Water Qual. Res. J. Canada, 2001
Volume 36, No. 2, 273–291

An Overview of Testing Procedures and Approaches for Identifying Endocrine Disrupting Substances

JOANNE PARROTT,[1]* MICHAEL WADE,[2] GARY TIMM[3] AND SCOTT BROWN[1]

[1]*National Water Research Institute, Environment Canada, 867 Lakeshore Road, Burlington, Ontario L7R 4A6*
[2]*Environmental and Occupational Toxicology Division, Healthy Environments and Consumer Safety Branch, Health Canada, Environmental Health Centre, Tunney's Pasture, Ottawa, Ontario K1A 0L2*
[3]*Office of Prevention, Pesticides, and Toxic Substances, U.S. Environmental Protection Agency, Washington, DC 20460*

Concern over the potential effects of exposure to endocrine disrupting substances (EDS) has resulted in recommendations for the development of specific endocrine and reproductive tests for assessment and regulation of industrial chemicals and effluents. This document consolidates and summarizes the current approaches taken by international agencies and scientific organizations for testing and screening EDS in mammals and wildlife. The material has been gathered from the Organization for Economic Cooperation and Development meeting reports, U.S. Environmental Protection Agency's Endocrine Disruptor Screening and Testing Advisory Committee and Endocrine Disruptor Screening Program reports, and summaries of various meetings prepared by Canadian and U.S. representatives. There are commonalties between programs of individual countries and the international scientific groups in which they participate. In general, these international scientific organizations have envisioned tiered groups of tests. The first tier is composed of less complex tests and short-term assays (acute exposures) that are very responsive to EDS (and have a low 'false-negative' rate). Higher tiers contain longer tests encompassing partial and full lifecycles of organisms with assessment of functional reproductive and developmental endpoints. Compounds are less expensively screened in the first tier, and move on to more complex and expensive higher-tier tests only if necessary. We document the proposed EDS tests in mammals and wildlife (birds, fish, amphibians and invertebrates) and overview the European Union's approach to EDS research, monitoring and risk assessment. We conclude with a summary of the main recommendations from Canada's interagency workshop to develop priorities and proposed actions for EDS.

Key words: **OECD, U.S. EPA, endocrine disrupting substances (EDS), bioassays, risk assessment**

Introduction

Over the past 5 years, endocrine disrupting substances (EDS) have emerged as a major environmental science and policy issue. Documented

* Corresponding author; joanne.parrott@cciw.ca

effects on fish and wildlife, and evidence from human epidemiology and experimental toxicology have led to an emerging hypothesis that chemicals may be affecting reproduction and development (Krimsky 2000). Concern that EDS exposure may adversely influence environmental and human health has led to calls for increased surveillance of commercial chemicals for endocrine toxicity and demands for study of environmental impact associated with disruption of endocrine function (Colborn and Clements 1992). International awareness of the issue was heightened with publication of *Our Stolen Future* in 1996 (Colborn et al. 1996). According to a 1997 report of the Organization for Economic Cooperation and Development (OECD 1997), no country had specific restrictions on the discharge of EDS. Moreover, the current suite of toxicological and bioaccumulation tests used for risk assessment were never developed to specifically identify problems associated with endocrine disruptors or their potential to cause reproductive and developmental effects. Due to the potential for adverse effects by EDS, the U.S. and OECD have begun to tackle the issue by forming working groups of experts with the aim of establishing internationally harmonized guidelines on screening and testing protocols.

The U.S. Environmental Protection Agency (U.S. EPA) and the OECD have invested considerable time and resources to develop tiered frameworks specifically for the testing and assessment of EDS. The U.S. has a legislative mandate to test and screen chemicals for endocrine disruption and has formed the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) and Endocrine Disruptor Screening Program (EDSP). The OECD endocrine disruptor expert groups are developing EDS screening and testing schemes that reflect the joint efforts of many countries. These screening and testing schemes are very specific, with recommended tests occurring in a defined progression from shorter, less complex tests to longer, more involved tests. The test results will be recommended for assessment and regulation of chemicals. Because these represent the most well-defined and developed testing schemes, we overview both the U.S. EPA and OECD EDS testing schemes. We also summarize the more general EDS strategy of the European Union (EU) countries (Commission of the European Communities 1999). Unlike the OECD and U.S. EPA, the EU strategy outlines their approach for EDS research, communication and implementation in risk assessment and regulation. We focused on the EU approach because, although not as specific as the tests set out in the OECD and U.S. EPA documents, it reflects harmonized opinion of many countries for EDS testing. We conclude with the major recommendations of a multi-stakeholder workshop, which discussed Canada's approach to EDS research, communication and implementation in risk assessment and regulation (Servos et al. 2000).

Endocrine Disruptor Testing in the United States

In response to concerns about exposure to EDS, the U.S. Congress passed the *Food Quality Protection Act* (FQPA) and the *Safe Drinking Water*

Act (SDWA) amendments of 1996. The FQPA requires the U.S. EPA to screen pesticides for estrogenic effects on human health and permits them to screen chemicals found in drinking water sources for other hormonal effects. Faced with implementing a new regulatory program, the U.S. EPA responded by sponsoring research to better understand the basic science of endocrine disruption and by establishing the EDSTAC to advise the agency on the design of a scientifically defensible endocrine disruptor screening program (EDSP).

