Monitoring of Constituents of Emerging Concern (CECs) in Aquatic Ecosystems – QA/QC Guidance

Nathan G. Dodder, Alvine C. Mehinto and Keith A. Maruya
Southern California Coastal Water Research Project Authority
Costa Mesa, CA 92626

1. Introduction

In 2009, the State of California Water Resources Control Board (SWRCB) tasked a scientific advisory panel ("Panel") to assess current scientific knowledge of the risks posed by CECs to freshwater, coastal and marine ecosystems, and to provide recommendations for CEC monitoring that will protect beneficial uses in these ecosystems. In their final report, the Panel utilized a risk-based screening framework to identify a list of CECs for monitoring in three representative receiving water scenarios, and recommended development of better CEC monitoring and assessment tools, including bioanalytical screening methods (Anderson et al. 2012).

In response to these recommendations, SWRCB staff tasked the Southern California Coastal Water Research Project Authority (SCCWRP) to generate a study plan to perform pilot monitoring of CECs statewide. The major elements of this pilot investigation are to (1) measure occurrence of CECs identified by the Panel in source and receiving waters and in appropriate matrices (i.e. discharged wastewater treatment plant (WWTP) effluent, waters receiving WWTP effluent and stormwater runoff, sediment, and/or tissue) and (2) evaluate alternative monitoring methods, including bioanalytical screening tools and whole organism toxicity tests that better target biological responses associated with CECs. A full description of the study plan elements is documented elsewhere (SCCWRP 2015).

By definition, CECs are not widely regulated and thus not routinely monitored. As a result, there is a likelihood of larger variation in data quality among laboratories, since available analytical methods may not be as robust as for historical (priority) pollutants. Statewide monitoring will include participation by multiple agencies, field crews, and laboratories; therefore, ensuring that results are comparable among different groups by maintaining consistency in field and laboratory operations is critical to success.

1.1. Scope

Since integrated statewide CEC pilot monitoring is not currently in the implementation phase, the level of detail available at the time of writing fall short of the information required in a Quality Assurance Project Plan (QAPP). In lieu of a QAPP, this document describes currently available QA/QC related information which should be used as guidance in generating a QAPP when the appropriate level of detail is made available for project implementation. A description of the necessary information is included in Section 7.

1.2. Objectives

The goal of this document is to ensure data quality and comparability among participating agencies, field crews, and laboratories, and to ensure data can confidently be compared to other surveys. Ensuring data quality consists of two distinct but related activities: quality assurance and quality control.

Quality assurance (QA) includes design, planning, and management actions conducted prior to field sampling to ensure appropriate types and quantities of data are collected. The goals of QA are to ensure that: 1) sample transport and processing, and laboratory analytical techniques will be applied consistently and correctly; 2) the number of lost, damaged, and uncollected samples will be minimized; 3) the integrity of the data will be maintained and documented from sample collection to entry into the data record; 4) data will be comparable; and 5) measurements can be reproduced. This will be achieved by:

1. Evaluation of laboratories' ability to conduct the analyses based on prior data, and the establishment of reporting levels (RLs),
2. Development of the project quality control procedures described below,
3. Evaluation of the comparability of analytical and bioassay methods through inter-laboratory evaluations, and
4. Development of a data management plan.

Quality control (QC) activities are implemented during the data collection phase of the project to evaluate the effectiveness of the QA procedures. These activities ensure that measurement error and bias are identified, quantified, and either accounted for or eliminated. This will be achieved by:

1. Standard procedures for sample collection and recording of field observations,
2. Standard procedures for sample shipment and storage, and
3. Adherence to a common set of measurement quality objectives (MQOs). The MOQ defines acceptance criteria based on calibration of the instrument, evaluation of blank concentrations, repeated measurements to establish method precision, and use of test samples to establish method accuracy.

2. Sample Collection, Handling and Preservation

Field personnel must strictly adhere to established protocols to insure the collection of representative, uncontaminated pilot study samples. Guidelines for sample storage are provided in **Table 2.1**. Changes and/or additions to these guidelines may be proposed by project participants if proper justification is provided.

- Field personnel must be thoroughly trained
    - in the proper use of sample collection gear,
    - in distinguishing acceptable versus unacceptable samples in accordance with pre-established criteria,
    - to recognize and avoid potential sources of sample contamination.

- Sampling equipment and utensils that come in direct contact with the sample should be made of non-contaminating materials and should be thoroughly cleaned between sampling stations.
- Sample storage containers should be of the recommended type and must be free of contaminants.
- Conditions for sample collection, preservation and holding times should be followed, and relevant field observations should be recorded.

On the day of sampling, field personnel should avoid contact with or consumption of products that contain the target analytes. This may include soaps, detergents, fragrances, sunscreen, and pharmaceuticals. Storage containers with Teflon should not be used to store samples that are slated for analysis of perfluorinated compounds (PFCs).

Table 2.1. Sample collection and holding time conditions.

