatory acceptance of the uterotrophic assay than it has to good scientific and logistical practice. This severely undermines its credibility in the validation process, so, in order for the OECD to be rightly perceived as an honest broker, it is recommended that the OECD should play no role in the validation of new or revised tests, until after they have been successfully validated, peer reviewed, and endorsed by the appropriate authorities, and are ready for test guideline development. With regard to the on-going OECD validation studies of other in vivo assays for endocrine disruptors, the OECD should take immediate steps to ensure full independence and transparency of their peer review."[35]

Given this controversy over the OECD peer review process, it is perhaps ironic that the actual Peer Review Report on the Uterotrophic test does not even support validation. In the Report's own words:

"This report reflects the peer review of a study done to validate a uterotrophic protocol for testing of endocrine disrupting activity. The protocol was developed by experts of an OECD Validation Management Group established by the OECD Endocrine Disruption Testing and Assessment Task Force.

The validation of this protocol included the testing of a number of substances by different laboratories for the evaluation of the reliability, reproducibility and relevance of the protocol. The peer review panel was asked to give its views on the validation report. The peer review panel did not reach a consensus but were able to agree a summary report."

***

"3. Regarding the overall validation exercise, the final conclusions and the views of the PRP are divided into broad groups and there were considerable differences expressed regarding the various components of the project. The PRP was unable to reach consensus on the issue of the validation status of the uterotrophic assay, and the differences in opinion between PRP members were significant. Some members considered the Uterotrophic Bioassay to be validated for the intended purpose of the assay, other

---

[35] *Peer review of validation studies: an assessment of the role of the OECD by reference to the validation of the uterotrophic assay for endocrine disruptors*, Altern Lab Anim. 2004 Jun;32(2):111-7, available online at **http://thecre.com/pdf/ATLA%20document.pdf**

members considered that further data, including on negative substances, was necessary to reach a decision on the validation status of the assay, whilst other members considered the efforts to date were not sufficient to validate the test method but to only be sufficient as a pre-validation study."[36]

EPA is relying on the OECD peer review/validation process for several tests in the Tier 1 Battery.[37] ECVAM's expose of the flaws in the OECD process for the Uterotrophic test leads inevitably to the following question. How does OMB know that the OECD process for these other tests is not also flawed and biased?

An August 2007 OECD report tried to rehabilitate the Uterotrophic test. Even this OECD report had to admit, however, that some chemicals with positive results in the test should have been submitted to the "enhanced 1-generation test," but weren't.

This OECD report also admitted the Uterotropic test results for one chemical are "considered questionable."[38]

This OECD report also acknowledges that "the ER binding assay is more sensitive than the uterotrophic assay for weak responses...."[39] The OECD report reemphasizes that "the ER binding assay is more sensitive than the uterotrophic assay and a weak response (RBA =<0.002%) may not be translated into a biologically meaningful signal in the uterotrophic and would be negative in that assay."[40]

---

[36] Number 68,SUMMARY REPORT OF THE UTEROTROPHIC BIOTEST GUIDELINE PEER REVIEW PANEL, INCLUDING AGREEMENT OF THE WORKING GROUP OF NATIONAL COORDINATORS OF THE TEST GUIDELINES PROGRAMAME ON THE FOLLOW-UP OF THIS REPORT, pages 14, 16, *available online at* http://www.olis.oecd.org/olis/2006doc.nsf/LinkTo/NT0000734A/$FILE/JT03219013.PDF

[37] The OECD tests include the Hershberger test, the 21-day fish test, the amphibian metamorphosis test, and the human cell stably transfected transcriptional activation (STTA) estrogen receptor binding test. The latest OECD action on these tests is not included on EPA's EDSP test validation website, http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/index.htm. EPA has not provided any other public link or access to the current OECD action on these tests.

[38] Page 18 of OECD report available online at http://www.epa.gov/endo/pubs/uterotrophic_OECD_validation_report.pdf

[39] Page 19 of OECD report available online at http://www.epa.gov/endo/pubs/uterotrophic_OECD_validation_report.pdf

[40] Page 19 of OECD document available online at http://www.epa.gov/endo/pubs/uterotrophic_OECD_validation_report.pdf.

As explained above, there is no ER binding test yet. Consequently, the practical utility of the entire EDSP Tier 1 Test Battery is very questionable because the Battery needs an ER binding test before the Battery could possibly be an effective endocrine effects screen.

OMB SHOULD NOT APPROVE AN ICR FOR THE FISH SCREENING TEST BECAUSE THIS TEST FAILED PEER REVIEW

EPA states that the validation of the Fish Screening Test was completed in January 2008, and that "[t]his assay is ready for use." [41] This statement conflicts with the Peer Review Report for the test.

Relevant portions of the Fish Screening test Peer Review Report are set forth below. The quoted portions of the Peer Review Report primarily state the peer reviewers' opinion that the Fish screening Test does not meet various validation criteria:

Pages 15-16
"Criterion 2 [for test validation]: *The relationship between the test method's endpoint(s) and the (biological) phenomenon of interest should be described."*

22. The consensus of the peer review panel was that this criterion has not been fully met."

Page 16
"Criterion 3: *A detailed protocol for the test method should be available."*

26. From this list, it can be seen that the peer review panel consider that this criterion has not been fully met."

Page 17
"Criterion 4: *The intra-, and inter-laboratory reproducibility of the test method should be Demonstrated"*

"30. The peer review panel considers that the criterion has not been fully met.

Page 18
"Criterion 8: *All data supporting the assessment of the validity of the test should be available for expert review***.**

"40. Thus, the peer review panel considers that this criterion has only been partially met, and that more effort needs to be put into producing clear, detailed protocols"

[41] EPA's EDSP December 2008 Status Filing, available online at http://thecre.com/pdf/EPA%20Wall%20letter.pdf

Page 19
"41. As a whole, the following conclusions are obtained.

