

APPENDIX H

Draft Strategy for Assessing Exposure to and Effects of Toxic Substances in the Great Lakes

Executive Summary

A key problem identified in the Great Lakes is the continued presence and emergence of toxic substances at levels considered unsafe for humans and wildlife. Evaluating the risks of potentially toxic substances has taken a largely empirical, exposure-based approach through monitoring of selected chemicals in various Great Lakes media and biological receptors. Therefore, more effects-based information needs to be developed along with a strategic emphasis on including new approaches to predict and evaluate toxicity of chemicals based on enhanced knowledge of biological systems. In addition to measuring effects at the individual and population level, the inclusion of sub-organismal endpoints, which can be incorporated into predictive toxicological methods, would enhance our ability to link chemical exposures to adverse effects relevant to risk assessment endpoints. The strategy proposed in this document is intended to provide an updated, independent scientific approach to the assessment of toxic substances in the Great Lakes. It incorporates both prospective and retrospective techniques for surveillance and effects monitoring. Prospective methods that incorporate improved predictive approaches are valuable in providing screening level information. Retrospective methods are important for diagnostic purposes and establishing causality between chemical exposure and adverse effects. The developed strategy is based on the premise that most of the priority issues pertaining to toxic substances in the Great Lakes can be addressed in three thematic areas: 1) Predictive and Empirical Assessment, which is also considered a prospective approach, 2) Effects Monitoring and Detection (prospective and retrospective approaches), and 3) Diagnostic Assessment (retrospective analysis). This strategy is further augmented by incorporating the Adverse Outcome Pathway (AOP) concept, a conceptual framework that portrays existing knowledge concerning the linkage between a direct molecular initiating event of a toxic substance (i.e., exposure) to an adverse outcome in a species of interest at a level of biological organization relevant to risk assessment.

Empirical testing approaches that include apical endpoints are the most direct with the lowest uncertainty when determining effects of interest resulting from exposure to a specific chemical. Predictive models and *in vitro* toxicological tools are also useful risk assessment tools that can take many forms and can range widely in their complexity. Quantitative Structure Activity Relationship (QSAR) Models relate physicochemical properties of chemicals to quantitative estimates of activity. Read-across is another predictive technique that uses chemical categories to fill data gaps for specific chemicals. The prerequisite to using this approach is that there must be a definable chemical category that can be ordered according to some measure or characteristic related to the potency of each chemical. Other types of species extrapolation models for sub-lethal effects use a homology-based approach, which assumes that phylogenetic relationships can be used to predict sensitivity toward a specific chemical. For example, the lack of a functional estrogen receptor in invertebrates explains the relative insensitivity of invertebrates to synthetic estrogens. Predictive modeling is anticipated to be improved in the future by incorporating results from *in vitro* and high throughput assays (HTA). *In vitro* results can be very useful as indicators of a specific activity or effect. Such activity determinations are most useful when linked to an AOP, where the role of that activity can be interpreted in a toxicological context. However, many uncertainties remain in extrapolating from

such assays, particularly the representativeness of individual and batteries of isolated *in vitro* system responses as compared to whole animal system responses. Evaluating the risks of chemicals of emerging concern (CECs) will need to rely on a weight of evidence approach, which can be informed from predictive models and *in vitro* testing, but these are unlikely to be sufficient tools in themselves to provide clear evidence for risk decisions.

Biological monitoring and surveillance provide important information on the current status of ecological populations and communities. There are a diverse variety of assays, endpoints, and sample types that could be useful for environmental monitoring. These include lab-based *in vitro* or *in vivo* assays using complex mixtures from the field, such as discharge or receiving waters, sediments/sediment fractions, and even tissue extracts. Samples also could be from organisms held *in situ* (e.g., caged fish studies), or animals collected from extant populations. Well-established monitoring endpoints associated with the latter types of studies are tissue-somatic indices such as the gonadal and liver somatic indices, fecundity, and condition index. Sub-organismal endpoints such as unnatural production of the egg-yolk protein vitellogenin and induction of ethoxyresorufin-O-deethylase (EROD) activity are also well-established endpoints. Other promising tools appear to require additional research and validation before use in monitoring. This would include global measurements of changes in expression of genes (transcriptomics), proteins (proteomics), and metabolites (metabolomics). These latter methods hold much promise both for identifying abnormal/perturbed biology and, through “fingerprinting” (pattern recognition) approaches, diagnosing causes of the perturbation.

Diagnostic assessments are based on the assumption that some type of adverse biological response has been observed, which may be associated with chemical exposure. It is recognized that the range of measures that might constitute a “biological response” is wide and varied. With this in mind, a strategy is presented for investigating the cause of observed effects with the goal of identifying the specific chemical or chemicals responsible for the observed response. In developing this strategy, several broad and specific recommendations are made to facilitate implementation. The overarching recommendation is the need to incorporate effects-based monitoring to complement chemical-based approaches. There is a clear need to maximize limited resources through leveraging with existing Great Lakes programs. To facilitate diagnostic assessments, prioritization of the most important toxicological pathways of concern is needed to support assay/endpoint selection and test development. The biological effects monitoring should include a mix of both sub-organismal (physiological) / individual based endpoints and, especially for CECs, biological pathway-based mechanistic assays. Another recommendation is to incorporate—where appropriate—tools that are currently ready for monitoring applications, but also support the continued development and validation of promising newer tools. Finally, it is recommended that two types of effects-based monitoring pilot studies be conducted to: (1) emphasize broad spatial coverage to help characterize baseline conditions, and (2) more intensively characterize impacted sites to explore utility of the methods as diagnostic tools in a field setting.

1.0 Problem and Need Statement

A key problem identified in the Great Lakes is the continued presence, as well as emergence, of toxic substances at levels which may be above those considered safe for humans and wildlife. “Legacy contamination” of chemicals from past practices continues to circulate through the Great Lakes system, warranting fish consumption advisories, causing drinking water concerns, and limiting recreational use of waters. In addition to these well-known toxicants, there are also chemicals of emerging concern that are now being released to the Great Lakes and subsequently detected in water, sediments, fish, and wildlife. In terms of potentially toxic compounds, 500-1000 new chemicals are added to the US Toxic Substances Control Act (TSCA) inventory each year, an inventory that lists about 83,000 chemicals. Classes of CECs, such as pharmaceuticals and personal care products, are being detected with increasing frequency in environmental samples from the Great Lakes. For example, in a review of the literature from 1997-2008, Klečka et al. (2010) found that 326 chemicals were measured in water, sediment, and biota sampled from the Great Lakes. This analysis was focused on CECs and excluded “legacy” contaminants. Nearly half of the chemicals detected were either current use pesticides or pharmaceuticals, and many were found at “trace” levels. Not surprisingly, wastewater treatment plants and agricultural operations were identified as important sources of contamination. Finally, a key conclusion of the paper was that, although anthropogenic chemicals have been detected in the Great Lakes basin, many have insufficient toxicological data associated with them to permit interpretation of their relative risk. Therefore, to better protect human health and the Great Lakes ecosystem, such substances must be better understood with respect to their potential risks. This creates an immediate need for data that supports screening level assessments and prioritization of chemicals for further consideration.

Within the Great Lakes, evaluating the risks of potentially toxic substances has taken a largely empirical, exposure-based approach through monitoring of selected chemicals in various Great Lakes media and biological receptors. This has resulted in an extensive knowledgebase regarding the spatial and temporal status of selected chemicals in the Great Lakes aquatic ecosystem. “Legacy” contaminants, such as PCBs, for example, have been studied for decades. More recently, monitoring chemicals of emerging concern (e.g., human and veterinary pharmaceuticals, personal care products, flame retardants, endocrine disrupting chemicals (EDCs), and other industrial chemicals, etc.) indicates that many are present in the Great Lakes aquatic ecosystem. However, reliance on chemical monitoring alone cannot provide a complete understanding of the impact of toxic substances to the Great Lakes. For example, detection of additive or synergistic interactions among contaminants can be difficult to identify from chemical monitoring data due to a lack of mechanistic understanding of the toxic mode of action. Prediction of biological effects at environmental exposure levels is also difficult due to a lack of dose-response information for the majority of contaminants. It is also recognized that chemical monitoring tends to focus on established contaminants and overlooks emerging chemicals and/or chemicals thought to be innocuous. Even for established contaminants, there is no guarantee that analytical detection limits are always adequate to detect at toxicologically important levels. Therefore, more effects-based information needs to be developed along with a strategic emphasis on including new approaches to predict and evaluate toxicity of chemicals based on enhanced knowledge of biological systems. In addition to measuring effects at the individual and population levels, the inclusion of sub-organismal endpoints, which can be incorporated into predictive toxicological methods, would enhance our ability to link chemical exposures to adverse effects relevant to risk assessment endpoints.

Other factors stimulating the development of a new, effects-based strategy for evaluating the impact of toxic chemicals in the Great Lakes are the advances in scientific knowledge and

capabilities and evolving programmatic strategies and priorities. Recent scientific advances in several disciplines and technologies have significantly impacted ecotoxicology. Molecular methods in biology have proliferated, extending and facilitating the application of genomic methods to a variety of species. The understanding of biological systems has expanded, bringing new insights to the mechanisms involved in toxicological processes. *In vitro* methods that provide medium to high throughput approaches to evaluating specific, toxicologically relevant endpoints continue to proliferate, providing more alternatives to organismal testing. Modeling approaches for predictive toxicology continue to improve and provide more meaningful information to the assessment process. Collectively, these changes must be taken advantage of in a new strategy to assess the effects of toxic substances.

With regard to programmatic strategies, several different stakeholders have undertaken efforts to develop strategies to address chemicals in the environment. Each of these strategies has focused on sector-specific chemicals and/or has emphasized the current programmatic objectives related to specific media. Some examples include: US Environmental Protection Agency Committee on the Environment and Natural Resources (CENR) sponsored interagency strategy for pharmaceutical chemicals (Bloom 2008), State of California strategy to assess the aquatic risks of CECs in coastal and marine ecosystems (SCCWRP 2010), WERF strategy for trace organic chemicals in wastewater (WERF 2010), Environmental Council of the States strategy for CEC (ECOS 2010), among others. These recent activities add to the existing programmatic strategies associated with pesticides, toxic substances, persistent toxic substances (Environment Canada 2008), and others (see Appendix B).

Important programmatic priorities specific to the Great Lakes include the US-Canadian Great Lakes Water Quality Agreement (WQA), which is undergoing renegotiation in 2011. The WQA was originally established in 1978 and has been updated periodically to reflect changes in programmatic priorities. The US has also recently adopted the Great Lakes Restoration Initiative (GLRI) and allocated significant funding to execute restoration activities detailed in the GLRI Action Plan for 2010-2014. Goals of the GLRI Action Plan include: preventing release of toxic substances, eliminating persistent toxic substances, reducing exposure to toxic substances from historic sources, reducing the levels of toxic chemicals in Great Lakes fish, and protecting the health and integrity of wildlife populations from adverse chemical and biological effects associated with toxic substances in the Great Lakes basin. These goals require a forward-thinking strategy to assess the effects of exposure to toxic chemicals that are known to be present and also those that are emerging as potentially harmful, and develop assessment approaches for use in identifying mitigation options.

