Comments by the Center for Regulatory Effectiveness ("CRE") to
Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel ("SAP") Considering and Reviewing the
Endocrine Disruptor Screening Program ("EDSP")
Tier 1 Screening Assays and Battery Performance,
Filed May 21, 2013, in Docket EPA-HQ-OPP-2013-0075,
www.regulations.gov

I. Executive Summary

The Amphibian Metamorphosis Assay ("AMA") failed validation peer review and should never have been used in the EDSP Tier 1 test battery.

EPA now admits that there are serious problems with the AMA based on the EDSP Tier 1 Test data the Agency has reviewed.

This SAP should advise EPA not to use the AMA test data for any purpose because there are too many unexplained problems in the assay resulting in the generation of flawed data.

This SAP should advise EPA not to use the AMA as part of the EDSP Tier 1 test battery.

There has never been SAP review of EPA's conclusion that **any** of the EDSP Tier 1 tests are validated, and many of the tests failed validation peer review.

This all seems to be a waste of time and money because EPA plans to abandon the current EDSP Tier 1 tests. EPA told Congress:

> "In FY2012 EPA will begin a multi-year transition from the [EDSP] to validate and more efficiently use computational toxicology methods and high-throughput assays that will allow the Agency to more quickly and cost-effectively assess potential chemical toxicity."[1]

[1] See, *e.g.*, U.S. Environmental Protection Agency Endocrine Disruptor Screening Program Comprehensive Management Plan Comprehensive Management Plan (June 2012), page 6, http://www.epa.gov/endo/pubs/EDSP-comprehensive-management-plan.pdf ; and Endocrine Disruptor Screening Program for the 21st Century: (EDSP21 Work Plan), page 2, available online at http://www.epa.gov/endo/pubs/edsp21_work_plan_summary%20_overview_final.pdf .

EPA should concentrate on validating other EDSP screening methods such as computational toxicology ("CompTox"), and move away from the current EDSP Tier 1 tests.

II. CRE's Data Quality Act Request for Correction was Accurate: The Amphibian Metamorphosis Test Should Not be Used in the EDSP

A) The AMA Flunked Validation Peer Review

In 2008, CRE filed a Request for Correction ("RFC") under the Information Quality Act ("IQA"). This RFC asked EPA to correct the Agency's publicly disseminated statements that the AMA had passed peer review and was properly validated. CRE's RFC pointed out that the AMA had received a negative peer review report and asked EPA not to use the AMA as part of the EDSP Tier 1 test battery.[2]
EPA used the AMA in the EDSP anyway.[3]

CRE's RFC pointed out that the AMA peer review report included the following criticisms of the test.

With regard to whether the AMA is reproducible, one peer reviewer stated in the Report: "This is a major flaw of the material provided ...."[4]

He also explained that the inter-laboratory inconsistencies obvious in just one table of the AMA validation study "would convince any reviewer for a reputable scientific journal to recommend rejection" of the validation study. [5]

He further stated "that the conclusions regarding inter-laboratory variability are not warranted and that it [the AMA test protocol] fails as a method for accomplishing the stated goal of the assay to be part of the Endocrine Disruptor Screening program (EDSP)."[6]

---

[2] CRE's AMA RFC is available online at http://www.epa.gov/QUALITY/informationguidelines/documents/08004.pdf
[3] EPA's response to CRE's RFC is available online at http://www.epa.gov/QUALITY/informationguidelines/documents/08004-response.pdf
[4]EPA's Response to the Peer Review Results for the Amphibian Metamorphosis Assay (hereinafter "EPA Response"), page 5, available online at http://www.epa.gov/endo/pubs/ama_peer_review_response_final.pdf .The Peer Reviewers' Report is entitled *Peer Review Results for the Amphibian Metamorphosis Assay* ("Peer Review Report"). It is available online at http://www.epa.gov/endo/pubs/ama_peer_review_121907.pdf ..
[5] Peer Review Report, page 3-7.
[6] Peer Review Report, page 2-1.

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." [7]

A second peer reviewer concluded in the Report:

> "One of the major concerns about the assay 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 assay system."[8]

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."[9]

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."[10]

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. [11]

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

[7] Peer Review Report, page 3-17
[8] EPA Response, page 5.
[9] EPA Response, page *5.*
[10] EPA Response, page 7.
[11] Peer Review Report, page 3-31.