Criterion 1[for test validation]: Considered fully met – but there may be a problem in acceptance of the rationale by some reviewers.

Criteria 2: There are some very fundamental suggestions/questions about the course of development of the screen (e.g. abandoning the zebrafish; re-introducing fecundity and histopathology as endpoints; reducing SSC endpoint to induction in females only). Testing of more compounds is also suggested (see also criterion 5).

Criteria 3 and 8: It seem very obvious from the number of comments that the reviewers felt starved of methodological information. All laboratories obviously need to provide far more detailed protocols.

Criterion 4: Intra-assay variability still needs to be resolved and some questions related to inter-assay variability need to be answered.

Criterion 5: More chemical testing is requested – mainly to resolve problems of interpreting concomitant changes in VTG and SSC. Taking all suggestions from all reviewers, the compounds that are recommended for inclusion are: androstenedione, octylphenol, pcp and methoxyethanol in medaka; another weak estrogen (that is not related to octylphenol), another non-aromatizable androgen, another aromatizable androgen, an AhR agonist and an ER antagonist in FHM and medaka.

Criterion 6: Better use of the existing literature for FHM.

Criterion 7: No major problem.

.

Criterion 8: See statement for Criterion 3 above."[42]

Upon information and belief, there was some OECD action on the Fish Screening test in March or April 2009. CRE and the rest of the public are not privy to this OECD action, and it is not discussed on EPA's test status site. This is still another example of the incomplete record and lack of transparency in EPA's EDSP test validation process.

[42] OECD Environment, Health and Safety Publications, Series on Testing and Assessment, No. 94, REPORT OF THE VALIDATION PEER REVIEW FOR THE 21-DAY FISH ENDOCRINE SCREENING ASSAY AND AGREEMENT OF THE WORKING GROUP OF THE NATIONAL COORDINATORS OF THE TEST GUIDELINES PROGRAMAME ON THE FOLLOW-UP OF THIS REPORT (July 25, 2008) Environment directorate http://www.olis.oecd.org/olis/2008doc.nsf/LinkTo/NT000035A6/$FILE/JT03249201.PDF

demonstrating a test method's specificity, as defined above by ICCVAM, is part of demonstrating that the test method is accurate, which is in turn essential to validating the test method. EPA admits in its response to critical peer review comments on the Male Pubertal Test that EPA has not demonstrated that the Male Pubertal Rat Test is specific:

> "The Agency agrees that specificity has not been shown yet...."[43]

Consequently, by EPA's own admission, the Male Pubertal Rat Test is not validated.

In its response to peer review comments, EPA also acknowledges that the Male Test failed inter-laboratory reproducibility:

> **Peer Reviewer Comment**
> "The extent of variability for many of the endpoints is troubling: all labs were out of compliance with pre-set performance criteria for 4 of 17 endpoints for one lab, 5 of 17 for two, and 6 of 17 for one. In other words, roughly one-fourth to one-third of the endpoints were more variable than was believed to be acceptable, a result that could compromise the resolving ability of the test guideline (as well as its reproducibility). These are issues that will need to be addressed in order for the test guideline to be used routinely to evaluate unknowns."
>
> **EPA Response**
> "The Agency agrees that the performance of the laboratories in the interlaboratory validation study was outside of the historical norms for a surprising number of endpoints."[44]

The American Chemistry Council filed comments on the Male Pubertal Rat Test which demonstrated that there are also sensitivity problems with the test:

> "There may be some question as to how sensitive the male pubertal test guideline is to weak thyroid-active compounds. During the validation work, the primary mode of action of Phenobarbital was not detected in the male pubertal test guideline. Phenobarbital,

---

[43] EPA Responses to peer review comments on Male Pubertal Rat Test guideline, page 49, *available online at* http://www.epa.gov/endo/pubs/pubertal_male_peer_review_response.pdf

[44] EPA's responses to peer review comments on MALE pubertal, page 9, *available online at* http://www.epa.gov/endo/pubs/pubertal_male_peer_review_response.pdf

which was used as a weak thyroid agent during validation of several Tier I test guidelines, did not alter thyroid weights, thyroid histopathology or serum T4 or TSH levels in the male pubertal test guideline.

Instead, Phenobarbital delayed preputial separation, and decreased reproductive and accessory sex tissue weights. Thus, Phenobarbital produced the same pattern of effects as the antiandrogens, linuron and flutamide. While the doses of Phenobarbital used in this test guideline did not achieve a maximum tolerated dose (10% change in terminal body weight), it is unlikely that Phenobarbital would have been identified as a primary thyroid toxicant when a variety of antiandrogenic effects were seen at lower dose levels."[45]

In sum, the Male Pubertal Rat Test lacks specificity, sensitivity and reproducibility. Therefore, it has not been validated. There are even more extreme problems with the Female Pubertal Rat Test.

In its response to peer review comments on the Female Pubertal Rat Test, EPA admits that this test may not be sensitive enough to detect some endocrine disrupting compounds:

Reviewer Comment
"The detection of estrogenic activity of methoxychlor at 12.5 mg/kg in the multi-dose study (Appendix 8) is cited as an example of the sensitivity of the test guideline (ISR page 47, lines 14-17), but the three laboratories in the interlaboratory comparison study did not detect activity at this dose. It is of interest to note that the study that did detect activity at 12.5 mg/kg appeared to use the diet with the highest phytoestrogen level, although it is certainly not clear what factors might have contributed to the discrepant result."