The strategy proposed in this document is intended to provide an updated, independent scientific approach to the assessment of toxic substances in the Great Lakes. The strategy addresses important questions facing stakeholders such as: How can we predict toxicity in the absence of established criteria? How can we determine if there are existing adverse effects? How can we diagnose the cause? These questions are addressed in the proposed strategy by using both prospective and retrospective techniques for surveillance and effects monitoring. Prospective methods are particularly valuable in providing screening level information. This also includes the use of improved predictive approaches that exploit recent advances in the relevant scientific fields. Retrospective methods are important for diagnostic purposes and for establishing causality between chemical exposure and adverse effects. Ultimately, the long-term purpose of the strategy is to provide a scientifically sound approach for effects-based monitoring that will guide restoration and research activities in the Great Lakes for several years.

2.0 Initial Strategy Development

The developed strategy is based on the premise that most of the priority issues pertaining to toxic substances in the Great Lakes can be addressed in three thematic areas: 1) Predictive and Empirical Assessment, which is also considered a prospective approach, 2) Effects Monitoring and Detection (prospective and retrospective approaches), and 3) Diagnostic Assessment (retrospective analysis). The scope of this strategy is limited to effects on aquatic species and associated wildlife, and it does not consider human health. It also does not consider sediment remediation, due to the advanced stages of planning currently underway and the specific nature of sediment remediation plans.

The strategy relies on the Guidelines for Ecological Risk Assessment (ERA) (USEPA 1998) as the overarching organizing principle. In general, the ERA approach requires a cogent problem statement, exposure information, and effects information, all of which feed into a risk characterization. The ERA approach has proven to be a durable approach which is germane to the issues associated with the presence of potentially toxic substances in the Great Lakes.

The strategy is further augmented by incorporating the Adverse Outcome Pathway (AOP) concept, a conceptual framework that portrays existing knowledge concerning the linkage between a direct molecular initiating event of a toxic substance (i.e., exposure) to an adverse outcome in a species of interest at a level of biological organization relevant to risk assessment (i.e., effect) (Ankley et al. 2010). The latter point emphasizes the need for the assessment endpoints to be linked to the risk management objective, or problem statement, in a way that supports regulatory and/or management needs and objectives. At the organismal level, assessment endpoints are most commonly considered survival, growth, development, and reproduction; endpoints that ultimately are relevant to population status via individual fitness. Unfortunately, effects on these endpoints are not often readily available, impeding completion of the ERA rationale. Therefore, a logical approach is needed which permits the use of a broader array of information in executing the ERA rationale. The AOP construct achieves this by establishing a rational and plausible framework for pathways of concern. Thus, AOPs provide the biological rationale for developing predictive toxicological tools, prioritizing and focusing empirical testing, facilitating inter-species and inter-chemical extrapolation, aiding exposure analysis and reconstruction, diagnosing effects/impairments, and addressing issues associated with understanding complex mixtures.

Ecological risk assessments must take into account both the exposure and the effects perspectives. With regard to potentially toxic chemicals in the Great Lakes, the available information tends to be relatively exposure-rich and effects-poor. Therefore, this document reflects an emphasis on the effects side of the problem. When considering CECs, it will be important to assess both the likelihood of future adverse effects (i.e., prospective) and evaluate the likelihood that effects are caused by past exposure to stressors (i.e., retrospective). In many cases, both approaches are included in a single risk assessment. For example, a retrospective risk assessment designed to evaluate the cause for amphibian population declines may also be used to predict the effects of future management actions. Combined retrospective and prospective risk assessments are typical in situations such as the Great Lakes, where ecosystems have a history of previous impacts and the potential for future effects from multiple chemical, physical, or biological stressors.

3.0 Prospective Assessments using a Tiered Assessment Strategy

Prospective assessments investigate the potential impacts of chemical contamination that has not yet occurred, as in the case of a newly registered chemical or at sites receiving new sources of contamination (e.g., newly built wastewater treatment plants, industrial plants). Prospective ecological risk assessments in several testing programs have utilized a tiered testing strategy. This approach uses a hierarchical sequence of tests or endpoints that increase in biological complexity (Balls and Fentem 1992). The objective of tiering is to balance the permissible uncertainties in a risk assessment situation with the resources required to reduce those uncertainties through incremental increases in time, effort, and money. Thus, the permissible uncertainties associated with a specific risk characterization should be articulated in the problem formulation, *a priori*, as this determination establishes the justification for advancing an assessment to a higher tier. As with other programs, tiered assessments are primarily designed to be exclusionary—that is, to prevent chemicals with low potential risk from advancing to higher tiers, which require more resources to complete.

Tiered assessment strategies typically utilize three or four tiers, the first being the screening level assessment. Screening level assessments require the least amount of data and can rely on predictive exposure and effects models, existing knowledge, or they may be based solely on limited empirical data. Screening helps to prioritize use of scarce resources by eliminating some chemicals from further consideration based on a lack of exposure potential (e.g., not detected in environmental samples) and/or lack of effects potential (e.g., determined to be practically non-toxic). While these appear, on the surface, to be simple arguments for exclusion, care should be exercised that analyses are appropriate. For example, highly hydrophobic compounds may appear to be at negligible concentrations in water, but they may also have a high bioaccumulative potential. Or, compounds which are practically non-toxic in an acute context may actually have important sub-lethal effects of concern.

Intermediate level assessments are typically used when screening level assessments cannot exclude the possibility that adverse effects could result from an exposure. Often, the intermediate level assessments are risk quotient-based, comparing limited exposure and effects (i.e., hazard) data to determine if the exposures could approach the magnitude necessary to elicit an adverse response.

The highest tier assessment, often considered “definitive” assessments, require the most exposure and effects data to complete. These assessments are typically limited to a specific context, which may be defined in terms of the chemical (or mixture), route of exposure, absorption, distribution, metabolism, elimination (ADME), mechanism of action, species, lifestage, and geospatial considerations.

It is common to include multiple, diverse endpoints or tests within some tiers, which are then viewed with equal weight (Combes 2007). This is important when endpoint sensitivity or response is selective towards a specific type of toxicant mode of action. The goal is to include a sufficient number of endpoints within a tier such that collectively, acceptable sensitivity is achieved for all relevant modes of action (Combes 2007). Figures 1 and 2 present conceptual views of tiering and the relationship between endpoints and biological scale.

3.1 Determining Potential for Effects

The major obstacle to assessing risk of CECs found or predicted to occur in Great Lakes surface waters is the lack of sufficient empirical hazard information. This situation is not unique

to the Great Lakes. Indeed, it is generally accepted that empirical approaches to hazard identification and dose response modeling on a chemical-by-chemical basis cannot address the backlog of existing chemicals nor keep up with the demands associated with the introduction of new chemicals. This is particularly true for CECs, since they often fall outside formal regulatory frameworks that require requisite data and are poorly defined in terms of exposure and effects. The intent of this section is to summarize the approaches that can be used to generate sufficient effects information, empirical and/or predicted, for screening level assessments.

This section will not address higher level assessments, such as the methods for establishing enforceable criteria for chemicals in surface waters in the US, as outlined in the 1985 ALWQC guidelines (USEPA 1985) and the 1995 Great Lakes Initiative (GLI) (USEPA 1995), both of which require significant effects data to permit criterion calculation. This section does, however, rely on a recent analysis of predictive tools and models conducted in support of an ongoing effort to work toward a common effects assessment approach that would provide for improved interpretation among the different critical values derived from the ALWQC and GLI, as well as benchmark values derived from acute and chronic risk assessments of pesticide active ingredients intended to support FIFRA pesticide assessments and label development (see EPA white paper, Predicting the Toxicities of Chemicals to Aquatic Animal Species, at www.regulations.gov, docket number EPA-HW-OW-2010-0818).

Effects of interest for ecological risk assessment purposes are those effects that impact or are expected to impact populations through reductions in individual fitness. In practice, these are generally considered to be effects on survival, growth, development, and reproduction. Although these effects are commonly measured at the organismal level, they are interpreted in the context of population and community level responses. (The notable exception is for species that are considered endangered, where individual responses are considered important.) Such effects can be assessed using both empirical and predictive approaches.

3.2 Empirical Toxicological Models

Empirical testing approaches that include apical endpoints are the most direct with the lowest uncertainty when determining effects of interest resulting from exposure to a specific chemical. There is a cadre of tests available in freshwater vertebrates and invertebrates designed to evaluate survival, growth, development, and reproduction (see Appendix D: Effects-Based Biomonitoring Tools for Ecological Risk Assessment). These tests are routinely used for a variety of assessment activities, including ambient water quality criterion, sediment assessment, and whole effluent testing. Although a variety of empirical toxicological assays are available and feasible, typically there are insufficient resources available to conduct the studies to obtain the necessary information for each assessment.

As an alternative to conducting empirical studies, values may be available in the scientific literature. These open literature studies utilize both standardized and non-standardized methods, often making it difficult to extract comparable data. Major compilations of existing toxicological information exist that are highly valuable. The ECOTOX database, for example, contains information on over 8,000 chemicals and nearly 6,000 aquatic and terrestrial species (<http://cfpub.epa.gov/ecotox/>). Data entered into this database undergoes critical review to ensure that the values reported meet a specified quality. And, it includes searching and reporting capabilities to allow users to select and display relevant data. Other sources are available as outlined in Appendix E. With regard to CECs, empirical effects information is often lacking for the species, effects, and chemicals or chemical mixtures of interest. Consequently,

with the exception of a relatively few well-studied chemicals, screening level assessments will likely need to use some form of predictive modeling.

3.3 Predictive Toxicological Tools and Models

Predictive toxicological tools and models useful to risk assessment can take many forms and can range in their complexity from conceptual diagrams to complex mathematical models, which estimate parameters of interest based on the conceptual construct. Underpinning all predictive models is a basic understanding of the biology and how chemicals perturb that biology. In theory, predictive models could be developed for any of the effects of interest. However, in practice, there are several barriers that prevent the development of a comprehensive set of predictive models. These barriers include: limited understanding of the basic biology, complex toxicological properties of chemicals, inadequate resources to develop the models, inherent complexities of systems involved, etc. Although the availability and applicability of predictive models to a specific risk question may be limited, the currently available models can still be extremely useful in screening level assessments based on acute and chronic lethality. For example, it is estimated that 70% of industrial organic chemicals exhibit non-specific acute toxicity, a mode of toxicity that can be readily estimated with relatively high confidence based on existing models (Bradbury et al. 2003). Furthermore, a relatively recent emphasis on the development of new predictive models and tools suggests that predictive capabilities will improve with time.