> 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."[12]

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 assays can be executed with adequate amounts of reproducibility."**[13]

Another peer reviewer noted:

> "This section [of EPA's AMA validation study under review] proclaims 'The reproducibility of the [A]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."[14]

There are many other peer review criticisms of the AMA.[15]

CRE's RFC correctly pointed out that EPA's public statements that AMA reproducibility has been demonstrated through peer review--and that the AMA is reproducible and validated—are inaccurate and misleading. Consequently, these statements violate EPA's IQA Guidelines because the IQA's Objectivity Standard requires that EPA ensure that information the Agency disseminates is reproducible and "accurate, reliable, and unbiased."[16]

Because they are inaccurate and reliable, these EPA statements also violate the IQA's utility requirement. Inaccurate, unreliable statements are not useful.

---

[12] EPA Response, page 7.
[13] Peer Review Report, page 2-33 (emphasis in the original).
[14] Peer Review Report, page 2-27.
[15] *E.g.*, Peer Review Report, 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.
[16] EPA IQA Guidelines, pages 15 and 22, available online at http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf .

B) EPA Now Admits that there are Serious Problems with the AMA Based on the EDSP Tier 1 Test Data the Agency has Reviewed

The obvious problems with the AMA test data include unexplained bent tails in the tadpoles. EPA has told this SAP that

> "Based on the initial data submissions, musculo-skeletal deformities (*e.g.*, spinal curvature or 'bent tail,' as it is typically confined to the tail region) have been reported in fifteen of the eighteen studies reviewed by the Agency, but the incidence varies widely across studies. The remaining three studies did not indicate whether spinal curvature was observed. The deformities have been reported both in studies which observe the guideline-recommended feeding regime and in studies which deviate from the guideline by reducing the amount of feed by approximately one-half. Therefore, a hypothesis of excessive feed as the primary factor in the development of spinal curvature does not appear to be supported in these cases. However, these observations are anecdotal; the Agency is unaware of any intra-lab comparisons from private laboratory facilities regarding feeding rates and musculo-skeletal deformities within a single population of *X. laevis*. There is no clear pattern in the metadata for the initial data submissions that reveals a single, predictive factor for the prevalence of spinal curvature (Table 32)."
>
> ***
>
> "With the exception of tadpole mortality in one replicate, all validity elements and performance criteria were met. Similar and apparently inverse concentration-related trends in the occurrence of spinal curvature were observed in some other studies; however, the significance and reproducibility of these findings is unclear."[17]

EPA has asked this SAP whether the AMA Tier 1 test data should be used given these mysterious problems with the test results:

> "**Charge Question 5.** Spinal curvature, usually manifesting as "bent tail" in *X. laevis* tadpoles, was reported in 15 of 18 AMA studies reviewed thus far. The anomaly appears to be first observed several days after study initiation, and prevalence increases with time. Overall, the prevalence of spinal curvature in these studies ranged from "a few per replicate" to 92% of a given treatment group by test termination. Experimental work by the EPA Office of Research and Development suggests that overfeeding can be a primary cause of spinal curvature in their *Xenopus* test populations; however, spinal curvature remained prevalent (range: 16-92%) in the five industry AMA studies in which feed was reduced by 50% compared to guideline recommendations. Overall, the incidence of spinal curvature appears to be highly variable. From a qualitative review of the data,

[17] Pages 118-119, http://www.regulations.gov/#!searchResults;rpp=25;po=0;s=%252C%252BEPA%252BID%252BEPA-HQ-OPP-2013-0075-0003;fp=true;ns=true (White Paper).

there appear to be no consistent differences in the incidence or variability of spinal curvature when studies using guideline versus reduced feeding regimes are compared. Please comment on whether the presence or prevalence of spinal curvature in test specimens, including controls, compromises the utility or validity of an AMA submission. If so, when does the prevalence of spinal curvature render the study unreliable? What technical guidance may be useful for laboratories in reducing the occurrence of spinal curvature and determining if, or at what point within the study, a study may be compromised because of this phenomenon?"[18]

Bent tails aren't the only problem with the AMA test results. EPA's Charge Question 6 to this SAP explains:

"a. In one study, the pathologist's report identified a lower incidence and severity of follicular cell hypertrophy when compared to the incidence and severity of this trait in control specimens. Similar trends have been observed in other studies. In this case, the pathologist concluded that the finding was potentially consistent with treatment-related delay of metamorphosis because thyroid follicular cells normally increase in height during tadpole development. Please comment on the validity of this conclusion.

b. What guidance may be given to better distinguish between histological changes in the thyroid associated with the normal progression of metamorphosis and treatment-related effects? Are there certain lesions or diagnoses which may, by their *absence or lessened severity* as compared to controls, be indicative of treatment-related HPT effects such as delayed metamorphosis?"[19]

C) The AMA Should Not Be used in the EDSP

This SAP should advise EPA not to use the AMA test data for any purpose because there are too many unexplained problems with the assay resulting in the generation of flawed data.