Matrix	**Container Type**	**Container size (mL)**	**Preservation Requirements**	**Maximum Holding Time**
Aqueous	Pre-cleaned amber glass	1000 (100% full)	Cold (4 °C), with preservative added as required	2 weeks
Sediment	Pre-cleaned amber glass	250 or 125 (80% full)	Frozen (-20 °C)	1 year
Tissue	Pre-cleaned amber glass	250 or 125 (80% full)	Frozen (-20 °C)	1 year

3. Laboratory Documentation of General Practices

All laboratories performing measurement of parameters specified in the pilot study plan (SCCWRP 2015) and as delineated in Sections 4-6 herein must have the appropriate facilities to store and prepare samples, and appropriate instrumentation and staff to provide data of the required quality within the specified time period. Laboratories are expected to conduct operations using good laboratory practices, including:

- A program of scheduled maintenance of analytical balances, laboratory equipment, static and flow through exposure apparatuses, and instrumentation.
- Checking and recording the composition of fresh calibration standards against the previous lot.
- Checking and recording of water or sediment quality parameters in toxicity tests.
- Monitoring and recording temperatures within exposure rooms, storage areas and freezer units.
- Acquisition of solvents, test cell lines/kits and other consumables of suitable quality.
- Dating and storing all samples safely upon receipt and use of a laboratory information management system to track the location and status of any sample.

Personnel shall be well versed in good laboratory practices, including standard safety procedures. It is the responsibility of the laboratory to ensure that safety training is mandatory for all personnel. The laboratory is responsible for maintaining a current safety manual in compliance with the Occupational Safety and Health Administration (OSHA), or equivalent state or local regulations. The safety manual should be readily available to laboratory personnel. Best safety practices should be followed at all times, including proper storage, handling, and disposal of chemicals; verification of fume hood operation; and use of supplies/equipment to prevent potential health hazards.

Laboratories shall be able to provide documentation of their ability to conduct analyses with the level of data quality specified herein. Specifically, the following documents and information must be available upon request:

- QA Plan: Policies and protocols specific to a particular laboratory including personnel responsibilities, procedures for determining the acceptability of results, and procedures for release of the data.
- Standard Operating Procedures (SOPs): Step-by-step instructions describing in detail implementation of the method, specific for the particular equipment and instruments used.
- Instrument performance information: Laboratories should collect ongoing data on instrument baseline noise, calibration standard response, detection limits, and laboratory blanks.

4. Analysis of Chemical Contaminants

4.1. General Approach

A performance-based approach to QA/QC is recommended. In this format, specific analytical methods are not prescribed, rather each laboratory may use methods of their choice as long as QA/QC requirement are met and acceptable performance is demonstrated. For CECs in particular, mass spectrometry based methods shall be used; e.g., gas chromatography/mass spectrometry (GC/MS) or liquid chromatography/tandem mass spectrometry (LC-MS/MS). Also, these methods shall employ spiked surrogate or internal standards to generate calibration curves. Standard addition methods shall not be used. Detailed criteria based on QA/QC guidelines adopted by the Southern California Bight Program (SCCWRP 2013), the Surface Water Ambient Monitoring Program (SWAMP 2008) and the U.S. Geological Survey (USGS 2004), are described in the following subsections:

Sec 4.2 Target, matrix and scenario specific reporting limits (RLs)

Sec 4.3 Performance in inter-laboratory comparison exercises

Sec 4.4 Sample completeness

Sec 4.5 Measurement quality objectives (MQOs)

4.2. Reporting Limits (RLs)

Recommended reporting limits (RLs) for pilot study CECs were set at 50% of monitoring trigger levels (MTLs) established by the Panel (Anderson et al. 2012) in order to allow for the collection of data that will be useful in evaluating CEC risk (SCCWRP 2015). These RLs are specified for each target compound (i.e. CEC), matrix and scenario, and thus may differ among scenarios (**Table 4.1**). In some cases, the Panel recommended RL is lower than what commercial services labs currently offer. As methods continue to improve and evolve, participating labs shall strive to achieve the recommended RLs, and shall in all cases meet the minimum achievable RLs.

4.3. Inter-Laboratory Comparison Exercise

All laboratories contributing analytical chemistry data for the pilot study shall participate in an inter-laboratory exercise to demonstrate comparability with all participants, including those considered as referee labs. The recent advent of commercially available services for many of the target CECs, coupled with the extremely low RLs required, necessitate an assessment of data comparability among participating labs. The inter-laboratory comparison will provide an opportunity to revise project MQOs, if warranted, based on group consensus. Additional value in participating in inter-laboratory exercises are: 1) laboratories not passing minimum performance criteria are made aware of methodological issues and can work with referee labs to resolve these issues, and 2) a quantitative assessment of among-laboratory variability will provide context for managers when comparing results to other CEC-related projects.

Table 4.1. Monitoring trigger levels (MTLs) and reporting limits (RLs) by scenario, compound and matrix. Recommended RLs are derived from MTLs as reported by the CEC Science Advisory Panel. Achievable RLs reflect the current state of art for commercial services laboratories. Recommended RLs for all CECs in wastewater treatment plant (WWTP) effluent and stormwater (MS4) influenced receiving waters are equivalent to Scenario 1 aqueous phase RLs; additional RLs for compounds that are otherwise measured only in sediment or tissues appear at the bottom of the table.