Several of the predictive approaches are summarized below. For more in depth descriptions of these approaches, including historical background and example applications, please refer to the EPA white paper, Predicting the Toxicities of Chemicals to Aquatic Animal Species (www.regulations.gov docket number EPA-HW-OW-2010-0818). Where applicable, references for specific tools and models point to the appendix, which contains access information, including online applications, similar to those in this document.

Predictive QSAR Models

Quantitative Structure Activity Relationship (QSAR) models relate physico-chemical properties of chemicals to quantitative estimates of activity. The available models are primarily useful only for lethality predictions. ASTER and ECOSAR models, for example, are widely used to estimate acute and chronic lethality of chemicals to fish and invertebrates. These predictive tools have primarily been designed to represent industrial organic chemicals and selected pesticide chemicals. Therefore, their application to other chemical classes of concern (e.g., EDCs, pharmaceuticals, personal care products, nanomaterials) may be inappropriate because the models do not consider the structure space nor the probable, relevant activities, particularly sub-lethal chronic activities, potentially associated with these chemicals. Therefore, currently available QSARs are not likely to contribute significantly to the assessment of CECs due to the lack of models for estimating sub-lethal toxicity.

Read-Across Using Chemical Categories

Read-across techniques using chemical categories are used to fill data gaps for specific chemicals. The prerequisite to using this approach is that there must be a definable chemical category that can be ordered according to some measure or characteristic related to the potency of each chemical. A chemical that lacks effects information is added to the series in accordance with the rank, based on the same measure or characteristic used to order the group. If the chemical of unknown activity lies between two chemicals of known activity, then an interpolative estimation is made. If the chemical of unknown activity lies outside the group, then an extrapolative estimation is made. In well-populated categories, QSAR may be derived if

trends in activity are related to the structural characteristic used to place and order chemicals within the same category.

Species Extrapolation Models

Extrapolation among species in an ecotoxicological context continues to be a major source of uncertainty. Unlike the situation in human health-based assessments, where extrapolation demands are bounded by the need to extrapolate from a few test species to only one species (i.e., humans), species extrapolation in ecotoxicology presents an entirely different challenge: extrapolating from a few test species to the entire spectrum of species across highly diverse taxa. Of course, in practice, the extrapolation space is necessarily constrained to species considered representative of various taxa.

Species sensitivity distribution and interspecies correlational models are mostly derived from lethality measurements. Since overt lethality is presumed to be relatively unimportant for most CECs, these types of models will not be addressed here.

Efforts are currently underway to develop extrapolation models for sub-lethal effects. Homology-based extrapolation techniques, for example, are rooted in phylogenetic differences and similarities, suggesting that closely related species are more likely to exhibit similar sensitivity toward a specific chemical than distantly related species. Homology-based extrapolations can be developed from different levels of information. Among highly divergent taxa, it may be possible to make estimations based on existing physiological information. For example, in a recent white paper on ALWQC for CECs, the lack of a functional estrogen receptor in invertebrates was used to explain the relative insensitivity of invertebrates to ethinylestradiol, an otherwise potent estrogen receptor agonist in vertebrates (Jukosky et al. 2008). However, as detailed analyses of post-genomic species emerge, it may be possible to use amino acid sequence data regarding the target protein to determine relative sensitivity. Already, a comparison of avian sensitivity to arylhydrocarbon receptor (AhR) agonists revealed that sensitivity to AhR agonists correlates with one or two amino acid sequence differences in the AhR ligand binding domain in an otherwise highly conserved, nearly identical sequence (Head et al. 2008).

One class of CECs that may be particularly amenable to this approach is that of pharmaceuticals. As a class, pharmaceuticals represent an activity sector within health and veterinary sciences, not a particular structural or functional class of compounds. However, since pharmaceuticals are intended to modulate physiological processes through targeted interaction with specific biomolecules, the mechanism of action of individual chemicals are relatively well defined in the target species (most often human). A comparison of biomolecular targets of some human pharmaceuticals across multiple species covering a broad taxonomic range demonstrate that the degree of basic homology is directly related to taxonomic distance and that some aquatic vertebrates have a high degree of amino acid sequence homology with the human drug target (Gunnarsson et al. 2008). The applicability of this approach is not limited to pharmaceutical chemicals, as it can aid in the screening level assessment of any chemical with a known target biomolecule.

In vitro and High Throughput Assays (HTA)

In recent years, *in vitro* methods representing specific biological functions and toxicological activities have proliferated. And, consistent with the NRC report (2007), their use for screening the toxicological activity of chemicals is rapidly increasing. This is due, in large part, to technological improvements that have enabled numerous assays to be used in a “high throughput” manner, meaning that numerous chemicals can be run through each assay in a

very short time. One of the largest and most comprehensive programs conducting high throughput testing is ToxCast (Dix et al. 2007). This program utilizes over 500 high throughput cell-based and cell-free assays, representing a broad array of toxicological mechanisms and targets, many of which could be applicable to CECs. Approximately 300 chemicals were assessed in the first phase of implementation, and an initial analysis of this complex dataset has been conducted (Judson et al. 2009). These data, which are primarily for pesticidal chemicals, are available online (http://www.epa.gov/ncct/toxcast/data_sets.html). Perhaps of more interest in the context of conducting screening level assessments of CECs, the second phase of the ToxCast program will screen a large number of chemicals representing a diversity of chemical classes and categories.

Whether derived from individually conducted studies or from high throughput programs, like ToxCast, *in vitro* results can be very useful as indicators of a specific activity. Such activity determinations are most useful when linked to an AOP, where the role of that activity can be interpreted in a toxicological context. However, many uncertainties remain in extrapolating from such assays, including: the role of metabolism in activating or inactivating chemicals, determining the dosimetry relationships between *in vitro* and *in vivo* systems, and the overall representativeness of individual and batteries of isolated *in vitro* system responses as compared to whole animal system responses. Clearly, in these relatively early stages of using *in vitro* assays to infer risk to the organism, considerable professional judgment will be required before such data can be used readily, even for screening level assessments.

Furthermore, the potential role of *in vitro* assays in assessing complex environmental samples (e.g., effluents, surface water, extracts from biological samples) is complicated. On the one hand, *in vitro* assays are typically developed to assess single chemicals within narrowly defined assay conditions (media composition, temperature, timing of assay, etc.), making their application to complex environmental samples complicated (as discussed further in Section 4.2). On the other hand, use of complex samples may be considered a more representative approach, taking into account the inevitable, multi-chemical characteristics of environmental samples. Although some *in vitro* assays have been successfully applied individually to environmental samples, significant technical challenges currently preclude application of high throughput methods to such samples.

Toxicity and Adverse Outcome Pathway Models

As mentioned in Section 2.0, the Adverse Outcome Pathway (AOP) is a conceptual framework that portrays existing knowledge concerning the linkage between a direct molecular initiating event of a toxic substance (i.e., exposure) to an adverse outcome in a species of interest at a level of biological organization relevant to risk assessment (Ankley et al. 2010). The AOP spans the biological hierarchy, including: sub-cellular, cellular, tissue, organ, organism, and population levels. Linkages are assumed to exist across these levels as molecular initiating events are transmitted upward through the biological hierarchy and expressed as “effects” at various levels. These effects could be the primary manifestation of toxicity or could reflect compensatory responses invoked through homeostatic mechanisms. Regardless, effects expressed at different levels represent opportunities to determine if a specific pathway has been affected by an exposure. Toxicity pathways, as defined by the NRC (2007), address cellular responses and thus, are nested components within an AOP.

One of the principal advantages of developing an AOP construct is that it serves to organize the problem in accordance with the risk issue and provides a framework for examining predictive relationships for different chemicals. The sources of information for the AOP may come from a variety of methods including those previously discussed (e.g., *in vitro* HTA and QSAR, Figure

2). For example, if one is confronted with monitoring information indicating that a pharmaceutical with androgenic properties is present, then one can take advantage of the existing therapeutic target information to establish possible targets for an initiating interaction, namely, the androgen receptor. Further, since androgens are integral to normal reproduction in vertebrates, then one can assume that the pathway of interest will include the cells, tissues, and organs associated with reproductive function. Conversely, if one is confronted with observations demonstrating estrogenic effects in organisms from the field, then one can work backward from the observation through a path associated with estrogen signaling.

Of course, a fully realized AOP may not always be obvious at the beginning of an assessment and will need to evolve as more information becomes available. Depending on the extent of the relevant biological and toxicological knowledge, it may not always be feasible to develop an AOP, and more empirical approaches may be needed. However, to the extent possible, the AOP construct should be used as a means to help develop a weight-of-evidence approach to the assessment. Furthermore, if an existing or novel AOP emerges as relevant to a specific situation, then it provides the connections to both defining effects-based measurements in the field and to diagnosing the cause of observed effects.

Conclusions

One characteristic that all CECs have in common is a high level of uncertainty. The uncertainty can be in either the potential for exposure or the potential for effects or both. Given this level of uncertainty, it can be assumed that sufficient data are lacking to establish criteria or benchmark concentrations. Therefore, the common, basic objective for CECs is to develop sufficient screening level assessment data to either advance the chemical to higher tiers of risk assessment (e.g., Figure 1) or to eliminate them from consideration as a likely risk.

Screening level decisions can be addressed from either the exposure side or the effects side of the problem. Conventionally, if exposure information supports the conclusion that exposure is unlikely, then the risk of a chemical can be dismissed without effects information. Conversely, if effects information supports the conclusion that a chemical has no likely effect on the target species, then the risk of the chemical can be dismissed without exposure information. However, most risk situations do not fall cleanly into either of those categories.

In general, the risk issues associated with CEC chemicals do not usually include lethality as a primary consideration. Their occurrence in the environment is typically at relatively low levels. And, their modes of action are highly variable, making it unlikely that predictive models capture the relevant mode of action (MOA). In specific instances where lethality is an issue, existing empirical approaches and predictive toxicology models may be appropriate applications to derive benchmark values.

Instead, evaluating risks of CECs must rely on weight of evidence approaches that use the most recent scientific knowledge and professional judgment. Many of the predictive tools currently available can inform the weight of evidence approach but are unlikely to be sufficient tools in themselves to provide clear evidence for risk decisions.

4.0 Effects Monitoring and Surveillance

The terms biological monitoring and biological surveillance are frequently used interchangeably, although the latter is sometimes used to refer to routine or baseline sampling performed with a randomized site selection process. For purposes of this document, we use the term biological monitoring for the systematic collection and analysis of samples (environmental media or

biological material) from predetermined sites, which may have been selected based on information obtained from other sources such as suspected chemical loadings, rare, vulnerable or vital habitats or anecdotal information on endangered species.

