

Original Article

The Use and Acceptance of Other Scientifically Relevant Information (OSRI) in the U.S. Environmental Protection Agency (EPA) Endocrine Disruptor Screening Program

Patricia L. Bishop[1,2]* and Catherine E. Willett[3]

[1]Regulatory Testing Division, People for the Ethical Treatment of Animals, Norfolk, Virginia
[2]PETA International Science Consortium, Ltd., London, UK
[3]Animal Research Issues Department, The Humane Society of the United States, Washington, D.C.

The U.S. Environmental Protection Agency (EPA) Endocrine Disruptor Screening Program (EDSP) currently relies on an initial screening battery (Tier 1) consisting of five in vitro and six in vivo assays to evaluate a chemical's potential to interact with the endocrine system. Chemical companies may request test waivers based on Other Scientifically Relevant Information (OSRI) that is functionally equivalent to data gathered in the screening battery or that provides information on a potential endocrine effect. Respondents for 47 of the first 67 chemicals evaluated in the EDSP submitted OSRI in lieu of some or all Tier 1 tests, seeking 412 waivers, of which EPA granted only 93. For 20 of the 47 chemicals, EPA denied all OSRI and required the entire Tier 1 battery. Often, the OSRI accepted was either identical to data generated by the Tier 1 assay or indicated a positive result. Although identified as potential sources of OSRI in EPA guidance, Part 158 guideline studies for pesticide registration were seldom accepted by EPA. The 93 waivers reduced animal use by at least 3325 animals. We estimate 27,731 animals were used in the actual Tier 1 tests, with additional animals being used in preparation for testing. Even with EPA's shift toward applying 21st-century toxicology tools to screening of endocrine disruptors in the future, acceptance of OSRI will remain a primary means for avoiding duplicative testing and reducing use of animals in the EDSP. Therefore, it is essential that EPA develop a consistent and transparent basis for accepting OSRI. *Birth Defects Res (Part B)* 00:1–20, 2013. © 2013 Wiley Periodicals, Inc.

Key words: ***animal testing; EDSP; endocrine disruptors; Endocrine Disruptor Screening Program; Other Scientifically Relevant Information; OSRI; pesticides***

INTRODUCTION

Creation of the Endocrine Disruptor Screening Program (EDSP)

Endocrine disruptors are substances that mimic, block, or otherwise disrupt the normal function of hormones. Concerns about the presence of endocrine disruptors in food and water and the potential risk to humans led Congress in August 1996 to pass amendments to the Federal Food, Drug, and Cosmetic Act (FFDCA; P.L. 104–170, August 3, 1996) and the Safe Drinking Water Act (P.L. 104–182), requiring the development by the U.S. Environmental Protection Agency (EPA) of an endocrine disruptor screening program (EDSP). The amendment to FFDCA, known as the Food Quality Protection Act , initially focused on the estrogen (E) hormone system in humans. Later the program was expanded to evaluate effects in wildlife, and also potential impacts on the androgen and thyroid (A, T) hormonal systems. EPA was required to implement the new EDSP, using validated test systems, by August 1999.

The agency spent more than 10 years developing the program, selecting and validating the assays to be used for evaluating potential endocrine effects, and determining the universe of chemicals to be tested. In October 2009, EPA issued the first EDSP test orders for 58 pesticide active ingredients and 9 high production volume (HPV) chemicals used as pesticide inert ingredients (74 FR 54422). The pesticides in this "List 1" were chosen based upon human exposure pathways, that is, food, drinking water, residential use, and occupational contact, and the more exposure pathways a pesticide had, the greater its priority for testing (70 FR 56449). The HPV/pesticide inert chemicals were selected based on specific pathways for human and ecologic exposure.

Parties receiving the test orders were directed to conduct a series of 11 tests known as the Tier 1 battery, which consists of five in vitro and six in vivo assays (Table 1). EPA describes its basis for selecting an assay

*Correspondence to: Patricia L. Bishop, Regulatory Testing Division, People for the Ethical Treatment of Animals, 501 Front Street, Norfolk, VA 23510. E-mail: PatriciaB@peta.org

Received 31 May 2013; Accepted 30 August 2013

Published online in Wiley Online Library (wileyonlinelibrary.com/journal/bdrb) DOI: 10.1002/bdrb.21077

Table 1
Tier 1 Battery of Assays (Based on USEPA, 2011c)

Assay	Number of animals used/test[a]	Modes of action: Receptor binding: E	Anti-E	A	Anti-A	Steroidogenesis: E	A	HPG axis	HPT axis
In vitro									
Estrogen receptor binding—rat uterine cytosol	13[b]	■	■						
Androgen receptor binding—rat prostate cytosol	10[b]			■	■				
Estrogen receptor α transcriptional activation		■							
Aromatase recombinant						■			
Steroidogenesis H295R						■	■		
In vivo									
Uterotrophic	18[c]	■							
Hershberger	48[d]			■	■		■		
Male pubertal	45[e]			■	■		■	■	■
Female pubertal	45[e]	■	■			■		■	■
Fish short-term reproduction (♂and ♀)	96[f]	■	■	■	■	■	■	■	
Amphibian metamorphosis	320[g]								■
Total	**595**								

[a]These are the numbers of animals used in the actual test and do not include animals used in range-finding and method optimization studies or animals that are not used in the study and eventually culled.
[b]Number of rats used to collect sufficient uterine or prostate cytosol for assay (LeBaron et al., 2013).
[c]Uterotrophic uses minimum of 6 animals per group with 2 test substance groups and 1 control group for a total of 18 animals.
[d]Hershberger test assumes both agonist and antagonist protocols performed. Agonist uses 6 animals per dose × 2 doses + control = 18 animals; the antagonist version uses 6 animals per dose × 3 doses + 2 controls = 30 animals, for a total of 48 animals.
[e]Both the male pubertal (MP) and female pubertal (FP) use 15 pups per treatment group, with 2 test substance groups and 1 control group for a total of 45 animals.
[f]FSTR assay uses 4 females and 2 males per tank, 3 test substance concentrations and 1 control, and 4 replicates per treatment for a minimum total of 96 animals.
[g]Twenty tadpoles per tank at 3 dose levels plus control and 4 replicates per treatment for a total of 320 animals.
Black boxes indicate the mode of action or actions evaluated by the assay. HPG, hypothalamic–pituitary–thyroidal axis; HPT, hypothalamic–pituitary–thyroidal axis.

to include in the battery as involving (1) the capacity of an assay to detect estrogen- and androgen-mediated effects by various modes of action including receptor binding (agonist and antagonist) and transcriptional activation, steroidogenesis, and hypothalamic–pituitary–gonadal (HPG) feedback; (2) the degree to which in vitro and in vivo assays complement one another in the battery; and (3) the capacity of rodent and amphibian in vivo assays to detect direct and indirect effects on thyroid function (hypothalamic–pituitary–thyroidal feedback) (USEPA, 2011c). Those chemicals found to have the potential to interact with E, A, or T hormonal systems in Tier 1, would presumably proceed to Tier 2 (putatively consisting of animal-intensive, multi-generation reproductive, and developmental toxicity studies in mammals, fish, birds, and amphibians) for further testing.

Use of Other Scientifically Relevant Information (OSRI) to Avoid Duplicative Testing

In creating the EDSP, FFDCA Section 408(p)(1) stated that EPA shall "…develop a screening program, using appropriate validated test systems and other scientifically relevant information to determine whether certain substances may have an effect in humans that is similar to an effect produced by a naturally occurring estrogen, or such other endocrine effect." EPA noted in its April 2009 EDSP Policies and Procedures for the initial screening of chemicals (74 FR 17566) that test order recipients have the option to "…cite or submit existing data (i.e., other scientifically relevant data) in lieu of developing new data, and ask EPA to determine whether the information could be used to satisfy part or all of the Tier 1 Order."

Before requesting any information on endocrine disruptors from test order recipients, EPA first was required to comply with the federal Paperwork Reduction Act by submitting an Information Collection Request (ICR) to the Office of Management and Budget (OMB). The ICR, among other things, had to demonstrate that the desired information "…is not duplicative of information otherwise accessible to the agency" and that there is a "practical utility," or benefit, in collecting the information (5 CFR § 1320.5(d)(ii) and (iii).). In its October 2009 "Terms of Clearance" (TOC), OMB approved the ICR, with certain caveats, one being that "…under the principles of the PRA, EPA should promote and encourage test order recipients to submit Other Scientifically Relevant Information (OSRI) in lieu of performing all or some of the Tier 1 assays, and EPA should accept OSRI as sufficient to satisfy the test orders *to the greatest extent possible* [emphasis

added]" (OMB, 2009). OMB also stated in the TOC that before issuing test orders for any additional chemicals beyond those on List 1, EPA must first provide a report on the submission and acceptance of existing data and OSRI, and descriptions of any instances in which submission of OSRI was deemed insufficient to satisfy the testing order.

The TOC issued by OMB was particularly relevant to the chemicals slated for testing in the initial round of EDSP because pesticides already have a wealth of data associated with them, having undergone extensive testing, including reproductive and chronic/life cycle studies in rodents, fish, and birds, as well as metabolism and pharmacokinetics studies, as part of the pesticide registration process (USEPA, 2013a) required by the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). Similarly, with respect to the pesticide inerts, EPA's HPV Challenge Program (USEPA, 2013b) already provided for the collection of chronic, reproductive, and developmental toxicity data in rodents and fish. Thus, when parties received the first test orders in October 2009, it appeared likely that many of the endpoints addressed in the Tier 1 battery, as well as some endpoints in the presumptive EDSP Tier 2 tests, had already been evaluated for the 58 pesticides and 9 HPV chemicals, and results could be submitted as OSRI in lieu of new testing.

OSRI is defined by EPA as " . . . information that informs the determination as to whether the substance may have an effect that is similar to an effect produced by a substance that interacts with the estrogen, androgen, and/or thyroid hormonal systems (e.g., information that identifies substances as having the potential to interact with the estrogen, androgen, and/or thyroid system(s); information demonstrating whether substances have an effect on the functioning of the endocrine system). Other scientifically relevant information may either be functionally equivalent to information obtained from the Tier 1 assays – that is, data from assays that perform the same function as EDSP Tier 1 assays – or may include data that provide information on a potential consequence or effect that could be due to effects on the estrogen, androgen or thyroid systems" (74 FR 17560). The agency produced a paper in March 2009 entitled *EPA's Approach for Considering Other Scientifically Relevant Information (OSRI) under the Endocrine Disruptor Screening Program* (USEPA, 2009f), which, though developed mainly to provide guidance to EPA staff and managers responsible for reviewing the responses to Tier 1 test orders, was made available to parties considering whether to submit OSRI in lieu of developing new test data. EPA noted that the paper provided general guidance and that when assessing submitted OSRI for appropriateness, it would do so on a case-by-case basis.

The OSRI guidance paper indicates that the quality of the data submitted would weigh heavily in EPA's evaluation. Sensitivity, specificity, confidence in the conclusions, and applicability of conclusions across taxa were also to be given consideration. The guidance paper also states that "[j]udgments will be made considering all factors, not just apparent similarity of endpoints and effects. Such factors may include route of administration, duration of dosing, rationale for the dose levels chosen, frequency of dosing, age at exposure, species, test system, vehicle used, number of test units, variability of data, other factors, and whether the final dataset available to the Agency provides a basis for conclusions about the potential for interaction with the endocrine system in both mammalian and non-mammalian systems" (USEPA, 2009f).

EPA noted that the types of studies that would likely be regarded as providing scientifically relevant information were as follows: studies that correspond to the assays in Tier 1 or Tier 2; metabolism studies; or other vertebrate or invertebrate toxicity studies, including but not limited to, developmental toxicity tests, carcinogenicity tests, toxicogenomic data, and reproductive toxicity tests (USEPA, 2009f). The OSRI guidance paper specifically mentions that guideline tests conducted under 40 CFR Part 158 data requirements to support pesticide registration (USEPA, 2013a) may be considered by EPA. While submitters were encouraged " . . . to provide a cogent and complete rationale for why they believe the OSRI is sufficient to satisfy part or all of the Tier 1 Order," the actual process by which EPA would judge the submitted OSRI and assign weights to the various studies was not described in the paper.

Weight of Evidence Guidance

In November 2010, more than a year after issuance of the first test orders, EPA published a draft document entitled Weight-of-Evidence (WoE) Guidance Document: Evaluating Results of EDSP Tier 1 Screening To Identify Candidate Chemicals for Tier 2 Testing (USEPA, 2010f), which according to the agency, contained the principles it would use in a WoE evaluation of the data from the Tier 1 screening battery and in its review of OSRI. As with the OSRI guidance paper (USEPA, 2009f), the draft WoE guidance lists several factors that EPA would consider in evaluating the data and acknowledges, in general terms, that some factors may be considered more important than others. But there was no explanation as to how these factors would be assessed, what relevancy weight would be assigned to each factor, and how a final result "score" would be determined. Lacking a clearly articulated Standard Evaluation Procedure, the document did not provide a transparent process with respect to decision making on acceptance or denial of OSRI that could ensure consistency among EPA staff reviewing the data.