Compound	Panel Freshwater MTL[1]	Recommended RL[2]	Achievable RL[3]
Aqueous Phase - Effluent dominated inland waterways (Scenario 1) (ng/L)			
Bifenthrin[4]	0.40	0.20	
Permethrin[4]	1.0	0.50	
Fipronil[4]	42	21	
Chlorpyrifos[4]	5.0	2.5	
Estrone	6.0	3.0	
Ibuprofen	100	50	
Bisphenol A	60	30	
17-beta-estradiol	2.0	1.0	
Galaxolide (HHCB)	700	350	
Diclofenac	100	50	
Triclosan	250	125	
Sediment Phase - Effluent dominated inland waterways (Scenario 1) (ng/g dw)			
Fipronil	0.090	0.045	1.0
Aqueous Phase - Coastal embayments (Scenario 2) (ng/L)			
Bisphenol A	6.0	3.0	
Bifenthrin	0.040	0.020	0.2
Permethrin	0.10	0.050	0.5
Fipronil	5.0	2.5	
Chlorpyrifos	1.0	0.50	
Estrone	0.60	0.30	2.0
17-beta-estradiol	0.20	0.10	0.4
Galaxolide (HHCB)	70	35	
Sediment - Coastal embayments (Scenario 2) (ng/g dw)			
Bifenthrin	0.052	0.026	0.20
PBDE-47	0.030	0.015	
PBDE-99	0.030	0.015	
Permethrin	0.073	0.036	0.40
Fipronil	6.5	3.25	
PFOS[5]	NA	0.1	

Sediment - Ocean discharge (Scenario 3) (ng/g dw)		
Bis(2-ethylhexyl) phthalate (BEHP)	130	65
p-nonylphenol	14	7.0
PBDE-47	0.30	0.15
PBDE-99	0.30	0.15
Butylbenzyl phthalate (BBP)	6.3	3.15
PFOS[5]	NA	0.1
Tissues (All Scenarios) (ng/g dw)		
PBDE-47	28.9	14.5
PBDE-99	28.9	14.5
PFOS	1000	500
WWTP Effluent and MS4 Receiving Water (ng/L) [6]		
Bis(2-ethylhexyl) phthalate (BEHP)		3.0
Butylbenzyl phthalate (BBP)		3.0
p-nonylphenol		22 [7]
PBDE-47		0.10
PBDE-99		0.10
PFOS		1.0

[1] Monitoring Trigger Level established by CEC Science Advisory Panel (Anderson et al. 2012).
[2] Set at 50% of MTL.
[3] Minimum RL reported by commercial services laboratories. Missing values indicate the achievable value is at or below the recommended RL.
[4] Scenario 1 pesticides are currently monitored by other programs. The recommended RLs are listed here for comparison purposes only.
[5] PFOS was recommended for Scenario 2 and 3 sediment monitoring to obtain information on sediment-biota transfer, not based on MTLs. The recommended RL was based on typical values observed in the literature and attainable values by laboratories.
[6] RLs for analytes otherwise measured in sediment or tissues only (no MTL values available). For all other analytes, RLs for WWTP Effluent and MS4 receiving water samples are the same as the aqueous RLs for Scenario 1.
[7] Estimated from the sediment RL (7.0 ng/g), an estimated sediment-water partitioning coefficient, and assuming 1% organic carbon content of the sediment.

A referee laboratory will be assigned to prepare reference materials representing the matrix and target analytes of interest (**Table 4.2**). Either materials with native levels of target CECs or representative matrices spiked with target CECs at concentrations at or above RLs will be used as reference materials. After division of the spiked reference material into multiple aliquots, the referee laboratory should verify the concentrations of the target analytes and establish sample homogeneity through within jar and between jar analyses. Participating laboratories should not have prior knowledge of target CEC concentrations, and should make repeated (e.g., triplicate) measurements of the reference material to assess within-laboratory variability. Standard reference materials (SRMs) that contain target CECs or analogs thereof, if available, may also be analyzed to test accuracy of methods employed for the reference material. The exact performance criteria should be decided by project participants based on their measurement knowledge for each analyte. However, laboratories should be assessed by comparing their results to a "target" value (e.g. ± 40% the group mean).

Table 4.2. Inter-laboratory comparison reference materials.

Reference Material	**Covers Scenario**
Freshwater	Scenario 1, Scenario 2, and Stormwater (MS4)
Effluent	WWTP effluent
Sediment	Scenario 2 and Scenario 3
Tissue	All scenarios

4.4. Completeness

Completeness is defined as the proportion of samples that are successfully collected, analyzed and that pass quality control (QC) validation. Losses may occur as a result of field conditions, logistical difficulties, or failure to achieve QC criteria. The MQO for completeness is 90% for each analyte. To achieve this criteria, the sampling design for the pilot study shall be sufficiently redundant to absorb the loss of up to 10% of the samples/analytes without compromising the pilot study goals, provided that the losses are not concentrated in a single subpopulation of interest.

4.5. Measurement Quality Objectives (MQOs)

The measurement quality objectives (MQOs) delineated in the following sections are intended to provide a common foundation for laboratory performance and should be considered as the minimum requirements for analyzing CECs in pilot study samples. Additional MQOs may be instituted by participating labs, as long as the MQOs presented herein are satisfied. Aqueous sample concentrations shall be reported using specific units (e.g., ng/L). Sediment sample concentrations shall be reported on a dry weight basis with the percent moisture of the corresponding sample also reported. Tissue sample concentrations should be reported on a wet weight basis, with percent moisture and percent lipid of the corresponding sample also reported. The methods for measuring percent moisture and percent lipids should be standardized among the participating laboratories.

4.5.1. Measurement Range and Sensitivity

Prior to the commencement of sample analysis, each laboratory should establish the working calibration range, and determine nominal method detection and reporting limits (MDLs and RLs, respectively) on an analyte- and matrix-specific basis. These steps are detailed in the following sections.