To assist the U.S. EPA in making decisions regarding EDS, a research agenda was developed for health and ecological effects in 1995 (Kavlock et al. 1996; Ankley et al. 1997). Major research questions focused on EDS exposure and adverse effects, the shape of the dose-response curve, effects in fish and wildlife, and types of test protocols needed to detect the effects of EDS. The U.S. Office of Science and Technology Policy also created a research inventory (*www.epa.gov/endocrine*) to coordinate research among 14 federal agencies and identify gaps in ongoing research that should be addressed in a federal research strategy. The U.S. research inventory served as a prototype for an international research inventory developed by the World Health Organization and the International Program on Chemical Safety.

U.S. Endocrine Disruptor Screening Program (EDSP)

In October 1996, the U.S. EPA established the EDSTAC, which consisted of 39 members representing the pesticide and chemical manufacturers, state and federal government, and environmental and public health advocacy organizations. While many viewed EDS issues as a problem with estrogenicity, the EDSTAC expanded the scope of its deliberations to include potential effects on the androgen and thyroid systems. Ecological effects were also deemed important, as these cases have provided the strongest evidence of endocrine disruption to date. EDSTAC also included chemicals other than pesticides and drinking water contaminants as candidates for screening and testing. The universe of candidate chemicals under consideration included approximately 900 pesticide active ingredients, 2,500 pesticide formulation inert ingredients, 75,500 industrial chemicals, and 8,000 cosmetics, food additives and nutritional supplements (EDSTAC 1998). The overall framework consists of four major steps: (1) initial sorting of chemicals, (2) establishment of screening priorities, (3) Tier 1 screening, and (4) Tier 2 testing.

For the establishment of screening priorities, EDSP considered both exposure- and effects-related information important. As well, mixtures, naturally occurring non-steroidal estrogens and specially nominated substances would be considered targeted priorities. EDSP recognized that few chemicals actually have data that directly measure their endocrine disrupting potential. Even chemicals that have data on reproduction and developmental effects were likely tested using conventional protocols that have not been specifically designed to detect endocrine effects.

Tier 1 Screening

EDSTAC built on the work of other expert subgroups in reviewing candidate assays for a screening system (Gray et al. 1997; Ankley et al. 1998). These groups reviewed a number of individual assays and concluded that a battery of assays was necessary to evaluate endocrine disruption potential. EDSTAC recommended a battery of three in vitro assays and five in vivo assays for Tier 1 screening (Table 1). The in vitro assays include an estrogen receptor binding or reporter gene assay, an androgen receptor binding or reporter gene assay, and a steroidogenesis assay using minced testes. In vitro assays are advantageous in that they are inexpensive, specific for a particular mode of action, and amenable to automation. However, they lack the metabolic and response complexity of intact animals. EDSTAC preferred functional assays to receptor binding assays because the former can be used to differentiate agonist from antagonist activity, and they tend to be more sensitive than the latter (EDSTAC 1998).

The in vivo Tier 1 assays are organized so that endpoint complementarity is incorporated into the screening battery, and the in vitro and in vivo assays work together as a whole. Thus, the ER reporter gene, uterotrophic and pubertal female assays screen for estrogenicity and anti-estrogenicity. The AR reporter gene and Hershberger assays screen for androgenicity and anti-androgenicity. The frog assay and pubertal female assays screen for thyroid (ant)agonism. The fish assay is present because fish are the class of vertebrates most distant from the mammals in terms of their metabolism and hormone systems. The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) suggest replacement of the fathead minnow gonadal recrudescence assay with a juvenile fish assay using reproductive biomarker endpoints (Hutchinson et al. 2000). Also ECETOC recommends supplementing the suite of screens with an assay using oviparous fish. Having diverse taxa in Tier 1 was to assess the consistency of results among organisms and to provide insight to overall variability. A weight of evidence approach will be used to evaluate the results in Tier 1 and in vivo results will outweigh in vitro results. Chemicals testing negative in Tier 1 would be regarded as having low potential for interaction with the estrogen, androgen or thyroid systems. Chemicals testing positive would proceed for more in depth evaluation in Tier 2.

Tier 2 Testing

The Tier 2 tests are meant to identify adverse effects due to endocrine disruption and to establish a relationship between dose and response. The criteria identified by EDSTAC for Tier 2 tests was that the most sensitive life stage be tested (*in utero* or *in ovo*), that the tests be multi-generational and that each major taxonomic group be represented. Thus EDSTAC recommended five multi-generation tests: one each in mammals, birds, fish, amphibians and invertebrates (Table 1). In theory, Tier 2 can be tailored based on both exposure and effects information. ECETOC suggest insertion of an intermediate fish assay, a partial-lifecycle test (fish develop-

Table 1. Summary of the proposed screens and tests for development as EDSTAC test guidelines