In the Great Lakes, there has been a multitude of monitoring programs associated with chemical occurrence. Some of these have emphasized long-term trends, while other programs have focused on shorter-term discovery-driven objectives. However, irrespective of the monitoring objectives, most efforts have relied solely on chemical measurements in different matrices (water, sediment, tissue). Some of the drawbacks to exclusive reliance of analytical monitoring for addressing current concerns in the Great Lakes was pointed out in Section 1.0 and include: (1) the possibility that “unknown” (and hence unmeasured) chemicals are responsible for adverse biological effects, (2) uncertainty that analytical detection limits are adequate with respect to biological effects, and (3) uncertainty of possible biological effects of chemical mixtures. Past regulatory/monitoring efforts have recognized these drawbacks and addressed them through the use of biological effects-based testing to complement chemical analyses. For example, almost 20 years ago, the National Pollution Discharge Elimination System (NPDES) permitting program administered under the auspices of the Clean Water Act started to incorporate toxicity limits into effluent permits. These limits were based initially on acute (lethal) and, later, on chronic effects of monitored effluents to aquatic invertebrate and fish species. Another notable permitting program that uses effect-based testing is the assessment of dredged materials, where the toxicity of complex mixtures (elutriates, sediments) to invertebrates and fish is used to complement chemical analyses for evaluating risk.

As monitoring programs in the Great Lakes move beyond a focus on regulated legacy pollutants to CECs, as well as more explicit consideration of chemical mixtures, it is essential that biological monitoring approaches be incorporated into the process. To effectively detect adverse effects associated with CECs, a biological monitoring strategy needs to incorporate both prospective and retrospective approaches. This will involve a combination of field monitoring and deployable bioassay test systems along with laboratory-based testing.

4.1 Biological Monitoring: Generic and Specific Endpoints

There are a variety of test/sample types that could be useful for environmental monitoring. These include lab-based *in vitro* or *in vivo* assays using complex mixtures from the field such as discharge or receiving waters, sediments/sediment fractions and even tissue extracts. Samples also could be from organisms held *in situ* (e.g., caged fish studies), or animals collected from extant populations. From these types of systems many different kinds of endpoints could be considered, depending upon monitoring objectives. Endpoints examined might be “generic” in that they could reflect multiple chemical stressors, or they could be “specific” in terms of capturing defined biological pathways of concern. An example of a generic biological endpoint would be acute lethality as measured, for example, via biological tests like those used for NPDES and dredged material assessment programs. An example of a more specific endpoint would be vitellogenin (VTG) expression in male fish, which is an indication of exposure of the animals to one or more estrogenic chemicals. There are a range of endpoint/assays along the continuum of generic to specific, and a number of considerations factor into which would be most appropriate for a monitoring program.

A very basic consideration relative to endpoint (and assay) selection involves the degree to which the monitoring program is hypothesis-based relative to stressor identification. For example, if there is considerable uncertainty as to the nature of the chemical mixture to which organisms are exposed, generic approaches focused on apical endpoints that integrate multiple

toxicity mechanisms/pathways are appropriate. Conversely, if there are biological pathways of known concern, biological monitoring tools can be tailored to reflect perturbation of these pathways. Needs of monitoring programs in the Great Lakes cannot be met solely with either generic or specific assays/endpoints. A robust, scientifically-defensible monitoring effort in such a large, biologically, chemically, and physically complex system almost by definition has to consider both types of tools. However, for the remainder of this document, we focus mostly on endpoints that would be regarded as more pathway/mechanism specific rather than apical. These types of endpoints largely are comprised of molecular, biochemical, histological and morphological responses. This emphasis was chosen, not to detract from the potential value of more apical responses for effects-based monitoring, but to highlight the utility of mechanistic approaches to guide prospective and retrospective (or diagnostic) assessments of the ecological risk of CECs.

4.2 Considerations in Endpoint/Assay Selection

There are a number of criteria to consider in the selection of effects-based responses for monitoring toxic chemicals in the environment. As discussed above, one is specificity of the response relative to the pathway(s) of concern. In this context, choosing those pathways of most concern presents a challenge that can be more complicated than might seem immediately apparent. For example, in virtually all programs focused on CECs, there is a desire to detect/monitor “endocrine-disrupting chemicals” (EDCs). One class of EDCs that has received substantial attention from a societal perspective is estrogens, but compelling scientific arguments—both in terms of occurrence and effects—can be made for monitoring other classes of EDCs, including androgens, anti-androgens, and substances that affect production of steroids. As one further expands the universe of biological pathways to consideration of those known, for example, to be impacted by CECs such as human pharmaceuticals/veterinary drugs, the possibilities in terms of the number of assays/endpoints to employ rapidly becomes daunting. To make the challenge of pathway (and hence assay) selection tractable, prospective/predictive tools are critical to prioritize pathways of greatest concern from potential exposure and effects perspectives.

As part of prioritizing pathways for effects-based monitoring, it is critical to understand the linkage of these pathways to adverse outcomes meaningful to ecological risk assessment (typically survival, growth/development, and reproduction). Molecular, biochemical, histological and morphological changes in organisms are ultimately useful to risk assessment only if they reflect in some manner health at the individual-, population-, and community-levels. As discussed previously, the AOP framework provides a basis for explicitly making the linkage between mechanism-specific responses and impacts of ecological concern (Ankley et al. 2010). Consideration of an AOP framework in the context of selection of effects-based monitoring tools will help ensure that monitoring results are applicable to risk assessment.

In this context it is important that endpoints/assays used for monitoring be interpretable relative to impacts on the biological system(s) of interest. For example, results from *in vitro* assays that detect specific chemical activities (e.g., agonists of the estrogen, androgen or Ah receptors) need to be calibrated against these same pathways in whole animals. Analogously, changes in expression of genes, proteins, and metabolites in whole animals from field or lab exposures need to be relatable to responses at higher biological levels of organization. For example, it is important to be able to put toxicant-induced changes in VTG gene or protein expression in fish in the context of possible alterations in reproductive capacity (Ankley et al. 2010). An additional challenge relative to interpretation of effects-based endpoints, especially in field-collected animals, involves identification of what is normal (or baseline) relative to the life history of the

organism being monitored. A relevant example of the type of challenge faced here involves interpretation of “intersex” in fish (simultaneous occurrence of gonad tissue reflective of both sexes in one animal). Numerous controlled lab studies with different classes of EDCs have produced intersex in fish. However, there is substantial cross-species variability in susceptibility to intersex, and there is evidence that non-chemical factors may contribute to occurrence of the condition. Further, it is unknown the degree to which intersex might interfere with reproduction to the extent that population-level impacts might occur in fish. Subsequently, although intersex in fish could be a potentially very useful effects-monitoring endpoint (and there have been many examples of it being used as such), uncertainties exist as to its overall utility.

Other factors that need to be considered relative to selection of effect-based tools involve more pragmatic concerns relative to actual measurement of endpoints of interest. First, measurement methods ideally should be relatively straightforward in the context of being able to move from a research to contract lab setting. Since most mechanistic biological endpoints have not been used for routine monitoring, the degree of validation is far less, compared, for example, to that of chemical analyses. Validation involves several components, including method optimization and standardization, and characterization of intra- and inter-laboratory precision, accuracy, and variability.

Another consideration that often is overlooked relative to the application of some types of mechanistic tools to monitoring involves compatibility of the assay systems with the types of samples collected from the field. This can be particularly important relative to some of the *in vitro* assays that could be useful for monitoring. The two primary technical obstacles to applying *in vitro* assays to environmental samples are maintaining media composition and preventing bacterial contamination. Stringent media composition requirements are often necessary for *in vitro* assays to perform properly. Direct addition of environmental samples or preparations of environmental samples can alter ionic composition and pH, for example. And, due to bacterial contamination, it often is not feasible to directly use effluent or surface water samples in cellular activity-based assays. There are ways to deal with this issue (e.g., filtration, sample extraction), but these techniques need to be characterized—often on a study-specific basis—in terms of how the manipulation both affects assay performance and (potentially) alters inherent properties of the test sample (e.g., changing contaminant bioavailability).

A final consideration relative to selection of effects-based monitoring tools involves how data from the work would/could be used for retrospective/diagnostic risk assessments. That is, once a problem has been identified via monitoring, it is necessary to ascertain the chemical(s) causing the response from what typically will be a very complex mixture of contaminants. Identifying chemicals responsible for observed effects can be critical to source identification and effective mitigation. Endpoints/assays which are relatively specific for a particular biological pathway (e.g., estrogens) could be very useful for quickly narrowing the universe of chemicals that would be considered as causative agents and may, in and of themselves, serve as effective tools for source tracking and/or biologically-based fractionation work.

4.3 Examples of Tools: Present and Future

There are a large number of potential assays and mechanistic endpoints that could be useful tools for effects-based monitoring of CECs in the Great Lakes. Some endpoints and assays take advantage of recent advances in molecular toxicological techniques and improved understanding of animal physiology while others have a longer history of use and are therefore more established. The latter type of tool would be more immediately useful while the more recently introduced techniques, although promising, require additional development and

validation. In this context, the endpoints/assays discussed below are intended to be illustrative rather than comprehensive. It is worth noting that in recent years, several different governmental agencies have implemented biological monitoring programs for various purposes. One of the longest, continuously active programs is the Canadian Environmental Effects Monitoring (EEM) program, which was originally developed for the pulp and paper sector in 1990 and has been expanded to the metal mining industry and municipal wastewater treatment plants. The EEM program is designed to assess environmental impacts of effluents discharged from pulp and paper and metal mining activities (Munkittrick et al. 2010). The focus of the EEM program is on fish populations and benthic invertebrate communities and emphasizes endpoints reflective of these biological levels of organization. Fish health at a given site is monitored by measuring the body weight and age, gonadal and liver somatic indices (GSI, LSI), condition index (CI), fecundity, and egg size. Invertebrate communities are monitored using four different endpoints: abundance, number of taxa (family level or higher), uniformity of taxa distribution (Simpson's evenness index), and community structure (Bray-Curtis index of similarity; Lowell et al. 2005).

Two other biological monitoring programs are also worth mentioning, the USGS Biomonitoring of Environmental Status and Trends (BEST) and NOAA fisheries Toxics-Focused Biological Observing System (TBiOS) programs. The BEST program operated from 1995 to 2004 with a focus on fish health in several major US river systems. A goal of this program was to incorporate endpoints from multiple levels of biological organization including cellular and sub-cellular endpoints such as hepatic EROD activity, circulating sex steroid levels, and vitellogenin (VTG) levels (Hinck et al. 2008). The TBiOS program is an ambitious, newly proposed framework for monitoring the effects of contaminants on marine organisms in Puget Sound, WA. A goal of the TBiOS program is to measure endpoints that can be used with population and community level mathematical models to predict future trends and status of populations in Puget Sound (e.g., prospective analysis). Included endpoints for the program's fish monitoring are the types and total amount of lipid, somatic indices (GSI, LSI, CI), plasma VTG, histopathology, stomach contents, and various biochemical indicators including sex steroids, thyroid hormones, and insulin-like growth factor. A summary comparison of the endpoints used by these monitoring programs is shown in Table 1. Additional description of these and other endpoints is provided below. See also Figure 2, which lists many of these endpoints in association with their level of biological scale.