EPA Response
"The reviewer raises a good point, that the sensitivity of the female pubertal test guideline of methoxychlor was not consistent between these two studies. This may raise concerns that very weak compounds (weaker than methoxychlor, which was reliably detected at 50 mg/kg) will not be detected consistently, even when tested at high dose levels. The Agency believes that future improvements to the test guideline may increase the consistency of detection of weak compounds, but that the Screening Program should not be delayed further to re-validate an improved protocol at this time." [46]

---

[45] ACC EDSP Comments dated March 20, 2008, pages 24-25, *available online at* http://www.regulations.gov/search/search_results.jsp?No=0&sid=11A556147894&Ne=2+8+11+8053+8054+8098+8074+8066+8084+8055&Ntt=EPA-HQ-OPP-2008-0012&Ntk=All&Ntx=mode+matchall&N=0&css=0

[46] EPA responses to peer review comments on the Female pubertal, page 72, *available online at* http://www.epa.gov/endo/pubs/pubertal_female_peer_review_response.pdf

In other words, EPA hopes that by requiring companies to perform the Female Pubertal Rat Test, EPA will obtain data that will enable EPA to demonstrate sensitivity and validate the test. As of now, there is no reason to believe the test will be able to detect all the endocrine disrupting substances that it is supposed to detect.

Comments submitted by the American Chemistry Council point out other sensitivity problems with the Female Pubertal Rat Test:

> "There is some question as to the ability of the female pubertal test guideline to detect weak aromatase inhibitors, particularly those with mixed endocrine activities. The female pubertal test guideline did not detect δ-testolactone, a moderately specific aromatase inhibitor, at doses high enough to cause antiandrogenicity in the male pubertal test guideline (Marty et al., 1999). In the female pubertal test guideline ISR (pp. 23-24), minimal effects were seen with the weak aromatase inhibitor, fenarimol, at doses that produced a significant decrease in due to thyroid changes, casting further doubt on the ability of the pubertal female test guideline to detect weak aromatase inhibitors."[47]

The ACC comments also note specificity problems with the Female Pubertal test guideline. For example,

> "The specificity of the test guideline remains a question due to the effect of 2-CNB. A potential negative control, 2-CNB, delayed vaginal opening and increased TSH levels...."[48]

In sum, the Female Pubertal Rat Test lacks specificity and sensitivity. Therefore, it has not been validated.

EPA's Science Advisory Panel also criticized the Male and Female Pubertal Rat Tests. EPA's SAP was not charged with reviewing the validation status of any Tier 1 test. Nevertheless, the SAP stated in its report that the Male and Female Pubertal Rat Tests lack specificity and that this "remains an issue for the validity of these assays":

> "It was noted repeatedly, and stressed as a major issue, that a negative control substance(s) has not been identified in this group of assays. This fact stands as a major limitation to the Tier 1 battery. Lacking demonstration of expected negative results remains an issue for the validity of these assays."[49]

---

[47] ACC EDSP Comments dated March 20, 2008, pages 24-25, *available online at* http://www.regulations.gov/search/search_results.jsp?No=0&sid=11A556147894&Ne=2+8+11+8053+8054+8098+8074+8066+8084+8055&Ntt=EPA-HQ-OPP-2008-0012&Ntk=All&Ntx=mode+matchall&N=0&css=0

[48] *Id.*

[49] Page 19 of SAP Minutes available online at http://www.epa.gov/scipoly/sap/meetings/2008/march/minutes2008-03-25.pdf

EPA itself admitted to the SAP that this lack of specificity was a major failing of the Tests:

> Dr. Delicos (SAP member)
> "I just have one question. I guess about a legal definition. I may be the only
> person confused here, but representing assays as validated...and some of the public
> commentors are saying these assays were not validated...for instance,
> if you did not have a...a demonstration of a chemical which you would expect to be
> negative and it's not demonstrated to be negative in these pubertal test guidelines,
> could you go forward with that program in August as you, as you plan,
> or do you have to stop and...and do that?
> Is that a legal requirement for the validation?"

> "DR. TIMM [from EPA]
> I think it's...it's clearly necessary to show that....
> Some people would like to...and we would like to, actually...have had a clear negative.
> We...we...we didn't choose well.
> I don't think that that means there isn't one out there.
> It just means we...we didn't make a very good choice." [50]

The ACC comments point out that this lack of specificity means that every substance tested in the Tier 1 battery may test positive and fail even if the substance is in fact not an endocrine disruptor:

> "As it now stands, it appears that every test substance evaluated in the
> pubertal assays could likely yield a positive response because these test methods require
> the highest dose tested to alter body weight gain, and this degree of systemic
> toxicity is sufficient to affect the apical endpoints measured in these test guidelines. The
> pubertal assays specifically fail to meet the ICCVAM validation criteria because the
> Agency has not demonstrated specificity of these test methods. [51]

The Peer Review Report for the Female Pubertal Rat Test cannot legitimately be used to support the validation of that test.[52]

Some critical statements from the Peer ReviewReport follow:

---

[50] Transcript, pages 43-44 *available online at* http://www.epa.gov/scipoly/sap/meetings/2008/march/transcript2008-03-26.pdf

[51] ACC comments at pages 24-25

[52] Document available online at http://epa.gov/endo/pubs/pubertal_female_peer_review_response.pdf

Page 7

"If the purpose of the assay is to quantify the effects of chemicals on pubertal development and thyroid function, then the procedures should optimize the chance of success and minimize confounds that would obscure the results. This reviewer sees a number of serious problems with the protocol that present confounds."

Pages 7-8

"In this reviewer's opinion, this assay will add very little to a battery."

Pages 8-9

"Although not the task I was assigned, I will make an unsolicited statement. This protocol is very disappointing from an endocrinological point of view, and although it addresses endocrinological questions, it appears to have been developed primarily by toxicologists without sufficient input from experts in endocrinology. My opinion is that the protocol could have benefited from the inclusion of at least reproductive and developmental endocrinologists in its development. My personal assessment is that, in its current form, it will provide scant information relative to the amount of work that will go into the experiments. As I have indicated, much of the work would not be publishable in a reputable endocrine journal. While probably not my place, I recommend that a group of scientists with diverse expertise (from toxicology to reproductive and thyroid physiology and endocrinology) be convened in the style of a scientific network to discuss this protocol from a wide range of perspectives. To do it serially, as is being done, slows down the process"

Pages 10-11

"I do not agree with the conclusion that 'The current study demonstrates that the female pubertal protocol is transferable and reproducible in contract laboratories.' As indicated above, day of vaginal opening was reasonably transferable, but many of the other □arameters were not. Not being a toxicologist, I do not know what level of replication from lab-to-lab is expected. From an endocrinological point of view, a well-controlled study should be entirely (or at least nearly entirely) repeatable from lab-to-lab."