After receiving public comment on the draft WoE guidance (see docket EPA-HQ-OPPT-2010-0877 at www.regulations.gov), EPA produced a final WoE guidance document in September 2011 (USEPA, 2011d), almost 2 years after the first test orders were issued. This document states again that sources of relevant scientific and technical information may include results from EPA Part 158 guideline studies or equivalent Organisation for Economic Co-operation and Development (OECD) test guidelines (OECD, 2013) and information from published or publically available peer-reviewed studies. Studies conducted under 40 CFR Part 158 subpart F for health effects (USEPA, 2013d), such as the mammalian two-generation reproductive toxicity study, the 90-day rodent and dog studies, 1-year chronic dog study, and chronic mouse and rat studies, are specifically noted as potentially providing useful information regarding estrogen-, androgen-, and thyroid-influenced endpoints. The 40 CFR Part 158 subpart G guideline studies for ecologic effects (USEPA, 2013c) are also identified as potential sources of OSRI.

Table 2
Endpoints Measured in Part 158 Studies Considered Influenced by the E, A, or T Hormonal Systems (USEPA, 2011d)

Endpoint	Influenced by
Age at vaginal opening; estrous cyclicity; reproductive organ weights and corresponding histopathology; fertility	Estrogen
Anogenital distance; age at preputial separation; hypospadias, epispadias, cleft phallus, and areola/nipple retention in male rodent pups; reproductive organ and accessory sex tissue weights and corresponding histopathology, spermatogenesis; fertility	Androgen
Thyroid organ weight and histopathology in 90-day subchronic toxicity studies in dogs, mice and rats as well as chronic toxicity studies in dogs, rats and mice; thyroid hormones levels (T3, T4, and TSH) in any study, although these are generally optional in Part 158 studies	Thyroid hormone

The final WoE guidance (USEPA, 2011d) lists specific endpoints measured in the Part 158 studies that are considered influenced by the estrogen, androgen, or thyroid hormone systems (Table 2). It also notes that Part 158 subpart G studies (avian reproduction, fish full-life cycle, and fish early-life cycle) supply measurements of endocrine-related endpoints including fecundity, reproductive success, egg development and embryo/larval survival, and growth that may be informative of endocrine-related effects, but are not considered diagnostic. While EPA discusses the use of Part 158 guideline studies in the WoE guidance in some detail, it then indicates that these studies were not specifically designed to test for the potential of a chemical to interact with the E, A, or T hormonal pathways, adding that results could be considered along with Tier 1 screening data in a WoE analysis.

A section not included in the draft was added to the final WoE guidance and discusses how the quality of the OSRI provided will be evaluated. EPA references its Information Quality Guidelines (USEPA, 2002), which "...set forth the Agency's policy and procedural guidance for ensuring and maximizing the quality of information, regardless of the source of information." EPA also states its intention to apply the following five General Assessment Factors, as recommended by its Science Policy Council (USEPA, 2003), to the OSRI: (1) soundness, (2) applicability and utility, (3) clarity and completeness, (4) uncertainty and variability, and (5) evaluation and review. EPA provides a general description for each of these General Assessment Factors and "illustrative considerations" when evaluating information, but does not describe the weight that will be assigned to each of these considerations or how they will be factored into WoE determinations.

Seemingly in contrast with previous discussions of the potential use of OSRI, language in Section 3.2 of the final WoE guidance (USEPA, 2011d) indicates that OSRI likely will be treated only as a secondary source under the EDSP. By the time this guidance document was issued in September 2011, most test order recipients were nearing their deadlines for submitting final data packages and had already submitted OSRI in 2010 with little direction as to which type of studies were most apt to be accepted and how EPA would evaluate them. EPA notes in the final WoE guidance that in their *initial* responses to the 2009 test orders "...test order recipients...often submitted existing scientific information to be considered by the Agency in lieu of the Tier 1 assays" and "EPA considered whether or not the submitted information could fulfill the test order requirements for one or more of the Tier 1 assays and informed the test order recipients accordingly" (USEPA, 2011d). With respect to *final* data submissions, however, the next paragraph states that "...[t]o comply with the test orders, *recipients must submit the results of EDSP Tier 1 screening* [emphasis added]. The submission may also include other scientifically relevant information. Sources of relevant scientific and technical information may consist of information that was previously submitted in the initial response to test orders or new or additional information. EPA will consider the additional information submitted and, based on the quality and relevance of that information, will consider it *along with the results of the Tier 1 screening assays* [emphasis added] in a WoE analysis to determine whether or not a chemical has the potential to interact with the E, A, or T hormonal pathways" (USEPA, 2011d). This appears to indicate that while OSRI might be used to *support* a determination of a chemical's potential to interact with the endocrine system, Tier 1 tests must still be conducted and their results will be the primary basis for deciding whether a chemical proceeds to Tier 2 testing.

Objectives of the Study

Nearly all of the respondents to the List 1 test orders submitted OSRI in an effort to reduce or eliminate new testing. The objectives of our study are as follows:

- To describe and enumerate the types of studies submitted as OSRI,
- To compare the OSRI submitted by test order respondents to what EPA allowed for in the guidance issued on OSRI and WoE,
- To review EPA's approach to evaluating OSRI and the reasons given by the agency for accepting and denying OSRI,
- To determine the effect of acceptance of OSRI on number of new animal tests used in Tier 1 screening of the List 1 chemicals,
- To provide an overall critique of the handling of OSRI by EPA and the implications for future use of this type of information in the EDSP.

METHODS

OSRI submissions and accompanying EPA responses were downloaded from EPA's EDSP web site (USEPA, 2011b). We reviewed the OSRI submitted by test order recipients; classified the studies cited as to type of study (Part 158 guideline study, special study, or peer-reviewed study published in the literature), when the study was conducted, whether it used an in vitro or in vivo assay,

and the assay methodology; and noted the Tier 1 tests for which waivers were requested. Interested stakeholders also submitted OSRI for some chemicals and we categorized these studies as well. We then reviewed EPA responses to the OSRI submissions and recorded the reasons given for accepting or denying OSRI. For each Tier 1 assay, we counted the number of times respondents cited each type of Part 158 study and each individual literature study, totaled the number of times each study was either accepted or denied, and summarized the reasons given by EPA for those acceptances/denials. Finally, we commented on EPA's treatment of OSRI in relation to the agency's published statements as to how OSRI could be used to satisfy the requirements of the EDSP and OMB's TOC.

When tallying Part 158 studies, each time a study was cited for a particular Tier 1 assay data requirement, we assigned it either (1) a "No" if it was not accepted to satisfy the data requirement, (2) a "Yes" if it was accepted alone to satisfy the data requirement, or (3) a "Supporting" if it was accepted along with other studies as part of a WoE approach to satisfying the data requirement. When respondents submitted OSRI to satisfy data requirements for a Tier 1 assay, especially in vivo assays, they typically cited several Part 158 studies: usually one or more reproduction toxicity tests in rats, one or more developmental toxicity tests, a combined carcinogenicity/chronic toxicity test in rats, a carcinogenicity test in mice, a subchronic toxicity test in rats, and a chronic toxicity test in dogs. Usually, the same Part 158 study was cited for more than one Tier 1 assay data requirement and we counted it each time. Additionally, more than one of a particular type of study, such as a developmental toxicity study conducted in both rats and rabbits, may have been cited for a particular chemical.

While Part 158 studies were unique to the pesticide tested, and most literature studies similarly tested only one pesticide, several literature studies evaluated more than one List 1 chemical. As with the Part 158 studies, literature studies were often cited by a respondent for more than one Tier 1 assay. Regardless of how many chemicals a study evaluated or how many Tier 1 assay data requirements it informed, each time a particular literature study was cited, it was counted with the same "Yes," "No," or "Supporting" designation as described above.

RESULTS

Initial Responses to Test Orders

Upon receipt of test orders, companies generally provided one or more of the following initial responses to EPA as allowed by Policies and Procedures developed for the EDSP (74 FR 17560):

1. In the case of multiple manufacturers/importers of a chemical, notification that a consortium would be formed to avoid duplicative testing and share the costs of producing data.
2. OSRI would be cited or submitted in support of waivers for some or all Tier 1 tests.
3. The Tier 1 tests, if any, the company or consortium would conduct.
4. The pesticide registration would be voluntarily canceled or the company was in the process of discontinuing the manufacture or import of the chemical.
5. The chemical was not and would not be used in pesticide products.
6. The product would be reformulated to exclude the chemical.
7. The test order recipient was not subject to the test order for various reasons allowed by EPA (see Section F, 74 FR 17560).

Initial responses for each of the 67 List 1 chemicals are summarized in Table 3. For 30 pesticide active ingredients and 2 pesticide inerts consortia were formed, while 22 pesticide active ingredients were each represented by a single manufacturer. Six pesticide active ingredient registrations were voluntarily cancelled and, as such, were no longer subject to EDSP testing. The initial response for seven inerts was either that the test order recipient was not subject to the order, or the manufacture or import of the chemical was discontinued, or the chemical was no longer being sold for use in pesticide products. With 6 pesticides and 7 inerts thus removed from the first list, 54 chemicals continued forward with EDSP test orders. Of these 54 chemicals, respondents for 47 indicated that OSRI would be submitted in lieu of some or all Tier 1 assays (Table 3). This study focuses on these 47 chemicals, except where noted. Pesticide registrations for 2 of the 54 chemicals (dicofol and endosulfan) were voluntarily cancelled after respondents submitted OSRI but are included in the analysis here.

Content of OSRI Reports

A comparison of OSRI reports suggests that test order respondents did not pursue Tier 1 assay waivers in a uniform manner. Based on a survey of respondents, Crop Life America calculated the total cost associated with preparing OSRI submissions for all List 1 chemicals to be $1,584,000 (CLA, 2010). Actual costs per chemical of preparing the OSRI submissions, as reported by survey participants, ranged from $8280 to $151,620. This disparity is reflected in the quality and length of OSRI reports, which varied considerably among respondents. While 3 reports were over 200 pages long, and another 5 contained 100 to 200 pages, 1 consisted of only 7 pages. The majority, however, were between about 30 and 80 pages long. Eighteen were authored by consultants, another 26 were prepared in-house by chemical companies, and 3 had unspecified authors.

Thirty respondents cited OSRI from all available sources including literature studies, Part 158 guideline studies, and ToxCast high-throughput assay screening data (e.g., Judson et al., 2010). Sixteen respondents cited OSRI from guideline studies only, or guideline studies combined with ToxCast data or a single literature study. One respondent cited OSRI from two literature studies only. Many respondents indicated in the OSRI reports that studies were assessed for reliability based on Klimisch criteria (Klimisch et al., 1997) and only studies categorized as "1" (reliable without restriction) or "2" (reliable with restrictions) were used in the OSRI submissions. Category 1 studies are considered GLP-compliant (GLP

Table 3
Initial Test Order Responses and Results of OSRI Submittals by Chemical

Chemical	Initial response	Number of assays for which OSRI submitted	Number of assays for which Waivers granted
PESTICIDES			
2,4-D	C	9	4
Abamectin	C	7	0
Acephate	C	5	0
Atrazine	C	11	10
Benfluralin	SM	2	0
Bifenthrin	C	11	1
Captan	C	8	0
Carbaryl	SM	11	6
Carbofuran	SM	11	0
Chlorothalonil	C	7	0
Chlorpyrifos	C	2	1
Cyfluthrin	SM	10	4[a]
Cypermethrin	C	11	3
DCPA (chlorthal-dimethyl)	SM	0	–
Diazinon	SM	7	1
Dichlobenil	SM	11	0
Dicofol	SM (later VC)	9	2
Dimethoate	C	8	5
Endosulfan	SM (later VC)	11	3
EPTC	C	11	0
Esfenvalerate	C	11	4
Ethoprop	SM	11	0
Fenbutatin oxide	SM	0	–
Flutolanil	SM	5	0
Folpet	SM	8	0
Gardona (tetrachlorvinphos)	C	9	0
Glyphosate	C	11	1
Imidacloprid	C	11	0
Iprodione	C	11	4
Linuron	SM	10	6
Malathion	SM	5	1
Metalaxyl	C	6	0
Methomyl	SM	11	0
Metolachlor	C	0	–
Metribuzin	C	11	1
MGK-264	SM	10	4
Myclobutanil	C	4	2
Norflurazon	SM	9	0
O-phenylphenol	C	10	1
Oxamyl	SM	11	0
Pentachloronitrobenzene (quintozene)	SM	6	3
Permethrin	C	11	6
Phosmet	SM	11	0
Piperonyl butoxide	C	0	–
Propargite	SM	5	0
Propiconazole	C	11	6
Propyzamide aka pronamide	C	0	–
Pyriproxyfen	C	7	2
Simazine	C	11	5
Tebuconazole	C	9	1
Triademefon	C	9	6
Trifluralin	C	6	0
Total	**28 C; 22 SM**	**412**	**93**
Disulfoton	VC		
Methamidophos	VC		
Methidathion	VC		
Methyl parathion	VC		
Propachlor	VC		
Resmethrin	VC		

(*Continued*)

Table 3
Continued.