This SAP should advise EPA not to use the AMA as part of the EDSP Tier 1 test battery because

1) the AMA failed validation peer review; and

[18] Page 3, http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2013-0075-0005 (charge questions).
[19] Page 4, http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2013-0075-0005 (charge questions).

2) the unexplained problems with the assay resulting in the generation of flawed test data demonstrate that the AMA is not sufficiently accurate, reliable and reproducible to be used as part of the EDSP Tier 1 test battery.

There is no point in trying to provide technical guidance on how to avoid or eliminate these AMA flaws and problems until and unless their cause is known.

III. No SAP Has Ever Reviewed EPA's Conclusion that the EDSP Tier 1 Tests are Validated, and Many of the Tests Failed Peer Review

EPA's White Paper for this SAP states:

> "It is important to note that all 11 Tier 1 assays are validated methods. This analysis is not a "revalidation" of the Tier 1 assays. Each of the Tier 1 assays was validated with a known set of reference compounds (FR Notice, October 21, 2009) for endocrine screening specific to the estrogen, androgen, and thyroid (EAT) pathways through an extensive process of test method development, prevalidation and interlaboratory validation which culminated in evaluation by a FIFRA SAP in 2008 (USEPA, 2008)."[20]

If EPA means that the 2008 SAP reviewed whether or not the individual EDSP Tier 1 tests have been validated, then EPA's statement is incorrect.

The 2008 SAP did not review the validation status of any of the EDSP Tier 1 tests.[21] No other SAP or external peer review has reviewed EPA's assertions that any of the EDSP Tier 1 tests are validated.

Attached to these CRE comments are CRE's earlier comments which discuss the negative peer review reports for EDSP Tier 1 tests in addition to the AMA test. There has been no expert external scrutiny of these peer review reports, or of EPA's assertions that they are validated.

IV. Validate CompTox

It's no secret that EPA plans to abandon the current Tier 1 tests and proposed Tier 2 tests, and EPA plans to use different tests. EPA told Congress:

---

[20] Page 20, at http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2013-0075-0003 .

[21] The 2008 SAP's report on its review, conclusions and recommendations is available online at http://www.epa.gov/scipoly/sap/meetings/2008/032508_mtg.htm#materials .

> "In FY2012 EPA will begin a multi-year transition from the [EDSP] to validate and more efficiently use computational toxicology methods and high-throughput assays that will allow the Agency to more quickly and cost-effectively assess potential chemical toxicity."[22]

Given this fact, and given the problems with the EDSP Tier 1 tests, this SAP should advise the Agency to concentrate its limited time and resources validating the tests that EPA intends to use.

For example, another SAP recently reviewed and reported on EPA's development of ComTox methods.[23]

We thank you for the opportunity to comment.

THE CENTER FOR REGULATORY EFFECTIVENESS,
http://www.thecre.com/

---

[22] See, *e.g.*, U.S. EPA, FY 2012, Justification of Appropriation Estimates for the Committee on appropriations, EPA-190-R-11-003, pages 60-61, http://nepis.epa.gov/Exe/ZyNET.exe/P100A4HZ.TXT?ZyActionD=ZyDocument&Client=EPA&Index=2011+Thru+2015&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C11thru15%5CTxt%5C00000002%5CP100A4HZ.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=p%7Cf&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL , quoted in U.S. Environmental Protection Agency Endocrine Disruptor Screening Program Comprehensive Management Plan Comprehensive Management Plan (June 2012), page 6, http://www.epa.gov/endo/pubs/EDSP-comprehensive-management-plan.pdf ; and in Endocrine Disruptor Screening Program for the 21st Century: (EDSP21 Work Plan), page 2, available online at http://www.epa.gov/endo/pubs/edsp21_work_plan_summary%20_overview_final.pdf .

[23] The CompTox SAP's report is available at http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2012-0818-0037 .