Integrated Environmental Assessment and Management — Volume 13, Number 2—pp. 317–330
Received: 8 August 2016 | Revised: 21 September 2016 | Accepted: 1 December 2016

Special Series

Population-Relevant Endpoints in the Evaluation of Endocrine-Active Substances (EAS) for Ecotoxicological Hazard and Risk Assessment

Mary S Marty,† Amy Blankinship,‡ Janice Chambers,§ Lisa Constantine,‖ Werner Kloas,# Anupama Kumar,†† Laurent Lagadic,‡‡ James Meador,§§ Daniel Pickford,‖‖ Tamar Schwarz,## and Tim Verslycke†††*

†Dow Chemical Company, Midland, Michigan, USA
‡US Environmental Protection Agency, Washington, DC
§Mississippi State University, Mississippi State, Mississippi, USA
‖Pfizer, New York City, New York, USA
#Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany
††Commonwealth Scientific and Industrial Research Organization (CSIRO), Adelaide, South Australia
‡‡Bayer AG CropScience Division, Monheim am Rhein, Germany
§§NOAA Fisheries, Northwest Fisheries Science Center, Seattle, Washington, USA
‖‖Syngenta UK, Cambridge, United Kingdom
##Centre for Environment Fisheries and Aquaculture Science (Cefas), Weymouth, Dorset, United Kingdom
†††Gradient, Cambridge, Massachusetts, USA

EDITOR'S NOTE:

This is 1 of 5 articles generated from the SETAC Pellston Workshop "Ecotoxicological Hazard and Risk Assessment Approaches for Endocrine-Active Substances (EHRA)" (February 2016, Pensacola, Florida, USA). The primary aim of the workshop was to provide objective advice, based on current scientific understanding, to regulators and policy makers, whether in industry, government, or academia. The goal is to make considered, informed decisions on whether to select an ecotoxicological hazard- or risk-based approach for regulating a given endocrine disrupting substance under evaluation.

ABSTRACT

For ecotoxicological risk assessment, endocrine disruptors require the establishment of an endocrine mode of action (MoA) with a plausible link to a population-relevant adverse effect. Current ecotoxicity test methods incorporate mostly apical endpoints although some also include mechanistic endpoints, subcellular-through-organ level, which can help establish an endocrine MoA. However, the link between these endpoints and adverse population-level effects is often unclear. The case studies of endocrine-active substances (EAS) (tributyltin, ethinylestradiol, perchlorate, trenbolone, propiconazole, and vinclozolin) from the Society of Environmental Toxicology and Chemistry (SETAC) Pellston Workshop® "Environmental Hazard and Risk Assessment Approaches for Endocrine-Active Substances (EHRA)" were used to evaluate the population relevance of toxicity endpoints in various taxa according to regulatory endocrine-disruptor frameworks such as the Organisation for Economic Co-operation and Development (OECD) Conceptual Framework for Testing and Assessment of Endocrine Disruptors. A wide variety of potentially endocrine-relevant endpoints were identified for mollusks, fish, amphibians, birds, and mammals, although the strength of the relationship between test endpoints and population-level effects was often uncertain. Furthermore, testing alone is insufficient for assessing potential adaptation and recovery processes in exposed populations. For this purpose, models that link effects observed in laboratory tests to the dynamics of wildlife populations appear to be necessary, and their development requires reliable and robust data. As our understanding of endocrine perturbations and key event relationships improves, adverse population-level effects will be more easily and accurately predicted. *Integr Environ Assess Manag* 2017;13:317–330. © 2017 The Authors. *Integrated Environmental Assessment and Management* published by Wiley Periodicals, Inc. on behalf of Society of Environmental Toxicology & Chemistry (SETAC)

Keywords: Population Fish Amphibian Bird Modeling

* Address correspondence to msmarty@dow.com

Published 4 January 2017 on wileyonlinelibrary.com/journal/ieam.

INTRODUCTION

Although testing to support ecotoxicological hazard or risk assessments is usually conducted at the level of the individual, the goal of these assessments typically is to protect the status of nontarget wildlife (i.e., terrestrial and aquatic) populations. This is consistent with the World Health Organization–International Programme on Chemical Safety (WHO/IPCS) definition of an endocrine disruptor, which implies that there must be reasonable evidence for a biologically plausible causal relationship between the endocrine activity and the induced adverse effects seen in an intact organism or a (sub) population (WHO/IPCS 2002). Therefore, it is important to establish that adverse effects observed in experimental test animals are mediated by an endocrine mode of action (MOA) and relevant to populations (EFSA 2013).

A number of toxicity assays include mechanistic endpoint data that are critical to identify whether a chemical exhibits a potential endocrine-mediated hazard. However, without a causal link to significant changes in population-relevant endpoints such as survival or reproduction (apical), mechanistic endpoint measures do not necessarily establish adversity. The relationship between mechanistic and apical endpoints is being defined using approaches such as adverse outcome pathways (AOP) (Kramer et al. 2011; Coady et al. this issue; Parrot et al. this issue). Ideally, an understanding of the dose–response relationship between the response variable and how it relates to a population-relevant outcome would be quantifiable.

Furthermore, the application of apical endpoints that measure individual fitness to models of population-level effects requires careful consideration of relationships among endpoints. For instance, apical endpoints related to survival and reproduction can be applied directly to population modeling. Other types of endpoints, such as measures of growth, behavior, development, and immune function, sometimes provide relevant information and can be linked to survival and reproduction using empirically based mathematical relationships. These data can feed into population models (Kramer et al. 2011). Finally, many endpoints measured at the cellular, tissue, or organ level can provide insights into chemical MOA but, at this time, generally cannot be linked readily (quantitatively) to other population-relevant endpoints.

Reversibility is particularly important when considering endocrine-mediated endpoint data in the context of evaluating population effects. The compensatory feedback mechanisms that typify endocrine systems can provide homeostatic capacity against various endocrine perturbations. Exposure to endocrine-active substances (EAS) may stimulate modulation in these feedback systems. If this modulation is temporary and/or within the homeostatic capacity of the endocrine system of the exposed organism, the effect of the substance on a certain endpoint might be considered "endocrine modulation" (EFSA 2013). Alternatively, if the body is unable to compensate for the induced changes within its limits of homeostasis, the threshold of adversity is crossed, and the observed changes are considered adverse.

While other chemicals have been shown to have adverse effects on wildlife populations from reproductive effects (e.g., DDT; Vos et al. 2000), the objective of the present paper is to illustrate strengths and limitations of invertebrate, fish, amphibian, avian, and mammalian toxicity endpoint data that may be collected to support ecotoxicological hazard and risk assessment of EAS, particularly where endocrine-related responses have population-relevant consequences in the field. Nonmammalian and mammalian toxicology data from 6 different EAS case studies were used to illustrate the issues of population relevance and reversibility of endpoint data. Discussion on how these factors may differentially affect the use of these data for hazard identification or risk assessment also is included. In addition, areas of ongoing research and critical data gaps are briefly reviewed.

These objectives were addressed using 6 case study chemicals evaluated for the Society of Environmental Toxicology and Chemistry (SETAC) Pellston Workshop® "Environmental Hazard and Risk Assessment Approaches for Endocrine-Active Substances (EHRA)" (Matthiessen et al. this issue). The 6 case study chemicals, each having a different endocrine MOA, were:

- 17α-ethinylestradiol (EE2), a synthetic steroid estrogen that acts via estrogen-receptor (ER) agonism.
- Perchlorate (PRC), a naturally occurring and manufactured anion, can competitively inhibit the sodium-iodide symporter (NIS) by blocking iodide uptake in the thyroid follicle, thereby lowering thyroid hormone production.
- Propiconazole (PPZ), a fungicide, can inhibit cytochrome P450 enzymes, which may alter steroid hormone levels.
- Tributyltin (TBT), a biocide formerly used in antifouling products, parasite control, and wood preservatives, is a ligand for the retinoid X receptor (RXR).
- Trenbolone (TRB) is a synthetic anabolic agent that mimics the actions of natural steroids (androgen agonist).
- Vinclozolin (VZ), a fungicide, is transformed into metabolites that are primarily androgen-receptor antagonists.

The details of study results for many of the chemical-specific examples cited here are provided in the Supplemental Data in the SETAC Pellston workshop overview paper (Matthiessen et al. this issue).

Use of screening data

In linking population-level effects to endocrine MOAs, study design and purpose should be considered in evaluating data utility. For example, numerous screening assays are available to evaluate potential endocrine activity (USEPA 2009a; OECD 2012). Although effects on apical endpoints have been reported in in-vivo screening assays, the use of these endpoints to infer population-level effects in a hazard or risk assessment warrants caution. Typically, in these studies, the dose spacing is quite large (to examine a broad range of exposure concentrations) and as such, the potential

to discern a dose–response relationship is limited (Coady et al. this issue). Dose levels may be higher in a screening assay to achieve a maximum tolerated dose over the limited duration of the assay, which may increase the potential for confounding effects (e.g., systemic toxicity) when evaluating whether the effects on apical endpoints are due to endocrine interactions or another MOA. This is a critical distinction if using effects on apical endpoints to conclude a potential adverse population effect that is endocrine mediated (Mihaich et al. this issue).

Ecotoxicity assays within endocrine-disruptor assessment frameworks

For many taxa, several different assay types are available for evaluating the toxic effect of a substance from the screening level to an entire life cycle. Each assay is designed to inform a particular set of questions with varying degrees of resolution about a substance's ability to interact with an endocrine pathway. Some ecotoxicity assays, which are used to support registration of plant protection products, measure apical endpoints of survival, growth, and/or reproduction but do not incorporate endocrine-relevant diagnostic endpoints. Other available assays, which are listed in the Organisation for Economic Co-operation and Development (OECD 2012) Conceptual Framework for Testing and Assessment of Endocrine Disruptors as well as the US Environmental Protection Agency (USEPA 2009a) Endocrine Disruptor Screening Program (EDSP) approach, include screening assays and longer-term tests. Screening assays (Level 3) are used to evaluate in-vivo mechanistic effects, whereas long-term in-vivo assays evaluate population-relevant endpoints and include endocrine-relevant diagnostic endpoints (for additional details concerning the 2 regulatory frameworks see Coady et al. this issue). Table 1 lists the available study types by taxa, along with their primary endpoints, and highlights the endpoints that are considered to be population relevant (surrogates) but not necessarily exclusively endocrine mediated. Note that the relevance of some endpoints to population-level effects is likely based on certain life-history strategies. For example, the impact of growth changes on biological processes is species dependent; reproduction is size based in some species, whereas in others, it is not. The impact on growth is further examined when evaluating effects on fish populations.

Invertebrates

Retinoid X receptor-mediated effects in mollusks exposed to TBT. Tributyltin has been shown to have endocrine-mediated population effects in both the laboratory and the field. Before it was banned, TBT was used mainly as a biocide in antifouling paints applied to ships' hulls, resulting in contaminated sediments. Tributyltin is an example of an environmental endocrine disruptor that is known to cause adverse effects; in mollusks, interaction of TBT with RXR seems to be the main initiating event for changes in the development of sexual organs in female snails, ultimately resulting in imposex (Castro et al. 2007). Other primary molecular mechanisms have been suggested, including the peroxisome proliferator–activated receptor (PPAR) pathways (Iguchi and Katsu 2008; Pascoal et al. 2013). It has been shown that *cis*-9-retinoic acid and rosiglitazone, which are known to bind vertebrate RXR and PPARγ, respectively, each induced imposex separately in prosobranch mollusks (Castro et al. 2007; Pascoal et al. 2013). The gonadotropin-releasing hormone (GnRH), its receptor (Castro et al. 2007), and the balance between free and esterified testosterone (LeBlanc et al. 2005) may also be involved in imposex, but their role and mechanistic links with molecular initiating events remain to be elucidated.

Alterations in reproductive organ responses in mollusks exposed to low concentrations of TBT (1–10 ng TBT/L and 10–100 ng TBT/g wet weight whole body tissue) have been observed, including penis development in female snails, abnormal testis histopathology, and sperm alterations (count, motility, morphology) (Horiguchi et al. 1994; Meador 2011). The development of male sexual organs, including penis, in TBT-exposed females follows a stepped process (i.e., imposex) that has been described in a number of stenoglossan snails (Gibbs et al. 1988; Gooding et al. 2003). Several studies indicate the threshold concentration for imposex induction starts at 1 ng TBT/L, with increasing sterilization as concentrations increase (Gibbs et al. 1988). Several studies indicate that TBT in the marine environment can impact populations of stenoglossan snails through female sterilization associated with imposex (Spence et al. 1990; Bailey et al. 1995). These population responses were associated with water concentrations in the 1 to 10 ng TBT/L, which is consistent with molecular studies characterizing the affinity of TBT for the RXR-PPARγ receptor. It is important to note that there is not always a linear relationship between imposex development and female sterility (Barroso et al. 2002); for example, female gametogenic activity is not affected by the occurrence of imposex in populations of *Buccinanops globulosus* in northern Patagonia, South America (Avaca et al. 2015). This shows that adverse effects in individuals do not necessarily translate into population-level changes.

The degree of imposex reversibility depends on the snail species. Some studies have shown that the imposex response in *Nucella lapillus* is largely irreversible for individuals (Bryan et al. 1993); however, snail populations have recovered worldwide after the reduction in use of TBT as an antifoulant (Birchenough et al. 2002; Matthiessen 2013; Nicolaus and Barry 2015). Similarly, Birch et al. (2014) found a correlation between the recovery of rock oyster (*Saccostrea glomerata*) populations and the reduction of TBT in estuaries with high densities of boat moorage. This conclusion is supported by additional studies that observed declining tissue concentrations in mollusks from this area over the same time period (Batley et al. 1992; Lewis et al. 2010).

Fish. For fish, additional non-apical endpoints (i.e., vitellogenin [VTG], secondary sex characteristics [SSC], gonadal histopathology) are included in several of the available tests to better elucidate a potential endocrine-mediated MOA. However, while work is ongoing (Watanabe et al. 2016), further validation

 DOI: 10.1002/ieam.1887

Table 1. Available in-vivo ecotoxicology tests that include endocrine- and/or population-relevant endpoints

Test title	Reference[b]	OECD framework level and USEPA EDSP tier[c]	Primary measured endpoints[d]
Fish			
Early life cycle (ELS)	TG 210; 850.1400	1	Growth[a] (body weight, length)
			Hatching success[a]
Fish short-term reproduction assay (FSTRA)	TG 229; 890.1350	3, 1	Fecundity[a]
			Growth[a] (body weight, length)
			Gonadosomatic index (GSI)
			Vitellogenin (VTG)
			Secondary sexual characteristics (SSC)
			Sex steroids (optional)
Fish screening assay	TG 230; NA	3	VTG and SSC
Androgenized female stickleback screening assay	GD 140	3	Spiggin production
Fish sexual development test (FSDT)	TG 234; NA	4	Phenotypical sex ratio[a]
			VTG
			Growth[a] (body weight, length)
			Hatching success[a]
Fish full life-cycle test (FFLC)	NA; 850.1500	5	Fecundity[a]
			Growth[a] (body weight, length)
			Hatching success[a]
			Offspring survival[a]
Medaka extended 1-generation reproduction test (MEOGRT)	TG 240; 890.2200	5, 2	Fecundity[a]
			Growth[a] (body weight, length)
			Sex ratio[a]
			Hatching success[a]
			Offspring survival[a]
			VTG
			Gonadal histopathology also includes histopathology on other organs (e.g., liver)
			SSC
			Sex steroids
Amphibians			
Amphibian metamorphosis assay (AMA)	TG 231; 890.1100	3, 1	Metamorphosis (development)[a]
			Thyroid histopathology
			Normalized hind-limb length
			Growth[a] (body weight and snout-vent length)
Larval amphibian growth and development assay (LAGDA)	TG 241; 890.2300	4, 2	Metamorphosis (development)[a]

(*Continued*)

Table 1. (*Continued*)

Test title	Reference[b]	OECD framework level and USEPA EDSP tier[c]	Primary measured endpoints[d]
			Sex ratio[a]
			Growth[a] (body weight and snout-vent length)
			Thyroid and gonadal histopathology also included histopathology on other organs (e.g., liver)
			Liver –somatic index (LSI)
			VTG (optional)
Birds			
Avian reproduction test	TG 206; 850.2300	4	Fecundity[a]
			Eggshell thickness or cracked eggs[a]
			Embryo viability and hatching[a]
			Offspring survival and weight[a]
			Parental body weight[a]
Avian 2-generation test with Japanese quail	NA; 890.2100	5, 2	Fecundity[a]
			Eggshell thickness or cracked eggs[a]
			Sex ratio[a]
			Embryo viability and hatching[a]
			Offspring survival and weight[a]
			Parental body weight[a]
			Time to sexual maturation[a]
			Histopathology
			Sex steroids
			Thyroid hormones
Aquatic invertebrates[e]			
Daphnia magna life-cycle test	TG 211; 850.1300	4	Nr offspring[a]
			Growth[a] (body weight and length)
Mysid shrimp life-cycle test[f]	NA; 850.1350	4	Nr offspring[a]
			Growth[a] (body weight and length)
Chironomid toxicity	TG 218, 219; 850.1735	4	Emergence[a]
			Sex ratio[a]
			Weight[a] (dry) (only for 850.1735)
Lumbriculus toxicity test using spiked sediment	TG 225; NA	4	Total nr organisms[a]
			Weight[a] (dry)
Mammals[g,h]			
Uterotrophic assay	TG 440; 890.1600	3, 1	Uterine weight
Hershberger assay	TG 441; 890.1400	3, 1	Weight of secondary sex organs

(*Continued*)