Chemical	Initial response	Number of assays for which OSRI submitted	Number of assays for which Waivers granted
INERTS			
Acetone	C	0	–
Isophorone	C	0	–
Butyl benzyl phthalate	NSO, DMI, or DWS		
Dibutyl phthalate	NSO, DMI, or DWS		
Diethyl phthalate	NSO, DMI, or DWS		
Dimethyl phthalate	NSO, DMI, or DWS		
Di-sec-octyl-phthalate	NSO, DMI, or DWS		
Methyl ethyl ketone	NSO, DMI, or DWS		
Tolulene	NSO, DMI, or DWS		
Total	**2 C; 7 NSO, etc.**		

C, consortium formed; SM, single manufacturer; VC, voluntary cancellation; NSO, not subject to order; DMI, discontinued manufacture or import of chemical; DWS, do not or will not sell for use in pesticides; PCNB, pentachloronitrobenzene.
[a]While OSRI was submitted for the Hershberger, the cyfluthrin consortium did not believe the data quality to be high enough and had agreed to do the test rather than seek a waiver; the OSRI was, nevertheless, accepted by EPA and a waiver granted.

implies Good Laboratory Practices; USEPA, 2012c) or comparable to a GLP-compliant study and category 2 studies, while not performed according to GLP, are considered well-documented, scientifically credible studies. Other respondents indicated that the studies cited (typically the guideline studies) were conducted using GLP but made no assessment of the literature studies. Still others made no mention of study reliability in their OSRI reports at all.

Waivers Granted and Effect on Animal Use

A complete Tier 1 battery conducted on the 47 chemicals would have required 517 assays to be performed; respondents sought waivers based on OSRI for 412 assays (Table 3), or 80% of the tests. Respondents for 20 chemicals proposed no new testing at all, requesting waivers for all 11 assays; respondents for another 10 chemicals proposed performing 1 or 2 new tests, seeking waivers for the rest. EPA granted 93 waivers, an overall OSRI acceptance rate of 23%. For 20 chemicals, EPA denied all OSRI and required the entire battery be performed. Of the remaining 27 chemicals, the number of assays for which OSRI was accepted was generally low, ranging between 1 and 6 per chemical. Only for the extensively studied chemical atrazine was the acceptance of OSRI high: in that case EPA granted 10 of the 11 waivers requested (Table 3).

Respondents requested waivers for 201 in vitro assays and EPA granted 44 of these or 22% (Table 4). In addition, waivers for 211 in vivo assays were requested and of these EPA granted 49 (23%). The lowest rate of OSRI acceptance was for the fish short-term reproduction assay at 4% (one waiver), followed by the amphibian metamorphosis assay (AMA), for which OSRI was accepted only 13% of the time (four waivers; Table 4) . The female pubertal had an acceptance rate of 19%, while the male pubertal, uterotrophic, and Hershberger assays had respective OSRI acceptance rates of 30, 31, and 39%.

Conducting a full Tier 1 battery on the 47 chemicals would have used 27,965 animals, consisting of 15,040 amphibians, 4512 fish, and 8413 rats, based on the number used per assay in Table 1. It should be noted that although classified as in vitro tests, the Tier 1 estrogen receptor (ER) and androgen receptor (AR) binding assays require respective collection of uterine and prostate tissues from rats; in the case of ER binding about 13 rats per chemical were reported used, and for AR binding, 10 rats per chemical were reported used (LeBaron et al., 2013). In initial test plans, respondents proposed conducting 105 Tier 1 assays using 8161 animals, consisting of 4800 amphibians, 2016 fish, and 1345 rats (Table 4). After EPA's review and rejection of much of the OSRI submitted, the number of animals needed to fulfill the data requirements of Tier 1 rose to 23,566. Testing on the 7 chemicals for which no OSRI was submitted, consumed another 4165 animals in Tier 1 assays, bringing the total used for the remaining 52 List 1 chemicals that were not exempted from testing, or voluntarily canceled before or after OSRI submissions, to 27,731 animals (15,360 amphibians, 4896 fish, and 7475 rats). The 49 in vivo waivers and 20 in vitro (ER and AR binding assays) waivers granted by EPA resulted in a savings of 3325 animals, about 60% of which were rats.

The animal numbers presented above do not include animals used in range-finding studies, initialization and optimization studies, and the culling of animals. The Tier 1 in vivo assays generally require establishment of a maximum tolerated dose (MTD) to be used as the high dose in the experiment. The MTD is supposed to "challenge" the animal but not cause overt toxicity, and can be estimated from existing data, such as Part 158 acute or chronic toxicity studies, or be determined through range-finding studies. Range-finding studies may be conducted with the same chemical for more than one assay to account for possible age and gender differences; for example, 3-week-old juvenile animals in used in pubertal studies may require a different MTD than 8- to 10-week-old ovariectomized adults in an uterotrophic study. A range-finding study for a single chemical that will be tested in both the male and female pubertals would typically use 3 males and 3 females per group, with 1 control group and 3 test substance dose groups, for a total of 24 animals. Results of EPA's preliminary evaluation of Tier 1 assay performance for 21 of the 52 List 1 chemicals that completed Tier 1 testing indicate that one or more range-finding studies were

Table 4
Tier 1 Assays Proposed in Initial Test Plans, OSRI Waivers Sought and Granted, Assays in Final Test Plans, and Comparison of Animal Use between Initial and Final Test Plans, for 47 Chemicals Seeking OSRI Waivers

Assay	Assays in initial test plans	OSRI waivers sought	OSRI waivers granted[a]	Assays in final test plans[b]	Number of animals used in initial test plans	Number of animals used in in final test plans
In vitro						
ER binding	2	45	10 (22)	36	26	468
AR binding	2	45	10 (22)	36	20	360
ERTA	3	44	12 (27)	35	–	–
Aromatase	10	37	6 (16)	40	–	–
Steroidogenesis	17	30	6 (20)	40	–	–
Subtotal	34	201	44 (22)	187	46	828
In vivo						
Uterotrophic	12	35	11 (31)	34	216	612
Hershberger	16	31	12 (39)[c]	34	768	1632
Male pubertal	3	44	13 (30)	32	135	1440
Female pubertal	4	43	8 (19)	38	180	1710
Fish short-term reproduction	21	26	1 (4)	44	2016	4224
Amphibian metamorphosis	15	32	4 (13)	41	4800	13,120
Subtotal	71	211	49 (23)	223	8115	22,738
Total	105	412	93 (23)	410	8161	23,566

[a]Percent of waivers requested that were granted in parentheses.
[b]Assays in final test plans do not include Tier 1 tests required by EPA for dicofol and endosulfan, two chemicals that were voluntarily canceled after submittal of OSRI.
[c]In the case of cyfluthrin, the respondent had agreed to perform the Hershberger assay due to deviations from the guideline protocol of an existing Hershberger study conducted by Zhang et al. (2008), which assessed only antiandrogenicity and showed that cyfluthrin had the potential to interact with the androgen system; EPA subsequently granted a waiver for the assay based on this study.

conducted for most of those chemicals, although Part 158 study results were used in some cases (USEPA, 2013e).

Some of the protocols for the Tier 1 in vivo assays also require culling of animals to reach the desired experimental grouping. For example, while a pubertal study (e.g., USEPA, 2009c) uses 45 animals in actual testing (3 dose groups of 15 animals), nearly 200 animals are discarded before the test even begins. The study typically starts with 15 to 20 pregnant dams producing an assumed average of 10 pups per litter (any litters with fewer than 8 pups are discarded). Because no more than 1 pup from each litter can be in a dose group, of the approximately 200 pups produced, 155 are discarded along with the 15 to 20 dams. In most cases, the excess pups cannot be used to test another chemical or to perform a pubertal on another gender because necropsies must all be done on the same one or 2 days, and most laboratories do not have enough experienced necropsy personnel to handle more than one study at a time. Other Tier 1 in vivo assays that cull large numbers of animals include the AMA and the fish short-term reproduction assay.

Finally, additional animals are likely to be used when preparing to run the assays. For example, LeBaron et al. (2013) report that uteri were collected from approximately 50 female rats for initial cytosol preparation, binding assays, radiometric detection, and other methodological optimization before actually testing the chemicals in the ER binding assay. The same authors, in preparing to run the AR binding assay, report using approximately 26 male rats to isolate the initial prostate cytosol preparation and optimize the method.

Due to these additional animals consumed outside of the actual testing of the chemicals, our estimate of nearly 28,000 animals used to evaluate the 52 List 1 chemicals is conservative, the total number of animals used most likely being much higher. Similarly, the number of animals saved by the OSRI waivers issued by EPA is probably somewhat more than the 3325 we estimate here.

Types of OSRI Submitted

Part 158 guideline studies. The results of Part 158 in vivo guideline studies, which are part of the data requirements for pesticide registration under FIFRA (USEPA, 2013a), were submitted as OSRI by 46 of 47 respondents in an attempt to satisfy data requirements for one or more Tier 1 assays. The most commonly cited tests for health effects included the following: two-generation reproduction toxicity (rats), developmental toxicity (rats and rabbits); chronic toxicity/carcinogenicity (rats); carcinogenicity (mice); subchronic toxicity, oral, dermal, and inhalation routes, (rats); chronic toxicity (dogs, rats); and developmental neurotoxicity (rats). For wildlife effects, avian reproduction toxicity, fish full life-cycle, and fish early life-stage studies were also cited. Other less frequently cited types of studies included the following: one- or three-generation reproduction toxicity (rats), subchronic oral toxicity (mice, dogs), and subchronic dermal toxicity (rabbits). Nearly all of the Part 158 studies submitted as OSRI were conducted before the current guideline protocols (USEPA, 2013b, 2013c), which include additional endpoints and updated procedures, were issued by EPA in 1998 or later. In fact, respondents for only three chemicals submitted data generated using the current two-generation reproduction toxicity test guideline (USEPA, 1998), which is also the proposed EDSP

Table 5
Number of Times Part 158 Guideline Studies Were Cited for Each Tier 1 Assay for the 47 Chemicals for Which OSRI Was Submitted

Assay	One-, two-, three-generation reprod[a]	Developmental toxicity	Developmental neurotoxicity	Chronic/ carcinogenicity	Carcino-genicity	Subchronic	Chronic	Other[b]
In vitro								
ER binding	41	68[c]	14	34	33	51	30	13
AR binding	40	67	14	33	32	50	31	7
ERTA	37	64	13	32	32	50	29	13
Aromatase	31	52	9	25	26	32	22	7
Steroidogenesis	30	46	10	23	24	28	21	10
Subtotal	179	297	60	147	147	211	133	50
In vivo								
Uterotrophic	39 (4)	58	10	31	29	43	29	7 (1)
Hershberger	35 (4)	54	9	27 (2)	23 (2)	34 (2)	25 (1)	9 (2)
Male pubertal	51 (7)	80	15	41 (4)	37 (3)	58 (3)	36 (1)	9 (2)
Female pubertal	51 (9)	81	15	40 (4)	38 (1)	56 (3)	35	9
Fish short-term reproduction	–	–	–	–	–	–	–	48
Amphibian Metamorphosis	30	37 (1)	7	31 (3)	24	45 (1)	32 (1)	9
Subtotal	206 (24)	310 (1)	56	170 (13)	151 (6)	236 (9)	157 (3)	91 (5)
Total	385 (24)	607 (1)	116	178 (13)	298 (6)	447 (9)	290 (3)	141 (5)

Indicated in parentheses are the number of times guideline studies were accepted alone for satisfying Tier 1 data requirements, or accepted as part of a WoE approach to satisfy data requirements (no parentheses indicates zero acceptance).
[a]All Part 158 guideline studies are in vivo mammalian tests except those noted under Other.
[b]Includes avian reproduction, fish early life–stage, fish full–life cycle, and several mammalian or fish special studies.
[c]For many pesticides, more than one particular type of study had been conducted and cited, for example, developmental toxicity in both rats and rabbits, and subchronic toxicity via oral and inhalation routes.