Calibration Range

The working calibration range for each target CEC must be established using a minimum of five concentrations, and acceptable performance should be demonstrated on an accuracy-based material (e.g., reference material described in 4.3). Only data resulting from quantification within the working calibration range may be reported by a laboratory without annotation. Samples with measured concentrations above or below the calibration range should be reanalyzed using appropriate sample mass and/or volume.

Reporting Level

The RL is the minimum concentration that can be reliably measured, and is also the minimum target concentration at which laboratories shall report data. By default, the RL is the lowest concentration in the calibration curve. If an alternate definition for RL is used, this shall be submitted to and approved by the Data Management Team (Section 7) prior to sample analysis.

Method Detection Limits (MDLs)

The method detection limit (MDL) represents a quantitative estimate of low-level response detected at the maximum sensitivity of a method. This level should be a concentration below the RL, and laboratories should describe the method they used to determine the MDL.

4.5.2. Ongoing Measurement Objectives

Following a successful setup phase, each laboratory must demonstrate maintenance of performance by repeating analysis of QC samples within each analytical batch. Descriptions of the QC samples are in the following sections, with the corresponding MQOs in **Table 4.3**. If control limits for any objective are not met, the laboratory shall take action to find and eliminate the problem before continuing with sample analysis. If a major unresolvable flaw is found, it may be necessary to repeat the analysis of the affected batch of samples.

Based on laboratory participant and project management consensus and the results of the inter-calibration exercise, it may be necessary to revise the MQOs for specific analytes prior to the collection of field data. The MQO criteria listed here should be viewed as a starting point for discussion with participation laboratories.

Initial Calibration and Continuing Calibration Verification

A new response factor or calibration curve should be established for each instrumental batch. A continuing calibration verification standard shall be analyzed at specified intervals (every 10 samples or 8 hours) to monitor temporal variability in the instrument. The continuing calibration verification standard should be at the mid-range calibration concentration, and must be within ±20% of the initial calibration response. An instrument blank should be included in the calibration curve to verify that the instrument is free of contamination or carryover.

Method Blanks

Method blanks assess laboratory contamination during sample preparation and analysis. One method blank should be run in each sample preparation batch, and it should be processed and analyzed using the same protocol used for samples. Blanks exceeding the MQO require corrective action to bring subsequent blanks into compliance. This may involve performing equipment maintenance, changing reagents and/or, as a last resort, modifying SOPs. Although acceptable laboratory blanks are important, improvements in analytical sensitivity and the pervasiveness of some contaminants result in situations where detection in laboratory blanks is unavoidable. The magnitude of the blank concentrations must be evaluated against the sample concentrations and the MQOs (see Table 4.3). Blank subtraction is allowed if the blank concentration is < 30% of the analyte concentration in the same batch.

Sample Duplicates

Analysis of sample duplicates is used to assess the precision of an analytical method and to check for sample heterogeneity. At least one sample per batch of 20 samples should be analyzed in duplicate (Table 4.3).

Matrix Spikes and Matrix Spike Duplicates

Matrix spike and matrix spike duplicates (MS/MSD) are laboratory-prepared samples of the matrix spiked with known levels of the target analytes, used to evaluate the effect of the sample matrix on analyte recovery, and additionally, to provide an estimate of analytical precision. The material to be spiked should represent the matrix of interest, i.e. be as similar as possible to the sample being analyzed. A minimum of one MS/MSD pair should be analyzed for every batch of 20 samples. The matrix spike solution should contain all the analytes of interest. The final spiked concentration of each analyte should be at least 3 times the RL. If the unspiked matrix contains background concentrations of any target analyte, the sample should be spiked with one to five times the preexisting concentration in the sample. Acceptance criteria for recovery of spiked analytes are provided in Table 4.3.

Standard Reference Materials or Laboratory Control Samples

Method accuracy is evaluated through the analysis of standard reference materials (SRMs) or laboratory control samples (LCS). Analyses of SRMs must yield values within the specified range of the certified (or reference) values provided by the supplier. Certified values have lower uncertainty than reference values, but in the absence of certified values, reference values are acceptable for assessment of accuracy. Due to the inherent variability in analyses near the MDL, criteria for accuracy will only apply to analytes having certified values that are >3 times the RL established by the laboratory. If a SRM for all target analytes is unavailable, an LCS can be substituted. An LCS is prepared by the laboratory using contaminant free water or an appropriate inert solid material spiked with the target analyte at a known concentration within the calibration curve. A minimum of one SRM or LCS should be analyzed per batch of 20 samples. Acceptance criteria for accuracy of target CECs are provided in Table 4.3.

Standards and Standard Recovery

Quantification standards are isotope-labeled or structurally similar analogs to the target analytes. Laboratories may refer to them as internal-, surrogate-, and/or isotope dilution standards, but the exact definition of these terms is inconsistently applied in the literature. These standards are used to

generate calibration curves and are added at known levels to field samples to monitor and adjust for extraction efficiency, sample losses, retention time shifts, instrumental drift, and ion suppression. The percent recovery of standards added prior to extraction and accounting for extraction and sample losses must be within control limits specified in Table 4.3.

Table 4.3. Ongoing Measurement Quality Objectives (MQOs) for target analytes in all matrices.