Table 1. (*Continued*)

Test title	Reference[b]	OECD framework level and USEPA EDSP tier[c]	Primary measured endpoints[d]
Subchronic			
Female pubertal assay	See GD 150; 890.1450	4, 1	Evaluation of vaginal opening and estrous cyclicity
			Thyroid hormones
			Reproductive organ weight
Male pubertal assay	See GD 150; 890.1500	4, 1	Evaluation of preputial separation
			Thyroid hormones
			Reproductive organ weight
Repeat dose (oral toxicity in rodents and nonrodents)	TG 407-8; 870. 3050, 3100, 3150	4	Body weight[a]
Reproductive or developmental toxicity screening test	TG 421; 870.3550	4	Reproductive success[a]
			Parental body weight[a]
			Offspring viability (survival and weight)[a]
			Organ weights
			Histopathology
Combined repeated dose toxicity study with reproduction or developmental toxicity screening test	TG 422	4	Reproductive success[a]
			Parental body weight[a]
			Offspring viability (survival and weight)[a]
			Behavioral measurements (potentially population-relevant)
			Organ weights
			Histopathology
Prenatal development	TG 414; 870.3700	4	Body weight[a]
			Developmental stage[a]
Developmental neurotoxicity	TG 426; 870.6300	4	Body weight[a]
			Behavioral measurements (potentially population-relevant)
			Organ weight
			Hormones
			Histopathology and/or morphometry
Chronic			
Chronic toxicity or carcinogenicity	TG 451-3; 870.4200, 4100	4	Body weight[a]
			Histopathology
Extended 1-generation reproduction test	TG 443; NA	5, 2	Reproductive success[a]
2-generation reproduction test	TG 416; 870.3800 (post 1998)	5, 2	Parental body weight[a]

(*Continued*)

Table 1. (*Continued*)

Test title	Reference[b]	OECD framework level and USEPA EDSP tier[c]	Primary measured endpoints[d]
			Offspring viability (survival and weight)[a]
			Organ weights
			Histopathology
			Behavioral measurements (potentially population-relevant)

EAS = endocrine-active substances; EDSP = USEPA Endocrine Disruptor Screening Programs; GD = guidance document; OCSPP = USEPA Office of Chemical Safety and Pollution Prevention; OECD = Organisation for Economic Co-operation and Development; TG = test guideline; USEPA = US Environmental Protection Agency.

[a] Endpoints generally considered population relevant, not necessarily endocrine specific. Individual-level apical endpoints generally form the basis for inferring adversity at the population level.

[b] First test guideline (TG) or guidance document (GD) number is for the OECD guidelines (available at: http://www.oecd-ilibrary.org/content/package/chem_guide_pkg-en); second TG number is for USEPA OCSPP guidelines. (available at https://www.epa.gov/test-guidelines-pesticides-and-toxic-substances/series-890-endocrine-disruptor-screening-program)

[c] First number is for the OECD framework level; second number is for the USEPA EDSP tier number when it is a relatively straightforward comparison. It is acknowledged that not all listed assays have a corresponding USEPA EDSP guideline number, but in general, OECD framework Levels 1 to 3 correspond to USEPA EDSP Tier 1, and OECD Level 5 corresponds to USEPA EDSP Tier 2. The placement of an OECD Level 4 test in the 2-tiered USEPA EDSP approach depends on the assay and measured endpoint because many assays may be used in the Tier 1 screen but may also provide relevant information in a Tier 2 evaluation (i.e., fish full life-cycle [OCSPP Guideline 850.1500]).

[d] All study types monitor survival, but these endpoints are typically not regarded as primary endpoints because concentrations or doses are intended to be optimized to evaluate sublethal effects. However, effects on survival are considered to be population relevant.

[e] At present, the available invertebrate assays solely involve apical endpoints that are able to respond to both EAS and non-EAS.

[f] A mysid multigenerational toxicity study (as well as a copepod reproduction test) was developed to support the USEPA EDSP but was not included as a finalized guideline.

[g] Many mammalian guideline studies are available to support human health risk assessments, and contain many different types of endocrine-relevant endpoints. Listed here are several that can be used to support evaluation of EAS. Due to the number of measured endpoints, specific endpoints are not listed here but rather types of endpoints measured are given.

[h] Although there are many different endpoints measured in subchronic and chronic mammalian assays that could be considered to be population relevant (i.e., body weight), before using in an ecological risk assessment to represent a population-relevant endpoint, consideration of the route of exposure should be given. For example, exposure via the diet is generally considered an environmentally relevant route, whereas repeat dose by oral gavage over a protracted time may not be as environmentally relevant.

is needed before these endpoints can be used as surrogates for population-relevant apical effects (Kramer et al. 2011).

The relationship between individual-level effects and potential population-level effects, as well as how these additional endocrine-mediated biomarkers relate to apical effects, were examined using data from the case study chemicals. For example, studies with TRB have reported alterations in sexual differentiation, which resulted in skewed sex ratio and all-male populations (e.g., Holbech et al. 2006; Morthorst et al. 2010). With regard to relevance to population-level impacts, in a study by Davis et al. (2000), treatment of catfish fry (sex undifferentiated at this age) with TRB acetate resulted in an all-male population (determined by gross pathology). In this study, mating with untreated females resulted in spawns, but the eggs were not viable (unfertilized). Other studies also reported masculinization that occurred at the same concentrations in which reduced fecundity was noted (Ankley et al. 2003; USEPA 2013a). The impact of effects on reproductive parameters, such as sex ratio or intersex, on reproduction, and ultimately on the population, will likely vary depending on severity and prevalence (Harris et al. 2011). Regarding growth, Ankley et al. (2003) noted an increase in the wet weight of female fathead minnows exposed to TRB for 21 d, starting at a concentration for which decreased egg production and changes in some mechanistic endpoints (e.g., decreased plasma VTG in females) were observed. Although it is not clear whether an increase in female weight would be considered an adverse outcome for a population, potential changes in behavior for fast-growing fish may include inappropriate timing for migration, season-inappropriate behavior (e.g., increased appetite during winter), increased activity resulting in higher metabolic demands, and increased risk taking to capture prey (Meador et al. 2011). Biro et al. (2004) demonstrated that fast-growing trout exhibited a 20% increase in specific growth rate, but their survival was reduced by up to 62%.

Similar to TBT, EE2 population-level effects have been observed in the laboratory using surrogate endpoints and in the field. Adverse effects in higher-tiered tests are observed and are linked to ER binding. Multiple life-cycle studies with different fish species exposed to EE2 have reported fertilization success (typically a surrogate for population-relevant effects) as the most sensitive endpoint (e.g., Nash et al. 2004; Schäfers et al. 2007; reviewed in Parrott et al. this issue). Induction of plasma VTG together with histological changes in the gonads and reduced larval growth were also shown to be valuable indicators for the long-term developmental and reproductive effects of EE2 (Länge et al. 2001). After EE2 exposure (concentrations of 5–6 ng/L) in a multi-year controlled field study, decreases in fish population have been observed (Kidd et al. 2007; Palace et al. 2009). In these trials, severe population declines were observed with shorter-lived species, such as fathead minnow (*Pimephales promelas*)

and pearl dace (*Margariscus margarita*), whereas for longer-lived fish, such as lake trout (*Salvelinus namaycush*), effects may have been due to impacts to food sources (loss of smaller fish). However, some species, such as the white sucker (*Catostomus commersonii*), did not appear to be affected, indicating that species-specific sensitivities and life-history strategies (e.g., reproductive, habitat preferences) can influence population-level impacts. In a fish screening assay with PPZ, alterations in endocrine-mediated biomarkers were observed (i.e., gonadosomatic index [GSI], estradiol, VTG, and/or gonadal histopathology), whereas effects on surrogate population-relevant endpoints (decreased fecundity and fertilization success) generally occurred at similar concentrations where these changes in biomarkers occurred (USEPA 2015a). Screening assays with fathead minnows exposed to VZ indicated potential reductions in fecundity as well as impacts to males (alterations of secondary sexual characteristics and gonadal weight and/or histopathology), which suggested potential androgen-related impacts (USEPA 2013a). Long-term (multigenerational) reproduction laboratory studies with Japanese medaka (*Oryzias latipes*) were also conducted with VZ (USEPA, 2013a) and generally, the most sensitive endpoint was reduction in SSC (anal fin papillae) at concentrations around 9 to 33 μg/L. Although this effect could be indicative of antiandrogenicity, there was a lack of other diagnostic evidence supporting the presumptive AOP for VZ. Whereas a reduction in SSC was observed in these studies with VZ, this effect did not appear to alter the reproductive output or development of offspring.