Tier 2 mammalian reproduction toxicity test for defining dose and response. In addition, one chemical had been evaluated using the recently OECD-validated extended-one-generation reproduction toxicity study (EOGRTS) (OECD, 2011). Both of these newer protocols include some measurements of endocrine-related endpoints that were lacking in the older, pre-1998 reproduction toxicity tests. Of the other types of mammalian Part 158 toxicity studies, that is, one-generation reproduction, subchronic, and developmental, data were submitted from only 12 studies conducted using current protocols. In addition, data generated using the current developmental neurotoxicity test method were submitted for 14 chemicals, although this test typically does not include measurements of E, A, or T-related endpoints.

The number of times Part 158 guideline studies were cited to satisfy Tier 1 data requirements is shown by assay in Table 5. Also indicated in parentheses is the number of times a guideline study alone was accepted for directly satisfying Tier 1 data requirements, or was accepted along with the results of other studies as part of a WoE approach to satisfy data requirements. EPA never accepted a guideline study alone or in a WoE approach to satisfy any of the Tier 1 in vitro mechanistic assays. Guideline studies were accepted alone or in a WoE approach to satisfy data requirements for Tier 1 in vivo assays however. Of the 24 times results from one-, two- or three-generation reproduction studies were accepted, only when the current 1998 two-generation reproduction toxicity (USEPA, 1998) or the EOGRTS (OECD, 2011) protocols had been used were the guideline studies alone accepted in lieu of performing one or more Tier 1 in vivo assays. The current two-generation reproduction toxicity guideline study was conducted with three chemicals, carbaryl, dimethoate, and MGK-264, and was accepted alone to waive the female and male pubertals, Hershberger and uterotrophic assays for dimethoate and MGK-264. For carbaryl, results of the two-generation study were used in a WoE approach with other studies to waive the female and male pubertals. The EOGRTS protocol (OECD, 2011) was conducted on one chemical, 2,4-D, and results were accepted alone to waive the female and male pubertals, Hershberger and uterotrophic assays. The remaining types of guideline studies, whether or not they were conducted using current protocols, were only accepted along with results from other guideline studies and/or literature studies in a WoE approach. Interestingly, mammalian studies that evaluated certain thyroid endpoints were accepted to satisfy the AMA data requirement for two chemicals: a combined chronic/carcinogenicity study together with a metabolism study following subchronic exposure, supported by two literature studies, satisfied the data requirement for carbaryl; and a combined chronic/carcinogenicity study together with a developmental toxicity study satisfied the requirement for metribuzin. In both cases, the results indicated the potential to interact with the thyroid system. Mammalian guideline studies were not accepted to waive the AMA when results were negative.

Literature studies. Peer-reviewed studies published in the literature were frequently cited for all Tier 1 assays, and included in vitro receptor binding, transactivation, cell proliferation, aromatase activity, and steroidogenesis studies as well as various in vivo studies, such as mammalian reproduction, chronic, subchronic, and developmental toxicity studies, fish sex effects studies, and

endocrine-related studies. Ten cited studies were actually the ones conducted by EPA or an EPA contractor to validate the Tier 1 assays (e.g., Battelle, 2005; Laws et al., 2000; Lech, 2006). Most cited in vivo studies were performed using a single chemical, but a few in vitro studies examined multiple chemicals, such as the study by Kojima et al. (2004), which used a Chinese hamster ovary cell reporter gene transactivation assay to screen 200 pesticides (including 33 List 1 pesticides) for estrogenicity and androgenicity. This study was cited as OSRI for 27 chemicals to satisfy the AR binding assay, for 20 chemicals to satisfy the ER binding assay, and for 28 chemicals to satisfy the ER transcriptional activation (ERTA) assay. The results of this study were accepted alone to satisfy the Tier 1 ERTA assay in the case of two chemicals, cyfluthrin and dicofol, both of which indicated positive results in the concentration range tested (10^{-8}–10^{-5} M). The study was also accepted in a WoE approach along with two other studies to demonstrate potential antiandrogenic activity of cyfluthrin.

Two other screening studies evaluated multiple chemicals: Nishihara et al. (2000), which evaluated the estrogenic activity of 514 chemicals (including 27 List 1 pesticides) using a yeast two-hybrid assay, and Blair et al. (2000), which tested the ER binding ability of 188 chemicals (including 8 List 1 pesticides and 3 List 1 HPV chemicals) using the same rat uterine cytosol (RUC) method as is used in the Tier 1 ER binding assay (USEPA, 2009b). Nishihara et al. (2000) was cited as OSRI to satisfy the Tier 1 ER binding assay for 8 chemicals, and to satisfy the ERTA assay for 11 chemicals. A negative response was indicated in each citation, but the study was not accepted by EPA in any of the cases. Blair et al. (2000) was cited for the ER binding assay for 6 pesticides, none of which showed any response. EPA rejected this study in 4 cases and indicated, for the other 2 chemicals, it would reconsider the results if concentration-response data were provided.

Other in vitro studies cited as OSRI for multiple chemicals included Andersen et al. (2002), Bauer et al. (2002), Chen et al. (2002), Fang et al. (2003), Hinfray et al. (2006), Kojima et al. (2005), Lemaire et al. (2006), Petit et al. (1997), and Soto et al. (1994). Andersen et al. (2002) tested 24 pesticides for interactions with the ER and the AR in transactivation assays and also investigated estrogen-like effects on MCF-7 breast cancer cell proliferation and effects on CYP19 aromatase activity in human placental microsomes. Of the 17 times this study was cited as OSRI for the Tier 1 AR binding, ER binding, ERTA, and aromatase assays, it was accepted only once, along with another ERTA study and several cell proliferation studies, as part of a WoE approach for satisfying the ERTA data requirement for endosulfan. Bauer et al. (2002) characterized receptor binding affinity of 29 pesticides (including 8 List 1 pesticides) using an immuno-immobilized recombinant human AR and was cited 4 times. EPA accepted the study once to satisfy the Tier 1 AR binding assay for a chemical that tested positive, rejected it twice for two negative chemicals, and indicated that it would reconsider the results if additional data were provided for another negative chemical. Chen et al. (2002), using three different in vitro assays to test for estrogenicity, including the RUC ER competitive binding method, was cited four times for ER binding and twice for ERTA. The study was not accepted for either of the two ERTA requested waivers, but for the ER binding assay, EPA accepted the study to satisfy data requirements for two chemicals that tested negative (dimethoate and malathion), accepted it in combination with another study to satisfy data requirements for one chemical that tested positive (cypermethrin), and rejected it for a weakly positive chemical (permethrin). Fang et al. (2003) was cited five times for AR binding; EPA accepted it twice for positive-testing carbaryl and endosulfan, rejected it twice for negative-testing atrazine and simazine, and indicated once that it would reconsider the results if concentration-response data were provided for negative-testing 2,4-D. Kojima et al. (2005), which used the E-Calux assay (a stably transfected ERTA assay with BG1 cells), was cited for four chemicals to satisfy the ERTA assay. EPA accepted it twice for two positive chemicals (diazinon, pyriproxyfen), rejected it once for negative-testing dicofol, and indicated once it would reconsider the results if additional data were provided for negative-testing captan. None of the other studies mentioned above, which were cited for multiple chemicals, was accepted to satisfy Tier 1 data requirements.

Table 6 summarizes the number of times literature studies were cited as OSRI and how often they were accepted to satisfy Tier 1 assay requirements either alone or in a WoE approach combined with other literature studies or Part 158 guideline studies. It should be noted that for a particular chemical often the same study was cited by respondents for more than one Tier 1 assay, or, as in the cases mentioned above, the same study was cited for more than one chemical. Therefore, the total number in Table 6 (780) does not represent the number of unique studies cited, which actually totaled 228, but rather the number of times literature studies were cited for all the Tier 1 assays combined. Importantly, even though literature studies were frequently cited to satisfy the Tier 1 data requirements, they were accepted alone by EPA only 56 times and in a WoE approach, only 74 times.

ToxCast data. Many of the OSRI reports cited results from in vitro screening assays conducted as part of the first phase of EPA's ToxCast Project (e.g., Judson et al., 2010). A total of 309 pesticide active ingredients, including 41 of the 47 pesticides for which OSRI was submitted, were tested in 467 assays, including endocrine-related, high-throughput cell-free assays and cell-based assays designed to assess chemicals at the molecular and pathway perturbation levels. For each of the Tier 1 in vitro assays, ToxCast data were cited for one-half to two-thirds of the chemicals. While ToxCast is currently being incorporated into EPA's proposed approach to future EDSP screening using computational toxicology and high-throughput assays known as EDSP21 (USEPA, 2011a), none of the ToxCast results cited for List 1 pesticides was accepted to waive Tier 1 testing.

Distribution of Waivers

The distribution of waivers among Tier 1 assays indicates that, overall, the greatest number of waivers was issued based on the results of a single literature study (Fig. 1). Waivers based on literature studies alone numbered higher for in vitro Tier 1 assays (34) than for in vivo assays (17). The reverse was true for literature

Table 6
Number of Times Literature Studies Were Cited as OSRI and Number of Times Studies Were Accepted to Satisfy Tier 1 Assay Data Requirements Either Alone or in a WoE Approach Combined With Other Literature Studies or Part 158 Guideline Studies for the 47 Chemicals for Which OSRI Was Submitted

Assay	Number of times studies cited	Number of times single study accepted to satisfy Tier 1 assay	Number of times studies accepted in a WoE approach to satisfy Tier 1 assay
In vitro			
ER Binding	125	10[a]	2
AR bnding	102	8[a]	14
ERTA	150	9	8
Aromatase	46	6[a]	2
Steroidogenesis	43	5	3
Subtotal	466	39	30
In vivo			
Uterotrophic	49	3	7
Hershberger	60	7	10
Male Pubertal	82	4	17
Female pubertal	49	2	2
Fish short-term reproduction	15	1	0
Amphibian metamorphosis	59	0	8
Subtotal	314	17	44
Total	780	56	74

[a]When a waiver was granted for that assay, in the case of one chemical more than one study cited was adequate to meet the data requirement. For the ER binding assay, 10 studies were accepted for 9 waivers, that is, the respondents for the pesticide atrazine cited 2 studies, each of which alone could have satisfied the Tier 1 assay data need, according to EPA. For the AR binding assay, 8 studies were accepted for 6 waivers, that is, the respondents for the pesticide linuron cited 3 studies, each of which alone could have satisfied the Tier 1 assay data need. Similarly, for the aromatase assay, 6 studies were accepted for 5 waivers, that is, each of 2 studies alone cited for atrazine would have satisfied the data requirement.

studies accepted together with other literature and/or guideline studies in a WoE approach: more waivers were granted for in vivo assays (16) than in vitro assays (10). Guideline studies alone accounted for 12 *in vivo* waivers, while guideline studies used in a WoE approach accounted for 4 in vivo waivers. None of the in vitro assay waivers was based on guideline studies alone or combined with literature studies or other guideline studies.

Bases for Lack of Acceptance of OSRI

In Vitro mechanistic assays. Although in vivo guideline studies do not directly measure the various endocrine mechanisms assessed by Tier 1 in vitro assays, that is, ER and AR binding, ERTA, steroidogenesis, and aromatase activity, results from these studies were often submitted by respondents to satisfy Tier 1 in vitro assay data requirements under the presumption that any perturbations to hormonal systems would result in a functional endocrine change in the intact animal and be evident in the apical endpoints evaluated in the higher tiered guideline studies. This reasoning assumes that if there is a lack of adverse biologic effect in a higher tiered study, further knowledge of any mechanistic action then provides little added value. EPA did not agree with this approach and often made statements in its review as to the deficiencies of the guideline studies presented, such as the following cited from the agency's review of carbofuran OSRI (USEPA, 2010c):

- "There was no substantive explanation explaining why lack of estrogenic effects in these studies should be considered evidence that binding to the estrogen receptor does not occur. A lack of effect on potentially receptor-mediated endpoints in the mammalian *in vivo* studies cited does not necessarily demonstrate absence of interaction with the receptor in other species."
- "The cited data do not permit the Agency to establish confident linkages between apical endpoints measured in whole animal studies and aromatase enzyme activity."
- "A lack of effects on potentially steroid-mediated endpoints in the mammalian *in vivo* studies cited does not necessarily demonstrate that steroidogenesis has not been affected."