Page 27

"Positive controls compounds. A positive control with results that are known with certainty should be used. This is essential to demonstrate that the laboratory has the expertise and laboratory conditions sufficient to support replicating a previous result. Although a high dose can be used as a secondary control, a low dose, positive control to demonstrate reliability of the laboratory should be included in the protocol"

Page 28

> "The lack of a positive control is a serious concern. Within the Integrated Summary Report, this omission is justified by the argument that it is highly unlikely that a single compound that will generate a positive result for all endpoints in the assay. This problem results from the inclusion of experimental endpoints designed to address two different and largely unrelated questions. By lumping pubertal endpoints, which assess estrogen action, together with thyroid endpoints, the choice of an appropriate positive control becomes complicated. It is readily apparent that the most salient and critical goal of this assay is to identify compounds that affect puberty. As such, a positive control that reliably and consistently advances puberty should be included, regardless of whether or not any thyroid endpoints are altered. Estradiol, DES, or estradiol benzoate would all be appropriate positive controls and at least one should be used by all labs for this purpose. Any labs not observing an effect with the positive control would then immediately know that they have a problem executing the assay properly."

Page 43

> "Given that no compound to date has tested negative in this assay it is difficult to evaluate the potential effects of test substance on outcome"

Page 46

> "A major current limitation, as pointed out by the EPA, is the lack of demonstrated specificity of the assay."

Page 46-47

> "The specificity of the test guideline remains a question due to the effect of 2-CNB. A potential negative control, 2-CNB, delayed vaginal opening and increased TSH levels. This issue is discussed in more detail below, but may be the result of increased metabolism of estradiol and/or thyroid hormones by the 2-CNB exposed liver."

Page 48

> "A number of elements within the protocol diminish the functional utility of the assay guideline. Ovarian and uterine weights are generally uninformative and complicated by cycle. Inclusion of the thyroid endpoints precludes the use of a needed positive control group for the pubertal measures. The duration of estrus monitoring is too short and the use of daily lavage will likely induce pseudopregnancy in some animals, potentially confounding the data. Failure to eliminate phytoestrogens introduces an unnecessary confound and increases the risk of inter-laboratory variability. Finally, only two doses are to be used, both of which are based on body weight and neither of which will approximate a "typical" human or wildlife exposure. The failure to include a dose within a reasonably physiological range is a considerable concern."

Page 55

"In discussion of specificity, the ISR mentions (page 82) that "a good faith effort was made to identify a chemical that was both toxic to other systems but without endocrine effects." It is not surprising that one could not be found, because a toxic compound will decrease body weight, and this seems to be required to demonstrate that the dose has exceeded the MTD. However, body weight loss due to toxicity would likely be accompanied by a decrease in nutrient intake. From a physiological point of view, food deprivation causes reproductive dysfunction. Unfortunately, approaching a problem like this from a toxicological view-point with little regard to the underlying endocrinology/physiology has problems. In short, any compound that compromises nutrition would be expected to have endocrine effects."

Page 59

"Finally, the use of high doses may explain why, to date, no compound has produced a negative result in this assay. The highest dose to be used is defined as a statistically significant reduction in body weight with "no clinical signs of toxicity." The acceptable 'signs of toxicity' are not identified or discussed in the protocol but should be, perhaps in an appendix. In general, the use of body weight to define dose is problematic for several reasons, most of which have already been addressed previously by Goldman et al, but again highlights the need for a positive control group within this assay. It is well established that estradiol administration significantly reduces body weight. Because a decrease in body weight of 10% or more can result in the disinclusion of subjects or treatment groups, the employment of a positive control group would help clarify whether or not the MTD was reached or exceeded, and whether or not the laboratory was conducting the test guideline properly."

OMB SHOULD NOT APPROVE THE AR BINDING TEST BECAUSE PEER REVIEW CONCLUDED THAT THIS TEST HAS REPRODUCIBILITY PROBLEMS AND IS OF QUESTIONABLE VALUE "AS A SCREENING TOOL"

EPA states that for the AR Binding Test, validation was complete in December 2007, and that "[t]his assay is ready for use."[53]

The Peer Review Report does not support these EPA statements.

A peer reviewer states in the Report that this test has reproducibility problems and is of questionable value "as a screening tool":

> "The intrinsic limitation of reproducibility of the assay in some labs is found throughout the study; it would appear that in experienced hands the assay works very well and is highly reproducible. Laboratories did not conduct the test guideline with similar precision using the same cytosol and chemicals. Reproducibility and quality of the data are problems related to solubility of chemicals and chemicals that bind weakly similar precision using the same cytosol and chemicals. Reproducibility and quality of the data are problems related to solubility of chemicals and chemicals that bind weakly to the androgen receptor; this may cause issues in the reporting of results between different laboratories. The variation between labs for low affinity binders puts into question the long-term value of this assay as a screening tool"[54]

[53] EPA's EDSP December 2008 Status Filing, available online at http://thecre.com/pdf/EPA%20Wall%20letter.pdf.

[54] Peer Review Panel Comments on the AR Binding Test guideline, available online at **http://epa.gov/endo/pubs/ar_binding_peer_review_response.pdf**

OMB SHOULD NOT APPROVE THE AMPHIBIAN METAMORPHOSIS TEST BECAUSE PEER REVIEW SAID THIS TEST ISN'T REPRODUCIBLE AND BECAUSE THE OECD PROCESS IS NOT TRANSPARENT

The AM test is an excellent example of EPA's general attitude toward peer review of tests in the Tier 1 Battery: pronounce a test validated once it has gone through peer review regardless of whether the peer review is negative, and then launder problems through the murky OECD process.