Measurement	**Frequency**	**Control Limit**
Initial Calibration	A new response factor or calibration curve should be established for each instrumental batch.	Relative standard deviation (RSD) of the response factor ≤ 25% Coefficient of determination r^2 ≥ 0.990 for linear and non-linear curves. First or second order curves allowed. Minimum of 5 points per curve.
Continuing Calibration Verification	Every 10 samples or 8 hours	Expected concentration ± 20%.
Method Blank	5% of total no. samples (1 per batch of 20 samples)	Less than the RL for target analytes.
Sample Duplicate	5% of total no. samples (1 per batch of 20 samples)	RPD ≤ 35%.
Certified Reference Material or Laboratory Control Sample	5% of total no. samples (1 per batch of 20 samples)	70-130% recovery if certified; otherwise, 50-150% recovery.
Matrix Spike/Matrix Spike Duplicate Pair	5% of total no. samples (1 per batch of 20 samples)	50-150% or based on historical laboratory control limits; RPD ≤ 25%.
Spiked Standard Recovery	All field and QC samples	50-150% or based on historical laboratory control limits.

5. Biological Testing

5.1. Bioanalytical Screening Tools

5.1.1. General Approach

The QA/QC criteria for these new monitoring tools were based on technical reports from EPA's Endocrine Disruptor Screening Program (USEPA 2013), and recently completed research projects on adapting in vitro bioassays (IVBs) for water quality screening (SCCWRP 2014; WRRF 2014). A performance-based approach is adopted where each laboratory may use their method of choice. General requirements are described in the following subsections:

Sec 5.1.2 In vitro bioassay (IVB) endpoints

Sec 5.1.3 Selection of reference toxicants

Sec 5.1.4 Measurement quality objectives (MQOs)

Sec 5.1.5 Performance in inter-laboratory comparison exercises

5.1.2. In Vitro Bioassay (IVB) Endpoints

Cellular (*in vitro*) bioassays will be used to screen chemicals and to determine their potential toxic effects. These tools will be applied for all four scenarios using water, sediment and tissue samples. The IVB endpoints described in the pilot study plan (SCCWRP 2015) and listed in **Table 5.1** can screen for endocrine disrupting chemicals (e.g. estrogens, androgens, progestins and glucocorticoid steroids) as well as dioxin-like chemicals.

Commercial Suppliers

In vitro bioassays selected for CEC monitoring are all commercially available. The existing suppliers are specified in Table 5.1.

Table 5.1. Recommended commercial suppliers for in vitro biossays (IVBs).

Endpoints	Bioassay, Supplier
Estrogen Receptor (ER)	GeneBLAzer ERα Division Arrested Assay, Life Technologies[1] ERα CALUX, BioDetection Systems[2]
Androgen Receptor (AR)	GeneBLAzer AR Division Arrested Assay, Life Technologies[1] AR CALUX, BioDetection Systems[2]
Glucocorticoid Receptor (GR)	GeneBLAzer GR Division Arrested Assay, Life Technologies[1] GR CALUX, BioDetection Systems[2]
Progesterone Receptor (PR)	GeneBLAzer PR Division Arrested Assay, Life Technologies[1] PR CALUX, BioDetection Systems[2]
Aryl Hydrocarbon Receptor (AhR)	AhR CALUX, BioDetection Systems[2]

[1] Madison, WI (USA); [2] Amsterdam, The Netherlands

Sample Processing

Samples to be screened by IVBs shall be collected and preserved following the methods described in Section 2. Samples will be extracted following the same protocols used for analytical chemistry *with one critical modification*. To prevent non-sample related interference in bioassay response, addition, fortification or spiking of chemicals of any kind (e.g. internal standards or recovery surrogates per section 4), except those specifically identified to evaluate IVB performance, shall not be performed.

5.1.3. Reference Toxicants

Reference toxicants used in the IVBs shall meet the following requirements:

- High affinity for the endpoint of interest
- Linear dose response shall have a dynamic range of 5-fold minimum
- Endpoint specific sensitivity thresholds reported in Table 5.1 shall be attained

Since there is limited information on the performance of alternative reference toxicants, it is recommended that all laboratories employ the reference toxicants listed in **Table 5.2**. The performance of these chemicals has been evaluated in recent studies that adapted bioassay protocols for water quality measurement (SCCWRP 2014; Escher et al. 2014).

Table 5.2. Recommended reference toxicants for in vitro bioassays (IVBs). Agonist mode (+); antagonist mode (-).

Endpoints	Reference Toxicant	Sensitivity Threshold (ng/L)
Estrogen Receptor (ER)	17-beta estradiol (+) 4-hydroxy-tamoxifen (-)	0.5
Androgen Receptor (AR)	Flutamide (-)	20
Glucocorticoid Receptor (GR)	Dexamethasone (+)	50 (TBR)
Progesterone Receptor (PR)	Levonorgestrel (+)	50 (TBR)
Aryl Hydrocarbon Receptor (AhR)	3,3',4,4',5-Pentachlorobiphenyl (PCB 126)(+)	50 (TBR)

TBR - to be resolved

5.1.4. Measurement Quality Objectives (MQOs)

The MQOs delineated in **Table 5.3** are intended to provide a common foundation for laboratory performance and should be considered as the minimum requirements for bioanalytical screening of pilot study samples. Additional MQOs may be instituted by participating laboratories, as long as the MQOs presented herein are satisfied. In vitro bioassay results shall be reported as bioassay equivalent concentrations (BEQs) in units of ng/L (as reference toxicant).

Table 5.3. Measurement quality objectives (MQOs) for *in vitro* bioassays (IVBs)

Measurement Parameter	Frequency of Analysis	Control Limits
Extract Cytotoxicity	Per sample extract	Dilutions of the extract shall not cause > 20% cell mortality (corrected for background).
Cell-Free Media Blank	Per assay plate	Average response for cell free blank (media only) shall be less than 75% of the solvent vehicle free blank response (cells and media). RSD of replicate wells shall be < 20%.
Vehicle Blank Response	Per assay plate	Average response of cells exposed to the solvent vehicle shall be within 15% RSD of the vehicle free response.
Initial Calibration	Per bioanalytical batch	Linear dose-response curve for reference toxicant; $r^2 > 0.95$. Minimum of 9 points per curve (one of them at or below sensitivity threshold (Table 5.2).
Calibration Verification	Per subsequent assay plates within a bioanalytical batch	Continuing calibration shall remain within 15% of mean response for initial calibration.
Spiked Sample	Per extraction batch	Assay response of sample spiked with reference toxicant shall be within 70 to 130% of expected response.
Reproducibility	Per sample	Differences among replicate bioassay responses shall be less than 20% RPD within and among laboratories.

5.1.5. Inter-laboratory Comparisons

All laboratories conducting IVBs shall participate in an inter-laboratory comparison exercise prior to sample testing. This exercise will include the analysis of spiked samples prepared by a referee laboratory, and un-spiked pilot study samples for each endpoint undertaken. Samples will be distributed blindly to the participating laboratories and analyzed in triplicate. Successful completion of this exercise will be evaluated based on attainment of MQOs (see sec 5.1.4), and data comparability among laboratories.

Data comparability will be based on the following acceptance criteria:

- Intra-laboratory reproducibility shall be 20 % relative percent difference (RPD)
- Percentage difference from the BEQ target value for each spiked sample shall be <30%
- Sensitivity and dose-response curve of reference toxicants shall be in accordance with MQOs (see Tables 5.2 and 5.3)

Laboratories unable to successfully complete the inter-laboratory comparison exercise will be asked to review their test procedures, make suggested changes, and retest the comparison samples. Failure to meet the inter-laboratory comparison criteria will result in the addition of a cautionary data qualifier

flag to that laboratory's data or exclusion from testing during the monitoring program. However, by participating in these exercises, laboratories not passing minimum performance criteria will be informed of methodological issues and shall be able to work with referee labs to resolve issues. In addition, the quantitative assessment of among-laboratory variability afforded by these exercises will provide context for managers when comparing results to other CEC-related projects.

5.2. In Vivo Toxicity Testing

5.2.1. General Approach

For each scenario (freshwater, embayment, ocean and stormwater), the toxicity of water and/or sediment samples will be evaluated using a whole organism (*in vivo*) test that include reproductive or developmental endpoints. To date, the 21-day reproduction test using fathead minnow (*Pimephales promelas*) is one of the most promising assays for detecting the effects of endocrine disrupting CECs. Thus, it is the only *in vivo* test to be evaluated in the pilot study at this time. This test will be conducted using aqueous freshwater samples (e.g. WWTP effluent, river water). Tests for other scenarios and matrices may be optimized and added to the pilot study plan at a later date.

5.2.2. Toxicity of freshwater samples using fathead minnow (*Pimephelas promelas*)

A short-term reproduction assay using *P. promelas* will be conducted on aqueous freshwater samplesaccording to USEPA (2007) and OECD (2012) guidelines. This test consists of a three to four week acclimation period, followed by a two week (minimum) pre-exposure period and a 21-day exposure to the test samples.

Clean water controls and freshwater samples will be tested in quadruplicate vessels under flow through conditions. Each test vessel will contain two males and four females fed daily with frozen blood worms.

The following test criteria are from the EPA and OECD fish reproduction protocols. These documents should be consulted for additional information on exposure conditions.

Selection of Organisms

Reproductively mature fish (namely, with visible secondary sexual characteristics) capable of actively spawning will be used. Fathead minnows should be preferably five to seven months old, and selected from a single laboratory population that has been cultured at 25 ± 2°C. If possible, the range of individual weights by sex should be kept within 20% of the mean weight of the same sex. For inter-calibration exercises or multi-laboratories studies, it is recommended a common supplier be identified to supply fish within a defined size (e.g. based on mass) range.

During the acclimation period, fish mortalities must be recorded and the following criteria applied:

- Mortalities less than 5% of fish population in seven days: accept the batch
- Mortalities greater than 10% of population in seven days: reject the entire batch
- Mortalities between 5 and 10% of population: acclimate for seven additional days; if more than 5% mortality during second seven days, reject the entire batch

Fish will not be treated for any disease during the holding period, pre-exposure period, or exposure period.

Toxicity Endpoints

Toxicity of test samples will be determined relative to the responses measured in the control vessels. The following endpoints will be measured over the course of the exposure or at termination of the test:

- Survival: Daily assessment. Dead fish will not be replaced in either control or treatment vessels.
- Behavior: Daily qualitative observations of changes in behavior such as uncoordinated swimming, loss of equilibrium, atypical feeding, and hyperventilation.
- Appearance: Daily qualitative assessment of secondary sex characteristics (e.g. size of males' fatpad, number and prominence of nuptial tubercles) and fish coloration conducted daily. Secondary sex characteristics are often affected by the presence of endocrine active chemicals.
- Egg production (fecundity): Number of eggs laid per surviving female per reproductive day.
- Fertilization success: Percentage of fertilized eggs, calculated as the number of embryos/ number of eggs x 100%.
- Vitellogenin (vtg) concentration: Vtg measurements in the plasma will be performed using a validated enzyme-linked immunosorbent assay (ELISA) method capable of detecting vtg in the low ng/mL range.
- Gonad condition measured as the gonadosomatic index (GSI; gonad weight/ body weight x 100%). Typical GSI values are 8 to 13% for reproductive females and 1 to 2% for reproductive males. CECs that affect egg production will also cause a reduction of the GSI in one or both sexes.
- Gonad histopathology (optional): Toxicity responses include intersex, decreased yolk formation, oocyte atresia, testicular degeneration, and hyperplasia.