Level	Purpose	Endpoints	Tests
Tier 1 screening	Priority setting for testing Mechanism of action (MOA) In vitro: inexpensive and specific for a particular MOA In vivo: metabolic response of intact animals	Estrogen/androgen (ant)agonism Assays on diverse taxa: rodent, amphibian, fish Covers sexual development (weight of sex organs), thyroid activity, steroid synthesis and estrogen/androgen (ant)agonism	Proposed in vitro screens: • Estrogen receptor binding or reporter gene assay • Androgen receptor binding or reporter gene assay • Steroidogenesis assay using minced testes Proposed in vivo screens: • Rodent 3-day uterotrophic assay • Rodent 20-day pubertal female assay with thyroid • Rodent 5–7 day Hershberger assay • Frog metamorphosis assay • Fish reproductive screen
Tier 2 testing	Identification and characterization of dose-response in vivo	• Most sensitive life stage (in utero or in ovo) • Multi-generational • Major taxonomic groups represented	Proposed tests Five multi-generation tests: • mammals • birds • fish • amphibians • invertebrates Tests are selected based on exposure and effects information.

mental test or extended ELS test) and a reproduction test, before requiring fish full lifecycle tests in Tier 2 (Hutchinson et al. 2000).

Standardization and Validation

The *Food Quality Protection Act* also states that the screens and tests be validated to ensure that they give reliable and repeatable results. A task force has been established in the U.S. to coordinate the standardization and validation effort among government, industry and public interest groups. The validation framework envisioned for endocrine disruptor screens and tests follows the general process and stages developed by the U.S. Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM) and will proceed in the following stages: 1) research and assay development, 2) demonstration of relevance and standardization, 3) inter-laboratory validation, 4) independent scientific peer review, and 5) agency regulatory approval (ICCVAM 1997).

A pilot demonstration of high throughput pre-screening (HTPS) was completed in March 1999 to determine the suitability of the reporter gene assay systems for ranking and identifying agonists and antagonists for estrogen, androgen and thyroid receptors. HTPS was initially viewed as a rapid, efficient means to provide preliminary endocrine effects data. The goal was never realized due to low signal to noise ratio in the U.S. EPA-sponsored demonstration. The U.S. EPA is still considering applying this technology but is also investigating the use of quantitative structure activity relationships (QSARs) to meet the objective of a quick way to provide an indicator of potential hormonal activity of unknown chemicals. Because QSARs predict a chemical's ability to interact with the active site of a hormone receptor, QSAR information is limited to this one aspect of a chemical's potential to interfere with endocrine systems. The U.S. EPA is considering several different QSAR approaches to assist in them in setting priorities in EDPSD for entry into Tier 1 screening.

The U.S. EPA will initially place a higher priority on standardizing and validating the mammalian assays used to assess health hazard and risk because the *Food Quality Protection Act* and the *Safe Drinking Water Act* are explicitly directed toward protecting food and drinking water to safeguard human health. Work is underway on the pre-validation phase of many of the proposed assays. A Tier 1 battery could be ready in 2003. Tier 2 tests are expected to complete validation between 2003 and 2005. The ecotoxicological screens or tests remain important because the U.S. EPA cannot evaluate ecological hazard and risk on the basis of mammalian tests alone. The U.S. EPA will develop and evaluate the ecotoxicological screens and tests using a phased approach over 2 to 5 years. Once the screens and tests are available, the U.S. EPA plans to implement screening and testing requirements in several phases. The U.S. EPA proposes the first phase to involve pesticide active ingredients and other pesticide formulation ingredients with high production volume. Testing of commercial chemicals and environmental contaminants will be addressed in subsequent phases.

International standardization and validation efforts are being conducted separately from U.S. domestic efforts. Through its Test Guidelines Program, OECD has established the Endocrine Disruptor Testing and Assessment Task Force. Laws, regulatory processes and regulatory procedures differ from country to country. Thus, while the U.S. EPA actively participates as a member of OECD, the OECD Test Guidelines Program and the OECD Task Force, separate domestic and international activities are necessary. Although international activities are distinct from the U.S. task force activities, there is overlapping membership to ensure appropriate liaison and communication, eliminate unnecessary duplication of effort and facilitate international harmonization, where appropriate.

OECD Endocrine Disrupter Screening and Testing

Background

Increased concern over the possibility that exposure to endocrine disrupting substances may be adversely influencing environmental and human health has necessitated increased surveillance of commercial chemicals for endocrine toxicity. As part of its work on Environmental Health and Safety, the OECD develops validated test protocols to determine the hazards of chemical preparations such as pesticides or pharmaceuticals. In 1996, the National Coordinators of the OECD Test Guidelines Program identified the need to develop internationally recognized assays to identify EDS. This need was further emphasized by an OECD survey of member states (OECD 1997) that indicated that most considered that existing tests in mammalian or non-mammalian models were insufficient to identify EDS.

In response to these concerns, in December 1997 the Endocrine Disrupters Testing and Assessment (EDTA) Task Force was formed jointly by the OECD Risk Assessment Advisory Body and the National Coordinators of the Test Guidelines Program. EDTA has the responsibility of identifying and prioritizing new and enhanced test guidelines and developing a comprehensive testing strategy to detect EDS. In addition, EDTA would serve as a forum for communication and cooperation among member countries on EDS-related activities. The main purpose of the EDTA was to ensure that testing and assessment approaches are harmonized as much as possible among member countries. Specifically, their objectives were to provide information and coordinate activities, to develop/revise test guidelines to detect EDS, and to harmonize hazard and risk characterization approaches for EDS.