For other case study chemicals, population-level effects in the field were not as readily observed; however, evaluation of individual-level endpoints in the laboratory were examined in the context of potentially endocrine-mediated effects. Tributyltin has been shown to cause reproductive effects in fish, including embryo malformation, hatchability, and sex ratio alteration at low environmental concentrations (0.1–100 ng TBT/L) (McAllister and Kime 2003; Zhang et al. 2011). Additionally, alterations in sperm and gonopodium structure and altered testis histopathology have been observed (Haubruge et al. 2000; Zhang et al. 2009). These responses occur in the same concentration range at which effects occur in mollusks, although there are no reports of population-level effects in wild fish.

For PRC, effects on apical endpoints, such as decreased reproductive output, reduced growth, altered sex ratio, and skeletal abnormalities, were observed in multiple fish species, including zebrafish (*Danio rerio*), eastern mosquitofish (*Gambusia holbrooki*), threespine stickleback (*Gasterosteus aculeatus*), and fathead minnows (Mukhi and Patiño 2007; Bernhardt et al. 2011). Additionally, PRC exposure (up to 100 mg/L) has resulted in infertility, altered sex steroid levels, altered sex ratio, and impaired testicular function (e.g., spermatogenesis) in fish, suggesting cross talk with thyroid hormones (Bernhardt et al. 2006; Sharma and Patiño 2013).

However, evidence of thyroid gland effects that did not affect apical endpoints or recovery after PRC exposure has been reported. In a study using zebrafish (Mukhi et al. 2005), although effects at the cellular and organ levels were observed, whole body thyroxine (T4) levels, growth, condition factor (weight/length), and spawning behavior were not affected. Furthermore, all thyroid gland effects reversed following a 12-week recovery in clean water with only residual effects on angiogenesis and colloidal T4 ring intensity. Similar thyroid histology changes in PRC-exposed female eastern mosquitofish at 0.1 to 1000 mg/L over 30 d had no effect on mortality or growth (Bradford et al. 2005). In a stickleback study, transfer of sexually mature F0 fish from PRC exposure to clean water resulted in reduced survival of F1; however, the morphological effects observed in F0 (bone development) at $\geq$12 ppm were not exhibited in the F1 (Bernhardt et al. 2011).

Amphibians. Current testing for thyroid disruption in amphibians includes the Amphibian Metamorphosis Assay (AMA) and the Larval Amphibian Growth and Development Assay (LAGDA), which rely on the obligate dependence of metamorphosis on endogenous thyroid hormones (OECD 2007; USEPA 2013b). The 21-d AMA, which spans late premetamorphic (prior to development of a functioning thyroid gland) and prometamorphic development, enables the detection of both thyroid agonism and inhibition.

When screening for thyroid toxicants, the most suitable endpoint is the Nieuwkoop and Faber (NF) (Nieuwkoop and Faber 1994) developmental stage distribution, because the AMA ends prior to completion of metamorphosis. Additionally, thyroid histopathology (e.g., colloid depletion, follicular cell hypertrophy and hyperplasia, diffuse thyroid gland hypertrophy) is a useful and sensitive biomarker in the AMA and has proven to be more sensitive to thyroid-inhibiting chemicals than gross developmental morphology (i.e., stage). However, thyroid histopathology often represents compensation to thyroid insufficiency and thus cannot be regarded as adverse at the individual level, much less at the population level, if no obvious delay of stage development is present. It may be assumed that developmental delays, as well as acceleration, could be predictive of adverse impacts on development and fitness of individuals.

Correlation between a given change in developmental stage distribution in the AMA and altered time or failure to complete metamorphosis could be addressed in higher-tier studies such as the LAGDA, which encompasses completion of metamorphosis and sexual development. In the AMA, benzophenone-2 (BP-2) exhibited mild thyroid-disrupting activity as displayed by alteration of thyroid histopathology and/or developmental stage (delay of 2 NF stages) at the top concentration of 6 mg/L. At the same concentration of BP-2 in the LAGDA, the time to metamorphosis was increased by 11 d. Small but statistically significant differences in developmental stage distribution at the 21-d time point in the AMA, therefore, may be predictive of a statistically significant increase in time to completion of metamorphosis.

The degree of delay or acceleration in completion of metamorphosis that can be considered to be a population-relevant adverse effect is less clear and is subject to

uncertainties in extrapolating from the laboratory to the field, environmental variability, and variation in species-specific life-history strategies of affected amphibians. For example, the test species for AMA and LAGDA is *Xenopus laevis*, which has an aquatic adult phase, whereas the majority of amphibians have a semiterrestrial adult phase; consequences of alterations in the larval development phase on adult reproduction and survival are not well understood. Accelerated metamorphosis can result in lower weight and size due to the complex remodeling processes of outer and inner organs, which reduces individual fitness and could affect later survival under natural conditions. Additionally, delay or failure of metamorphosis, potentially caused by antithyroidal EAS, might cause an increase or complete mortality of a tadpole population in the environment due to an inability to respond to a change in pond status (drying up, freezing, etc.). Thus, altered timing of metamorphosis may result in population-level effects. However, the wide range of possible outcomes at the population level would need information from models as well as more field studies on basic ecology and population dynamics of amphibians.

In order to demonstrate the potential of EAS to affect reproduction of amphibians and thereby populations, more basic research is needed. Thyroid hormones (TH) are permissive for gonadal development and gametogenesis in vertebrates (Swapna and Semikhutaran 2007), but in amphibians, only preliminary data exist. It is likely that in amphibians, as in teleosts and mammals, TH contribute markedly to fecundity and fertility of males and females due to the conservation of reproductive physiology among vertebrates. Therefore, antithyroidal compounds such as PRC also may diminish reproductive success in amphibians with potential population-level effects.

In *Xenopus laevis,* male mating call is sensitive to EAS exposure (e.g., EE2 and VZ), decreasing the number of advertisement calls and even affecting the sound of that call so that it becomes less attractive for females (Hoffmann and Kloas 2010, 2012). Estrogens and antiandrogens also can affect gamete quality of adult *Xenopus*, as shown by altered spermatogenesis and testicular oocytes in males and atretic oocytes in females with EE2 (Cevasco et al. 2008). The antiandrogen, flutamide, also caused reduced spermatogenesis in male *Xenopus*. However, the degree of change in altered mating calls and gamete quality that would affect amphibian populations is unclear.

Birds. It has been suggested that birds have several unique characteristics that may make them vulnerable to endocrine disruption. Examples are flight (with associated adaptations for high metabolic rate and reduced body mass during migration), a primarily estradiol-mediated reproductive system, and a range of different types of development from altricial to precocial with associated different breeding and nest behaviors.

Two regulatory chronic avian toxicity test methods are available to evaluate potential adverse effects at the individual level (Table 1). The avian reproduction test (OECD TG 206; OCSPP 850.2300, OECD 1984; USEPA 2015b) is an apical test, but it does not contain endpoints that solely respond to EAS. As such, this test by itself cannot be used to screen EAS, but it can be used to investigate whether chemicals with an endocrine MOA (based on prior screening) can affect avian development, growth, or reproduction. The other test method is the Japanese Quail Two Generation Toxicity Test (JQTT; USEPA 2013c), which covers 4 key life stages:

- embryonic development,
- post hatch growth and development,
- sexual maturation, and
- adult.

Results of interlaboratory validation studies with VZ revealed a relatively low response in quail (attributed to the relative insensitivity of birds to antiandrogens) as well as relatively large differences in endpoint responses within each assay and between laboratories (attributed to strain differences, although the influence of control variability cannot be ruled out). In both of these tests only precocial species are tested, and neither include behavioral endpoints (e.g., mating behavior), which may be sensitive to EAS (Coady et al. this issue). Current results with the JQTT do not provide a basis for identifying non-apical endpoints that can be quantitatively linked with population-relevant endpoints. Future testing may provide additional information on the potential usefulness of mechanistic endpoints measured in the JQTT for use in predicting population-level effects and informing risk assessment. Additionally, developing a better understanding of characteristics unique to birds and their underlying mechanisms may help in predicting population-level effects.

Researchers have also used subchronic studies to evaluate the impact of EAS on birds; again, laboratory data for the case study chemicals suggest potential endocrine-mediated effects although field data are, in many cases, lacking. For example, several studies have evaluated the effects of PRC on avian thyroid hormone levels, thyroid gland histopathology, organ and body weight, egg production and morphometry and behavior (Hooper et al. 2003; Gentles et al. 2005). Non-apical endpoints used in these studies (e.g., thyroid gland histopathology, plasma and thyroid gland triiodothyronine [T_3] and/or thyroxine [T_4] levels) can help elucidate differences or similarities in MOA and sensitivity between wildlife models. For example, the mechanism of PRC toxicity is substantially conserved among birds and mammals, with similar endpoint responses seen in avian and mammalian tests. However, the doses at which these endpoints respond in birds vary by orders of magnitude from 0.05 mg/L (in drinking water provided ad libitum to quail) for changes in thyroidal T4, 0.1 mg/L for thyroid histopathological effects, 500 mg/L for changes in thyroid weight, to 4000 mg/L for changes in tibia and femur length, which may have potential population relevance but is unclear. Additionally, the concentration of 4000 mg/L is above environmentally relevant concentrations (i.e., $<$1–5 μg/L; Kalkhoff et al. 2010).