Near-Term

There are several mechanism/pathway-based tools that are being used for different types of monitoring programs, and could be easily adapted to Great Lakes applications in the near-term (summarized in Table 2). Both the EEM and BEST programs emphasized fish health assessments as the focus of biomonitoring efforts. These programs primarily incorporated endpoints at the individual – population level such as GSI, LSI, fecundity, and CI. These endpoints were felt to be most important due to their ecological relevance (Munkittrick et al. 2010). However, potential problems with reliance on these endpoints have been noted such as difficulties in establishing normal variation and confounding impacts of other stressors such as temperature (Hinck et al. 2008). These concerns can be partially mitigated by careful study design and species selection focusing, for example, on sexually mature individuals with a sedentary (nonmigratory) life style (Munkittrick et al. 2010).

The value of incorporating more diagnostic endpoints has also been noted in monitoring programs, particularly when such endpoints can be closely linked to toxicant class or mode of action (Hinck et al. 2008). In this regard, there are a number of potential biochemical and physiological endpoints that could be incorporated into a Great Lakes monitoring program for

both prospective and retrospective assessments. A variety of *in vitro* assays have been developed and used in environmental monitoring to detect chemicals in complex mixtures with the potential to exert toxicity through activation of the Ah, estrogen, and androgen receptors. For example, an assay with the H4IIE cell line to detect Ah receptor agonist, which includes a variety of legacy chemicals (PCBs, TCDD), was developed some years ago (Tillitt et al. 1991), and has been used in a number of environmental applications, including assessments in the Great Lakes (Whyte et al. 2004). There also are a number of *in vitro* assay systems to detect estrogens in environmental samples. A comparative analysis of five different *in vitro* assays relative to detecting estrogen mixtures in effluent, surface and ground water, and drinking water recently was presented by Leusch (2008). *In vitro* assays to detect androgenic activity also have been developed and used for environmental monitoring (e.g., Parks et al. 2001; Wilson et al. 2002), and subsequent studies have been conducted to help validate the methods for this application (e.g., Blake et al. 2010).

A number of biochemical measures in animals either exposed to environmental samples or collected from the field also are at a point where they could be readily used for an effects-based monitoring program in the Great Lakes. For example, hepatic ethoxyresorufin-O-deethylase (EROD) activity has been used for many years to detect the occurrence of chemicals that interact with the Ah receptor (Whyte et al. 2000). Other types of useful biochemical measures are those associated with the chemical defense (Goldstone et al. 2006) and would include proteins and small molecules that protect against oxidative stress. Example endpoints are glutathione status (GSH/GSSG; ratio of reduced vs. oxidized glutathione) and the activities of several enzymes such as superoxide dismutase, catalase, glutathione S-transferases (GSTs), glutathione peroxidase and glutathione reductase (GR). Other commonly used markers of oxidative stress reflect oxidative changes to macromolecules such as lipids. One of the more widely applied assays is TBARS (2-thiobarbituric acid reactive substance). For example, Oakes et al. (2005) demonstrated elevated hepatic TBARS in rainbow trout exposed to kraft pulp-mill effluents. Induction of VTG (either mRNA or protein) in male fish has been widely used to detect estrogenic chemicals in the environment (Wheeler et al. 2005), and recent studies have shown that depression of VTG in females can be equally useful to detect other classes of EDCs, such as chemicals that antagonize the estrogen receptor or inhibit sex steroid production (Miller et al. 2007). Effects-based monitoring programs, for example by Environment Canada, have for many years measured sex steroid (estradiol, testosterone) concentrations in fish as an indication of endocrine-active chemicals (e.g., McMaster et al. 2006).

Another important type of endpoint extensively used in both the US and elsewhere for environmental monitoring is gonad histology in fish from extant populations. In fact, the observation of intersex gonads in fish in the UK (Purdom et al. 1994) was a primary “driver” in terms of the current emphasis on estrogens in environmental monitoring programs throughout the world. With the caveats discussed above (e.g., need to understand normal biology relative to interpretation), gonad histology is another endpoint that could be viewed as having applicability in the near-term.

Longer-Term

There are also a number of endpoints/assays that, with some additional research and subsequent validation, are promising for application to effects-based monitoring in the future (e.g., 3 to 5 years; see Table 3). For example, the technology now exists to easily develop cell-based, receptor-reporter gene activity assays, similar to those used to detect estrogens and androgens, for a variety of different biological pathways. Key here is prioritization of those pathways of most concern in order to focus development of specific pathway-based assay systems.

There are ongoing research efforts by several groups with a variety of ecologically-relevant species to develop the knowledge base necessary to build “tool boxes” of gene expression assays that would enable detection of classes of chemicals with different mechanisms of action in a complex mixture (e.g., Ankley et al. 2009). The goal here would be to examine multiple gene expression (via polymerase chain reaction technology) in organisms either exposed to environmental samples in the lab or *in situ* to ascertain perturbations in specific toxicity pathways of concern. Analogous approaches could focus on changes in expression of multiple proteins reflective of different pathways, although the measurement technology to achieve this is somewhat less well developed/defined than for gene products.

Substantial attention has been given to the role of genomics in ecological risk assessment, including effects-based monitoring (Ankley et al. 2006). Global measurements of changes in expression of genes (transcriptomics), proteins (proteomics), and/or metabolites (metabolomics) hold much promise both for identifying abnormal/perturbed biology and, through “fingerprinting” (pattern recognition) approaches, diagnosing causes of the perturbation. Genomic approaches that have received the most attention to date in terms of possible applicability to environmental monitoring are transcriptomics (via microarray technology), proteomics and metabolomics (Van Aggelen et al. 2010; Lemos et al. 2010). Initial studies have been promising, for example, in differentiating populations of stressed versus unstressed animals. Unfortunately, due to the complex nature of the datasets generated, and—in most cases—a lack of baseline understanding of the biological processes underlying observed changes, diagnostic utility of genomic tools relative to complex mixtures of stressors is limited. However, progress is being made in this area. For example, Ekman et al. (2008) recently described the use of NMR-based metabolomics to characterize direct effects, compensation and recovery in fish exposed to the pharmaceutical ethinylestradiol. This type of information can be used not only for identifying causative stressors, but also for developing exposure reconstruction profiles.

4.4 Design Strategies for Effects-Based Monitoring

The implementation of an effects-based monitoring program in the Great Lakes will require effective cataloguing and integration of what has been done in the past and what currently is being done. This is critical to not only integrating data effectively, but also promoting efficiency relative to collecting and sharing different types of samples. Monitoring for effects of chemicals should be incorporated into routine monitoring programs, including the cyclical binational monitoring programs currently being implemented on each of the Great Lakes on a 5-year rotation schedule. Demonstration of effects-based monitoring pilot studies could (and probably should) be conducted on a couple of different scales. For most of the endpoints discussed above, there is little information as to baseline/background and any regional or subregional pattern over a broad spatial scale (e.g., Where, how frequently, and to what degree is estrogenicity detected in surface water using *in vitro* bioassays? How often are water samples toxic? Does intersex [in a given fish species] appear to bear any relationship to known point-source discharges?, etc.). A probabilistic monitoring effort to answer these types of baseline conditions would ideally utilize a smaller set of well-characterized effects-based endpoints at a relatively large number of sites stratified, for example, by watershed-oriented land use patterns (including occurrence of known chemical inputs).

In addition to broader-scale effects-based monitoring, there is a need for more focused, hypothesis-driven studies that could be conducted, for example, at specific sites of concern closely associated with a pollution source such as wastewater treatment effluents. These more focused studies ideally would include a more robust temporal component than would be

possible in broad spatial sampling. In addition, it would be feasible to employ a larger number of effects-based endpoints/assays, in part to develop the critical knowledge base needed to move a promising tool from a research setting to field application. Designing these studies around known/suspected contaminant sources/types also would provide important information about the utility of effects-based endpoints to aid in source identification. For example, studies in the UK have shown that measurement of VTG in caged fish from relatively small, effluent-dominated rivers can be related spatially to wastewater treatment plant discharges. However, comparable studies have not been done in systems more characteristic of Great Lakes rivers/harbors—large water-bodies that may not be effluent-dominated. Information derived from studies of this type is critical to the validation and broader use of mechanism-based effects endpoints for monitoring, particularly in the context of helping define cause and effect relationships.

5.0 Diagnostic Assessment Strategy

Invoking a diagnostic assessment is based on the assumption that some type of biological effect has been observed, suggesting the possibility that a chemical exposure was inducing a biological response. This is also the basis for performing a retrospective risk assessment, which causally (diagnostically) investigates the impacts of chemical contamination that has already occurred or is ongoing at a specific site. It is recognized that the range of measures that might constitute a “biological response” is wide and varied. It could include measurements at any level of biological organization, from biochemical to population. It could also be some change measured in organisms living in the environment, such as reduced recruitment of lake trout or vitellogenin induction in male fish, or an effect measured in an environmental sample collected from the field, such as toxicity in whole sediment toxicity test. While the nature of response may be varied, there is a common possibility that responses are induced by exposure to one or more chemicals present in the Great Lakes.

The objective of this section is to outline a framework for investigating the cause of observed effects with the goal of identifying the specific chemical or chemicals responsible for the observed response. With sufficiently advanced analytical technology, a great many chemicals may be detected in the environment. However, it is likely that most are present at concentrations below those that will cause adverse effects. Because resources to address environmental problems are finite, it is very important that we be able to connect observed responses with the chemicals that cause them, to ensure that the most resource-effective approaches are taken to mitigating or eliminating the problem. As a simple example, both ammonia and DDT are frequently detected in aquatic sediments, and both are capable of causing toxicity to benthic invertebrates. However, the appropriate management actions to address these problems would be dramatically different; problems with ammonia might be addressed through controls on nutrient release (e.g., reducing inputs of agricultural runoff or nutrient-rich wastewater) while DDT contamination of sediment might be managed through dredging or capping. Therefore, making a causal connection between observed effects and the chemicals that cause them is clearly essential to effective management of the Great Lakes.

5.1 Diagnostic Assessment Flowchart

An approach to assigning causes to observed effects is outlined in flowchart form in Figure 3. As a prelude to actual diagnostic assessment, it is worth considering whether evaluation should proceed directly to determining cause (box 1), or if it is more appropriate to more thoroughly understand the observed effect and its potential environmental consequences—in other words, whether the observed response is sufficiently serious that it intrinsically demands investigation, or if the connection between the observed effects and serious environmental consequences has

some uncertainty. As an example, dramatic changes in field populations of fish can be viewed as “intrinsically adverse,” meaning that the seriousness of the effect, and the importance of determining its cause, is of little doubt. However, the environmental significance of some observed biological effects might be less clear. For example, certain changes in gene expression might be demonstrated to be different among fish in different locations, but the certainty that this change is indicative of significant environmental effects is limited. In such cases, it may be worth initially emphasizing additional studies with an emphasis on better understanding the potential consequences of the observed effects, to determine whether diagnostic evaluation is warranted (box 2).