Peer review said the AM test is not reproducible. EPA inexplicably said that peer review said this test is reproducible.

CRE raised the peer review criticisms of the AM test in a Request for Correction under the Information Quality Act. EPA refused to acknowledge that CRE had filed an IQA petition, but EPA finally responded to CRE's petition anyway.

EPA's response essentially claimed that the reproducibility problems noted by the peer reviewers were solved in subsequent versions of the AM developed by OECD.

EPA's response did not provide any link to the supposedly corrected AM test. Nor does EPA's response state whether the OECD has actually adopted the AM test.

Consequently, CRE does not have the available information to comment on EPA's response to CRE's IQA request for correction of the AM test.

Based on the AM's Peer Review Report and the other information that is available, this test is not validated and is not ready for use.

For example, with regard to whether the AM is reproducible, one peer reviewer stated in the Report : "This is a major flaw of the material provided ...."

He further stated "that the conclusions regarding inter-laboratory variability are not warranted and that it [the AMA test] fails as a method for accomplishing the stated goal of the test guideline to be part of the Endocrine Disruptor Screening program (EDSP)."[55]

He advised EPA that "[b]efore the AMA can be used as a screening tool that is open to contract laboratories, the issues raised above should be addressed. The bottom line is that the AMA is not suitable as a screening tool for endocrine disrupting compounds." [56]

---

[55] Page 5 of AMA Peer Review Report, available online at http://www.epa.gov/endo/pubs/AMAa_peer_review_response_final.pdf

[56] Pages 34-35 of *id.*

A second peer reviewer concluded in the Report:

> "One of the major concerns about the test guideline is the degree of inter-laboratory consistency.... while overall trends are observed (ie T4 accelerates, perchlorate and IOP delay), there is surprising inconsistency among the laboratories....Based on these observations, the consistency of findings across laboratories remains a major concern for the future viability of the test guideline system."[57]

A third peer reviewer was more positive, but even she concluded that "there was some variation and testing may need to be conducted independently in at least two separate labs."[58]

A fourth peer reviewer concluded that "Concerning was that not all aspects were always controlled for. Moreover, when conducting the inter-laboratory study using weak thyroid modulators, it seems that the consistency was lost."[59]

This peer reviewer also commented on the inconsistency of test result interpretation among laboratories performing the AMA:

> "A much stronger guideline for data interpretation within the AMA Test Method Documents is necessary.... *In summary, this phase trial demonstrated that data interpretation across the validation studies needs to be consistent, and guidelines need to be carefully developed to facilitate this interpretation.* In fact, in the AMA Test Method, there is no section on data interpretation, and in the overall ISR [Integrated Summary Report] , there are no clear guidelines for how many parameters need to be significantly different from controls before a compound is to be interpreted as thyroid disrupting. Such guidelines are essential and should be provided clearly in the final AMA Test Method Protocol, along with appropriate summary tables."[60]

The fifth and final peer reviewer concluded:

> "One of my greatest concerns in the AMA documentation is the high variance in reproducibility of the results obtained from the various labs during the various test phases. I am disquieted by the little attention given to the variance between the labs, when their protocols were (supposedly) identical. Most of the chemicals used in these studies were well known inhibitors or accelerators of metamorphosis. The fact that inhibition and acceleration were seen in the test results is, of course, exactly what one expected. I did not expect, however, the variance in the reports between the different labs. It is bothersome that more effort was not made to explain the inter-laboratory variance."[61]

---

[57] Page 5 of *id*.
[58] Page 5 of *id.*
[59] Page 7 of *id*.
[60] Page 12 of *id*.
[61] Page 7 of *id.*

This reviewer also explained:

> "My greatest concerns about the AMA center on the document "Draft Method for the AMA." Various laboratories should be able to follow the methodology of this essential document and achieve identical results. **There is simply not enough detail in this methodology to be confident that the test guidelines can be executed with adequate amounts of reproducibility."**[62]

These are just some examples of the Peer Reviewer's concern with the reproducibility of the AM., test, and with the reproducibility of AM test result interpretation. There are many other examples in the Peer Review Report.[63]

As one peer reviewer noted:

> "This section [of EPA's AM validation study under review] proclaims "The reproducibility of the MA, for screening purposes, has been well-demonstrated using several representative thyroid-active chemicals across geographically diverse laboratories." However, if the variation between the labs cannot be explained, then one cannot feel as confident about this proclamation as the author of the review."[64]

---

[62] Page 4 of *id.*

[63] 18 *E.g.*, *Id*. at pages 2-8 to 2-11, 2-14 to 2-15, 2-21 to 2-24, 2-25 to 2-26, 2-27, 2-67 to 2-70, 3-1, 3-7, 3-8, 3-17, 3-25, 3-26, 3-27, 3-31, 3-44, 3-56, to 3-58, 3-59, 3-66, 3-67, 3-69, 3-70, 3-72, 3-80

[64] Id. at 2-27.

EPA'S EDSP BURDEN ESTIMATES HIGHLIGHT THE PROGRAM'S COST TO THE FEDERAL GOVERNMENT

There are three economic conclusions and one policy recommendation to be drawn from EPA's EDSP Information Collection Request. The conclusions are:

- EPA has not included relevant cost components in their burden estimates;
- The Agency's burden estimation scheme places the "cart-before-the-horse" by using the Tier 1 test order itself as a *de facto* pilot project to develop the information that should have been included for the ICR; and
- EPA's burden in administering the Tier 1 test order and evaluating the data received would require substantial agency resources.

The policy recommendation drawn from the above conclusions is that the agency is not yet ready to proceed with the Tier 1 test program.

A. *Need for EPA to Include All Relevant Costs in Burden Estimates*

Aside from other burden estimation concerns, as discussed below, EPA has not included the burden associated with two specific cost components:

- Pointless use of all Good Laboratory Practices (GLPs); and
- Costs associated with ensuring the accuracy and completeness of information submitted to EPA.