In vitro literature studies cited as OSRI for the Tier 1 in vitro assays were commonly rejected based on one or more of the following reasons:

- The reported results were negative but no supporting information, such as raw, replicate, concentration-response, limit of solubility, and/or cytotoxicity data, was provided;
- The method was acceptable, but a single concentration was used or too narrow a concentration range was tested or the substance was not tested to its limit of solubility;
- In the case of ER or AR binding, negative results from yeast cell-based assays could not be accepted because there were questions about the ability of some chemicals to fully penetrate the yeast cell wall, potentially resulting in false-negatives;

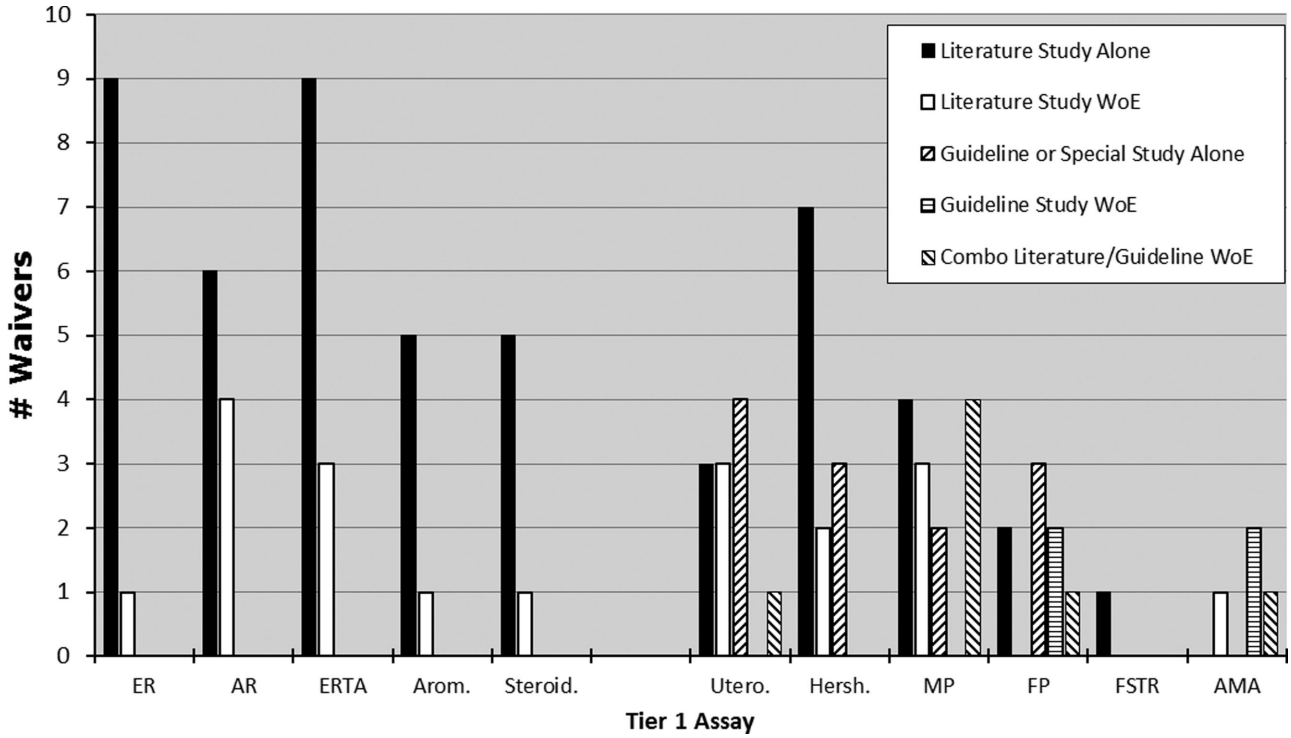

Fig. 1. Number of waivers per assay based on (1) a literature study alone, (2) two or more literature studies taken together in a WoE approach, (3) a single guideline or special study alone, (4) two or more guideline studies taken together in a WoE approach, or (5) a combination of literature studies and guideline studies in a WoE approach.

- For ERTA, cell proliferation assays were cited instead of transcriptional activation assays;
- For ER or AR binding, transcriptional activation assays or cell proliferation assays were cited instead of binding assays;
- ToxCast data were considered potentially appropriate for use in priority setting but were not considered an acceptable alternative for the Tier 1 assays.

In Vivo assays. With the exception of the current two-generation mammalian reproduction toxicity study (USEPA, 1998) and the EOGRTS (OECD, 2011), the Part 158 guideline studies cited by nearly all respondents were not considered by EPA to contain measurements of all of the requisite endpoints that may be indicative of a possible endocrine effect. Data that EPA typically noted as missing from or being deficient in the guideline studies for each in vivo assay are listed below.

Female/Male Pubertal

- Age and weight at vaginal opening/preputial separation not measured;
- Offspring in the developmental toxicity study were evaluated for age of vaginal opening/preputial separation but exposure had ceased by the time the endpoint was evaluated, unlike in the pubertal assay where exposure continues through the critical peripubertal period;
- Estrous cyclicity/male secondary organ weights, testosterone levels not measured;
- Other organ weights and histopathology obtained in adults not juveniles;
- Thyroid hormone levels not evaluated;
- While thyroid histopathology may have been conducted, it was not clear that follicular cell height and colloid area were evaluated: Tier 1 pubertal assay requires use of a five-point grading system of these endpoints.

Uterotrophic/Hershberger

- Cited studies used intact animals, not ovariectomized/castrated ones;
- No measurements made of uterine weight in females and ventral prostate and secondary sex organ (seminal vesicle, Cowper's gland, LABC muscle complex and glans penis) weights in males.

Amphibian Metamorphosis

- Lack of effects on the thyroid axis in mammalian in vivo studies did not necessarily rule out effects in other species. The role of the thyroid axis in anuran metamorphosis is complex and could be impacted on many different levels, some of which may not be apparent in existing mammalian assays.

When Part 158 studies were cited as OSRI for the Tier 1 in vivo assays to show no effect, they were generally rejected by EPA as insufficient because not all endocrine-sensitive endpoints were measured; however, there were cases when positive results on certain endpoints were sufficient to lead to Tier 1 test waivers.

EPA waived the AMA for carbaryl based on observed thyroid follicular cell hypertrophy/adenomas in a 1993 combined chronic/carcinogenicity guideline study and a 1997 metabolism/subchronic oral toxicity study, both performed on rats (USEPA, 2010b). Similarly, the AMA was waived for metribuzin based on a 1986 developmental toxicity study in rats, which showed increased thyroid weight and decreased levels of T3 and T4; and a 1993 combined chronic/carcinogenicity study in rats, which also showed increased thyroid weights and reduced T3 and T4 levels as well as thyroid follicular cell hyperplasia (USEPA, 2010j). The female pubertal was waived for dicofol based on results of a 1997 one-generation reproduction toxicity study in rats, which included measurements of estrous cyclicity, age, and weight of vaginal opening (no effects); and a 1989 combined chronic/carcinogenicity study in rats, which showed no effects on T3, T4, and TSH over a 3-month treatment period, and a 1986 subchronic toxicity study in rats, which indicated hypertrophy of thyroid follicular epithelium in two highest dose groups (USEPA, 2010e). It is not clear from EPA's review if it considered the overall results positive or negative.

DISCUSSION

Inconsistent Treatment of OSRI from In Vitro Studies

Our findings do not necessarily suggest that instances where EPA judged individual studies to be inadequate for ruling out interaction with endocrine pathways are suspect; for some chemicals the deficiencies and limitations of OSRI that EPA considered reason for denying Tier 1 waivers, as described above, may have been justified based on the category of chemical being evaluated and its mode of action, the dose range, the type of test system used, and other factors. However, EPA's conclusions are often not transparent and, in some cases, appear inconsistent when the stated limitations of a particular study were used to justify new testing if results were negative, but were apparently overlooked if the study showed positive effects, that is, that a chemical had the potential to interact with the endocrine system. As an example, the Bauer et al. (2002) study noted above directly evaluated the AR binding potential of 29 pesticides, including 8 (propiconazole, tebuconazole, dimethoate, metalaxyl, cypermethrin, permethrin, metribuzin, and fenvalerate) of the 47 chemicals discussed here. The pesticides were tested for their ability to displace the standard AR ligand, [^{3}H]-DHT, bound to the recombinant human AR, up to their maximal soluble concentration; inhibition constants (K_i), and relative binding affinities (RBA) in comparison to DHT, were calculated from the IC_{50} (50% inhibition of DHT binding). Tebuconazole and propiconazole tested positive, with respective RBAs of 0.0060 $\pm$ 0.0001 and 0.0018 $\pm$ 0.0001%. The remaining six List 1 chemicals were unable to displace [^{3}H]-DHT, even at maximal soluble concentrations, and were considered nonbinders with no K_i determined. This study was cited by the respondents for propiconazole, cypermethrin, permethrin, and metalaxyl, and was accepted by EPA for the positive-testing propiconazole (USEPA, 2010m) but not for the latter three nonbinders. The agency stated that for the three nonbinders, even though "…the methodology was a binding study and was well described," the fact that the highest concentration of cypermethrin and permethrin tested, or in the case of metalaxyl, the concentration-response data, were not reported "…render[ed] this study unable to satisfy the requirements of the Test Order requirement for the Androgen Receptor Binding Assay using Guideline 890.1150" (USEPA, 2010d, 2010i, 2010k). With regard to the reported nonbinders, EPA also took issue with the use of Scatchard analysis (Scatchard, 1949) to determine the dissociation constant (K_d) for DHT binding to the AR (part of the equation to calculate K_i), noting that it "…is subject to considerable error" (e.g., USEPA, 2010k). Use of Scatchard analysis apparently was not a cause for concern, however, with the positive determination of propiconazole as a potential binder to the AR.

The results of three other studies (two transactivation studies and a Hershberger assay) cited as OSRI for both permethrin and cypermethrin provided some evidence that these chemicals may act as weak androgen antagonists, although it was uncertain whether or not the effects were due to binding since this was not directly measured. Even though the negative results in direct measurement of binding by Bauer et al. (2002) potentially conflicted with the other OSRI results showing weak antagonist activity, EPA went on to waive the AR binding assay for cypermethrin and permethrin based on the positive findings of the other OSRI. It is not clear how this waiver will be interpreted and if EPA will require these two chemicals to undergo Tier 2 testing. It may have made more sense to require the Tier 1 AR binding study to see whether it corroborated the findings of Bauer et al. (2002) or the other OSRI. Somewhat in contrast, EPA required the Tier 1 AR binding assay be performed for metalaxyl, despite the negative Bauer et al. (2002) finding and no indications in any other studies (including a transactivation study) of effects of this chemical on the androgen hormone system.

Additionally, while esfenvalerate/fenvalerate was tested in the Bauer et al. (2002) study, the respondents for this pesticide did not submit the results as part of their OSRI for AR binding, citing instead a competitive AR binding study by Sumitomo (2001), which, like the Bauer et al. study, also demonstrated that fenvalerate did not bind to the AR. The Sumitomo et al. study was accepted by EPA to satisfy the Tier 1 data requirement (USEPA, 2010h), and as such, lends additional credibility to the negative results of the Bauer et al. study. The respondents for reported nonbinders, dimethoate and metribuzin, and reported binder, tebuconazole (Bauer et al., 2002), also failed to cite the Bauer et al. study, instead using results from Part 158 studies alone or in combination with studies that did not directly measure AR binding, and thus were required by EPA to conduct the Tier 1 assay. Had the Bauer et al. (2002) study been accepted universally by EPA and cited by respondents for all List 1 chemicals tested in the study, 4 more Tier 1 AR binding assays could have been avoided and at least 40 animals saved.

Similar treatment by EPA occurred with other studies. Fang et al. (2003) evaluated 202 substances using a recombinant rat AR (PanVera) competitive binding assay. This study was funded in part by EPA and the Food and Drug Administration, and the dataset generated was the largest

and most diverse at the time, becoming known as the (Food and Drug Administration) National Center for Toxicological Research AR dataset. The authors validated the PanVera assay by comparing results to other AR binding assays, including one using a rat prostate cytosol method (Waller et al., 1996) similar to the Tier 1 rat prostate cytosol assay (USEPA, 2009a). The results indicated that the PanVera and cytosol assays were generally comparable for chemicals with RBA >0.001. For the chemicals with lower RBAs, the PanVera assay was reported to be more sensitive. Through structure–activity relationship (SAR) analysis, the authors also were able to define general chemical structure requirements for binding to the AR (Fang et al., 2003).