The initial meeting of the EDTA in March 1998 brought together representatives from 19 OECD member countries in addition to a delegation from the European Union, representatives of industry, and the World Wildlife Fund (OECD 2000b). The Task Force has placed the large number of potential testing needs into a conceptual framework (Table 2, OECD 2000a) to help prioritize the work and to describe the levels of testing and type of test that may contribute to hazard identification and characterization. The approach has three tiers comprised of (a) initial assessment of

Table 2. Summary of the proposed screens and tests for development as OECD test guidelines (modified from OECD 1999)

Level	Purpose
Initial assessment	
Information collection	Existing information collection and analysis. Before using any of the tools in the tool box existing information should be collected and its relevance should be analyzed. No new test data is generated.
Sorting	Sorting based on simple test and measurements. This compartment contains tools to provide basic information, predict data where testing gaps may be present and help set priorities for further work. May contain in vitro tests and high throughput screening in future.

Level	Purpose	Endpoints	In vivo tests
Screening Short-term tests	Mechanism of action Absorption, distribution, metabolism and excretion	Male/female sexual development (weight of sex organs) in pubertal animals to cover thyroid activity steroid synthesis and estrogen/androgen (ant)agonism	Mammalian assays • uterotrophic assay • Hershberger assay • 407 repeated dose toxicity Ecotoxicological assays • juvenile fish assay (TG 204 or TG 215) • adult fish assay • frog metamorphosis assay (proposed) (There is currently no avian screen and invertebrates require further research.)

(continued)

Table 2. *(continued)*

Level	Purpose	Endpoints	In vivo tests
Testing Longer-term tests	Identification/ characterization	Long-term testing: focus on reproduction and development	Mammalian assays • two-generation reproduction assay (enhanced TG 416) • developmental neurotoxicity (TG 414)
	Long-exposure period, complex endpoints	In vivo: sexual development, thyroid, activity steroid synthesis and estrogen/ androgen (ant)agonism	Ecotoxicological assays • fish developmental assay (enhanced TG 210) • fish developmental assay (enhanced TG 210) • fish reproduction test • fish full-life-cycle test (FFLC) • avian reproduction one- and two-generation (enhanced TG 206) (There is currently no definitive amphibian test.)

existing information to set priorities for further evaluation; (b) screening tier to detect chemicals which may be endocrine disrupters in humans and wildlife (i.e., estrogen, androgen and thyroid active substances); and (c) testing tier to confirm, characterize and quantify the nature of the endocrine disrupting activity identified at screening tier (e.g., dose-response relationship, type of organisms at risk, sensitive life stage). Each tier contains assays with increasing biological complexity in both mammalian and ecotoxicological areas (OECD 2000a,d). For example, the complexity of tests varies from those addressing interaction with hormone receptors, such as in screening level tests, to tests in whole animals that involve exposure at different stages of an animal's life cycle and multi-generational effects.

These tiers follow sequentially with results from one tier determining the extent to which subsequent tiers are utilized. It was also recognized that frameworks would exist in parallel with some separate assays for mammalian (human) and non-mammalian or environmental (ecotoxicological) EDS effects. Some screening tests, however, will be relevant to both frameworks due to the high degree of similarity between sex steroid and thyroid hormone physiology between mammals and other vertebrate species (OECD 2000a).

Validation Management Committees

The EDTA recommended the establishment of two validation management groups (one for mammalian testing and one for ecotoxicological testing) to coordinate the development and validation work (OECD 2000b). It was recommended that the validation management committees be comprised of relevant experts from member countries and include representatives of Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and European Centre for Validation of Alternative Methods (ECVAM) to provide objective review with a view to facilitating regulatory acceptance. The roles of the EDTA and its subordinate validation management groups are distinct. Within the framework and management structure of the Test Guidelines Program, EDTA focuses on identifying and prioritizing testing and assessment needs in both the mammalian and ecotoxicological areas. The validation management groups act on the recommendations from EDTA and ensure that any validation work is conducted acceptably and in accordance with the agreed approach.

OECD Mammalian Test Methods

At an Expert Consultation Meeting in Washington in August 1998, at which specific EDS screening and testing protocols were discussed, three mammalian test methods were identified as highest priority for validation as screening methods (OECD 1997). These were two test methods to examine activation or inhibition of estrogen or androgen receptor in vivo (rodent uterotrophic and Hershberger assays, respectively) and one existing subchronic toxicity OECD test guideline (TG 407 — 28 day subchronic toxicity assay; Table 2).

Rodent uterotrophic assay

Uterine weight in rats or mice has long been recognized as a sensitive and specific response to estrogen administration. Estrogenic or anti-estrogenic activity of a chemical (dosed via subcutaneous injection or oral exposure route) is determined as increased or decreased uterine weight after 3 days. The test uses either weanling animals whose ovaries are immature or mature ovarectomized females, as endogenous estrogens compromise the sensitivity of the assay. In a pre-validation exercise, 19 different laboratories tested the effects of ethinylestradiol, comparing the sensitivity of immature and mature ovarectomized models and oral versus subcutaneous exposure in the immature model. The assay was conducted 41 separate times with remarkably consistent results. Details of this exercise will be reported shortly in the scientific literature, and the validation will continue with a variety of weakly estrogenic compounds.