1. Use of GLP in Testing But Not Test Validation

In their response to CRE's Information Quality Act (IQA) Request for Correction, EPA notes, based on OECD guidance, that while "it would have been ideal if the studies used in the validation exercise for the AMA had been performed in compliance with all components of GLP....compliance with all components of GLP was not essential to the AMA validation exercise and all the laboratories that participated in the studies complied with the components of GLP that the OECD Guidance Document sets forth as essential." Thus, EPA asserted that their compliance with only select components of GLP was adequate for validation even though their own validation standards, as summarized in their AMA Integrated Summary Report, states that the data should be obtained in compliance with GLPs.

In their EDSP Policies and Procedures Federal Register notice, EPA stated that "order recipients generating data must adhere to the good laboratory practice (GLP) standards described in 40 CFR part 160 when conducting studies in response to a FFDCA section 408(p) test order." The

ICR Supporting Statement notes that any deviation from GLPs must be described "in detail." Thus, EPA is requiring that the test order recipients "must" adhere to all GLP components contained in 40 CFR part 160 even though the agency determined that use of only select GLP components was sufficient for test validation.

EPA's requiring that companies use higher standards for carrying out tests than the Agency used for validating some of the tests is analogous to requiring use of a high resolution camera to take pictures through a smudged window. From a PRA standpoint, requiring test order recipients to use all GLP components rather than just the subset used in the validation process needlessly increases burden and is inconsistent with the requirement that EPA reduce "to the extent practicable and appropriate the burden on persons who shall provide information to or for the agency."

Requiring use of quality control procedures (GLP components) for tests that were not used by the agency and their contractors in validating those same tests is not appropriate. For each Tier 1 test, test order recipients should not be required to adhere to any GLP component not used in the validation of that same test. It should be noted that CRE does not object, per se, to requiring use of GLP, only to requiring the private sector to use those GLP components that the government did not adhere to in developing and validating the tests.

An additional concern with respect to EPA's use of only select GLP components concerns their methodology for estimating testing cost. As EPA explains in ICR Attachment F, "The estimated cost of the fish screen and frog assay are based on professional judgement and expertise considering the prices EPA paid to have these tests conducted, and adjusted to reflect the expected protocol details that impact cost." [Emphasis added] Since EPA did not pay to have all GLP components employed, their costs, and thus their burden estimate, are lower than the costs of test order recipients who need to adhere to all GLP components.

There is no indication that adjustment for "protocol details" that impact cost mentioned by EPA refers to the comprehensive use of GLP. If the costs have been adjusted to reflect use of all GLP components by test order recipients, then EPA should specifically state how much higher testing costs will be as a result of their decision to require private sector use of GLP components that the Agency deemed unnecessary for themselves. Thus, either EPA has not calculated the higher costs that test order recipients will incur in adhering to all GLPs, or EPA should detail and provide for public comment the higher costs.

2. Cost of Certifying Data Accuracy and Completeness

In their ICR Supporting Statement, EPA relates that their estimate of the Tier 1 paperwork burden is "percent-based estimate of paperwork associated with conducting a test was initially established in consultation with OMB in the 1980's in an effort to provide a reasonable estimate of the burden associated with the paperwork component of data generation, which may vary based on the complexity of the test performed."

EPA's methodology for determining paperwork burden is, therefore, based on a methodology that existed prior to the 1995 revisions to the Act. EPA's rule-of-thumb estimation technique falls far short of the PRA's requirement for the agency to provide "a specific, objectively supported estimate of burden."

In EPA's "Response to Comments on the Public Review Draft" of the EDSP ICR, the agency cites OMB's PRA implementing guidance – an excellent source for information on PRA compliance. This 1999 document discusses the extensive efforts the public needs to invest in providing accurate, reliable data in response to federal information collections. In language essentially unchanged from the 1997 preliminary draft implementation guide, OMB explains:

> "*All time, effort and other resources which need to be expended to certify the accuracy and/or reliability of information developed, submitted, disclosed, disseminated or retained, or to certify compliance with any statutory or regulatory provision, represent paperwork burden. Such certifications generally require intensive scrutiny, whether by an individual (such as with respect to a tax filing), senior officers or managers (where the respondent is a firm), or senior elected or appointed officials (where the respondent is a government or agency), and cannot legally or practically be delegated. This burden is rarely, if ever, small or trivial and generally entails a comprehensive audit by the certifier of all components of the information or declaration which need to be certified. Audits entail additional indirect burdens on subordinates, partners, associates, consultants, counsel and other experts necessary to fully replicate and document the process used to derive the response, as well as the response itself, in a manner that the certifier can understand and credibly affirm to be accurate and valid irrespective of the degree of technical detail involved. Such certification burden should be evaluated within the context of the legal consequences to the respondent of an improper or false certification.*" [Emphasis added]

EPA's estimate of 36 managerial hours per chemical to "compile and review the final data for submission" does not include "intensive scrutiny" needed by companies to ensure that the information provided under the EDSP test order is "accurate and valid irrespective of the degree of technical detail involved."

It is important to note that company and consortia officials are required to formally certify, under penalty of law, that the information contained in their Initial Response Forms (Attachments D-1 and D-2) is true, accurate and complete. As EPA's Initial Response Form for Individual Order Recipients states,

"*3.1. Certification. I certify that the statements made on this form and all attachments are true, accurate and complete. I acknowledge that any knowingly false or misleading statement may be punishable by fine or imprisonment or both under applicable law.*"

It is important to note that the personal penalties for any false certification are in addition to the substantial economic sanctions companies face for not complying with the test order, including possible de-registration of their products.