Measurement Quality Objectives (MQOs)

Measurement quality objectives (MQOs) for the fathead minnow reproduction assay are summarized in **Table 5.4**.

Table 5.4. Measurements quality objectives (MQOs) for fathead minnow assay

Parameter	Acceptance Criteria
Survival	≥ 90% survival in clean water and/or solvent control vessels at the end of the exposure.
Egg Production	Spawning of 50 to 250 eggs every 4 days minimum during the pre-exposure. Parameters shall be maintained in the control vessels during the exposure.
Fertilization Success	Control fertilization shall be ≥ 95%.
Vitellogenin Concentration	Calibration curve with 6 points minimum, $r^2 \geq 0.98$). Absorbance of duplicate blank samples shall be ≤ 5% of the maximum calibration standard absorbance with a RPD < 20%.
Water Chemistry	≥ 60% air saturation; temperature 25°C± 1°C
Spiked Chemical Exposure	Concentrations shall be maintained within ± 20% of the mean measured value throughout the exposure period.

Note: Water chemistry parameters (e.g. temperature, dissolved oxygen, pH, conductivity) should be recorded daily and reported with the test results. If a parameter falls outside of the MQO for one replicate on a given day, best professional judgment should be used to determine the validity of the test.

6. Data Management Plan

The following sections describe the roles and responsibilities, formatting, verification and quality assessment, and reporting requirement for all pilot study monitoring data.

6.1 Roles and Responsibilities

Pilot study data shall be submitted by all participating entities to a single Data Management Team (DMT). The DMT is responsible for coordinating receipt of the data, developing and maintaining a data repository for the project, verifying data quality, and providing information to stakeholders and the State data repository. The DMT is responsible for coordinating the development of a common submission format.

6.2 Data Submission Format

The data submission formatting will align as closely as possible to the California Environmental Data Exchange Network (CEDEN) data submission templates, with additional fields to include project specific information as needed. Ultimately, the complete data set must be submitted to CEDEN by the DMT.

6.3 Data Submission

Data will be submitted by electronic spreadsheet, with an accompanying narrative describing any issues that should be brought to the attention of the DMT. Upon receipt and evaluation by the DMT, the analytical laboratory must be notified of any additional information or corrective actions deemed necessary. Following satisfactory resolution of all "corrective action" issues, the final action is to notify the laboratory in writing that the submitted results have been officially accepted as complete. Evaluation of the data by the DMT should begin as soon as possible following its receipt, since delays increase the chance that information may be lost. The following steps are to be followed and documented: 1) checking data completeness, 2) assessing data quality, and 3) QC reporting. All instrumental data and calculations leading to the submitted results should be retained by the laboratories in case a detailed inspection is required.

6.4. Data Completeness

Upon receipt of data, the DMT will verify it has been supplied in the correct format and enter it in to the repository. Checks will be performed to verify results have been reported for all expected stations, samples, and analytes, and all QC data has been included. The field crew or laboratory will be contacted to request any missing data. Significant revisions may require resubmission of the entire data set. Raw data (e.g., chromatograms or original quantitation reports) are not required for submission but must be maintained by the laboratories and made available if requested.

6.5. Assessing Data Quality

Data quality will be validated by the DMT as follows:

1. A check to verify that all reporting units and number of significant figures are correct.
2. A check to verify that all calculated percent recovery values and relative percent difference values are correct.
3. All QC data should be compared against the established MQO criteria.

There are several possible courses of action to be taken if the reported data are deficient during the assessment of data quality. First, the laboratory's narrative explanation should be consulted to determine if the problems were satisfactorily addressed. If there were minor MQO criteria exceedances in isolated cases, then it is appropriate for the laboratory to report the results along with appropriate qualifiers for those cases. Pervasive violations of MQO criteria, however, will result in one of the following courses of action. 1) All associated results will be qualified as estimated values. For example, if an analyte had minor QC violations in 3 of 5 analytical batches, the results from all 5 batches may be qualified as estimated. 2) In the most extreme situation, all associated data will be rejected and deleted from the repository.

Because some degree of expert judgment and subjectivity is typically necessary to evaluate QA/QC results and assign data qualifiers, validation will be conducted only by qualified personnel. Data which are qualified as estimates because of minor MQO violations are still usable for most assessment and reporting purposes. However, all QA/QC data will be available in the repository, so interested users may make their own determination of data quality.

6.6. Reporting

The DMT will produce reports documenting the results of QC reviews. These documents will summarize all conclusions concerning data acceptability and should note all significant quality assurance problems. These reports provide data users with a written record of QC concerns and a documented rationale for why certain data were accepted as estimates or were rejected. The following items should be addressed in the QA report:

1. A statement on the completeness of the data set relative to the original objectives.
2. A summary of overall data quality, including a description qualified data and rationales
3. Brief descriptions of analytical methods and the method(s) used to determine reporting and detection limits.