 DOI: 10.1002/ieam.1887

Therefore, these observed changes in the thyroid-mediated endpoints (i.e., T4, thyroid weight) following PRC exposure confirm a thyroid MOA but do not result in observable changes in population-relevant endpoints. Furthermore, the storage capacity of the thyroid gland and the potential for cyclic response patterns adds complexity to the interpretation of thyroid endpoint responses in birds. Feedback activation of the hypothalamic-pituitary-thyroid (HPT) axis leads to the release of stored hormones, which provide at least temporary compensation and restoration of euthyroid levels of circulating TH in birds (Delange and Ermans 1996; Taurog 1996). The potential for reversibility, compensation, and the associated cyclic patterns observed for several thyroid-related endpoint responses in birds make it difficult to assign adversity on the basis of these endpoints.

A number of limitations identified on the basis of the PRC case study may be relevant areas for further research. There were some indications that PRC may affect zebra finch behavior (i.e., increased begging activity, reduced flight attempts at $\geq$10 mg/kg; Rainwater et al. 2008). Other studies have similarly suggested that behavior may be a sensitive endpoint in birds, which may have population relevance (USEPA 2013c). Smith et al. (2001) reported preliminary data suggesting that PRC in food appeared to induce changes in thyroid-related endpoints more frequently than PRC administered via drinking water. All avian exposure studies reviewed as part of the case study relied on oral dosing via water or food. As such, dosing was estimated on the basis of assumed water or food ingestion in the case of ad libitum feeding or on administered volume in the case of individual dosing. In addition, control variability of most endpoints has not been well documented, limiting current understanding of sensitivity, repeatability (within a laboratory between studies), and transferability (between laboratories). This issue is not unique to the PRC avian studies that were reviewed; however, further proof of concept and validation work is needed before these avian non-apical endpoints can be reliably used to inform hazard or risk assessment.

Mammals. Mammalian laboratory study data are used to evaluate potential human health impacts of chemical exposure; however, in addition, these data are used to evaluate potential effects in mammalian wildlife populations. These data are used differently for these applications because population-level effects are the focus of wildlife risk assessments; therefore, the emphasis is typically on endpoints that affect growth development, reproduction, and survival.

As one example, the mammalian data in the PRC case study indicated numerous effects of PRC on thyroid endpoints in rats and deer mice, including decreased thyroid hormones, increased thyroid stimulating hormone (TSH), thyroid histopathology (decreased colloid and follicular cell hypertrophy continuing to hyperplasia). Thyroid follicular cell hyperplasia would be significant for human health hazard characterization because this effect may indicate increased risk of thyroid cancer. However, follicular cell hyperplasia is less biologically meaningful for wildlife populations because cancer is often a disease in older animals, a subpopulation that has reached reproductive senescence. Thus, cancer incidence generally does not affect population maintenance. If the toxicant induced tumors in animals of reproductive age or affected critical elements of TH-reproductive system cross-talk, this could be a relevant adverse effect at the population level.

A number of endocrine-relevant endpoints are evaluated in standardized repeated-dose laboratory studies in rodents that are designed to be used for regulatory purposes (Table 1). Some of these endpoints address reproductive function, which is population relevant (e.g., number of offspring and offspring viability, body weight) and would be relevant to EAS, although these studies may not establish an endocrine MOA. Furthermore, these studies also contain many non-apical endpoints, which at this time cannot be quantitatively linked to a population-level effect (i.e., histopathology, organ weight).

For example, consider the reproductive toxicity data for VZ. Vinclozolin and/or its metabolites have been shown to produce antiandrogenic effects in male rats, which are particularly sensitive during in utero development. Vinclozolin exposure (3–6 $mg \cdot kg^{-1} \cdot d^{-1}$) has been associated with numerous developmental changes (e.g., decreased prostate weight, nipple and/or areola retention, decreased anogenital distance); however, the population relevance of these effects is questionable. At higher doses ($\geq$50 $mg \cdot kg^{-1} \cdot d^{-1}$), more pronounced effects were seen, including low ejaculated sperm and decreased fertility. Altered development also was seen in male offspring exposed to $\geq$50 $mg \cdot kg^{-1} \cdot d^{-1}$ VZ during critical windows of development, resulting in ectopic testes, vaginal pouches, hypospadias, cleft phallus, other structural changes in the urogenital system, and accessory sex tissue hypoplasia or agenesis (Ostby et al. 1999). In addition, delayed puberty onset and decreased fertility occurred, which are population-relevant endpoints. Vinclozolin also produced Leydig cell tumors, but these effects were seen in older rats that were no longer critical for reproduction. In utero and lactational exposure to VZ also affected male sexual development and mating behavior in rabbits (Palmer et al. 2000; Veeramachaneni 2000). In an ecological assessment based on evaluations of endocrine and reproduction endpoints, the mammalian chronic risk assessment concluded low risk from VZ at typical use rates (USEPA 2000), which was supported by an experimental field study in voles (Caslin and Wolff 1999). However, in a subsequent Endangered Species Act evaluation (USEPA 2009b), potential exposures to VZ and its metabolites were judged to exceed risk levels of concern for direct, indirect, and habitat modification for the Californian red-legged frog (CRLF), including concern for potential chronic effects on small mammals serving as CRLF prey.

Population adverse effects and recovery

Field data, while complex, can examine population recovery, whereas laboratory studies typically are conducted under

conservative exposure assessments that limit the opportunity for observations of recovery. Population recovery is an important aspect of risk assessment and requires comprehensive information about demography, environmental variability, and other factors. The population vulnerability of some potentially sensitive species to endocrine disruption may be influenced by characteristics of their exposure to an EAS, their sensitivity to it, and the ability of that population to recover. To a certain extent, a population is able to tolerate some level of loss of individuals (survival) or impaired reproduction before persistence of that population becomes unsustainable. However, commercially harvested populations or those that are threatened or endangered may have limited capacity to tolerate individual loss. Recovery can occur through internal processes (e.g., survival or persistence of reproduction in unaffected or unexposed individuals or subpopulations and/or reversibility of the effect in the individual itself), and external processes (i.e., immigration from unaffected areas).

Internal recovery processes may exhibit density dependence, and the speed of recovery may be influenced by environmental fate and behavior of the substance (e.g., rapid degradation), the MOA of the chemical (e.g., latency), and the life-history strategy of the affected organism (e.g., generation time, reproductive strategy). For example, fish species with only 1 reproductive phase at the end of their life (e.g., lampreys, eels, Pacific salmon) may exhibit lower recovery potential because internal recovery processes will be slow or nonexistent, while the half-life of an EAS in the environment can predict whether reduction of fertilization rate would result in population-level effects. Other factors are important when considering population recovery. The number of generations required to recover from an acute mortality event depends on the fraction of the original population that survives the event and on the per-generation growth rate of the population. Factors important for population recovery time have been discussed in greater detail in previous studies (e.g., Barnthouse 2004; Ibrahim et al. 2014).

External recovery processes may be affected by pattern of use, environmental fate, and behavior of the chemical (persistence and mobility), life-history characteristics of the affected organism (e.g., dispersal ability), landscape characteristics (e.g., habitat isolation), and seasonal conditions that influence dispersal and recolonization processes. These latter factors, in addition to the potential presence of other chemical contaminants, represent multiple stressors, which may affect the adaptive capacity of individuals and may impair populations. However, the degree of additional stress on a given population is hard to measure or predict. While safety of chemicals is conducted on a chemical-by-chemical basis, the action of additional stressors can potentially be addressed in assessment or uncertainty factors applied during risk assessment for certain regulatory actions.

Future research needs and approaches

In the case study examples, available data were accumulated almost exclusively on individual organisms or experimental groups of test animals; adverse impacts on apical endpoints in individuals are typically suggestive of potential adverse changes in natural populations. Furthermore, the association of adverse changes for in-vitro endpoints or in-vivo non-apical endpoints with population changes has rarely been demonstrated and is the focus of ongoing research. Developing AOPs allows for the potential to predict (qualitatively with a goal of quantitatively) causal linkages between perturbation of a molecular initiating event (MIE), receptor responses, and subsequent individual- or population-level impacts (Coady et al. this issue). However, it will always be very difficult, if not impossible, to determine whether or not endocrine-active chemicals (or any chemicals) adversely affect wildlife populations if we know very little about 1) wildlife populations and 2) the factors influencing the sizes of these populations (Sumpter 2009). These major gaps in our knowledge severely limit our ability to link endocrine activity observed in toxicity tests with (falling) populations. To decrease these uncertainties, definitive data are therefore needed on a variety of population-relevant parameters from experimental studies, and some of these needs are indicated in the bulleted points below.