As in the Bauer et al. (2002) study, Fang et al. (2003) conducted Scatchard analysis of binding data to determine the K_d for [^{3}H]-R1881, the standard AR ligand used in both this assay and the Tier 1 assay, although EPA made no comment about Scatchard analysis in the review of this study. Test chemicals were run in duplicate tubes with at least two replicates at concentrations ranging from 4.28×10^{-9} to 4.28×10^{-4} M at one log unit concentration intervals. The IC_{50} and RBA were calculated for each chemical competitor; chemicals that failed to compete with [^{3}H]-R1881 in binding were designated as nonbinders, while chemicals that showed binding but did not reach 50% inhibition at maximum concentration were designated as slight binders. Five List 1 chemicals were included in this assessment and displayed the following RBAs: endosulfan (0.0133), linuron (0.0056), carbaryl (0.0008), atrazine (nonbinder), simazine (nonbinder), and 2,4-D (nonbinder). EPA accepted the Fang et al. (2003) study for endosulfan, adding that "Although this assay does not use the full length receptor, it is considered to be a valid assay that meets the requirements of the test order" (USEPA, 2010g). The agency also accepted the study for carbaryl, stating that "Although positive results were reported, the limitations of this study are that only the ligand binding domain was used which may not give the same results as the full receptor, and no concentration response data were provided," adding that positive findings for carbaryl's interaction with the androgen system were demonstrated in certain in vivo studies (USEPA, 2010b), which lends support and confidence in Fang et al. (2003) results. The respondent for linuron did not cite the positive results of Fang et al. (2003), but this chemical's ability to interact with the androgen system was demonstrated in three other AR binding studies accepted by EPA, including the validation study for the EDSP Tier 1 assay (USEPA, 2007). Atrazine was used in the EDSP validation study (USEPA, 2007) as a negative (nonbinding) chemical and, as with linuron, the Tier 1 test was waived based on its inclusion in the validation study. However, the Fang et al. study was also cited in the atrazine OSRI, to which EPA responded with the following: "Fang et al. (2003) reported atrazine as a nonbinder but provided no data. Therefore, this study could not be considered as satisfying the requirements of the test order" (USEPA, 2010a). Additionally, Fang et al. (2003) was cited in the OSRI for simazine, reporting this chemical to be a nonbinder, but EPA responded by rejecting this evidence and stating that the PanVera protein is a "...rat receptor ligand-binding domain only, not a full receptor. Simazine was reported to be negative, but no data were provided. In addition, there is some concern that the ligand-binding domain in a chimeric construct does not always give the same results as the full receptor, and may provide false positives and negatives" (USEPA, 2010n).

The Fang et al. (2003) study provided an analysis of SAR and bases for explaining the AR affinity of various classes of chemicals. The three nonchlorinated triazine pesticides assessed in the study, atrazine, simazine, and prometon, all tested as nonbinders, a result the authors noted was consistent with findings in a report on the development of the male pubertal assay (Stoker et al., 2000) that showed these chemicals act as endocrine disruptors, not by binding to the receptor but through direct effect on the central nervous system. It is unreasonable that the Fang et al. (2003) study, which was a valid and well-documented binding study, funded by two government agencies, and co-authored by several staff from one of those agencies, was rejected in the case of negative findings but accepted for positive findings. It is also unreasonable to expect that concentration-response data from nonbinders would be included in a journal article covering 202 chemicals, most of which tested negative, that was published at a time when inclusion of supplemental data was certainly less prevalent. Presumably raw concentration-response data from a study that was partially funded by EPA would be obtainable by the agency if it was considered necessary for a waiver determination. Finally, it is disconcerting that the same study could be treated so differently depending on the chemical in question, and points not only to variability in assessment methods by EPA staff, but also to a narrow view of what constitutes OSRI and a lack of clear standards for accepting studies submitted as OSRI.

Yet another study that was treated differently depending on the chemical response is the one by Kojima et al. (2005), which evaluated ERTA using the E-Calux assay, an apparent earlier version of the BG1Luc (Lumi-Cell) method developed by Xenobiotic Detection Systems and recently validated by the U.S. Interagency Coordinating Committee on the Validation of Alternative Methods for use in the EDSP (NTP, 2013). The BG1Luc ERTA assay is at least equivalent to the Tier 1 ERTA assay, and was actually found to have several advantages over the existing method (NTP, 2013). Twelve List 1 chemicals were tested over a range of concentrations, but only diazinon and pyriproxyfen displayed measurable estrogenic activity. For these two pesticides, an EC_{10}, the 10% response in relation to the E2 maximal response as 100%, was calculated and concentration-response curves with SD bars were shown in the paper. EPA granted a Tier 1 ERTA waiver for diazinon based on the Kojima et al. (2005) study alone, and for pyriproxyfen based on this study and the results of a cell proliferation assay study using a rat pituitary cancer cell line (Manabe et al., 2006), both of which showed estrogenic activity. Of the remaining 10 List 1 chemicals, only the respondents for captan and dicofol cited the Kojima et al. (2005) study, which reported negative results for both of those chemicals but provided no supporting data. For captan, negative results from several other studies including an ERTA study using Chinese hamster ovary cells with a luciferase reporter (Kojima et al., 2004), a cell proliferation study using the

MCF-7 cell line (Okubo et al., 2004), and a yeast two-hybrid system (Nishihara et al., 2000), were also cited as OSRI supporting the nonestrogenic finding of Kojima et al. (2005). Yet, despite this apparent weight of evidence, EPA required the Tier 1 ERTA assay to be conducted for captan. Interestingly, for dicofol, conflicting results were reported in the literature: Kojima et al. (2004) observed a positive estrogenic response and the study was accepted by EPA to satisfy the Tier 1 requirement, while Kojima et al. (2005) reported a negative response but was rejected for a lack of data presented, and Vinggaard et al. (1999), using a cell proliferation assay with estrogen-responsive MCF-7 breast cancer cells, also reported dicofol to be negative but was rejected for being a cell proliferation study (USEPA, 2010e). Kojima et al. (2004) and Kojima et al. (2005) were each often noted in the EPA reviews of OSRI to be adequate studies using proper controls, and were accepted in the few cases when chemicals were positive, but they were universally rejected when chemicals tested negative for reasons of not enough information provided, and in the case of Kojima et al. (2004), not a wide enough range of concentrations tested and the lack of cytotoxicity data provided.

These examples underscore the need for a clear method of evaluating OSRI, one that determines the validity and applicability of a study based on a priori criteria, regardless of whether the chemical exhibited positive or negative effects. This is particularly important in those cases where multiple chemicals were evaluated in a single study; unless there are identifiable factors associated with certain chemicals or types of chemicals that preclude proper evaluation by that test system, it would seem that either all results from that study are deemed valid or all are not.

Lack of a WoE Approach

The 2011 EPA guidance on WoE (USEPA, 2011d) indicates that the process of supporting a hypothesis on a chemical's effect or lack of effect involves several steps, including evaluating the data for quality, relevance, and agreement; integrating the different lines of evidence; applying professional judgment; and determining whether there is corroborating evidence of effects (or no effects) at different levels of biologic organization. However, EPA's response to some of the OSRI submitted appears to contradict this approach. In the case of evaluating the chemical 2,4-D for the potential to interact with the estrogen system, 13 in vitro mechanistic studies offered as OSRI by the respondent indicated a lack of effects (Neal et al., 2010), including an RUC binding study that tested at two high concentrations (Blair et al., 2000), several yeast screens (e.g., Hurst and Sheahan, 2003; Jung et al., 2004; Orton et al., 2009), two MCF-7 cell proliferation studies (Soto et al., 1995; Lin and Garry, 2000), a competitive binding study using ER derived from alligator oviducts (Vonier et al., 1996), a study that used three different assays and tested over a range of concentrations (Petit et al., 1997), an ERTA study (Kojima et al., 2004), and several others. Rather than consider the negative results of the 13 studies holistically in a WoE approach, EPA reviewed each study separately and found a reason to reject each one, indicating that none of the data submitted satisfied the test order. In contrast, both the uterotrophic assay and the female pubertal assay, which are designed to evaluate perturbations to the estrogen system in an animal through measurement of apical endpoints, were waived for 2,4-D based on the availability of a recently conducted EOGRTS, which showed no effects on the estrogen pathway but did show potential effects on the thyroid system at a dose level resulting in nonlinear toxicokinetics. It seems illogical for EPA to require the in vitro ER binding and ERTA assays when the results of numerous in vitro studies taken together, support a finding of no effect, and further, when the agency waived two higher tiered animal tests based on results from an acceptable in vivo study that also showed no effect on the estrogen system. Responding to the submitted OSRI in this manner is more indicative of a check-box approach to evaluation of OSRI rather than a WoE approach.

Another case in which a WoE approach may have been somewhat lacking or was inconsistently applied is EPA's response to OSRI for the pesticide pentachloronitrobenzene. EPA granted a waiver for the ER binding assay based on the respondent's citation of Ashby et al. (2005), which reported pentachloronitrobenzene to be negative in an RUC ER binding assay, showing concentration-response curves over a range of 10^{-10} to 10^{-4} M. EPA also waived the uterotrophic assay based on a nonguideline uterotrophic assay conducted by Ashby et al. (2005) over a dose range of 100 to 800 mg/kg/d for 3 days that produced negative results. EPA rejected the studies cited for the Tier 1 ERTA assay, which showed negative effects, and included the ERTA study by Kojima et al. (2004), the cell proliferation study by Vinggaard et al. (1999), and an antiestrogenic yeast transcriptional activation assay conducted as part of Ashby et al. (2005). But then, seemingly in contrast to the agency generally requiring that it have information concerning both binding of the ER and direct activation of ER-controlled DNA transcription, EPA waived the ERTA assay, indicating that "...this test order requirement is satisfied by ER binding assay [*sic*] which was part of the Ashby et al. (2005) study when considered with the results of the uterotrophic assay" (USEPA, 2010l). This waiver for the ERTA assay may indicate a WoE approach because all ER-mediated effects would have presumably been evaluated in the studies cited for the waiver, but the agency still required the female pubertal assay be conducted because of "...the deficiencies and unanswered questions for the estrogen system" (USEPA, 2010l). The female pubertal can detect effects on estrogen signaling through alternate modes of action, such as via the hypothalamic–pituitary–gonadal axis and inhibition of estrogen biosynthesis, but a female pubertal study conducted by Ashby et al. (2005), which dosed animals at 800 mg/kg/d for 12 days beginning on PND 25, was submitted as OSRI. The Tier 1 female pubertal protocol requires that dosing of animals start on PND 22 to ensure that onset of puberty in the control group has not begun (USEPA, 2009c). While the Ashby et al. (2005) study began on PND 25, it was noted that sexual maturation in the controls had not yet commenced at the start of the experiment, lending confidence that this deviation from the Tier 1 protocol likely did not impact the result. Because the Tier 1 protocol doses for 20 days, though, EPA stated that the 12-day exposure period in the Ashby et al. study was not long enough to provide confidence in the

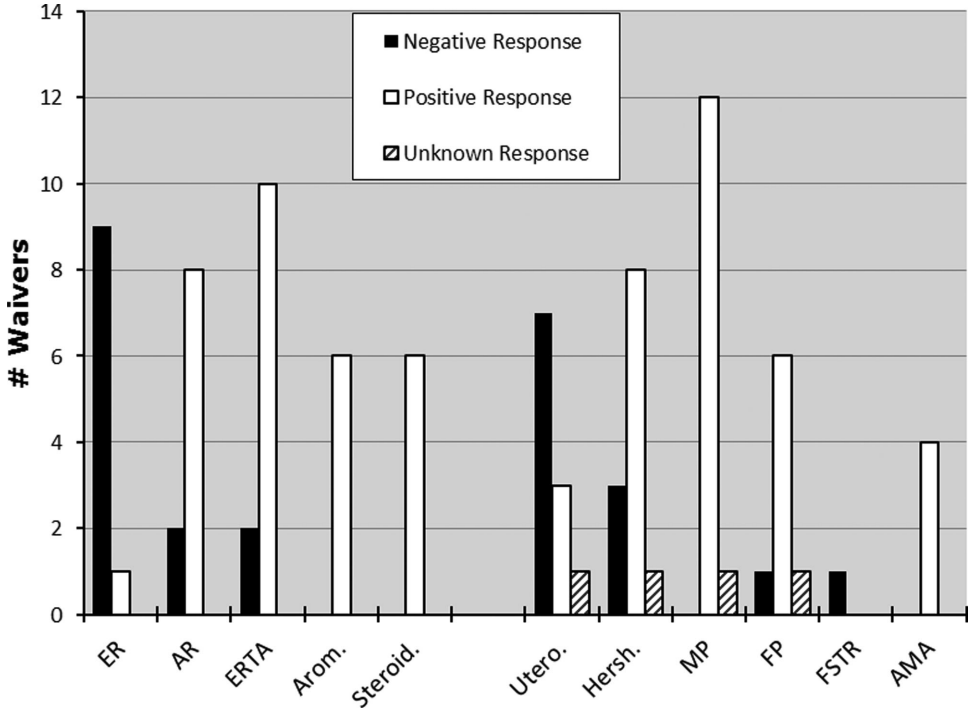

Fig. 2. Waived assays based on OSRI showing a negative, positive (potential interaction with the endocrine system), or unknown response. Note with respect to unknowns: the uterotrophic and Hershberger assays were waived for carbaryl based on two studies that were submitted separately from the respondent's OSRI report and the results for which were not described in EPA's response (USEPA, 2010b); it was unclear in EPA's response for 2,4-D with regard to the pubertal tests, whether the waivers were based on positive or negative effects.

negative result (USEPA, 2010l). It is difficult to see how a WoE approach applied to this situation could have resulted in a requirement to conduct another female pubertal assay when all ER-mediated Tier 1 assays had been waived based on OSRI showing no effect, and a female pubertal study was available, albeit one that did not use the exact same protocol as the Tier 1 assay. It was clearly stated in EPA's response (USEPA, 2010l) that the agency was requiring the assay because of uncertainty about potential effects on the estrogen system, not potential effects on the thyroid system, which had already been demonstrated in several Part 158 studies.