7. ADDITIONAL INFORMATION

Information in this document is intended to be serve as guidance in generating a project QAPP for statewide CEC pilot monitoring data collection. This QAPP should follow EPA Guidance for Quality Assurance Project Plans (EPA QA G-5), which requires information that can only be known in the implementation phase of the project, i.e. once the organization and scope of the various project components are finalized. This includes 1) project management information such as the names of key personnel, 2) data generation information such as exact sampling and analytical methods, and 3) an assessment plan to ensure the QA Project Plan is being implemented as approved.

References

Anderson, P.D., Denslow, N.D., Drewes, J.E., Olivieri, A.W., Schlenk, D., Scott, G.I., Snyder, S.A. 2012. Science Advisory Panel for Southern California Coastal Water Research Project (SCCWRP). Monitoring Strategies for Chemicals of Emerging Concern (CECs) in California's Aquatic Ecosystems, Technical Report 692, Southern California Coastal Water Research Project Authority, Costa Mesa, CA, 215 pp.

Escher, B.I., Allinson, M., Altenburger, R., Bain, P.A., Balaguer, P., Busch, W., Crago, J., Denslow, N.D., Dopp, E., Hilscherova, K., Humpage, A.R., Kumar, A., Grimaldi, M., Jayasinghe, B.S., Jarosova, B., Jia, A., Makarov, S., Maruya, K.A., Medvedev, A., Mehinto, A.C., Mendez, J.E., Poulsen, A., Prochazka, E., Richard, J., Schifferli, A., Schlenk, D., Scholz, S., Shiraishi, F., Snyder, S., Su, G., Tang, J.Y., Burg, B.V., Linden, S.C., Werner, I., Westerheide, S.D., Wong, C.K., Yang, M., Yeung, B.H., Zhang, X., Leusch, F.D. 2014. Benchmarking organic micropollutants in wastewater, recycled water and drinking water with in vitro bioassays. *Environ Sci Technol* 48:1940-1956.

Organization for Economic Cooperation and Development (OECD). 2012. Fish Short Term Reproduction Assay. OECD Environmental Health and Safety Publications, Series on Testing and Assessment No. 229, October 2, 2012, OECD, Paris.

Southern California Coastal Water Research Project (SCCWRP). 2013. Bight '13 Contaminant Impact Assessment Committee, Southern California Bight 2013 Regional Marine Monitoring Survey (Bight '13) Quality Assurance Manual, Southern California Coastal Water Research Project Authority, Costa Mesa, CA.

SCCWRP. 2014. Development of Bioanalytical Techniques for Monitoring of Constituents/Chemicals of Emerging Concern (CECs) in Recycled Water Applications for the State of California, Final Report, Agreement 10-096-250, Southern California Coastal Water Research Project Authority, Costa Mesa, CA, 281 pp.

SCCWRP. 2015. Monitoring of Constituents of Emerging Concern (CECs) in Aquatic Ecosystems – Pilot Study Guidance, Southern California Coastal Water Research Project Authority, Costa Mesa, CA, 72 pp.

SWAMP Quality Assurance Team. 2008. Surface Water Ambient Monitoring Program Quality Assurance Program Plan, Version 1.0, September 1, 2008.

U.S. Environmental Protection Agency (USEPA). 2007. Validation of the Fish Short-Term Reproduction Assay: Integrated Summary Report, December 2007, accessed 9/29/2014, at http://www.epa.gov/endo/pubs/att-f_fish_assay_protocol.pdf.

USEPA. 2013. Endocrine Disruptor Screening Program. Sap Review of EDSP Tier 1 Screening: Assay and Battery Performance, May 21-23, 2013.

WateReuse Research Foundation (WRRF). 2014. Development of Bioanalytical Techniques to Assess the Potential Human Health Impacts of Recycled Water, WRRF, Arlington, VA, 316 pp.

U.S. Geological Survey (USGS). 2004. Techniques of Water-Resources Investigations, book 9, chapter A5, Wilde, F.D., Radtke, D.B., Gibs, Jacob, Iwatsubo, R.T., eds., Processing of water samples (ver. 2.2), April 2004, accessed 9/3/2014, at http://pubs.water.usgs.gov/twri9A5/.

ABBREVIATIONS

AhR – aryl hydrocarbon receptor
AR – androgen receptor
BEQ – bioassay equivalent concentration
CEC - constituents of emerging concern
CEDEN - California Environmental Data Exchange Network
DMT – data management team
ER – estrogen receptor
GC/MS – gas chromatography/mass spectrometry
GR – glucocorticoid receptor
GSI – gonadosomatic index
IVB – in vitro bioassay
LC-MS/MS – liquid chromatography/tandem mass spectrometry
LCS – laboratory control sample
MDL – method detection limit
MQO – measurement quality objective
MS/MSD – matrix spike/matrix spike duplicate
MS4 – municipal separate stormwater sewer system
MTL – monitoring trigger level
OECD – Organization for Economic Cooperation and Development
OSHA – Occupational Health and Safety Administration
PBDE – polybrominated diphenyl ether
PFC – perfluorinated compound
PR – progesterone receptor
QA – quality assurance
QC – quality control
QAPP – Quality Assurance Project Plan
RL – reporting limit
RPD – relative percent difference
RSD – relative standard deviation
SCCWRP – Southern California Coastal Water Project Authority
SOP – standard operating procedure
SRM – standard reference material
SWAMP – Surface Water Ambient Monitoring Program
SWRCB – State Water Resources Control Board
TBD - to be determined
TBR - to be resolved
Vtg – vitellogenin
WWTP – wastewater treatment plant