Assuming the observed effect is sufficient to warrant diagnostic analysis, the next decision is whether it is likely that toxic chemicals are the cause of toxicity (box 3). This distinction may not be of great importance to the overall management of the Great Lakes, as the undesirability of the effect is probably not influenced by the nature of its cause. However, because the focus of this effort is on toxic chemicals, and the methods used to evaluate such chemicals as a cause of effects may differ from those used to evaluate other factors (e.g., loss of spawning habitat, nutrient enrichment), it is important to note the potential divergence in methods, and the potential importance of factors other than toxic chemicals in producing adverse effects. In many cases, both toxic chemicals and other factors may be potential causes, and investigations will need to proceed along both lines (boxes 4 and 5).

Box 6 involves determining whether there is a known exposure that is plausibly connected to the observed effect. This apparent connection may be related to the documented presence of a chemical exposure known to produce the types of effects observed, or perhaps a spatial relationship between the observed effect and known sources of contamination. The importance of considering suspected causes is that it is often true that a more focused question, such as “chemical x the cause of effect y,” can be more easily resolved than a broader investigation. However, investigation of an initially suspected cause does carry the risk that alternatives will not be fully considered and effectively ruled out—sometimes, it can be an impediment to suspect an answer before the proof is developed. The development of evidence to support or refute the suspected cause must be structured to ensure that it is sufficiently rigorous, and be subject in the end to the same critical evaluation (boxes 6 and 15).

Another consideration designed to narrow the scope of investigation is determining whether the nature of the effect suggests a particular “adverse outcome pathway” (AOP) (box 7). Similar to the previous question, the emphasis here is on determining whether the pathway through which the effect occurs can effectively narrow the range of possible causes, and thereby narrow the scope of the investigation. A clear example here would be the induction of estrogenic effects, whose very nature suggest particular pathways of toxicity and effects. Also similar to the previous question, care must be taken that suspicions regarding particular causes do not inadvertently blind the investigator to alternative explanations, and in the end, the investigation must pass the same critical evaluation (boxes 8 and 15).

To this point, the evaluation has focused on determining whether existing information can be used to steer the investigation toward investigation of a candidate cause, or at least a subset of possible causes. Lacking that, it is likely that new data will be required to help determine candidate causes. In box 9, one considers whether the nature of the observed effect lends itself to experimentation, such as biologically-directed fractionation or toxicity identification evaluation (TIE). These methods generally rely on effects that can be demonstrated in an experimental system, such as a laboratory toxicity test conducted on an environmental sample or preparation thereof. Responses such as a fish population decline or a depauperate benthic community do

not lend themselves to this type of investigation. In such cases, it may be necessary to develop a surrogate response that is plausibly linked to the observed effect, but involves an experimental system that lends itself better to TIE-type approaches (box 12). As an example, a whole sediment toxicity test (which is amenable to TIE) might be a logical alternative to explore in a case where the observed effect is a depauperate benthic community.

If the observed effect is amenable to experimentation, or if a suitable surrogate is identified, then a variety of TIE-type tools can be applied to help define a physical/chemical profile of the likely causative chemicals, and from there an identification of those chemicals (boxes 11 and 13). Published TIE methods are available for both water and sediment (including interstitial water) samples (USEPA 2007). However, it is important to remember that the basic approach of biologically-directed fractionation is a much more generalizable approach, and has been applied successfully with many more matrices and bioassays than just those described by USEPA for water and sediment (Custer and Burton 2008; Picone et al. 2009; Lebo et al. 2000). If toxicant identification studies are pursued using a surrogate bioassay, it is very important that confirmatory studies are able to not only identify the cause of response in the surrogate bioassay, but also show that the candidate cause is also the likely cause for the original effect being evaluated (box 13). Differences could arise not only if the original and surrogate responses have different sensitivities to chemicals present, but also if the translation of effect to the surrogate bioassay creates an artifact that confounds interpretation. For example, toxicity of ammonia to the amphipod *Hyalomma* can be shown in sediment interstitial water that is not expressed in bulk sediment toxicity tests, because the organisms in bulk sediment tests can behaviorally avoid exposure (Whiteman et al. 1996). This is not intended to discourage the use of surrogate bioassays, just to emphasize the importance of confirming that the putative cause of effects in surrogate bioassays must be confirmed with respect to the effect originally observed.

Although TIE-type approaches have been shown to be very effective diagnostic tools in many instances, the emergence of new tools has opened possibilities for other types of diagnostic assessment (box 10). In particular, a variety of genomic (and other “omic”) tools, as well as various biochemical or *in vitro* assays might be used effectively to indicate the involvement of particular adverse outcome pathways (toxic mechanisms) in producing a particular effect. For example, evaluation of gene expression patterns in exposed organisms might be able to suggest physiological pathways that are being affected by exposure. This information may be able to suggest a specific AOP involved in producing the observed effect, and allow the investigation to be narrowed accordingly (see box 8 and accompanying text).

Where none of the above approaches are effective or even tenable, the investigation is left to more general inferential approaches (box 14) rather than more experimental approaches such as TIE (though the two need not be mutually exclusive). These inferences may involve approaches such as establishing temporal or spatial correlations between exposure and effect, and may parallel epidemiological approaches often applied in human medicine.

Regardless of the path by which a candidate cause of an effect is identified, it is very important that the evidence supporting the candidate cause is assembled and critically analyzed en masse (box 15). Both Hill (1965) and Fox (1991), among others, provide a framework for such an analysis. Not only do these approaches help organize evidence into a logical evaluation, but they may also suggest other types of supporting evidence that could be collected to help solidify the case for a particular cause.

6.0 Recommendations

The overarching recommendation is the need to incorporate effects-based monitoring to complement chemical-based approaches. In developing a strategy for effects-based monitoring of toxic substances, several broad areas of need were identified and associated with knowledge gaps in the toxicity of CECs: the need to incorporate more diagnostic tools in monitoring efforts and to select sites to focus initial monitoring efforts. It is also recognized that there is a need to maximize limited resources by leveraging existing Great Lakes programs. Based on the discussion in the preceding sections, several specific recommendations, listed below, can be made to further the implementation and use of effects-based studies.

- Conduct a comprehensive literature review of tools currently available/suitable for effects-based monitoring, and evaluate these tools as to their suitability, using well-defined criteria related to biological significance/interpretation, availability of standardized/validated methods, etc. (See Appendix D.)
- To avoid duplication of monitoring efforts in terms of collections and measurements, it is important to catalogue and then coordinate with existing monitoring efforts conducted by different federal and state agencies (this has been started, see <http://www.glc.org/monitoring/greatlakes/>).
- Evaluate historic and current chemical monitoring efforts in terms of target analytes, and consider dropping chemicals consistently not detected and/or with no biological significance. Another consideration for inclusion is whether analytical detection limits are sufficient to detect a chemical in the low or sub part-per-trillion level.
- To facilitate diagnostic assessments, prioritize pathways of greatest concern to support assay/endpoint selection and test development. Closely related to this recommendation is the need to formally list toxicant modes of action of highest concern.
- The biological effects monitoring should include a mix of both physiological/individual based endpoints (Tiers 3-4; e.g., apical endpoints) and, especially for CECs, pathway-based mechanistic assays.
- It will be important to select assays/endpoints that are directly relatable to adverse impacts on processes of concern (an aspect of this involves defining baseline/normal conditions).
- Incorporate—where appropriate—tools that are currently ready for monitoring applications, but also support the continued development and validation of promising newer tools.
- Conduct two types of effects-based monitoring pilots to: (1) emphasize broad spatial coverage to help characterize baseline conditions, and (2) more intensively characterize impacted sites to explore utility of the methods as diagnostic tools in a field setting.

Table 1. Comparison of Major Endpoints of Three Monitoring Programs

Endpoint	EEM	BEST	TBios
Age	X	X	X
Growth	X	X	X
Condition	X	X	X
GSI	X	X	X
LSI	X	X	X
SSI		X	
Fecundity	X		
Lipid			X
VTG		X	X
Hormones		X	X
Histopath		X	X
Macrophage		X	
Morphology		X	
Diet: direct			X
Diet: indirect	X		

Table 2. Examples of tests and endpoints suitable for near-term use in biomonitoring

Test Method	Target Organism	Test Endpoint	Relevance / Importance
<i>Cellular Responses</i>			
Vitellogenin	Egg-laying vertebrates / some invertebrates	Males- induction of protein. Females – decreased production	Males – Biomarker for exposure to estrogenic substances. Females – Decreased fecundity
EROD	Vertebrates / some invertebrates	Induction of CYP 1A activity	Ah receptor mediated; Dioxin-like MOA
Oxidative stress	Vertebrates / invertebrates	TBARS, SOD, GSSG	DNA & macromolecular damage, ↑chronic disease susceptibility
Gill ATPase	Fish /some invertebrates	Na ⁺ ,K ⁺ -ATPase	Osmoregulation, membrane integrity
<i>Organ Responses</i>			
Tissue-somatic indices	Vertebrates	GSI, LSI	Enlarged or decreased relative mass indicative of altered physiology
Circulating sex steroids	Vertebrates	↑↓E2, T, 11-KT levels in plasma	Indicative of altered steroidogenesis & excretion
Tissue histopathology	Vertebrates / invertebrates	Microscopic examination of fixed tissue	Can detect necrosis & other evidence of tissue damage
<i>Whole Organism Responses</i>			
Fecundity	Vertebrates / invertebrates	Egg production	Reflective of reproductive performance
External morphology	Vertebrates / invertebrates	Unnatural sexual dimorphism / tumors	Indicative of endocrine disruption / exposure to genotoxic agents
Condition Index	Fish	Length-weight relationship	Indicator of general health and normal maturation

CYP = cytochrom-P-450; TBARS – ThioBarbituric Acid Reactive Substances; SOD = superoxide dismutase; GSSG = oxidized glutathione; GSI / LSI = gonadal / liver somatic index;

Table 3. Examples of tests and endpoints requiring additional testing before use in biomonitoring

Test Method	Target Organism	Test Endpoint	Relevance / Importance
<i>Sub-cellular / Cellular Responses</i>			
Genomics (transcriptomics)	vertebrates / invertebrates	Changes in gene expression	May be associated with molecular initiating event.
Proteomics	Vertebrates / invertebrates	Changes in protein expression	May be associated with molecular initiating event.
Metabolomics	Vertebrates / invertebrates	Changes in metabolite abundance	May be associated with molecular initiating event.
<i>Organ Responses</i>			
Proteomics	Vertebrates / invertebrates	In biological fluids (plasma, urine, saliva)	Indicative of departure from normal homeostasis; disease; acute / chronic toxicity
Metabolomics	Vertebrates / invertebrates	In biological fluids (plasma, urine, saliva)	Indicative of departure from normal homeostasis; disease; acute / chronic toxicity
<i>Whole Organism Responses</i>			
Behavior	Vertebrates / invertebrates	Changes in feeding or predator avoidance behavior	Directly relevant to growth and survival
Field deployable non-invasive imaging	Vertebrates / invertebrates	Ultrasound / MRI analysis of internal tissue morphology & body fat	Relevant to growth, survival and reproduction

Tier 1 Screening / Prioritization

- Existing knowledge and empirical data
 - QSAR
 - Read across
 - Exposure modeling
- 

Tier 2 Intermediate

- Chemicals which cannot be excluded from Tier 1 tests
 - May be risk quotient-based
 - May include in vitro testing or short-term in vivo testing
- 

Tier 3 Definitive

- May be specific to pollutant source, chemical, exposure route and organism.
 - May include a suite of diagnostic endpoints / tests
 - May include multi-generational testing
- 

Figure 1 Tiering in ecotoxicological testing.