Since EPA is requiring company and consortium officials to be personally responsible for the accuracy of information submitted just in the Initial Response Forms, it seems likely that the agency will require a similar level of responsibility from the companies/consortia when it comes to submitting the actual test results. Therefore, the agency needs to include in its burden estimate the time it will take company/consortium officials to understand the technical details of each assay and ensure that the work done by either the company or contract laboratory is complete and accurate.

B. *EPA's Cart-Before-the-Horse Cost and Burden Estimation Methodology*

With respect to concerns that the agency has not accurately estimated the costs of conducting the tests, EPA's "Response to Comments on the Public Review Draft," states on multiple occasions that their cost will revised after the tests are conducted.

For example, in a response to Commenter #12, EPA states "Once these tests are available on the market, these costs will be adjusted as appropriate." In a response to Commenter #14, EPA states "Once these tests are available on the market, the Agency's estimated costs will be adjusted as appropriate." Similarly, in response to Commenter #15, EPA states "Although not possible for this ICR, the Agency does intend to consult with the recipients of these first Tier 1 Orders about their experiences, costs and burdens. This consultation will be used to revise the estimates presented in this ICR for the ICR renewal or future ICRs related to the EDSP."

In short, the agency is stating that they will develop a reasonable estimate of the costs, after the tests are conducted. The PRA, however, requires that the agency develop "a specific, objectively supported estimate of burden" **before** the information collection is reviewed by OMB.

It is important to note that EPA takes a similar retrospective approach to meeting their small business protection obligations under the PRA. In response to concerns that EPA has not adequately characterized the burden on small businesses and ensure that they are not unduly affected, the agency states that "EPA will attempt to identify which Tier 1 Order recipients might qualify as a small business so as to consult specifically with them about potentially disproportionate burdens that they experienced. Upon renewal of this ICR or for the subsequent ICR, EPA will use revise the ICR to reflect this consultation."

EPA has it backwards. The agency is required by the Paperwork Reduction Act of 1995 and the PRA amendments contained in the Small Business Paperwork Relief Act of 2002 to determine and mitigate to the extent practical the burden on small business before not after the information collection is conducted.

EPA's admission that they have not yet determined which Tier 1 Order recipients are small businesses means that they could not have possibly complied with the PRA's requirement that they have reduced "to the extent practicable and appropriate the burden on persons who shall provide information to or for the agency, including with respect to small entities" and also that they have made "efforts to further reduce the information collection burden for small business concerns with fewer than 25 employees."

How can EPA have made further efforts to reduce the burden on entities with fewer than 25 employees when they do not even know if any such companies will be served with test orders?

The agency's ramshackle, retrospective approach to estimating burdens on companies large and small demonstrates that they are not yet ready to submit a Tier 1 EDSP ICR to OMB.

C. *54,236.5 Hours: EPA's Estimate of Agency Resources to Conduct the ICR*

EPA's ICR Supporting Statement states that it will take the agency more than 54,000 hours, over 29 person-years, to issue and administer the test orders and process the data received. This burden on the agency is in addition to the extensive resources involved in developing and validating the Tier 1 tests, and in developing and eventually validating the Tier 2 tests.

CRE is not questioning the agency's internal burden estimate. To the contrary, CRE commends EPA for being forthright about the extensive level of agency resources that would be required to conduct the Tier 1 information collection. CRE does, however, seriously question whether the extensive internal burden of administering the Tier 1 assay as currently construed is a worthwhile use of the agency's highly skilled personnel. Put simply, CRE believes that over 54,000 hours of agency staff time could be put to better uses than administering an initial screening battery of tests that, as we have explained in detail, have not been properly evaluated and would produce information of, at best, ambiguous quality.

CRE strongly recommends that OMB evaluate the burden on the resource-constrained agency of the proposed Tier 1 battery in light of EPA's more pressing and important responsibilities to protect the environment.

REQUESTED ACTIONS

For the reasons stated above, OMB should deny EPA's ICR for Tier 1 of the EDSP.

Several actions should occur before OMB approves any subsequent ICR. These actions should include SAP review of two questions:

1) Whether EPA has validated each of the Tier 1 tests; and

2) Whether all of the proposed Tier 1 tests are necessary in order to provide a useful Tier 1 screening battery for endocrine effects.

These actions should also include full EPA disclosure and publication of the record for each of the tests in the EDSP Tier 1 Battery.

We once again thank OMB for the opportunity to submit these comments, and we request a meeting with OMB on them before OMB makes a determination on this ICR.

Scott Slaughter
Center for Regulatory Effectiveness
1601 Connecticut Avenue, NW
Suite 500
Washington, D.C. 20009
202/265-2383
Slaughter@MBSDC.com

APPENDIX
PEER REVIEW CRITICISMS OF THE H295R TEST
(from *Disposition to Peer Review Panel Comments on the H295R Steroidogenesis Assay*, finally available online at http://www.epa.gov/endo/pubs/assayvalidation/h295r_pr.htm)

Comment
"It is important to stress that chemicals that generate a negative result in the H295R steroidogenic test guideline could be false negatives and they should not be considered safe without a complete evaluation of them with the other Tier 1 battery test guidelines. This in vitro system lacks that ability to study complex interactions that could occur in vivo such as metabolism of tested compounds, biodistribution, interaction with other endocrine systems that may modulate sex hormones steroidogenesis, etc."

Comment
"One major question is whether any small change in hormone production in an isolated in vitro system has any relevance for the health outcome of an exposed organism. This remains unaddressed in the documents available for review."

Comment
"As the steroidogenesis assay only looks at one final outcome, namely the amount of estradiol and testosterone secreted, it is not possible to make biologically meaningful statements on the relevance of any observed disruption for the organisms as a whole. There are so many factors not directly related to steroidogenesis that could influence the assay system as it is currently described and intended to be used, that the issue of 'false positives' is likely to be an important concern, particularly once dealing with unknown complex environmental samples."