Because the data needs are so great across so many compounds and so many taxa, development of population modeling may be a practical approach to making judgments of likely population effects in the absence of definitive data. Furthermore, greater understanding of fundamental biological responses in conjunction with input from population dynamic experts is a potential path forward to predict the relative impacts of EAS from the knowledge of a target species' natural population drivers. Numerous examples of modeling approaches developed for fish are available. Miller and Ankley (2004) developed a fathead minnow population model and used data from Ankley et al. (2003) to predict population size from exposure to TRB; this analysis provides a critical link between laboratory-based fecundity data and population-level effects of TRB. These changes may be amenable to life-cycle modeling in fish, which can be used to predict alterations to demographic traits, the species' population growth rate, and the likelihood of population decline (Spromberg and Meador 2005). Sex ratio can be affected by endocrine-active chemicals (including TRB and TBT) and is an apical population-relevant endpoint; models for population effects have been discussed for chemicals that affect sex ratios (e.g., Hazlerigg et al. 2014). Mechanistic models are currently developed to infer population-relevant effects from physiological changes (e.g., Coupled Dynamic Energy Budget-DEB and Individual-Based Models-IBM; Martin et al. 2012), or even initiating molecular events through predictive systems models (Forbes and Calow 2012). Parameterization of such models requires reliable and robust data. However, robustness is the weakness of much ecotoxicology research (Harris et al. 2014). This is an issue that needs to be addressed to prevent the development of models based on poor data.

 DOI: 10.1002/ieam.1887

Additionally, while not an exhaustive list, important data gaps and research needs include the following:

- The impact on populations of behavioral changes and/or other neurodevelopmental changes (behavioral or structural) related to survival, growth, and reproduction
- The effect of altering non-apical endpoints (e.g., thyroid, estrogen, or androgen hormone levels) on survival, growth, and/or reproductive success in populations
- The impact that decreased growth and/or development (e.g., larval or juvenile) have on achieving reproductive age or reproductive success in a population
- The impact of estrogenic and/or antiandrogenic EAS on amphibian gametogenesis and subsequent changes in mating behavior on reproductive success of populations
- The acquisition of definitive data on a variety of population-relevant parameters (e.g., fundamental biological responses, population dynamics).

Further research also should concentrate on the realism of laboratory findings with respect to what is actually happening to wildlife in the environment. This research should be supported by sound ecotoxicological investigations resulting in robust and unbiased science-based environmental risk assessment (Harris et al. 2014; Staveley et al. 2016).

CONCLUSIONS

Population-level impacts from exposure to EAS have been demonstrated for a few cases of our case study chemicals, such as gastropod population declines following TBT exposure and fish population declines following EE2 exposure. Although EAS also have been shown to impact a wide range of subcellular- to individual-level endpoints in laboratory studies, the population relevance and endocrine control of these endpoints is often unknown, as illustrated by our case study chemicals. As with any endpoint, the impact on a population will likely vary depending on the severity and prevalence of the response for exposed individuals and their life-history traits. Additionally, the ability for recovery, adaptation, or reversibility of an individual or population to a particular perturbation should be considered. As a result, the availability of reliable population-relevant endocrine-mediated endpoints in current ecotoxicity test methods is an important data gap. Additional enhancements to population modeling, using information routinely measured in current ecotoxicology studies, along with newer endpoints and techniques (e.g., suborganism effects) should allow for greater predictability of potential impacts in the absence of data on population-level alterations. However, the development of models that link effects observed in laboratory tests to the dynamics of wildlife populations requires a more robust knowledge of the factors regulating these dynamics. As our understanding of 1) endocrine perturbations and key-event relationships and 2) environmental regulation of population dynamics improves, adverse population-level effects should be more easily and accurately predicted.

Disclaimer—The views and statements expressed in this paper are those of the authors alone. The views or statements expressed in this publication do not necessarily represent the views of the organizations to which the authors are affiliated, and those organizations cannot accept any responsibility for such views or statements. This paper was developed as part of a SETAC Pellston Workshop "Environmental Hazard and Risk Assessment Approaches for Endocrine-Active Substances (EHRA)" and the authors have no conflicts of interest to report. The peer review process for this article was managed by the Editorial Board, without the involvement of J Meador.

Data availability—This review is based on published data, which are publicly available.

REFERENCES

Ankley GT, Jensen KM, Makynen EA, Kahl MD, Korte JJ, Hornung MW, Henry TR, Denny JS, Leino RL, Wilson VS, et al. 2003. Effects of the androgenic growth promoter 17-β-trenbolone on fecundity and reproductive endocrinology of the fathead minnow. *Environ Toxicol Chem* 22:1350–1360.

Avaca MS, Martin P, van der Molen S, Narvarte M. 2015. Comparative study of the female gametogenic cycle in three populations of *Buccinanops globulosus* (Caenogastropoda: Nassariidae) from Patagonia. *Helgoland Mar Res* 69:87–99.

Bailey SK, Davies IM, Harding MJC. 1995. Tributyltin contamination and its impact on *Nucella lapillus* populations. In: Dunnet GM, McIntyre AD, editors. Proceedings of the Royal Society of Edinburgh. Section B. Biological Sciences. Cambridge England: Royal Society of Edinburgh, Cambridge University Press (CUP). 103.p 113–126. DOI: 10.1017/S0269727000005960

Barnthouse LW. 2004. Quantifying population recovery rates for ecological risk assessment. *Environ Toxicol Chem* 23(2):500–508.

Barroso CM, Moreira MH, Bebianno MJ. 2002. Imposex, female sterility and organotin contamination of the prosobranch *Nassarius reticulatus* from the Portuguese coast. *Mar Ecol Progr Ser* 230:127–135.

Batley GE, Scammell MS, Brockbank CI. 1992. The impact of the banning of tributyltin-based antifouling paints on the Sydney rock oyster, *Saccostrea commercialis*. *Sci Tot Environ* 122:301–314.

Bernhardt RR, von Hippel FA, Cresko WA. 2006. Perchlorate induces hermaphroditism in threespine stickleback. *Environ Toxicol Chem* 25(8):2087–2096.

Bernhardt RR, von Hippel FA, O'Hara TM. 2011. Chronic perchlorate exposure causes morphological abnormalities in developing stickleback. *Environ Toxicol Chem* 30(6):1468–1478.

Birch GF, Scammell MS, Besley CH. 2014. The recovery of oyster (*Saccostrea glomerata*) populations in Sydney estuary (Australia). *Environ Sci Pollut Res Int* 21(1):766–773.Epub 2013 Sep 24. DOI: 10.1007/s11356-013-2168-x

Birchenough AC, Evans SM, Moss C, Welch R. 2002. Re-colonisation and recovery of populations of dogwhelks *Nucella lapillus* (L.) on shores formerly subject to severe TBT contamination. *Mar Pollut Bull* 44(7):652–659.

Biro PA, Abrahams MV, Post JR, Parkinson EA. 2004. Predators select against high growth rates and risk-taking behaviour in domestic trout populations. *Proc R Soc Lond* 271:2233–2237.

Bradford CM, Rinchard J, Carr JA, Theodorakis C. 2005. Perchlorate affects thyroid function in eastern mosquitofish (*Gambusia holbrooki*) at environmentally relevant concentrations. *Environ Sci Technol* 39(14):5190–5195.

Bryan GW, Burt GR, Gibbs PE, Pascoe PL. 1993. *Nassarius reticulatus* (Nassariidae: Gastropoda) as an indicator of tributyltin pollution before and after TBT restrictions. *J Mar Biol Assoc UK* 73:913–929.

Caslin TM, Wolff JO. 1999. Individual and demographic responses of the gray-tailed vole to vinclozolin. *Environ Toxicol Chem* 18:1529–1533.

Castro LFC, Lima D, Machado A, Melo C, Hiromori Y, Nishikawa J, Nakanishi T, Reis-Henriques MA, Santos MM. 2007. Imposex induction is mediated through the Retinoid X Receptor signaling pathway in the neogastropod *Nucella lapillus*. *Aquat Toxicol* 85:57–66.

Cevasco A, Urbatzka R, Bottero S, Massari A, Pedemonte F, Kloas W, Mandich A. 2008. Endocrine disrupting chemicals (EDC) with (anti)estrogenic and (anti)androgenic modes of action affecting reproductive biology of *Xenopus laevis*: II. Effects on gonad histomorphology. *Comp Biochem Physiol C Toxicol Pharmacol* 47(2):241–251.

Coady KK, Biever RC, Denslow ND, Gross M, Guiney PD, Holbech H, Karouna-Renier NK, Katsiadaki I, Krueger H, Levine SL, et al. 2017. Current limitations and recommendations to improve testing for the environmental assessment of endocrine active substances. *Integr Environ Assess Manag* 13:302–316.

Davis KB, Morrison J, Galvez JI. 2000. Reproductive characteristics of adult channel catfish treated with trenbolone acetate during the phenocritical period of sex differentiation. *Aquaculture* 189(3):351–360.

Delange FM, Ermans A-M. 1996. Iodine deficiency. In: Braverman LE, Utiger RD, editors. Werner and Ingbar's The Thyroid. 7th ed., Philadelphia (PA): Lippincott-Raven. p 296–316.