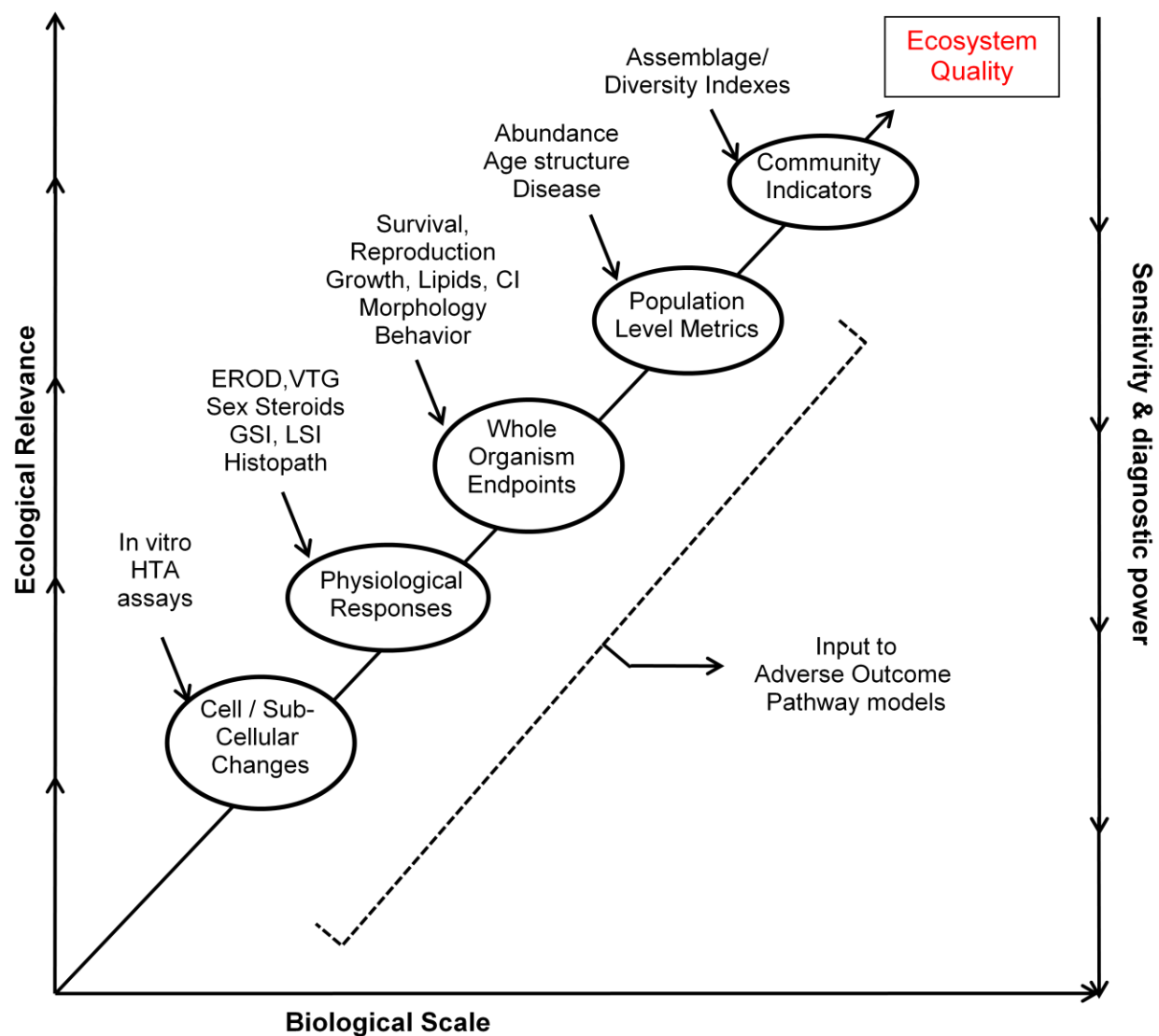
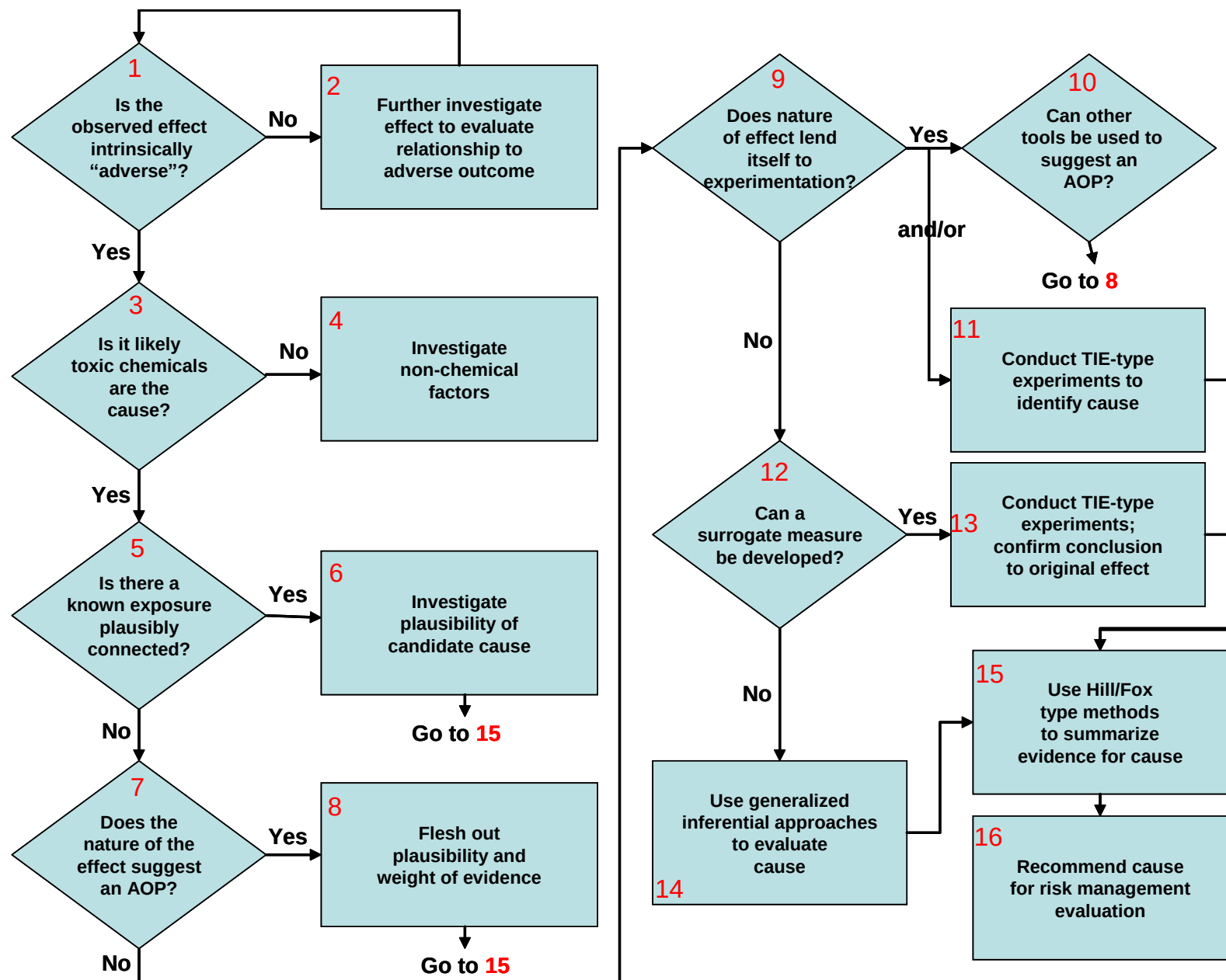


Figure 2. Conceptual description of monitoring tools, biological scale and connection to adverse outcome pathway modeling.

Figure 3. Diagnostic Assessment Flowchart



References

- Ankley, G.T., G.P. Daston, S.J. Degitz, N.D. Denslow, R.A. Hoke, S.W. Kennedy, A.L. Miracle, E.J. Perkins, J. Snape, D.E. Tillitt, C.R. Tyler and D. Versteeg. 2006. Toxicogenomics in regulatory ecotoxicology. *Environ. Sci. Technol.* 40, 4055-4065.
- Ankley, G.T., D. Bencic, M. Breen, T.W. Collette, R. Connolly, N.D. Denslow, S. Edwards, D.R. Ekman, K.M. Jensen, J. Lazorchak, D. Martinovic, D.H. Miller, E.J. Perkins, E.F. Orlando, N. Garcia-Reyero, D.L. Villeneuve, R.-L.Wang, and K. Watanabe. 2009. Endocrine disrupting chemicals in fish: Developing exposure indicators and predictive models of effects based on mechanisms of action. *Aquat. Toxicol.* 92, 168-178.
- Ankley, G.T., R.S. Bennett, R.J. Erickson, D.J. Hoff, M.W. Hornung, R.D. Johnson, D.R. Mount, J.W. Nichols, C.L. Russom, P.K. Schmieder, J. A. Serrano, J. E. Tietge and D.L. Villeneuve. 2010. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.* 29(3): 730-741.
- Balls, M. and J.H. Fentem. 1992. The use of basal cytotoxicity and target organ toxicity tests in hazard identification and risk assessment. *Alternatives to Laboratory Animals.* 20(3): 368-388.
- Blake, L.S., D. Martinovic, L.E. Gray, V.S. Wilson, R.R. Regal, D.L. Villeneuve and G.T. Ankley. 2010. Characterization of the androgen-responsive MDA-kb2 cell line for assessing complex environmental mixtures. *Environ. Toxicol. Chem.* In Press.
- Bloom, R.A. 2008. Pharmaceuticals in Water Regulatory Landscape: FDA Perspective. Research Triangle Environmental Health Collaborative 2008 Summit: Pharmaceuticals in Water. Available at <http://environmentalhealthcollaborative.org/images/RonBloomPresentation.pdf>.
- Bradbury, S.P., C.L. Russom, G.T. Ankley, T.W. Schultz, and J.D. Walker. 2003. Overview of data and conceptual approaches for derivation of quantitative structure-activity relationships for ecotoxicological effects of organic chemicals. *Environ. Toxicol. Chem.* 22:1789-1798.
- Combes, R. 2007. Developing, validating and using test batteries and tiered (hierarchical) testing schemes. *Alternatives to Laboratory Animals.* 35(4): 375-378.
- Custer, K.W. and G.A. Burton Jr. 2008. *Isonychia spp.* and macroinvertebrate community responses to stressors in streams utilizing the benthic in situ toxicity identification evaluation (BiTIE) method. *Environ. Pollution*, 151(1): p. 101-109.
- Dix, D. J., K. A. Houck, M. Martin, A. M. Richard, R. W. Setzer, And R. J. Kavlock. 2007. The toxcast program for prioritizing toxicity testing of environmental chemicals. *Toxicol Sci.*, 95(1):5-12.
- Ekman, D.R., Q. Teng, D.L. Villeneuve, M.D. Kahl, K.M. Jensen, E.J. Durhan, G.T. Ankley and T.W. Collette. 2008. Investigating compensation and recovery of fathead minnow (*Pimephales promelas*) exposed to 17 α -ethynylestradiol with metabolite profiling. *Environ. Sci. Technol.* 42, 4188-4194.