Hershberger assay

The Hershberger assay compares weights of the sexual accessory tissues (SAT) of castrated male rats (Hershberger et al. 1957). Androgens cause increased SAT weights, while anti-androgenic activity is determined as reduction in weight of SAT after co-administration of a reference androgen with the test compound. A pre-validation exercise was conducted using male rats castrated at 6 weeks of age to determine the optimum dose of testosterone propionate (TP) to be used to stimulate regrowth of SAT and to examine inter-laboratory variability. A subsequent exercise will determine the dose-response of flutamide in reducing SAT weights stimulated by an optimum dose of TP.

TG 407 — 28-day subchronic toxicity assay

One of the responsibilities of EDTA was to determine if any of the existing OECD test guidelines could be modified to detect EDS activity. The 28-day repeated dose toxicity test, TG 407, was seen as the most suitable to modify, with added endpoints to examine changes in circulating thyroid hormones and thyroid stimulating hormone, measures of epididymal sperm numbers and morphology, the weights of reproductive and accessory sex tissues, and the weights of the pituitary and thyroid glands. The chemicals initially tested include six strongly acting substances that affect estrogen, androgen or thyroid physiology (ethinylestradiol, tamoxifen, methyltestosterone, flutamide, thyroxin and propylthiouracil) and four weakly acting substances (genistein, a phytoestrogen; nonylphenol, a estrogenic industrial compound; DDE, an antiandrogenic environmental contaminant; and fenarimol, an aromatase inhibitor).

OECD Ecotoxicological (Wildlife) Test Methods

OECD has prepared a detailed review paper to appraise the current testing methods available for the detection of chemicals that may disrupt the sex hormone system. In addition, a questionnaire to member countries

was used to assess priority needs and issues of concern. Discussion has taken place among experts on potential avian, amphibian, fish and invertebrate tests for endocrine disrupters.

Testing for endocrine disruption in birds

At the meeting of the OECD avian expert group, held in May 1999, there was general agreement on the development of a both one-generation and two-generation reproduction toxicity test in quail for the assessment of endocrine disrupting effects in birds (OECD 2000e). The drafts have been circulated to national coordinators early in 2000, and decisions regarding validation will be made at upcoming OECD avian expert group meetings.

Testing for endocrine disruption in amphibians

The U.S. EDSTAC proposes a frog metamorphosis assay for their Tier 1 screening battery and an amphibian development and reproduction test for Tier 2 testing. The rationale for a frog metamorphosis assay is that this test can also be used as surrogate for detecting potential thyroid effects in humans. The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) questions the need for the frog metamorphosis assay in Tier 1, as they feel that the mammalian screens will detect (anti-)thyroid activity (Hutchinson et al. 2000). In Germany, an amphibian metamorphosis test to identify disruption of the hypothalamus-pituitary-thyroid axis is under development (OECD 2000d). German regulating authorities within their pesticides registration processes require the test when hormonal effects (i.e., pituitary, hypothalamus or thyroid gland) have been demonstrated in mammals and when additional information on endocrine disrupting effects in aquatic organisms is needed. Consequently the OECD EDTA has established an expert group on "Endocrine Disrupters Testing in Amphibians" to further develop a suitable method and draft a proposal for a frog metamorphosis assay using protocols developed by Germany and the U.S., respectively (OECD 2000d).

Testing for endocrine disruption in fish

A testing scheme consisting of three tiers was proposed (OECD 2000c,f). Screening contains short-term tests (3–8 weeks) aimed at (i) identifying first signs of endocrine activity in vivo, (ii) setting priority for testing, and (iii) investigating mechanisms of action. The OECD expert group on EDS testing in fish recommends drafting two protocols for Tier 1 screening assays using juvenile (based on OECD Test Guideline 215/204) and adult fish. Tests should be applicable to different species and examine three core endpoints: gross morphology (such as hepatosomatic index, gonadosomatic index, secondary sex characteristics); vitellogenin in blood, whole body or liver tissue; gonad histopathology (assessing stage of gonadal development, and abnormalities such as ovipositor formation or evidence of intersex) (OECD 2000c). In addition to these core endpoints, supplementary endpoint(s) might be included optionally (e.g., fecundity, fertility, etc.). Generally, the standard OECD tests species, such

as fathead minnow, Japanese Medaka, juvenile carp, zebra fish and rainbow trout, are being investigated. Testing Tier 2 involves extended developmental and reproductive endpoints aimed at the identification and characterization of endocrine disrupting effects. Further research and prevalidation are being carried out for three fish tests: a developmental early life stage test (based on OECD Test Guideline 210), a reproduction test and a full life-cycle test (OECD 2000c). Enhancements to OECD Test Guideline 210 for a developmental test are being investigated by the European Chemical Industry Council (CEFIC) within the Endocrine Modulator Steering Group (EMSG) research project. The assay exposes larval fish during sexual differentiation and early gametogenesis, and assesses hatching times and larval survival, embryo larval development, weight and length, abnormal appearance or behaviour, and sex ratio (based on histopathology). Reproduction tests and partial life cycle tests are being investigated within research projects at the U.S. EPA for fathead minnow and in Japan for Medaka. These expose sexually mature adult fish and assess time to first spawning, spawning frequency, number of eggs per hatch and per female, number of fertilized eggs and hatching success of F1 generation. Several research programs include the development of a full lifecycle test for fathead minnow (U.S. EPA), zebra fish (Nordic Joint Research Project, Germany) and Medaka (Japan: CERI and Kyushu University; Nagoya University). The endpoints that are thought important include time to sexual maturity, secondary sex characteristics, sex ratio, spawning frequency/behaviour, egg production, fertilization success, embryo hatching and viability, larval survival/growth/development, histopathology, gamete maturation, and the selected biochemical markers vitellogenin, steroids and steroid enzymes (OECD 2000c).