Use of a clear decision tree that evaluates a stated hypothesis, applies a priori criteria, and gives varying weights to different types of data (e.g., in vitro vs. in vivo) would help to avoid confusing cases like those described above. Not only would this allow EPA staff to consistently and transparently evaluate OSRI, it would provide a framework for test order respondents to prepare and present their OSRI submissions, thereby increasing the chances for acceptance.

Accepted OSRI Often Used Tier 1 Protocols or Showed a Positive Response

Despite EPA's statement in its Policies and Procedures for the first list of chemicals that, in lieu of testing, it would consider OSRI that was functionally equivalent to information obtained in Tier 1 assays, that is, data from assays that perform the same function as Tier 1 assays (74 FR 17560), studies that deviated from the Tier 1 protocols were generally not accepted unless they indicated a potential to interact with the endocrine system. With the exception of the ER binding and the uterotrophic assays, most waivers were based on observation of a positive effect, and overall, waivers based on OSRI showing a positive response were much more common than waivers based on a negative response (Fig. 2). In the relatively few cases where a waiver was granted based on a chemical displaying negative results, the assay methodology was almost always the same as the Tier 1 assay, or the study was a Tier 1 validation study. Six of the nine negative studies, and the one positive study, accepted to waive the ER binding assay used an RUC method as called for in the Tier 1 method (USEPA, 2009b). Six of the seven uterotrophic assay waivers showing negative effects were based on uterotrophic assays conducted using the same protocol as the Tier 1 test (USEPA, 2009e) or one very similar. The remaining uterotrophic assay waiver showing negative effects and the three waivers with positive results were based on studies conducted using the current Part 158 two-generation reproduction toxicity test guideline (USEPA, 1998), the presumptive Tier 2 reproduction assay, or the EOGRTS method (OECD, 2011). Similarly, of the three waived Hershberger assays with negative results, one was based on a study that used the current two-generation reproduction toxicity guideline (USEPA, 1998), one was based on results from an EOGRTS study, and one was based on results from two Hershberger studies using the same, or a very

similar, protocol as the Tier 1 assay (USEPA, 2009d). One of each of the two waivers associated with negative AR binding and ERTA results were based on Tier 1 assay validation studies. All waivers for the AMA, aromatase, and steroidogenesis assays were based on positive responses. For the male pubertal, 12 waivers were based on positive results for either the androgen or thyroid hormone systems or both; however, in the case of one chemical that was tested using the EOGRTS method (OECD, 2011), it was unclear whether EPA considered the reported thyroid effects to be positive or artifacts of other effects. Similarly, for the female pubertal, six waivers were based on positive findings for the estrogen or thyroid hormone systems, or both; one waiver was granted for negative results based on the current two-generation reproduction toxicity test guideline (USEPA, 1998), and for the same chemical noted above that was tested using the EOGRTS method (OECD, 2011), it was again unclear whether EPA considered the reported thyroid effects to be positive or artifacts of other effects.

Now that the Tier 1 results from the List 1 chemicals are available for comparison to submitted OSRI, EPA can reevaluate the types of studies and data it considers "functionally equivalent" and broaden its acceptance, where possible, of study methodologies to those that may deviate somewhat from Tier 1 assays but provide equivalent information for decision making. Again, identification of the types of studies EPA considers acceptable for submission as OSRI *beforehand* will assist both test order respondents in preparing their OSRI reports and EPA staff reviewing those reports, and lead to a more efficient and transparent process.

OSRI Not Utilized to its Fullest Extent

Despite language in the laws creating the EDSP that provides for the use of OSRI, direction by OMB in its ICR TOC to accept OSRI to the greatest extent possible, and EPA's own discussion of potentially relevant studies and possible methods for assessing their quality in its EDSP Policies and Procedures and earlier guidance documents, the final WoE document issued in September 2011 (USEPA, 2011d) was clear in stating that Tier 1 tests must be conducted and their results used for deciding whether a chemical has the potential to interact with the endocrine system. Our analysis shows that most OSRI accepted in lieu of new testing indicated either a positive result or the study method used was identical or very similar to a Tier 1 or Tier 2 assay, substantiating that EPA did, in fact, rely heavily on new testing and apply a rather narrow approach to acceptance of OSRI. It is likely that a consistently applied WoE approach, rather than an endpoint check-box approach, would have resulted in more waivers based on OSRI, and further reduced the number of tests performed and animals used. Borgert et al. (2011) in their evaluation of the Tier 1 screening battery and interpretation of results, stress the importance of considering all data relevant to evaluating EAT modes of action, particularly effects on reproduction and development, which are the health endpoints of greatest concern for EAT pathways. These endpoints have to a large extent already been evaluated for the List 1 chemicals through the pesticide registration process and, thus, existing data would seem to be essential for informing the decision of whether or not a chemical has the potential to interact with the endocrine system. However, as Borgert et al. (2011) point out, a clear stepwise approach to evaluating OSRI, which is objective and transparent, should be applied along with consideration of whether the totality of the data available is consistent with endocrine activity for each particular hormonal pathway. While it is understandable that the agency wants to be conservative and identifies all potential endocrine disruptors, there were many cases where the existing information appeared to provide absolutely no basis for requiring further testing because either the chemical's mode of action was already known, and/or higher tiered in vivo studies had shown no indication of an endocrine effect, and/or one or more functionally equivalent assays had already been performed with negative results. Once the results of the Tier 1 testing that EPA required become publicly available, it will prove informative to retrospectively compare these results to the submitted OSRI to determine how often the two data sources disagreed.

The apparent requirement in the final WoE guidance that Tier 1 tests must be conducted, and the low acceptance rate of OSRI for the data-rich List 1 chemicals, directly contradicts OMB's TOC charge to accept OSRI to satisfy test orders to the greatest extent possible. The demonstrated check-box approach to OSRI and apparent inconsistencies in EPA's evaluation of OSRI must be addressed before the next list of chemicals to ensure that duplicative testing is avoided and existing data are used fully to inform decisions. EPA should consider changing from a two-tiered strategy to a chemical-specific, integrated testing strategy that takes full advantage of OSRI before any new testing is considered. Willett et al. (2011), using the chemical atrazine as an example, demonstrated how a multilevel testing framework and an iterative WoE analysis conducted at each level could reduce the number of tests required under the EDSP, thus saving money and animal lives. Before any new testing is performed, this type of approach utilizes all available information, such as physicochemical properties, structure–activity relationships, existing in vitro and in vivo test data, and known toxicity modes of action; establishes defined criteria for prioritization; and establishes clear off-ramps based on information goals.

Future Use of OSRI in the EDSP

Even with EPA's planned integration of computational toxicology methods and high-throughput in vitro assays into EDSP for prioritization and screening purposes, as described in the agency's EDSP21 work plan (USEPA, 2011a) and the EDSP 5-year comprehensive management plan (USEPA, 2012a), it is essential that OSRI be utilized to the greatest extent possible to avoid duplicative data collection and reduce the number of animals used in testing and in pretesting. With more than 10,000 chemicals on EPA's recently released *Universe of Chemicals* (USEPA, 2012b) that are potentially subject to EDSP screening, existing data from guideline and literature studies will be increasingly important for informing decisions about potential endocrine activity. Many of these chemicals are pesticides, HPV chemicals, and common drugs and

supplements, which like the List 1 pesticides, already have copious amounts of information associated with them. OSRI will be invaluable for characterizing, categorizing, and, in some cases, ruling out chemicals for any further consideration as potential endocrine disruptors.

In addition, we note that test order recipients for 7 chemicals (5 pesticides and 2 inerts) did not submit any OSRI at all and conducted all 11 Tier 1 assays. For some of these chemicals, studies may have been older, containing few measurements of endocrine-sensitive endpoints, but it seems unlikely that there was not any existing information available that could have been used to seek waivers and reduce animal use. Therefore, in support of the three Rs (reduction, refinement, replacement of animals in testing), we recommend that EPA require test order respondents to investigate the use of OSRI to avoid duplicative testing to the greatest extent possible, and/or conduct its own literature search to ensure all available information is fully considered.

CONCLUSIONS

Pesticides have a wealth of data associated with them due to the extensive testing required for registration and most respondents for the first list of pesticides evaluated under the EDSP submitted OSRI in an effort to avoid duplicative testing and secure waivers for some or all Tier 1 assays. Despite indications in the guidance provided by EPA to List 1 test order recipients that Part 158 guideline studies would be considered as potential OSRI for satisfying Tier 1 EDSP data requirements, results from these studies, while frequently cited, were rarely accepted by EPA. Acceptance of literature studies using assays that were functionally equivalent to Tier 1 assays was generally erratic and low unless the study showed a potential to interact with the endocrine system. Many of the literature studies that were accepted used the same protocol as the Tier 1 assay, or a very similar one; this was particularly true when waivers were granted for chemicals showing no effects. Overall, only 23% of the requested waivers were granted, reducing the number of animals used in testing by 3325. This is a small reduction considering 27,731 animals were eventually used to test the 52 remaining List 1 chemicals and additional animals were used in preparing to run the Tier 1 assays. It is essential that EPA retrospectively evaluate its handling of List 1 OSRI in light of Tier 1 assay results on a case-by-case basis to determine if testing provided any new information, to characterize more definitively the type of data it considers to be "functionally equivalent" to the data generated in Tier 1 assays, and to revise its approach to and practices for OSRI submissions and WoE accordingly. Due to the large number of chemicals potentially subject to future EDSP testing, it is critical that EPA support and develop explicit guidance for the use of OSRI and standardize and clearly articulate its own evaluation procedures to avoid duplicative testing, even with its plans to advance and eventually incorporate 21st-century toxicity-testing tools in the EDSP.

ACKNOWLEDGMENTS

The authors would like to thank Joseph Manuppello for his assistance in compiling the data and Kristie Sullivan for her critical editorial review. The authors declare that they have no competing financial interests.

REFERENCES

Andersen HR, Cook SJ, Waldbillig D. 2002. Effects of currently used pesticides in assays for estrogenicity, androgenicity, and aromatase activity in vitro. Toxicol Appl Pharmacol 179:1–12.

Ashby J, Odum J, Burns A, Lefevre P. 2005. The reported in vitro anti-estrogen pentachloronitrobenzene enhances the estrogenic activity of estradiol in vivo in the rat. Environ Toxicol Pharma 20(2Q05):199–208.

Battelle. 2005. Draft final report WA 2–29 on multi-chemical evaluation of the short-term reproduction assay with the fathead minnow. Report to U.S. Environmental Protection Agency, Endocrine Disruptor Screening Program. Available at: http://www.docstoc.com/docs/7850282/Assay-Status-Draft-Final-Report-on-Multi-Chemical-Evaluation-of-the-Short-Term-Reproduction-Assay-with-the-Fathead-Minno. Accessed 16 March 2013.

Bauer ERS, Bitsch N, Brunn H, Sauerwein H, Meyer HHD. 2002. Development of an immuno-immobilized androgen receptor assay (IRA) and its application for the characterization of the receptor binding affinity of different pesticides. Chemosphere 46(7):1107–1115.

Blair RM, Fang H, Branham WS, Hass BS, Dial SL, Moland CL, Tong W, Shi L, Perkins R, Sheehan DM. 2000. The estrogen receptor relative binding affinities of 188 natural and xenochemicals: structural diversity of ligands. Toxicol Sci 54:138–153.

Borgert CJ, Mihaich EM, Quill TF, Marty MS, Levine SL, Becker RA. 2011. Evaluation of EPA's Tier 1 Endocrine Screening Battery and recommendations for improving the interpretation of screening results. Reg Toxicol Pharmacol 59(3):397–411.

Chen H, Xiao I, Hu G, Zhou J, Xiao H, Wang X. 2002. Estrogenicity of organophosphorus and pyrethroid pesticides. J Toxicol Environ Health A 65:1419–1435.

Crop Life America (CLA). 2010. EDSP Tier 1 screening: preliminary burden report November 2010. Endocrine Policy Forum, Crop Life America. Available at: http://www.croplifeamerica.org/sites/default/files/node_documents/EPF-Burden%20Tracking-final.pdf. Accessed 23 February 2013.

Fang H, Tong W, Branham WS, Mass BS, Moland CL, Dial SL, Hong H, Xie Q, Perkins R, Owens W, Sheehan DM. 2003. Study of 202 natural, synthetic, and environmental chemicals for binding to the androgen receptor. Chem Res Toxicol 16:1338–1358.