[EFSA] European Food Safety Authority. 2013. Scientific opinion on the hazard assessment of endocrine disruptors: Scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. [cited 2016 Mar 9]. http://www.efsa.europa.eu/en/efsajournal/pub/3132

Forbes VE, Calow P. 2012. Promises and problems for the new paradigm for risk assessment and an alternative approach involving predictive systems models. *Environ Toxicol Chem* 31:2663–2671.

Gentles A, Surles J, Smith EE. 2005. Evaluation of adult quail and egg production following exposure to perchlorate-treated water. *Environ Toxicol Chem* 24(8):1930–1934.

Gibbs PE, Pascoe PL, Burt GR. 1988. Sex change in the female dog-whelk, *Nucella lapillus*, induced by tributyltin from antifouling paints. *J Mar Biol Assoc UK* 68:715–731.

Gooding MP, Wilson VS, Folmar LC, Marcovich DT, LeBlanc GA. 2003. The biocide tributyltin reduces the accumulation of testosterone as fatty acid esters in the mud snail (*Ilyanassa obsoleta*). *Environ Health Perspect* 111(4):426–430.

Harris CA, Hamilton PB, Runnalls TJ, Vinciotti V, Henshaw A, Hodgson D, Coe TS, Jobling S, Tyler CR, Sumpter JP. 2011. The consequences of feminization in breeding groups of wild fish. *Environ Health Perspect* 119:306–311.

Harris CA, Scott AP, Johnson AC, Panter GH, Sheahan D, Roberts M, Sumpter JP. 2014. Principles of sound ecotoxicology. *Environ Sci Technol* 48(3):3100–3111.

Haubruge E, Petit F, Gage MJ. 2000. Reduced sperm counts in guppies (*Poecilia reticulata*) following exposure to low levels of tributyltin and bisphenol A. *Proc Royal Soc London B: Biol Sci* 267 (1459):2333–2337.

Hazlerigg CRE, Tyler CR, Lorenzen K, Wheeler JR, Thorbek P. 2014. Population relevance of toxicant mediated changes in sex ratio in fish: An assessment using an individual-based zebrafish (*Danio rerio*) model. *Ecol Model* 280:76–88.

Hoffmann F, Kloas W. 2010. The endocrine disrupting antiandrogen vinclozolin affects the calling behavior of male South African clawed frogs (*Xenopus laevis*) at environmentally relevant concentrations. *Horm Behav* 58:653–659.

Hoffmann F, Kloas W. 2012. Estrogens can disrupt amphibian mating behavior. *PlosOne* 7:e32097.

Holbech H, Kinnberg K, Petersen GI, Jackson P, Hylland K, Norrgren L, Bjerregaard P. 2006. Detection of endocrine disruptors: Evaluation of a fish sexual development test (FSDT). *Comp Biochem Physiol C: Toxicol Pharmacol* 144(1):57–66.

Hooper MJ, Millam JM, Rainwater TR, McNabb FMA, Bounds R, Kendall RJ. 2003. Perchlorate, thyroid function and brain development in a granivorous passerine. In: Kendall R, editor. Ecological risk assessment of perchlorate in avian species, rodents, amphibians and fish. p 88–127. SERDP Project ER-1235.

Horiguchi T, Shiraishi H, Shimizu M, Morita M. 1994. Imposex and organotin compounds in *Thais clavigera* and *T. bronni* in Japan. *J Mar Biol Assoc UK* 74:651–669.

Ibrahim L, Preuss TG, Schaeffer A, Hommen U. 2014. A contribution to the identification of representative vulnerable fish species for pesticide risk assessment in Europe—A comparison of population resilience using matrix models. *Ecol Model* 280:65–75.

Iguchi T, Katsu Y. 2008. Commonality in signaling of endocrine disruption from snail to human. *BioScience* 58:1061–1067.

Kalkhoff SJ, Stetson SJ, Lund KD, Wanty RB, Linder GL. 2010. Perchlorate data for streams and groundwater in selected areas of the United States, 2004: U.S. *Geological Survey Data Series 495*. 43 p with appendix.

Kidd KA, Blanchfield PJ, Mills KH, Palace VP, Evans RE, Lazorchak JM, Flick RW. 2007. Collapse of a fish population after exposure to a synthetic estrogen. *Proc Natl Acad Sci* 104(21):8897–8901.

Kramer VJ, Etterson MA, Hecker M, Murphy CA, Roesijadi G, Spade DJ, Spromberg JA, Wang M, Ankley GT. 2011. Adverse outcome pathways and ecological risk assessment: Bridging to population-level effects. *Environ Toxicol Chem* 30(1):64–76.

Länge R, Hutchinson TH, Croudace CP, Siegmund F, Schweinfurth H, Hampe P, Panter GH, Sumpter JP. 2001. Effects of the synthetic estrogen 17α-ethinylestradiol on the life-cycle of the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 20:1216–1227.

LeBlanc GA, Gooding MP, Sternberg RM. 2005. Testosterone-fatty acid esterification: A unique target for the endocrine toxicity of tributyltin to gastropods. *Integr Comp Biol* 45:81–87.

Lewis JA, Baran IJ, Carey JM, Fletcher LE. 2010. A contaminant in decline: Long-term TBT monitoring at a naval base in Western Australia. *Aust J Ecotox* 16:17–34.

Martin BT, Zimmer EI, Grimm V, Jager T. 2012. Dynamic energy budget theory meets individual-based modeling: A generic and accessible implementation. *Meth Ecol Evol* 3:445–449.

Matthiessen P. 2013. Detection, monitoring, and control of tributyltin – An almost complete success story. *Environ Toxicol Chem* 32:487–489.

Matthiessen P, Ankley GT, Biever RC, Bjerregaard P, Borgert C, Brugger K, Blankinship A, Chambers J, Coady KK, Constantine L, et al. 2017. Recommended approaches to the scientific evaluation of environmental hazards and risks of endocrine-active substances. *Integr Environ Assess Manag* 13:267–279.

McAllister BG, Kime DE. 2003. Early life exposure to environmental levels of the aromatase inhibitor tributyltin causes masculinisation and irreversible sperm damage in zebrafish (*Danio rerio*). *Aquat Toxicol* 65(3):309–316.

Meador JP. 2011. Organotins in aquatic biota: Occurrence in tissue and toxicological significance. In: Beyer WN, Meador JP, editors. Environmental contaminants in biota: Interpreting tissue concentrations. Boca Raton (FL): Taylor & Francis. p 255–284.

Meador JP, Sommers FC, Cooper KA, Yanagida G. 2011. Tributyltin and the obesogen metabolic syndrome in a salmonid. *Environ Res* 111:50–56.

Mihaich EM, Schäfers C, Dreier DA, Hecker M, Ortego L, Kawashima Y, Dang Z-C, Solomon K. 2017. Endocrine active substances: Direct and indirect action, a risk and hazard framework. *Integr Environ Assess Manag* 13:280–292.

Miller DH, Ankley GT. 2004. Modeling impacts on populations: Fathead minnow (*Pimephales promelas*) exposure to the endocrine disruptor 17β-trenbolone as a case study. *Ecotoxicol Environ Safe* 59(1):1–9.

Morthorst JE, Holbech H, Bjerregaard P. 2010. Trenbolone causes irreversible masculinization of zebrafish at environmentally relevant concentrations. *Aquat Toxicol* 98(5):336–343.

Mukhi S, Carr JA, Anderson TA, Patino R. 2005. Novel biomarkers of perchlorate exposure in zebrafish. *Environ Toxicol Chem* 24(5):1107–1115.

Mukhi S, Patiño R. 2007. Effects of prolonged exposure to perchlorate on thyroid and reproductive function in zebrafish. *Toxicol Sci* 96(2):246–254.

Nash JP, Kime DE, Van der Ven LT, Wester PW, Brion F, Maack G, Stahlschmidt-Allner P, Tyler CR. 2004. Long-term exposure to environmental concentrations of the pharmaceutical ethynylestradiol causes reproductive failure in fish. *Environ Health Perspect* 112(17):1725–1733.

Nicolaus EEM, Barry J. 2015. Imposex in the dogwhelk (*Nucella lapillus*): 22-year monitoring around England and Wales. *Environ Monit Assess* 187:736.DOI: 10.1007/s10661-015-4961-0

Nieuwkoop PD, Faber J. 1994. Normal Table of Xenopus Laevis. New York: Garland Publishing.

[OECD] Organisation for Economic Co-operation and Development. 2007. Series on testing and assessment, validation of the amphibian metamorphosis assay as a screen for thyroid-active chemicals: Integrated summary report. Paris (FR): Environmental Health and Safety.