Invertebrate testing for endocrine disruption

Various work groups have been discussing the development of the invertebrate bioassays. The international SETAC/OECD expert workshop on "Endocrine Disruption in Invertebrates: Endocrinology, Testing and Assessment" (EDIETA) was held in December 1998. The group noted that there are existing methods for lab testing of aquatic and terrestrial invertebrates (e.g., crustaceans, insects) and that these tests may provide a good foundation for detecting EDS (OECD 2000d). Many invertebrate toxicology tests include endpoints that may reflect changes in endocrine function (e.g., growth, development, reproduction). However, the roles for many hormones and their potential target sites have not yet been established for invertebrates. There needs to be considerable basic research on invertebrate endocrinology to select the best species and endpoints for invertebrate assays. Because of this, the OECD felt it was too early to recommend an invertebrate test for screening or testing (OECD 2000d). ECETOC agree that it is too early to recommend invertebrate tests (Hutchinson et al. 2000), as the science must first address the serious gaps in our basic knowledge of endocrine function in these and other invertebrate taxa.

OECD Reference Chemicals for Research and Development Studies

The OECD expert groups discussed selection of reference chemicals recommended for research and development purposes. It was agreed that thyroid active substances should not be included, as more research is needed with respect to chemical interaction with thyroid hormone systems. The following list of five chemicals was agreed as a priority core set of reference compounds (OECD 1999): ethinylestradiol ('strong' estrogen), ZM-189,154, (anti-estrogen), methyltestosterone (androgen), flutamide (anti-androgen) and 4-t-pentylphenol, ('weak' estrogen). A second list of four chemicals to be considered as additional series of reference compounds for further studies was also suggested: diethylstilbesterol (estrogen), methoxychlor ('weak' estrogen), fadrozole, (aromatase inhibitor) and genistein, (phytoestrogen) (OECD 1999). Researchers and test validation labs can access standard batches of test compounds from a repository. This allows for comparisons among labs and countries, and removes the possibility of chemical impurities and batch-to-batch variability influencing the results.

Strategies for Managing EDS Testing and Monitoring

The Commission of the European Countries

In addition to the above detailed tiered screening and testing schemes of the U.S. EPA and OECD, several countries and international organizations have developed general testing and research strategies to deal with EDS. The European Union (EU) is one such organization, and we outline their EDS program (Commission of the European Communities 1999). In view of its role in protecting the EU citizens and the environment, and following a precautionary approach, the Commission of the European Communities issued a Community Strategy for Endocrine Disruptors to the Council and the European Parliament in November 1999. This document outlined their proposed strategies and actions to assess and control potential EDS, and takes into account existing consumer health and environmental protection policies (Commission of the European Communities 1999). The EU strategy includes actions in the short-, medium- and long-term, based the time frame within which results can be achieved, i.e., 1–2 years for short-term actions, 2–4 years for medium-term actions and more than 4 years for long-term action.

Short-term goals include compiling a list of substances for 'priority' testing and further evaluation of their role in endocrine disruption. However, until agreed test methods and an effective screening and testing strategy are available, many substances may escape attention when compiling lists of potential ED substances. Therefore, the EU will support research on selection criteria for identifying substances and on development and validation of test methods for the identification of EDS. The Commission realizes they must identify knowledge gaps to guide research and monitoring efforts, and support the development and validation of environmental monitoring tools (Commission of the European Communities 1999).

As well, the Commission proposes to document the quantities of these substances in the environment, including production, consumption and import/export volumes. The Commission will identify specific cases of consumer use (e.g., babies, children), for special consideration from a consumer policy point of view. The Commission will implement a public communication program in the interests of transparency and public confidence and continue to support international cooperation and coordination in order to make best use of available resources and to avoid duplication of efforts (Commission of the European Communities 1999).

European Community action plans to take into account the wider international context, the compliance costs and potential impact for any given policy action. In the short-term, legislative actions are to include establishment of monitoring programs to estimate exposure to and effects of the substances on the EDS priority list and collection of data on observed effects in order to understand potential causal links to specific chemicals. Medium-term actions of the Commission will support the development of agreed test methods within the framework of OECD and for an appropriate EU testing strategy. Results of ongoing research projects need to be fed into the policy process. In addition, identification of substitute substances and consideration of voluntary initiatives will be an important action in order to eliminate or find replacements for chemicals of concern (Commission of the European Communities 1999). The EU's long-term actions focus on implementation of EDS legislation by adaptation or amendment of the EU legislation currently covering environmental and human health aspects of chemicals (such as existing directives on classification, regulation and risk assessment).