Comment
"The implications of the presence of these other pathways (aldosterone, cortisol synthesis) may be far reaching for the reliable application of the proposed H295R steroidogenesis assay, as all these pathways are interconnected (at least in adrenocortical cells, not necessarily in gonadal cells). There is no critical discussion of the potential drawbacks of choosing an adrenocortical cell line to study effects of chemicals on gonadal testosterone and estradiol production. There is no scientifically supported discussion of the possible differences in regulation of steroidogenesis in adrenocortical cells and gonadal cells, yet it is known these are qualitatively and quantitatively very different. Several of the above points have been discussed in detail in several publications from my own lab in recent years (Sanderson and van den Berg, 2003; Sanderson, 2006)."

Comment
"The reproducibility of the test system appears to be relatively poor. This may be partly due to the variability inherent in the use of cell lines in culture, but is also likely to be due to the various immunoassay-based hormone analysis methods used. The latter influence may be reduced by selecting a single method of detection, preferably not immunoassay guideline based."

Comment
"For the most part the assay is sufficiently repeatable and reproducible. However, I am concerned with the high CV among laboratories and also within laboratories. The within lab CV is particularly high for prochloraz and this could be because it is inhibiting the basal steroid production. As the constitutive levels are being inhibited this may lead to error as the levels may differ due to autoregulation that is inherent in this system. I would recommend using a test group where the inhibition is tested using acute-stimulated (forskolin or 8bromocamp) steroid production as a model. This might reduce the variability and make the data set more comparable between the laboratories. For instance there is a large variability in EC50 for foskolin between the different labs (Table 10.3)."

Comment
"The way the H295R cell system is being proposed to be used is like a black box. It will be difficult to interpret the meaning of any outcomes that may be observed on testosterone and estradiol levels, and this is further compounded by the drawbacks of using immunoassay-based detection methods. A more focused definition of the purpose of a tier 1 assay for steroidogenesis would be recommendable; allowing for the development of a H295R cell-based steroidogenesis test guideline that would provide less ambiguous information about the steroidogenesis disruption potential of chemicals or unknown environmental extracts."

Comment
"The huge CV reported for between laboratory comparisons may have to do with the difference in basal hormone production and associated differences in the magnitude of response to know inducers and inhibitors as well as test substances."

Comment
"The implications of the presence of these other pathways (aldosterone, cortisol synthesis) may be far reaching for the reliable application of the proposed H295R steroidogenesis assay guideline, as all these pathways are interconnected (at least in adrenocortical cells, not necessarily in gonadal cells). There is no critical discussion of the potential drawbacks of choosing an adrenocortical cell line to study effects of chemicals on gonadal testosterone and estradiol production. There is no scientifically supported discussion of the possible differences in regulation steroidogenesis in adrenocortical cells and gonadal cells, yet it is known these are qualitatively and quantitatively very different. Several of the above points have been discussed in detail in several publications from my own lab in recent years (Sanderson and van den Berg, 2003; Sanderson, 2006)."

Comment
"Little is known about the impact of most of the test substances on steroid production. The lack of response to a known inducer of sex steroid production in gonadal tissue, for instance human chorionic gonadotropin (hcG), suggests that this system has limitations because of the type of tissue involved (adrenal carcinoma)."

Comment
"One of the most important aspects of the H295R steroidogenesis assay, the analysis of testosterone and estradiol, is poorly defined in the provided documents. The choice of analysis

method is left entirely to the implementing laboratory. It is known that ELISAs and RIAs can have very different outcomes dependent on the sample dilution, kit and antibodies used, not to mention the numerous confounding factors (solvent, cross-reactive components). The issues of cross-reactivity, how to deal with conjugated metabolites, and how to reliably compare between hormone levels determined by RIA or LC-MS are left undiscussed. It is highly inconsistent that there is an elaborate protocol for the 'consistent' use of a standard method such as the LIVE/DEAD cytotoxicity kit while no detailed attention is given to the crucial hormone analysis methodology."

Comment
"For instance this assay will only detect changes that happens post-receptor activation. This is a drawback to this cell system because in vivo the steroidogenic cells secrete steroids in response to trophic hormone stimulation. This assay completely bypasses the receptor signaling which is an essential step in steroid biosynthesis. So substances that can affect steroid production by altering trophic hormone signaling will not be evaluated by this cell system. Also, the high constitutive production of the hormone is abnormal in vivo as this usually happens only in response to trophic hormone stimulation. So it is unknown whether the changes seen with the test substances can be mimicked in vivo to the same extent (or may be even greater) and will require confirmation with animal models or other relevant cell or tissue systems.

Also, the high constitutive levels of steroids, for instance testosterone, may deplete the precursor available for steroid synthesis and may be limiting the steroid biosynthetic capacity in response to test (inducer) substances. The changes in the magnitude of steroid synthesis with forskolin, smaller change for testosterone because basal secretion is high and higher for E2 because of lower basal secretion, clearly support this contention. This requires testing perhaps by supplementing the medium with cholesterol"

Comment
"Analysis of sex hormones. The greatest weakness in the protocols is the lack of detail on sex hormone analysis methodology. This reviewer is of the opinion that LC-MS would be, by far, the preferred analysis tool for the detection of testosterone and estradiol. LC-MS would avoid the problems that will be (and already have been) encountered with inappropriate cross-reactivity of test samples/chemicals with the antibodies used in sex steroid ELISAs and RIAs. Please see also comments on trenbolone under point 7. The validation of a sensitive LC-MS method should be a logical part of the H295R steroidogenesis test guideline as currently defined. Furthermore, a single LC-MS analysis could detect a number of steroids in addition to estradiol and testosterone at little additional effort/expense, thus improving the 'expandability' of the H295R tool for other hormone endpoints"

Comment
"There is a brief discussion of strengths and weaknesses, but lacks detail and supporting scientific references. The main strength mentioned in the interim report is that the H295R cell line is a pluripotent cell lines that expresses all the enzymes necessary for the production of testosterone and estradiol. However, the fact that numerous other steroid hormone synthesis pathways are also present, although acknowledged, is not discussed."