- Environment Canada. 2008. Canada's Chemical Management Plan, presented by Suzanne Easton at the GLBTS Substance Working Group on April 8, 2008. Available at <http://www.epa.gov/glnpo/bns/integration/April2008/index.html>.
- Environmental Council of the States (ECOS). 2010. State Experiences with Emerging Contaminants: Recommendations for Federal Action. Available at http://www.ecos.org/files/3959_file_January_2010_ECOS_Green_Report.pdf. (Accessed 04 Oct 2010).
- Fox, GA. 1991. Practical Causal Inference for Ecoepidemiologists. *J Toxicol Environ Health*, 33:359-373.
- Goldstone, J.V., Hamdoun, A., Cole, B.J., Howard-Ashby M., Nebert D.W., Scally M., Dean M., Epel D., Hahn M.E., Stegeman, J.J. 2006. The chemical defensome: Environmental sensing and response genes in the *Strongylocentrotus purpuratus* genome. *Developmental Biology* 300:366–384.
- Gunnarsson L, Jauhiainen A, Kristiansson E, Nerman O, Larsson DG. 2008. Evolutionary conservation of human drug targets in organisms used for environmental risk assessments. *Environ Sci Technol*. 42(15):5807-13.
- Head JA, Hahn ME, Kennedy SW. 2008. Key amino acids in the aryl hydrocarbon receptor predict dioxin sensitivity in avian species. *Environ Sci Technol*. 42(19):7535-41.
- Hill, AB. 1965. The environment and disease: association or causation?. *Proc Royal Soc Med* 58: 295–300.
- Hinck, J.E., Schmitt, C. J., Ellersieck, M. R. and Tillitt, D.E. 2008. Relations between and among contaminant concentrations and biomarkers in black bass (*Micropterus spp.*) and common carp (*Cyprinus carpio*) from large U.S. rivers, 1995–2004. *J. Environ. Monit.*, 10:1499–1518
- Judson, R., K. A. Houck, R. J. Kavlock, T. B. Knudsen, M. T. Martin, H. M. Mortensen, D. Reif, D. Rotroff, I.A. Shah, A.M. Richard, And D.J. Dix. 2009. In Vitro screening of environmental chemicals for targeted testing prioritization The ToxCast Project. *Environ Health Perspect* 118(4): 485-492
- Jukosky, J.A., Watzin, M.C. and J.C. Leiter. 2008. Elevated Concentrations of Ethinylestradiol, 17 β -Estradiol, and Medroxyprogesterone have Little Effect on Reproduction and Survival of *Ceriodaphnia dubia*. *Bull. Environ. Contam. Toxicol*. 81:230-235.
- Lebo, J.A., et al. 2000. Selective removal of organic contaminants from sediments: a methodology for toxicity identification evaluations (TIEs). *Chemosphere*. 40(8): p. 811-819.
- Lemos, M.F.L., Soares, A., A.C., Correia, A.C. and A. C. Esteves. 2010. Proteins in ecotoxicology – How, why and why not? *Proteomics* 10: 873–887.
- Leusch, F.D.L. 2008. Tools to detect estrogenic activities in environmental waters. Global Water Research Coalition Report.

- Lowell, R.B., B. Ring, G. Pastershank, S. Walker, L. Trudel and K. Hedley. 2005. National Assessment of Pulp and Paper Environmental Effects Monitoring Data: Findings from Cycles 1 through 3. National Water Research Institute, Burlington, Ontario. NWRI Scientific Assessment Report Series No. 5. 40 pp.
- McMaster, M.E., L.M. Hewitt and J.L. Parrott. 2006. A decade of research on the environmental impacts of pulp and paper mill effluents in Canada: Field studies and mechanistic research. *J. Toxicol. Environ. Health B*, 9, 319-339.
- Miller, D.H., K.M. Jensen, D.L. Villeneuve, M.D. Kahl, E.A. Makynen, E.J. Durhan and G.T. Ankley. 2007. Linkage of biochemical responses to population-level effects: a case study with vitellogenin in the fathead minnow. *Environ. Toxicol. Chem.* 26, 521-527.
- Munkittrick, K., T.J. Barrett and M.E. McMaster. 2010. Guidance for Site-Specifically Assessing the Health of Fish Populations with Emphasis on Canada's Environmental Effects Monitoring Program. *Water Quality Research Journal of Canada*, Vol. 45 (2): 209-221.
- Fate and Effects of Pulp and Paper Mill Effluents 2010: Select Papers from the 7th International Conference
- National Research Council (NRC). 2007. Toxicity Testing in the 21st Century: A Vision and a Strategy, 2007: Committee on Toxicity Testing and Assessment of Environmental Agents, National Research Council of the National Academies. The National Academies Press, Washington, DC. 216 pp.
- Oakes, K.D, Tremblay, L.A., and G.J. Van Der Kraak. 2005. Short-term Lab Exposures of Immature Rainbow Trout (*Oncorhynchus mykiss*) to Sulfite and Kraft Pulp-mill Effluents: Effects on Oxidative Stress and Circulating Sex Steroids. *Environ. Toxicol. Chem.* 24, 1451-1461.
- Parks, L.G., C.S. Lambricht, E.F. Orlando, L.J. Guillette, G.T. Ankley and L.E. Gray. 2001. Masculinization of female mosquitofish in kraft mill effluent contaminated Fenholloway River water is associated with androgen receptor agonist activity. *Toxicol. Sci.* 62, 257-267.
- Picone, M., et al. 2009. Sequential toxicity identification evaluation (TIE) for characterizing toxicity of Venice Lagoon sediments: Comparison of two different approaches. *Ecotox. Environ. Safety.* 72(2): p. 538-545.
- Purdom, C.E., P.A. Hardiman, V.J. Bye, N.C. Eno, C.R. Tyler and J.P. Sumpter. 1994. Estrogenic effects of effluents from sewage treatment works. *Chem. Ecol.* 8, 275-285.
- Southern California Coastal Water Research Project (SCCWRP). 2010. Research Area: Contaminants of Emerging Concern.
<http://www.sccwrp.org/ResearchAreas/Contaminants/ContaminantsOfEmergingConcern.aspx>.
- Tillitt, D.E., J.P. Giesy and G.T. Ankley. 1991. Characterization studies of the H4IIE rat hepatoma cell bioassay for use in the analysis of planar halogenated hydrocarbons (PHHs) in environmental samples. *Environ. Sci. Technol.* 25, 87-92.

- USEPA. 1985. Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and their Uses. United States Environmental Protection Agency. Stephan, C.E., D.I. Mount, D.J. Hansen, J.H. Gentile, G.A. Chapman and W.A. Brungs. PB85-227049. National Technical Information Service, Springfield, VA.
- USEPA. 1995. Great Lakes Water Quality Initiative Technical Support document for Wildlife Criteria. EPA-820-B-95-009.
- USEPA. 1998. Guidelines for Ecological Risk Assessment. Federal Register 63(93):26846.
- USEPA. 2007. Sediment Toxicity Identification Evaluation (TIE): Phases I, II, and III Guidance Document. EPA/600/R-07/080.
- Van Aggelen, G., G.T. Ankley, W.S. Baldwin, D.W. Bearden, W.H. Benson, J.K. Chipman, T.W. Collette, J.A. Craft, N.D. Denslow, M.R. Embry, F. Falciani, S. G. George, C.C. Helbing, P. F. Hoekstra, T. Iguchi, Y. Kagami, I. Katsiadaki, P. Kille, L. Liu, P.G. Lord, T. McIntyre, A. O'Neill, H. Osachoff, E.J. Perkins, E.M. Santos, R.C. Skirrow, J.R. Snape, C.R. Tyler, D. Versteeg, M.R. Viant, D.C. Volz, T.D. Williams and L. Yu. 2010. Integrating omic technologies into aquatic ecological risk assessment and environmental monitoring: Hurdles, achievements and future outlook. Environ. Health Perspect. 118, 1-5.
- Water Environment Research Foundation (WERF). 2010. Diagnostic Tools to Evaluate Impacts of Trace Organic Compounds: Prioritization Framework for Trace Organic Compounds. Report number CEC5R08a.
- Wheeler, J.R., S. Gimeno, M. Crane, E. Lopez-Juez and D. Morritt. 2005. Vitellogenin: A review of analytical methods to detect (anti) estrogenic activity in fish. Toxicol. Mech. Methods 15, 293-306.
- Whiteman, FW, Ankley, GT, Kahl, MD, Rau, DM, and MD Balcer. 1996. Evaluation of interstitial water as a route of exposure for ammonia in sediment tests with benthic macroinvertebrates. Environ. Toxicol. Chem. 15:794-801.
- Whyte, J.J., R.E. Jung, C.J. Schmitt and D.E. Tillitt. 2000. Ethoxyresorufin-O-deethylase activity in fish as a biomarker of chemical exposure. Crit. Rev. Toxicol. 30, 347-570.
- Whyte, JJ., C.J. Schmitt and D.E. Tillitt. 2004. The H4IIE cell bioassay as an indicator of dioxin-like chemicals in wildlife and the environment. Crit. Rev. Toxicol. 34, 1-83.
- Wilson, V., K. Bobseine, C. Lambright and L.E. Gray. 2002. A novel cell line, MDA-kb2, that stably expresses an androgen- and glucocorticoid-reponse reporter for the detection of hormone receptor agonists and antagonists. Toxicol. Sci. 66, 69-81.

Appendices

Appendix A: Literature Search on “Chemicals of Emerging Concern” (CECs) and Related Terms

Appendix B: Summary of Existing Strategies for Ecological Risk Assessment of Chemicals of Emerging Concern

Appendix C: Resources for Exposure Assessment

Appendix D: Literature Search on Effects-Based Biomonitoring Tools for Ecological Risk Assessment

Appendix E: Resources for Effects Assessment