The Canadian Agenda for Assessing Endocrine Disrupting Substances

The development of internationally accepted and validated screening and testing methodologies is an important proactive research and regulatory goal. The complexity of this issue necessitates an integrated international effort, not only on screening and testing, but also on the broader issues of conducting scientifically defensible risk assessments and developing risk management strategies. The close linkage made between the EDS issue and the application of the Precautionary Principle (e.g., EU communications on EDS and Precautionary Principle, Commission of the European Communities 2000), also emphasizes the urgency and importance of addressing this issue. The *Canadian Environmental Protection Act, 1999* (CEPA 1999), as well as making research on EDS a requirement, makes application of the Precautionary Approach and a weight of evidence approach mandatory in decision making. For the immediate future, how these concepts are applied to the EDS issue in other jurisdictions and in Canada may have large implications for the management of chemicals.

The implications of these international activities on screening and testing for EDS were considered as part of a multi-stakeholder workshop on the environmental assessment of EDS in the Canadian environment hosted by the federal government's Five Natural Resources Departments

(5-NR) EDS Working Group, February 2000 (Servos et al. 2000a,b [the 5-NR are comprised of Environment Canada, Health Canada, Agriculture and Agri-Food Canada, Fisheries and Oceans Canada, and Natural Resources Canada]). The conclusions and recommendations of this workshop have been incorporated as components of the Canadian government's agenda on scientific assessment of EDS in the environment (Servos et al. 2001). The workshop participants identified four main EDS areas that the federal government should support: 1) national leadership and communication on the EDS issue, 2) better knowledge of exposure and effects in the Canadian environment, 3) national and international harmonization of screening and testing protocols, and 4) enhanced scientific assessment and action on priority substances.

The Canadian agenda recognizes that international activities will continue to have major implications for the development of the EDS issue in Canada. The major activities in the U.S. EPA, EU and OECD related to screening and testing of EDS will influence both public perception and development of Canadian policy. As screening and testing develops in other countries, there will be increasing public pressure to apply and integrate these results within a Canadian context. For example, the development and publication of lists for priority screening and testing in other jurisdictions, such as the EU, will influence Canadian programs and priorities. To ensure that the interests of Canadians and the Canadian environment are considered, Canada needs to play an active role internationally by communicating and contributing to the effort to develop, harmonize and validate internationally accepted screens and tests (Servos et al. 2000a). The results of this process and the resulting EDS tests should also be assessed for their applicability to the Canadian situation and existing regulatory and risk management frameworks. In the immediate future, programs such as the New Substances Program and Domestic Substance Program under the *Canadian Environmental Protection Act*, and the assessment of pesticides by the Pest Management Review Agency will need to address how these tests will be used with a Canadian context.

There was also a general consensus among workshop participants that Canada should focus its efforts and resources on areas that would not exactly duplicate the efforts of others regarding the development of these tests. Emphasis of research in Canada should be to determine the ecological significance of the tests to better assist in the application of risk assessment. There is a need to relate the results of standardized tests to possible impacts in the environment or human populations, especially with regard to low-dose effects on early growth and development. Canadian field research such as monitoring of the effects of organochlorines in birds in the Great Lakes and reproductive effects on fish exposed to pulp mill effluents has contributed significantly to the identification and development of the EDS issue (Servos and Luce 1997). Validating the applicability of the standardized tests to assessment of the potential adverse impacts on Canadian populations and environments should continue as an emphasis.

Although it is expected that tests will be applicable to the general Canadian population, there may be selected populations that are unique and may require special consideration, e.g., aboriginal populations with traditional diets (Servos et al. 2000a). Canadian environments, such as the Arctic, may require special consideration because of unique species and food chains and the cold climate that may influence the fate, exposure and bioavailability of EDS. A variety of other environments and species (e.g., migratory birds, marine mammals, endangered or threatened species) may also be at increased risk because of unique characteristics. Assessing the exposure of Canadians and the Canadian environment to the vast array of EDS will be a critical component of applying the results of the effects testing within a risk management framework. Links between standardized testing results and potential for impacts on the environment will greatly reduce uncertainty associated with risk assessments and development of risk management strategies.

References

Ankley GT, Johnson RD, Detenbeck N, Bradbury SP, Toth G, Folmar L. 1997. Development of a research strategy for assessing the ecological risk of endocrine disruptors. Review Environ. Toxicol. **1:**231-267.

Ankley G, Mihaich E, Stahl R, Tillit D, Colborn T, McMaster S, Bantle J, Campbell P, Denslow N, Dickerson R, Folmar L, Fry M, Giesy J, Gray L E, Guiney P, Hutchinson T, Kennedy S, Kramer V, LeBlanc G, Mayes M, Nimrod A, Patino R, Peterson R, Purdy R, Ringer R, Thomas P, Touart L, Van Der Kraak G, Zacharewski T. 1998. Overview of a workshop on screening methods for detecting potential (anti-) estrogenic/androgenic chemicals in wildlife. Environ. Toxicol. Chem. **17:**68-87.

CEPA 1999. *Canadian Environmental Protection Act*. 1999. The House of Commons of Canada, Bill C-32, as passed by the House of Commons, June 1, 1999.

Colborn T, Dumanoski D, Myers JP. 1996. Our stolen future. A Dutton book, New York.