Hinfray N, Porcher JM, Brion F. 2006. Inhibition of rainbow trout (Oncorhynchus mykiss) P450 aromatase activities in brain and ovarian microsomes by various environmental substances. Comp Biochem Physiol C Toxicol Pharmacol 144(3):252–262.

Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, Reif DM, Rotroff DM, Shah I, Richard AM, Dix DJ. 2010. In vitro screening of environmental chemicals for targeted testing prioritization: the ToxCast project. Environ Health Persp. doi:10.1289/ehp.0901392. Available at: http://www.epa.gov/ncct/toxcast/data.html. Accessed 26 March 2013.

Kojima H, Katsura E, Takeuchi S, Niiyama K, Kobayashi K. 2004. Screening for estrogen and androgen receptor activities in 200 pesticides by in vitro reporter gene assays using Chinese hamster ovary cells. Environ Health Persp 112:524–531.

Kojima M, Fukunaga K, Sasaki M, Nakamura M, Tsuji M, Nishiyama T. 2005. Evaluation of estrogenic activities of pesticides using an in vitro reporter gene assay. Int J Environ Health Res 15:271–280.

Laws SC, Ferrell JM, Stoker TE, Schmid J, Cooper RL. 2000. The effects of atrazine on female Wistar rats: an evaluation of the protocol for assessing pubertal development and thyroid function. Toxicol Sci 58:366–376.

LeBaron M, Coady KK, Marty MS, O'Connor JC, Nabb DL, Markell LK. 2013. EDSP Tier 1 in vitro assays. Birth Defects Res B Dev Reprod Toxicol.

Lech JW. 2006. Interlaboratory validation of the 15-day adult intact male rat assay with linuron and phenobarbital. Available at: http://www.epa.gov/endo/pubs/interlab_charlesriver_final.pdf. Accessed 16 March 2013.

Lemaire G, Mnif W, Mauvais P, Balaguer P, Rahmani R. 2006. Activation of α- and β- estrogen receptors by persistent pesticides in reporter cell lines. Life Sci 79:1160–1169.

National Toxicology Program (NTP). 2013. Validation of the BG1Luc estrogen receptor transactivation test method. Available at:

http://iccvam.niehs.nih.gov/methods/endocrine/end_eval.htm. Accessed 21 March 2013.

Neal B, Bus J, Marty MS, Williams A, Coady KK. 2010. Evaluation of functional equivalence of other scientifically relevant information (OSRI) to EPA's Tier 1 screening battery of evaluation the potential estrogen, androgen or thyroid (EAT) effects of 2,4-dichlorophenoxyacetic acid (2,4-D). Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0212. Accessed 21 March 2013.

Nishihara T, Nishikawa J, Kanayama T, Dakeyama F, Saito K, Imagawa M, Takatori S, Kitagawa Y, Hori S, Utsumi H. 2000. Estrogenic activities of 517 chemicals by yeast two-hybrid assay. J Health Sci 46:282–298.

Office of Management and Budget (OMB). 2009. Tier 1 screening of certain chemicals under the Endocrine Disruptor Screening Program (EDSP), terms of clearance, OMB control no: 2070-0176, ICR reference no: 200904-2070-001, Oct. 2, 2009. Available at: http://www.reginfo.gov/public/do/PRAViewICR?ref_nbr=200904-2070-001. Accessed 26 March 2013.

Organisation for Economic Co-operation and Development (OECD). 2011. Test no. 443: extended one-generation reproductive toxicity study. Available at: http://www.oecd-ilibrary.org/environment/test-no-443-extended-one-generation-reproductive-toxicity-study_9789264122550-en. Accessed 31 January 2013.

OECD. 2013. Testing of chemicals. Available at: http://www.oecd.org/env/ehs/testing/. Accessed 13 March 2013.

Okubo T, Yokoyama Y, Kano K, Soya Y, Kano I. 2004. Estimation of estrogenic and antiestrogenic activities of selected pesticides by MCF-7 cell proliferation assay. Arch Environ Contam Toxicol 46:445–453.

Petit F, LeGoff P, Cravded JP, Valotaire Y, Pakdel F. 1997. Two complementary bioassays for screening the estrogenic potency of xenobiotics; recombinant yeast for trout estrogen receptor and trout hepatocyte cultures. J Mol Endocrinol 19:321–335.

Stoker TE, Laws SC, Guidici DL, Cooper RL. 2000. The effect of atrazine on puberty in male Wistar rats: an evaluation in the protocol for the assessment of pubertal development and thyroid function. Toxicol Sci 58:50–59.

U.S. Environmental Protection Agency (USEPA). 1998. Health effects test guidelines: OPPTS 870.3800 reproduction and fertility effects. EPA Office of Prevention, Pesticides and Toxic Substances. 712-C-98-208, August 1998. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2009-0156-0018. Accessed 26 March 2013.

USEPA. 2002. Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by the Environmental Protection Agency, EPA 260R-02-008, Office of Environmental Information. Available at: http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf. Accessed 26 March 2013.

USEPA. 2003. A summary of general assessment factors for evaluating the quality of scientific and technical information. Science Policy Council, Washington DC. EPA/100/B-03/001. Available at: http://www.epa.gov/osa/stpc/pdfs/assess3.pdf. Accessed 26 March 2013.

USEPA. 2007. Integrated summary report for the validation of an androgen receptor (AR) binding assay as a potential screen in the EDSP. US Environmental Protection Agency, Nov. 7, 2007. Available at: http://www.epa.gov/scipoly/oscpendo/pubs/ar_binding_isr.pdf. Accessed 14 February 2013.

USEPA. 2009a. Endocrine Disruptor Screening Program test guidelines—OPPTS 890.1150: androgen receptor binding (rat prostate cytosol) [EPA 640-C-09-003], October 2009. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2009-0576-0003. Accessed 26 March 2013.

USEPA. 2009b. Endocrine Disruptor Screening Program test guidelines—OPPTS 890.1250: estrogen receptor binding assay using rat uterine cytosol (ER-RUC), EPA 740-C-09-005, October 2009. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2009-0576-0005. Accessed 16 March 2013.

USEPA. 2009c. Endocrine Disruptor Screening Program test guidelines—OPPTS 890.1450: pubertal development and thyroid function in intact juvenile/peripubertal female rats [EPA 740-C-09-009], October 2009. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2009-0576-0009. Accessed 26 March 2013.

USEPA. 2009d. Endocrine Disruptor Screening Program test guidelines—OPPTS 890.1400: Hershberger Bioassay [EPA 740-C-09-008], October 2009. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2009-0576-0008. Accessed 26 March 2013.

USEPA. 2009e. Endocrine Disruptor Screening Program test guidelines—OPPTS 890.1600: uterotrophic assay [EPA 740-C-09-0010], October 2009. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2009-0576-0012. Accessed 26 March 2013.

USEPA. 2009f. EPA's approach for considering Other Scientifically Relevant Information (OSRI) under the Endocrine Disruptor Screening Program. March 17, 2009. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2007-1080-0032. Accessed 26 March 2013.

USEPA. 2010a. Atrazine—report of the OPP endocrine disruptor review team—test order # EDSP-080803-1; -2; -3; and -4. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0215. Accessed 21 March 2013.

USEPA. 2010b. Carbaryl—Report of the OPP endocrine disruptor review team—test order # EDSP-056801-57. Available at: http://www.regulations.gov/search/Regs/contentStreamer?objectId=0900006480b7d6b5&disposition=attachment&contentType=pdf. Accessed 16 March 2013.

USEPA. 2010c. Carbofuran—report of the OPP endocrine disruptor review team—test order # EDSP-090601-61. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0199. Accessed 16 March 2013.

USEPA. 2010d. Cypermethrin—report of the OPP endocrine disruptor review team–test order # EDSP-109702-75; -76, and -78. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009--0634-0198. Accessed 21 March 2013.

USEPA. 2010e. Dicofol—report of the OPP endocrine disruptor review team—test order # EDSP-010501-101. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0180. Accessed 16 March 2013.

USEPA. 2010f. Draft weight of evidence (WoE) guidance: evaluating results of EDSP Tier 1 screening to identify candidate chemicals for Tier 2 testing. Available at: http://www.epa.gov/endo/pubs/EDSP_Draft_WoE_Paper_for_Public_Comments_2010.pdf. Accessed 15 March 2013.

USEPA. 2010g. Endosulfan—report of the OPP endocrine disruptor review team—test order # EDSP-079401-102. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0217. Accessed 21 March 2013.

USEPA. 2010h. Esfenvalerate–report of the OPP endocrine disruptor review team—test order # EDSP-109303-80; -81; and -82. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009--0634-0188. Accessed 21 March 2013.

USEPA. 2010i. Metalaxyl—report of the OPP endocrine disruptor review team—test order # EDSP-113501-124 – EDSP-113501-126. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0184. Accessed 21 March 2013.

USEPA. 2010j. Metribuzin–report of the OPP endocrine disruptor review team–test order # EDSP-101101-127. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0164. Accessed 16 March 2013.

USEPA. 2010k. Permethrin–report of the OPP endocrine disruptor review team—test order # EDSP-109701-83; -84; -85; -87; -89; -90; and -91. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0186. Accessed 21 March 2013.

USEPA. 2010l. Pentachloronitrobenzene (PCNB)–report of the OPP endocrine disruptor review team—test order # EDSP-056502-105. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0192. Accessed 21 March 2013.

USEPA. 2010m. Propiconazole—report of the OPP endocrine disruptor review team—test order # EDSP-122101-133. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0172. Accessed 21 March 2013.

USEPA. 2010n. Simazine—report of the OPP endocrine disruptor review team—test order # EDSP-080807-130; -131; and -132. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2009-0634-0230. Accessed 21 March 2013.

USEPA. 2011a. Endocrine Disruptor Screening Program for the 21st century (EDSP21 work plan): the incorporation of in silico models and in vitro high throughput assays in the Endocrine Disruptor Screening Program (EDSP) for prioritization and screening. Available at: http://www.epa.gov/endo/pubs/edsp21_work_plan_summary%20_overview_final.pdf. Accessed 16 March 2013.

USEPA. 2011b. Status of EDSP OSRI: order recipient submissions and EPA responses. Available at: http://www.epa.gov/endo/pubs/EDSP_OSRI_Response_Table.pdf. Accessed 15 March 2013.

USEPA. 2011c. Tier 1 screening battery. Available at: http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/tier1battery.htm. Accessed 26 March 2013.

USEPA. 2011d. Weight-of-evidence: evaluating results of EDSP Tier 1 screening to identify the need for Tier 2 testing. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2010-0877-0021. Accessed 15 March 2013.

USEPA. 2012a. Endocrine Disruptor Screening Program Comprehensive Management Plan, June 2012. Available at: http://www.epa.gov/endo/pubs/EDSP-comprehensive-management-plan.pdf. Accessed 22 March 2013.

USEPA. 2012b. Endocrine Disruptor Screening Program universe of chemicals. Available at: http://www.epa.gov/endo/pubs/edsp_chemical_universe_list_11_12.pdf. Accessed 22 March 2013.

USEPA. 2012c. Good laboratory practice standards. Available at: http://www.epa.gov/oecaerth/monitoring/programs/fifra/glp.html. Accessed 8 February 2013.

USEPA. 2013a. Data requirements for pesticide registration. Available at: http://www.epa.gov/opp00001/regulating/data_requirements.htm. Accessed 26 March 2013.

USEPA. 2013b. High Production Challenge Program. Available at: http://www.epa.gov/HPV/. Accessed 21 August 2013.

USEPA. 2013c. OCSPP harmonized test guidelines, series 850—ecological effects test guidelines. Available at: http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series850.htm. Accessed 31 January 2013.

USEPA. 2013d. OCSPP harmonized test guidelines, series 870—health effects test guidelines. Available at: http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series870.htm. Accessed 31 January 2013.

USEPA. 2013e. U.S. Environmental Protection Agency Endocrine Disruptor Screening Program SAP review of EDSP Tier 1 screening: assay and battery performance, May 21–23, 2013. Available at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2013-0075-0003. Accessed 20 May 2013.

Vinggaard AM, Breinholt V, Larsen JC. 1999. Screening of selected pesticides for oestrogen receptor activation in vitro. Food Addit Contam 16(12):533–542.

Waller CL, Juma BW, Gray LE Jr, Kelce WR. 1996. Three-dimensional quantitative structure-activity relationships for androgen receptor ligands. Toxicol Appl Pharmacol 137:219–227.

Willett CE, Bishop PL, Sullivan KM. 2011. Application of an integrated testing strategy to the U.S. EPA Endocrine Disruptor Screening Program. Toxicol Sci 123(1):15–25.

Zhang J, Zhu W, Zheng Y, Yang J, Zhu X. 2008. The antiandrogenic activity of pyrethroid pesticides Cyfluthrin and beta-cyfluthrin. Reprod Toxicol 25:491–496.