[OECD] Organisation for Economic Co-operation and Development. 2012. Conceptual framework for testing and assessment of endocrine disruptors (as revised in 2012). [cited 2016 Mar 9]. http://www.oecd.org/env/ehs/testing/OECD%20Conceptual%20Framework%20for%20Testing%20and%20Assessment%20of%20Endocrine%20Disruptors%20for%20the%20public%20website.pdf

[OECD] Organisation for Economic Co-operation and Development. 1984. Series on testing and assessment. http://www.oecd-ilibrary.org/environment/test-no-206-avian-reproduction-test_9789264070028-en

Ostby J, Monosson E, Kelce WR, Gray Jr LE. 1999. Environmental antiandrogens: Low doses of the fungicide vinclozolin alter sexual differentiation of the male rat. *Toxicol Ind Health* 15(1–2):48–64.

Palace VP, Evans RE, Wautier KG, Mills KH, Blanchfield PJ, Park BJ, Baron CL, Kidd KA. 2009. Interspecies differences in biochemical, histopathological, and population responses in four wild fish species exposed to ethynylestradiol added to a whole lake. *Can J Fisher Aquat Sci* 66:1920–1935.

Palmer JS, Kane CM, Higuchi TT, Bodensteiner KJ, Pau K-YF, Veeramachaneni DNR. 2000. Gestational and lactational exposure to antiandrogenic pesticides, p,p'DDT and Vinclozolin, impairs reproductive function in male rabbits [Abstract]. *Biol Reprod* 62 (Suppl 1):185.

Parrott JL, Bjerregaard P, Brugger KE, Gray LE, Jr, Iguchi T, Kadlec SM, Weltje L, Wheeler JR. 2017. Uncertainties in biological responses that influence hazard or risk approaches to the regulation of endocrine active substances. *Integr Environ Assess Manag* 13:293–301.

Pascoal S, Carvalho G, Vasieva O, Hughes R, Cossins A, Fang Y-X, Ashelford K, Olohan L, Barroso C, Mendo S, et al. 2013. Transcriptomics and in vivo tests reveal novel mechanisms underlying endocrine disruption in an ecological sentinel, *Nucella lapillus*. *Mol Ecol* 22:1589–1608.

Rainwater TR, Wood MB, Millam JR, Hooper MJ. 2008. Effects of perchlorate on growth and behavior of granivorous passerine, the zebra finsh (*Taeniopygia guttata*). *Arch Environ Contam Toxicol* 54:516–5240.

Schäfers C, Teigeler M, Wenzel A, Maack G, Fenske M, Segner H. 2007. Concentration- and time-dependent effects of the synthetic estrogen, 17α-ethinylestradiol, on reproductive capabilities of the zebrafish, *Danio rerio*. *J Toxicol Environ Health Part A* 70(9):768–779.

Sharma P, Patiño R. 2013. Regulation of gonadal sex ratios and pubertal development by the thyroid endocrine system in zebrafish (*Danio rerio*). *Gen Comp Endocrinol* 184:111–119.

Smith EE, Stafford J, Williams L, Birdwell B, Kendall R. 2001. Response of mallards to perchlorate contaminated food and/or water following oral exposure. In: Kendall R. editor. Ecological risk assessment of perchlorate in avian species, rodents, amphibians and fish. SERDP Project ER-1235.

Spence SK, Bryan GW, Gibbs PE, Masters D, Morris L, Hawkins SJ. 1990. Effects of TBT contamination on *Nucella populations*. *Funct Ecol* 4:425–432.

Spromberg JA, Meador JP. 2005. Population-level effects on chinook salmon from chronic toxicity test measurement endpoints. *Integr Environ Assess Manag* 1:9–21.

Staveley J, Wentsel R, Sumpter J, Harris CA, Henry T, Pease A, Knopper LD, Breton RL, Moore DRJ, Hall T, Bowers L. 2016. How can we improve the quality of ecotoxicology research to increase relevance and use in regulatory decision making? *Environ Toxicol Chem* 35:14–19.

Sumpter JP. 2009. Protecting aquatic organisms form chemicals: the harsh realities. *Phil Trans Royal Soc A* 367:3877–3894.

Swapna I, Senthikumaran B. 2007. Thyroid hormones modulate the hypothalamo-hypophyseal-gonadal axis in teleosts: molecular insights. *Fish Physiol Biochem* 33:335–345.

Taurog A. 1996. Hormone synthesis: Thyroid iodine metabolism. In: Braverman LE, Utiger RD, editors. Werner and Ingbar's The Thyroid. 7th ed. Philadelphia (PA): Lippincott-Raven. p 47–81.

[USEPA] US Environmental Protection Agency. 2000. Reregistration eligibility decision for vinclozolin. Washington (DC): Office of Chemical Safety and Pollution Prevention (formerly Office of Prevention, Pesticides and Toxic Substances, OPPTS). EPA 738-R-00-023, October 2000. [2016 Nov 20]. https://www3.epa.gov/pesticides/chem_search/reg_actions/reregistration/red_PC-113201_1-Oct-00.pdf

[USEPA] US Environmental Protection Agency. 2009a. Series 890 – Endocrine disruptor screening program test guidelines. Endocrine Disruptor Screening Program, Washington (DC). [cited 2016 Jul 11]. https://www.epa.gov/test-guidelines-pesticides-and-toxic-substances/series-890-endocrine-disruptor-screening-program

[USEPA] US Environmental Protection Agency. 2009b. Risks of vinclozolin use to federally threatened California Red-legged frog (*Rana auroro draytonii*). Pesticide effects determination. Washington (DC): Environmental Fate and Effects Division, Office of Pesticide Programs. October 14, 2009. [cited 2016 Nov 20]. https://nepis.epa.gov/Exe/ZyNET.exe/P1007YPJ.TXT?ZyActionD=ZyDocument&Client=EPA&Index=2006+Thru+2010&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C06thru10%5CTxt%5C00000019%5CP1007YPJ.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL

[USEPA] US Environmental Protection Agency. 2013a. Validation of the medaka multigeneration test: Integrated summary report. Endocrine Disruptor Screening Program, Washington (DC). 28 May 2013. [cited 2016 Mar 9]. http://www.oecd.org/env/ehs/testing/MMT%20ISR%20final.pdf

[USEPA] US Environmental Protection Agency. 2013b. Validation of the larval amphibian growth and develoment assay: Integrated summary report. Endocrine Disruptor Screening Program. Washington (DC). 28 May 2013. [cited 2016 Mar 9]. http://www.oecd.org/chemicalsafety/testing/LAGDA-Integrated-Summary-Report.pdf

[USEPA] US Environmental Protection Agency. 2013c. Validation of the Japanese quail two-generation test: Integrated summary report. Endocrine Disruptor Screening Program. Washington (DC). [cited 2016 Aug 9]. https://www.regulations.gov/document?D=EPA-HQ-OPPT-2009-0576-0017

[USEPA] US Environmental Protection Agency. 2015a. EDSP weight of evidence conclusions on the Tier 1 screening assays for the List 1 chemicals – Propiconazole. Memorandum from Greg Akerman and Amy Blankinship, Endocrine Disruptor Screening Program. Washington (DC). June 29, 2015. [cited 2016 Aug 9]. https://www.epa.gov/sites/production/files/2015-06/documents/propiconazole-122101_2015-06-29_txr0057144.pdf

[USEPA] US Environmental Protection Agency. 2015b. Avian two-generation toxicity test in the Japanese quail. Washington (DC): Endocrine Disruptor Screening Program https://www.epa.gov/test-guidelines-pesticides-and-toxic-substances/series-890-endocrine-disruptor-screening-program

Veeramachaneni DN. 2000. Deteriorating trends in male reproduction: Idiopathic or environmental? *Anim Reprod Sci* 60-61:121–130.

Vos JG, Dybing E, Greim HA, Ladefoged O, Lambre C, Tarazona JI, Brandt I, Vethaak AD. 2000. Health effects of endocrine-disrupting chemicals on wildlife, with special reference to the European situation. *Crit Rev Toxicol* 30:71–133.

Watanabe KH, Mayo M, Jensen KM, Villeneuve DL, Ankley GT, Perkins EJ. 2016. Predicting fecundity of fathead minnows (*Pimephales promelas*) exposed to endocrine-disrupting chemicals using a MATLAB (R)-based model of oocyte growth dynamics. *PloS One* 11(1). DOI: 10.1371

[WHO/IPCS] World Health Organization/International Programme on Chemical Safety. 2002. Global assessment of the state-of-the-science of endocrine disruptors. [cited 2016 Mar 9]. http://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en/

Zhang JL, Zuo ZH, He CY, Cai JL, Wang YQ, Chen YX, Wang CG. 2009. Effect of tributyltin on testicular development in *Sebastiscus marmoratus* and the mechanism involved. *Environ Toxicol Chem* 28:1528–1535.

Zhang J, Zuo Z, Wang Y, Yu A, Chen Y, Wang C. 2011. Tributyltin chloride results in dorsal curvature in embryo development of *Sebastiscus marmoratus* via apoptosis pathway. *Chemosphere* 82(3):437–442.