Colborn T, Clement C. 1992. Chemically induced alterations in sexual development: the wildlife/human connection, *In* Advances in Modern Environmental Toxicology, Vol. 21, Colborn T, Clement C (ed), Princeton Scientific, Princeton, NJ.

Commission of the European Communities. 1999. Community strategy for endocrine disrupters: a range of substances suspected of interfering with hormone systems of humans and wildlife. Communication from the Commission to the Council and the European Parliament. Brussels. 32 p.

Commission of the European Communities. 2000. Communication from the Commission on the Precautionary Principle. Brussels, 29 p. *http://www.eu.org*.

EDSTAC (Endocrine Disruptor Screening and Testing Advisory Committee). 1998. Endocrine Disruptor Screening and Testing Advisory Committee: Final report. Environmental Protection Agency: Washington, DC. *http://www.epa.gov/scipoly/oscpendo/history/* and *www.epa.gov/endocrine*.

Gray E, Kelce W, Weis T, Tyl R, Gaido K, Cook K, Klinefelter G, Desaulnier D, Wilson E, Zacharewski T, Waller C, Foster P, Laskey J, Giesy J, Laws S,

Mclachlan J, Breslin W, Cooper R. DiGiulio R, Johnson R, Purdy R, Mihaich E, Safe S, Colborn T. 1997. Endocrine Screening Methods Workshop Report: Detection of estrogenic and androgenic hormonal and anti-hormonal activity for chemicals that act via receptor or steroidogenic enzyme mechanisms. Reproductive Toxicol. **11:**719-750.

Hershberger LG, Shipley EG, Meyer RK. 1957. Myotrophic activity of 19-nortestosterone and other steroids determined by modified levator ani muscle method. Proc. Soc. Expt. Biol. Med. **83:**175-180.

Hutchinson TH, Brown R, Brugger KE, Campbell PM, Holt M, Länge R, McCahon P, Tattersfield LJ, van Egmond R. 2000. Ecological risk assessment of endocrine disrupters. Environ. Health Perspect. **108:**1007-1014.

ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). 1997. Validation and regulatory acceptance of toxicological test methods. A report of the Ad Hoc Interagency Coordinating Committee on the Validation of Alternative Methods, NIH Report No. 97-3981, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, March 1997.

Kavlock RJ, Daston GP, DeRosa C, Fenner-Crisp P, Gray LE, Kaattari S, Lucier G, Luster M, Mac MJ, Maczka C, Miller R, Moore J, Rolland R, Scott G, Sheehan DM, Sinks T, Tilson HA. 1996. Research needs for the risk assessment of health and environmental effects of endocrine disruptors: a report of the U.S. EPA-sponsored workshop. Environ. Health Perspect. **104(Suppl. 4)** 715-740.

Krimsky S. 2000. Hormonal chaos: the scientific and social origins of the environmental endocrine hypothesis. The John Hopkins University Press, Baltimore, MD.

OECD. 1997. Draft detailed review paper: Appraisal of test methods for sex-hormone disrupting chemicals. Further publication as OECD Environment, Health and Safety Publication, Series on Testing and Assessment No. 20, Paris, 2000.

OECD. 1999. Screening and testing for endocrine disruption in fish. Document prepared for the fourth meeting of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA), Paris, May 17, 2000.

OECD. 2000a. Current and expected assessment approaches: revision of the conceptual framework. ENV/JM/EDTA(2000)8. Document prepared for the fourth meeting of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA), Paris, May 17, 2000.

OECD. 2000b. OECD chemicals programme. Coordination of endocrine disrupters assessment. *http://www.oecd.org/ehs/endocrin.htm.*

OECD. 2000c. Record from the second OECD expert consultation on endocrine disrupters testing in fish (EDF2), Tokyo, March 15–16, 2000. *TG\Fish\2000.02 _EDF2 record.*

OECD. 2000d. Status report on environmental effects testing in amphibians and invertebrates. ENV/JM/EDTA(2000)3. Document prepared for the fourth meeting of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA), Paris, May 17, 2000.

OECD. 2000e. Status report on environmental effects testing in birds. ENV/JM/EDTA(2000)2. Document prepared for the fourth meeting of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA), Paris, May 17, 2000.

OECD. 2000f. Status report on environmental effects testing in fish. ENV /JM/EDTA(2000)1. Document prepared for the fourth meeting of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA), Paris, May 17, 2000.

Servos M, Luce S. 1997. Proceedings of Environment Canada's Workshop on Endocrine Disruptor Compounds: Identifying Research Needs and Priorities. Niagara Falls, Ontario, Oct. 23-24, 1997. 115 p

Servos M, Delorme P, Fox G, Sutcliffe R, Wade M. 2000a. Proceedings of the 5-NR Workshop: Establishing a National Agenda for the Scientific Assessment of Endocrine Disrupting Substances. 340 p.

Servos M, Delorme P, Fox G, Sutcliffe R, Wade M. 2000b. Establishing a National Agenda for the Scientific Assessment of Endocrine Disrupting Substances: Workshop executive summary. 26 p.

Servos M, Delorme P, Fox G, Sutcliffe R, Wade M. 2001. A Canadian perspective on endocrine disrupting substances in the environment. Water Qual. Res. J. Canada, **36:**331–346.

U.S. Environmental Protection Agency. 1997. Final report of the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC).