Baseline Testing for *Ceriodaphnia dubia* Toxicity Testing Laboratory Standardization

Study Plan and Logistics

June 13, 2022

Second Draft

TABLE OF CONTENT

1. BACKGROUND

The *Ceriodaphnia dubia (C. dubia)* chronic reproduction toxicity test is an established whole effluent toxicity (WET) test method (U.S. EPA 2002a, b, c; U.S. EPA 2016), commonly used in regulatory programs including the Toxicity Provisions recently adopted by the State of California. However, regulators and stakeholders have recognized that some laboratories may need to improve their implementation of the *C. dubia* method to reduce intra-laboratory (within-laboratory) variability and increase inter-laboratory (amongst-laboratory) comparability. The present study commissioned by the California State Water Resources Control Board, in collaboration with stakeholders and laboratories, aims to (1) evaluate laboratory performance among those accredited by the State of California Environmental Laboratory Accreditation Program, (2) investigate factors that can lead to intra-and inter-laboratory test variability and decrease confidence in assessments of toxicity, and (3) provide specific laboratory techniques guidance to improve laboratory performance and reduce intra- and inter-variability.

To standardize test methods and parameters that may contribute to intra- and inter-laboratory variability, the Expert Science Panel and Stakeholder Advisory Committee have recommended that an intercomparison exercise be conducted by all California-accredited laboratories.

Two key questions were identified:

- Which laboratory technique(s) should be standardized to reduce intra- and inter-laboratory variability?
- Does standardizing laboratory techniques improve consistency and comparability in *C. dubia* test results?

2. GENERAL APPROACH

To address these questions, a three-step approach was proposed. During Step 1, all laboratories will participate in an intercomparison exercise using their current protocols and provide additional data that may not be routinely collected/reported by all laboratories. Based on the results of Step 1 and discussions among the Expert Science Panel and the Stakeholder Advisory Committee, Step 2 will aim to standardize select *C. dubia* test parameters. Finally, Step 3 will consist of another intercomparison exercise amongst all laboratories using split samples and the standardized *C. dubia* toxicity testing protocol developed from step 2. This document describes the approach, overall methodology and logistics that will be used to conduct Step 1 baseline testing intercomparison exercise. Detailed description of the subsequent steps will depend on the analyses and group discussions of the results of Step 1 baseline testing.

The document below aims to describe the key elements and steps for the baseline study. A separate quality assurance project plan (QAPP) will be produced to provide all the detailed instructions for the laboratories.

3. BASELINE TESTING PROCEDURE

The specific objective of the baseline testing is to collect additional test samples data and a more complete/consistent lab technique dataset across all California-accredited laboratories. The laboratories will participate in an intercomparison exercise consisting of several split water samples tested in multiple batches. Three testing options (Table 1) are proposed to ensure that a minimum of seven (7) control

datasets are generated by each laboratory. This sample size was determined based on a power analysis conducted by the project biostatistician. This was determined using a statistical approach to assess the width of the confidence interval for the mean control neonate production for different sample size. More detailed information is provided in Appendix A.

Testing option #1: Three separate test batches will be analyzed, with each batch including: (a) one unspiked sample, consisting of dilution water to be tested at full strength (i.e., 100%); (b) one spiked sample, using sodium chloride, to be tested at 5 different concentrations, and (c) one duplicate sample to be determined by SCCWRP. A total of 315 split-samples will be prepared and tested. Using this testing option, each laboratory will generate twelve laboratory control datasets and three reference toxicant datasets, as well as three unspiked, three spiked, and three duplicate sample datasets per laboratory.

Testing option #2: Two separate test batches will be analyzed, with each batch including: (a) three unspiked samples, consisting of different dilution water recipes to be tested at full strength; (b) one spiked sample, using sodium chloride, to be tested at 5 different concentrations, and (c) one duplicate sample, to be determined by SCCWRP. A total of 270 split-samples will be prepared and tested. Using this testing option, each laboratory will generate twelve test control datasets and two reference toxicant datasets as well as six unspiked, two spiked and two duplicate sample datasets, and per laboratory.

Testing option #3: Two separate test batches will be analyzed, with each batch including: (a) two unspiked sample, consisting of dilution water to be tested at full strength (i.e., 100%); (b) one spiked sample, using sodium chloride, to be tested at 5 different concentrations, and (c) one duplicate sample to be determined by SCCWRP. A total of 240 split-samples will be tested. Using this testing option, each laboratory will generate ten laboratory control datasets and three reference toxicant datasets, as well as four unspiked, two spiked and three duplicate sample datasets.

Standard Operating Procedures

Participating laboratories will perform three test batches within a 4 to 6-week window, depending on the testing option chosen, using their own standard operating procedures for the *C. dubia* chronic toxicity test. A summary of standard operating procedures, test acceptability criteria and measurement expectations are provided in Table 2. In addition, all laboratories will be required to do the following specifications:

- All tests will be carried out to 8 days (i.e., 192 hours).
- Each sample will be tested with a separate laboratory control using their standard control dilution water.
- A concurrent reference toxicant will be run with each test batch using the laboratory's own reference chemical.
- Each sample/dilution will be tested using 10 replicate chambers.
- Test set-up will be randomized using blocking by known parentage.

Additionally, participating laboratories will be required to report data that may not be currently documented/reported including:

- Number of males, unhealthy and dead adults, and dead neonates in the brood board
- Specific beginning and end time window for age of neonates at test initiation
- Renew test solutions daily within a 24 +/- 1 hour window to enhance comparability of neonate counts among laboratories. Specific time of renewal (hours and minutes) shall be recorded
- Water quality parameters (temperature, pH, DO, conductivity) at test initiation, termination and before and after daily renewal, measured in surrogate test chambers
- Sub-samples for ionic composition at test initiation that must be shipped to SCCWRP on day 0. SCCWRP will ship all samples to the same laboratory for analysis.
- Subsample of laboratory reference toxicant stock solution to be shipped to SCCWRP. SCCWRP will ship all samples to the same laboratory for analysis. This will ensure that data on measured concentrations (not just nominal) of the stock are documented for all laboratories.
- Light intensity and air temperature within the testing area at the time of the experiments

Sample Preparation and Distribution

All split-samples will be prepared in the SCCWRP laboratories using large sample containers and thoroughly mixed on a large-capacity stirrer to ensure that the samples are homogenous. Samples will be allowed to equilibrate for up to 48 hours prior to subsampling. Subsampling will be conducted while continuously mixing the samples. Subsample cubitainers (volume TBD (L)) per sample per laboratory) will be filled using a peristaltic pump and pre-cleaned (inside and outside) sampling hose kept in constant motion within the large sample container. The laboratory technician responsible for handling the sampling hose will ensure that the hose remains between 30 and 80 percent of the depth of the water column and does not touch the bottom of the water container. All samples will be kept in the walk-in fridge at 4 ℃ up to [TBD] days before shipping them to the participating laboratories.

To ensure that all subsamples are representative of the original test samples, each cubitainer will be subsampled to measure conductivity, alkalinity and hardness in triplicate. A subset will be sent for ion composition analyses and another subset of samples will be archived in the SCCWRP laboratories.

A total of xx split-samples will be shipped to each laboratory every other week starting mid-July according to the schedule agreed upon with the participating laboratories. Samples will be shipped on wet ice using priority overnight (OnTrac or FedEx) service to the laboratories to the addresses in Table 3. The shipments will also include chain-of-custody (COC) forms completed by SCCWRP and a copy of the study plan and testing instructions. SCCWRP will notify the laboratories via email once the samples are in transit and provide a tracking number. It is the responsibility of the laboratories to contact SCCWRP if they have not received the samples by the following day 2:00 pm.

Upon delivery, temperature, conductivity, hardness, alkalinity, temperature, pH and dissolved oxygen must be measured and recorded for each sample to verify stability of the sample before testing is initiated. The cubitainers must be kept at 4℃ up to 48 hours before first use.

Data Submission

SCCWRP will provide an Excel data submittal form and culture/bench sheet templates to the participating laboratories. All test data in electronic format and scanned copies of the culture/bench sheets must be submitted to the SCCWRP data portal no later than [date TBD]. Data required include:

- Laboratory information
- Sample information upon receipt
- Testing conditions including dilution water and food recipe
- Brood board health data
- Bench water quality, survival and reproduction counts
- Control charts for reference toxicant tests for the last 12 months

4. COMMUNICATION AND SCHEDULE

Coordination with Participating Laboratories

Participating laboratories and other stakeholders will meet with SCCWRP and the Expert Science Panel advising on this project to finalize the study plan, discuss logistics and review the results. A minimum of three remote meetings will be scheduled to provide a forum for discussion and clear communication among the project team and participants. Additional communication via email will be encouraged throughout the study. For more information on the overall study design and coordination meetings, please contact Alvina Mehinto alvinam@sccwrp.org. For questions regarding samples shipping from and to SCCWRP and data submission, please contact Darrin Greenstein darring@sccwrp.org.

The first meeting, held remotely on May 24, 2022, and attended by the stakeholders aimed to review the first draft of the testing approach (including sample preparation and shipping, test measurements and data reporting) and discuss the timeline for testing and data submission. The second meeting to be held on June 24, 2022 among members of the Expert Science Panel, stakeholders and laboratories will aim to finalize the study design, review the QAPP and the logistics. The third meeting will focus on providing training for data collection and data submission.

Schedule

- **May 17:** Draft study plan sent to all stakeholders for review
- **May 24:** Stakeholder Committee meeting, held via Zoom, to discuss the first draft of the study plan
- **June 14:** SCCWRP will send the revised study plan to the Expert Science Panel
- June 24: Public meeting with Expert Science Panel and participating laboratories to finalize the study plan and approve the QAPP
- July date TBD: Meeting with participating laboratories to provide training on data collection and submission

- July dates TBD: First batch of split samples prepared by SCCWRP
- July date TBD: Cubitainers containing first batch of split samples shipped to the laboratories.
- Dates TBD: First batch of *C. dubia* toxicity tests performed
- July dates TBD: Second batch of split samples prepared by SCCWRP
- Dates TBD: Cubitainers containing second batch of split samples shipped to the laboratories.
- Dates TBD: Second batch of *C. dubia* toxicity tests performed
- August dates TBD: Third batch of split samples prepared by SCCWRP
- August date TBD: Cubitainers containing third batch of split samples shipped to the laboratories.
- August dates TBD: Third batch of *C. dubia* toxicity tests performed
- August date TBD: Deadline for data submission

5. CONTINGENCIES

Lost Sample

If a sample is not delivered to a laboratory on the expected arrival date or if the sample has spilled during shipment, the laboratory must contact SCCWRP promptly. SCCWRP will ship new cubitainers that same day. However, this second batch of samples sent must be tested within 24 hrs to ensure that holding times are comparable to other laboratories.

Failed Test Acceptability Criteria

A laboratory will be given the opportunity to retest up to two test batches if acceptability criteria are not met. A laboratory planning to retest must contact SCCWRP within 24 hrs of knowing that a test failed the acceptability criteria. Laboratories are encouraged to retest with remaining sample; however, arrangements might be made to re-test with archived samples. Laboratories that fail to provide data for Step 1 baseline testing may still be considered to participate in the confirmation testing in Step 3.

Late Data Submission

All data must be submitted to the SCCWRP data portal and pass the QA checkers by [date TBD]. If a laboratory experiences some delays, SCCWRP must be contacted no later than 48 hours before the deadline. Laboratories will be granted an additional three (3) days to submit all their data. Past this new deadline, SCCWRP cannot guarantee that the data will be used in subsequent data analyses.

6. REFERENCES

U.S. EPA. 2002a. Short-term methods for estimating the chronic toxicity of effluents and receiving water to freshwater organisms. EPA-821-R02-013. U.S. Environmental Protection Agency. Washington DC.

U.S. EPA. 2002b. Guidelines establishing test procedures for the analysis of pollutants: Whole effluent toxicity test methods; Final rule. Environmental Protection Agency. 40 CFR Part l36 [FRL-7408-6].

U.S. EPA. 2002c. Guidelines establishing test procedures for the analysis of pollutants: Whole effluent toxicity test methods; Final rule. 67 Fed. Reg. 69952-69972 (November 19, 2002).

U.S. EPA. 2016. Whole effluent toxicity methods errata sheet. Office of Water, Environmental Protection Agency. 821-R-02-012-ES.

Table 1. Total number of spit-samples based on testing option. Note that each test batch will include a laboratory control and reference toxicant.

Test sample type	**No. of test batches**	**No. of test sample**	**No. of dilutions per test sample**	**No. participating labs**	**Total no. test samples to prepare**
Testing option #1					
Unspiked	3	1	1	15	315
Spiked		1	5		
Duplicate		1	1		
Testing option #2					
Unspiked	2	3	1	15	270
Spiked		1	5		
Duplicate		1	1		
Testing option #3					
Unspiked	2	2	1	15	240
Spiked		1	5		
Duplicate		1	1		

Table 2. Summary of test conditions and acceptability criteria for the *Ceriodaphnia dubia* survival and reproduction test.

Parameter[1]	**Description**
Test organism	*Ceriodaphnia dubia*
Protocol	EPA/821/R-02-013, EPA 2002 Chronic Manual
Exposure	Static, daily renewal
No. replicate test chambers	10 replicates per sample/dilution
Sample holding time[2]	Up to 48 hrs before test initiation
Test duration	8 days, i.e., 192 hours
Endpoints	Survival and reproduction
Laboratory control	One laboratory dilution water control per test sample
Reference toxicant	One laboratory reference toxicant per test batch, tested in a serial dilution plus separate control
Water quality measurements	Daily: temperature in℃ and pH reported with 0.1 precision; conductivity in µS/cm; dissolved oxygen in mg/L Upon receipt and test termination: hardness and alkalinity in mg/L $CaCo_3$ Once during test in testing area: light intensity in foot-candles; air temperature in ℃ (0.1 ℃ precision)
Test Acceptability Criteria	80% or greater survival and an average of 15 or more neonates per surviving female in the controls

[1] Parameters and test conditions used in this study are suitable for investigative/non-compliance testing but not suitable for NPDES permit testing.

[2] This is a deviation from the promulgated method.

Table 3. Laboratory contact information and shipping address. **DRAFT TABLE BASED ON LABS THAT HAVE PROVIDED HISTORICAL DATA. NO LABS HAVE FORMALLY COMMITTED TO PARTICIPATING YET**.

Laboratory	Name	Contact information	Shipping Address
49er	Shane Burr	209-418-3175 Shane@49erwaterlab.com	245 New York Ranch Rd, Ste A, Jackson, CA 95642
Aquatic Bioassay & Consulting Labs, Inc	Joe Freas	805-643-5621 x18 joe@aquaticbioassay.com	29 N Olive St., Ventura, CA 93001
Aquatic Testing Laboratory	Joe LeMay	805 650-0546 jlemay12@pacbell.net	4350 Transport Street, Unit 107 Ventura, CA 93003
~~Aquatic Toxicity Lab (UCD)~~	~~Marie Stillway~~	~~530-752-0772~~ ~~Mstillway@ucdavis.edu~~	~~UC Davis AHP. Institute of Ecology CABA, Bldg. #5. Garrod Road West, Davis, CA 95616~~
AquaScience	Kimberly Miller	530-753-5456 Kimberley@aqua-science.com	630 Cantrill Dr., Davis, CA 95618
City of Los Angeles	Stacee Karnya	310-648-5923 stacee.karnya@lacity.org	12000 Vista del Mar, Playa del Rey, CA 90293
EcoAnalysts	Brian Hester	360-297-6040 x6045 bhester@ecoanalysts.com	4770 NE View Dr., Port Gamble, WA 98364
Enthalpy	Peter Arth	858-587-7333 ext. 214 Peter.arth@enthalpy.com	4340 Vandever Avenue, San Diego, CA 92120
GEI*	Natalie Love	303-264-1070 Nlove@geiconsultants.com	4601 DTC Boulevard, Suite 900, Denver, CO, 80237
Inland Empire	Sushmitha Reddy	909-993-1813 Sreddy@ieua.org	Water Quality Laboratory, Building C, 6075 Kimball Ave., Chino 91708

McCampbell	Drew Gantner	925-252-9262 Drew.gantner@mccampbell.com	1534 Willow Pass Road Pittsburg, CA 94565-1701
MBC Applied Environmental Sciences	Sonja Beck	714-850-4830 x225 Smbeck@mbcaquatic.com	MBC is 3000 Redhill Ave., Costa Mesa CA, 92626
Pacific Ecorisk	Stephen Clark	707-207-7760 Slclark@pacificecorisk.com	2250 Cordelia Road. Fairfield, CA 94534
Sanitation Districts of Los Angeles County	Josh Westfall	562-908-4288 x2815 Jwestfall@lacsd.org	San Jose Creek Biology Lab. 1965 Workman Mill Rd. Whittier, CA 90601
TetraTech*	Marcus Bowersox	410-902-3142 Marcus.Bowersox@tetratech.com	10711 Red Run Blvd., Suite 105, Owings Mills, MD 21117
Wood	Steve Carlson	858-299-5368 Steve.carlson@woodplc.com	4905 Morena Blvd. Ste. 1304, San Diego, CA 92117

* Participating laboratory is no longer accredited in California.

7. APPENDICES

Appendix A – Statistical Approach for Sample Size Determination

By Dr Jing Zhang (5/4/2022)

The objective of this analysis to find the number of tests needed in a lab in order for the standard error of mean number of neonates per female in the control group to achieve a desired level.

Assumptions made here include:

1. The standard error we need to consider here are from the observed data of labs. Labs with coefficient of variation below 0.10 or above 0.20 are excluded in the consideration here. They are Labs B, C, D, J, K, M and N.

2. This sample size calculation is assumming normality of the mean neonate numbers of each test.

3. Since the objective is to find out the additional number of tests needed in order to obtain reliable mean neonates from the control group, i.e. it is of interest to find the sample size needed to achieve certain prevision level in estimation, we will not pursue the power and size sample size determination approach, which aims at finding the sample size that achieves a desired power for a given size (significance level) in hypothesis testing. The sample size calculation here aims at finding the minimum sample size that will achieve the desired width of the CI of mean neonates (or equivalently, certain desired level of the standard error of the mean neonates), rather than certain level power in the testing of hypothesis that the mean neonates is equal to a given number.

4. The desired width of the confidence interval, is chosen based on the mean neonates reported by all the labs that are considered, more specifically, it is chosen to be the inter quartile range of the mean neonates per female across all the different labs, which is roughly 6.69.

The first thing we try is to select the labs that are associated with a value of the coefficient of variation below 0.10 or above 0.20 for the mean neonates. Here is the list of selected labs and their relevant summary of control group information.

```
##      lab mean_test_length num_of_rep_NValue mean_num_of_neo_per_female
## 1  LAB-A         6.506329               79                  29.48101
## 2  LAB-E         7.000000               79                  24.95570
## 3  LAB-F         6.453333               75                  25.75467
## 4  LAB-G         6.724138               29                  23.82069
## 5  LAB-H         6.176471               17                  18.89412
## 6  LAB-I         6.083333               60                  31.84833
## 7  LAB-L         7.035088               57                  25.61053
## 8  LAB-O         6.166667               60                  35.88667
## 9  LAB-P         6.287037              108                  33.10556
## 10 LAB-Q         6.062500               48                  24.43958
```

```
##    sd_num_of_neo_per_female cv_num_of_neo_per_female
## 1                  5.278138                0.1790352
## 2                  3.819375                0.1530462
## 3                  4.357451                0.1691907
## 4                  3.997533                0.1678177
## 5                  2.277738                0.1205527
## 6                  5.248518                0.1647973
## 7                  4.627391                0.1806832
## 8                  6.200204                0.1727718
## 9                  4.542527                0.1372134
## 10                 4.468792                0.1828506
```

After that we compute three standard deviation estimates for the population of mean neonates:

1. minimum standard deviation among all the remaining 10 labs
2. maximum standard deviation among all the remaining 10 labs
3. pooled standard deviation across all the remaining 10 labs

pooled_sd	min_sd	max_sd
4.697968	2.277738	6.200204

Note that these information were obtained from tests that are mainly using 10 replicates, what if we increase the number of replicates? This will reduce the population standard deviation of the mean neonates per female in the control group for each test, since the standard deviation of the sample mean of neonates in each test is $\frac{\sigma^2}{n_{rep}}$, where σ^2 is the variance of number of neonates each female has produced in the control group and n_{rep} is the number of replicates used in each test.

We can use this information to get three rough estimates of the standard deviation estimates for the population of mean neonates when the number of replicates are different. We are interested in the 95% CI of control mean neonates, so a t-confidence interval can be used, with the standard deviations adjusted according to the number of replicates used in a given experiment.

Here the desired width of the confidence interval of the mean neonates in the same lab is chosen as 10. Combining the desired width with the adjusted estimates of the standard deviations of the control sample mean neonates, we could obtain the desired number of tests when different number of replicates are used, using the “precision()” in “emon” package.

Here are the number of tests needed to achieve a width of 10 for the 95% CI of mean neonates, when different number of replicates are used in each test, assuming three varying levels of standard deviations of the sample mean neonates: pooled standard deviation, minimum standard deviation among all the remaining 10 labs, and maximum standard deviation among all the remaining 10 labs.

Number.reps.per.test	Sample.size.pooled.SD	Sample.size.optimistic	Sample.size.conservative
10	6	4	9
11	6	4	8
12	6	4	8
13	6	3	8
14	5	3	7
15	5	3	7
16	5	3	7
17	5	3	6
18	5	3	6
19	5	3	6
20	5	3	6

We can see here when we assume the pooled standard deviation and a desired width of 10 for the 95% CI of the mean neonates, we need 6 new tests at the lab, each with 10 replicates or more.

Alternatively, we can compare the width of the Wald CI for the population of mean neonates when different numbers of replicates are considered. Here is the plot of width of the CI vs. number of tests when 10 replicates are used in each test: