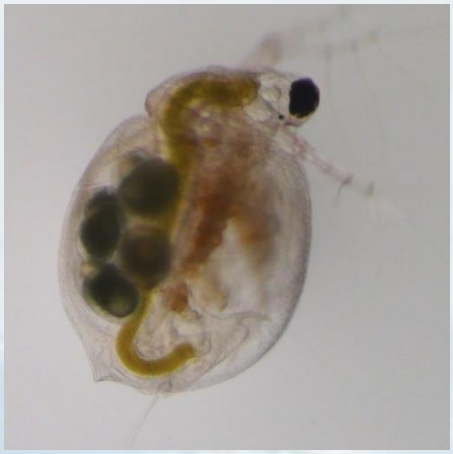
Ceriodaphnia dubia Quality Assurance Guidance Recommendations



Robert Brent

Howard Bailey

Teresa Norberg-King

Leana Van der Vliet

A. John Bailer

SOUTHERN CALIFORNIA COASTAL WATER RESEARCH PROJECT

Technical Report 1341

Ceriodaphnia dubia Quality Assurance Guidance Recommendations

Robert Brent[1], Howard Bailey[2], Teresa Norberg-King[3], Leana Van der Vliet[4], and A. John Bailer[5]

[1]*James Madison University, Harrisonburg, VA*

[2]*Nautilus Environmental, Burnaby, British Columbia, Canada*

[3]*Formerly U.S. Environmental Protection Agency, Duluth, MN*

[4]*Environment and Climate Change Canada, Gatineau, Quebec, Canada*

[5]*Miami University, Oxford, OH*

September 2023

Technical Report 1341

Preface

This is a guidance manual with recommendations for improving the quality assurance of the Ceriodaphnia dubia survival and reproduction test. This document does not promulgate a new method or a formal change to an existing method. Instead, the recommendations are supported by a two and a half-year study analyzing data from current and formerly accredited toxicity testing laboratories for the State of California. This project was facilitated by the Southern California Coastal Water Research Project (SCCWRP) under contract to the State Water Resources Control Board (SWRCB) and the California Association of Sanitation Agencies (CASA). The project governance included a decision-making body, the Expert Science Panel, a five-member team of North American (non-Californian) experts familiar with performing *C. dubia* toxicity tests and representing the fields of aquatic toxicology and chemistry, statistics, and quality assurance programs related to environmental testing laboratories. The Expert Science Panel was assisted by a Stakeholder Advisory Committee representing 12 different sectors who utilize the *C. dubia* toxicity test for environmental management decision making, which includes a variety of regulated dischargers, state and federal regulators, non-governmental organizations, and toxicity testing laboratories. All laboratory-specific data collection and analyses are anonymous as a condition of their participation in the study.

Members of the Expert Science Panel:

Robert Brent (James Madison University), Expert in Freshwater WET Testing

Howard Bailey (Nautilus Environmental), Expert in WET Testing

Teresa Norberg-King (Formerly US EPA), Expert in Freshwater WET Testing

Leana Van der Vliet (Environment and Climate Change Canada), Expert in quality assurance in single-species toxicity tests

A. John Bailer (Professor Emeritus, Miami University), Expert in methods for designing and analyzing environmental toxicology studies

Members of the Stakeholder Advisory Committee:

Katie Fong, Representative for State Water Resources Control Board

Steven Boggs, Representative for California Environmental Laboratory Accreditation Program

Amelia Whitson (EPA Region IX), Representative for Federal Government

Rochelle Cameron (EPA Region IX), Alternate representative for Federal Government

Veronica Cuevas (Los Angeles Regional Water Quality Control Board), Representative for NPDES permits

Mitch Mysliwiec (LWA), Representative for California Associations of Sanitation Agencies (CASA)

Paul Bedore (Robertson-Bryan, Inc.), Alternate representative for CASA

Jian Peng (Orange County Public Works), Representative for California Stormwater Quality Association

Sarah Lopez (Central Coast Water Quality Preservation, Inc.), Representative for agricultural associations

Peter Arth (Enthalpy), Representative for private toxicity testing laboratories

Josh Westfall (Sanitation Districts of Los Angeles), Representative for public toxicity testing laboratories.

Annelisa Moe (Heal the Bay), Representative for Non-Governmental Organizations

John Wheeler, Former representative for State Water Resources Control Board

Debra Denton (EPA Region IX), Former representative for Federal Government

Jeff Miller (Aqua-Science), Former representative for private toxicity testing laboratories

Kaitlyn Kalua (California Coastkeeper Alliance), Former representative for Non-Governmental Organizations

SCCWRP Project Team

Alvina Mehinto, Toxicologist, Co-Principal Investigator

Ken Schiff, Toxicologist, and Co-Principal Investigator

Darrin Greenstein, Quality Assurance Officer

David Gillett, Biostatistician

1. Executive Summary

The State of California (State) recently promulgated Toxicity Provisions for regulated dischargers (SWRCB 2020), part of which requires testing for the Ceriodaphnia dubia (*C. dubia*) survival and reproduction toxicity test (Method 1002.0, EPA 2002a,b,c), a test with decades of use since EPA ratified the approval of several test procedures for measuring the toxicity of effluents and receiving waters based on robust scientific evidence. Members of the regulated community expressed concerns about variability associated with the *C. dubia* test method during the Toxicity Provision adoption process. While variability is an inherent property of test systems, high levels of variability could impair the usefulness of the *C. dubia* survival and reproduction test data for monitoring water quality and assessing compliance with discharge limits. Therefore, this project focused on evaluating and possibly reducing sources of variability in control samples and reference toxicants as a means to improve consistency and comparability of *C. dubia* toxicity testing results. This document summarizes the results of a two-and-a-half-year investigation into the level and sources of variability associated with the C. *dubia* survival and reproduction aquatic toxicity test. It also includes a series of recommendations to the state regulatory agency, the regulated parties, and the toxicity testing laboratories for maintaining and, in some cases, improving data quality for the *C. dubia* survival and reproduction toxicity test.

There are many potential sources of variability impacting toxicity testing results. For the *C. dubia* survival and reproduction toxicity test, these may include inherent differences in the response of individual test organisms, testing conditions (e.g., temperature, water chemistry and diet) that can alter organism condition and toxicant bioavailability, health of the brood stock from which neonates are obtained, the skill and experience of staff performing laboratory techniques, and data quality evaluation and data analysis. To investigate the sources of variability in *C. dubia* test results, the following activities were performed:

- Inventory of laboratory techniques used by accredited laboratories
- Compilation and statistical analysis of historical data
- Baseline laboratory interlaboratory comparison study
- Laboratory visits and roundtable workshop
- Second laboratory interlaboratory comparison study

Inventory of laboratory techniques and historical data from accredited laboratories

At the start of this project, 18 laboratories were accredited by the State. Three of the accredited laboratories were public agencies, 12 were private laboratories, and two were academic laboratories. One academic laboratory did not have sufficient historical data to participate in the project. Therefore, data and laboratory documentation were collected from 17 laboratories. The inventory of laboratory techniques was created by 1) reviewing each laboratory's Standard Operating Procedures (SOPs) and Quality Assurance Plans (QAPs), 2) electronic questionnaires sent to each laboratory, and 3) one-on-one phone interviews with key laboratory staff.

The evaluation of the laboratory documentation revealed that no two laboratories performed the *C. dubia* test in exactly the same manner. While most method differences were within the acceptable range of laboratory techniques described in the method, a third of the laboratories had at least one or two techniques that was not described or in conflict with the EPA method. For example, when evaluating how laboratories create their dilution water, 12 out of 17 laboratories reported using one of the recipes in the EPA manual. Five laboratories used diluted mineral Perrier® or Evian® water and seven used reconstituted moderately hard water (MHW) or hard water as described in the EPA method manual (EPA, 2002a). The others use a modified version of the MHW recipe or a completely different recipe that is not described in the EPA method. Lack of similarity could be observed for other laboratory techniques such as food and feeding, test set up (e.g., test chamber material and volume, light intensity, and photoperiod, etc.), trigger and method to terminate the test, culturing and brood boards, and reference toxicants. The most challenging element of evaluating sources of variability was unrecorded or missing data and the lack of standard forms to document laboratory techniques. For example, assessment of brood board health was implemented very differently among laboratories. Some used a detailed quantitative approach while others used a qualitative approach.

In total, laboratories voluntarily provided control data for 551 environmental sample tests along with 452 reference toxicant tests from the last three to six years (2016-2021). A comprehensive database was created to summarize (by test and by laboratory) key metrics, including percent survival, mean number of young produced by the controls, relative measures of variability including the coefficient of variation (CV) for mean number of young produced in the controls, inhibition concentration (IC) associated with the 25% or 50% reduction in reproductive output, IC25 and IC50, and the percent minimum significant difference (PMSD). Water quality data and available brood board health data were also documented for each test.

Evaluation of laboratory performance

To characterize and evaluate laboratory performance over time, the Panel focused on three categories of performance metrics, biological metrics for survival and reproduction, variability in reproduction (CV) and percent minimum significant difference (PMSD, that reflects test uncertainty), and toxicity potency (IC). The Panel proposed multiple thresholds (based on target value and frequency of meeting such target) for each criterion to assess laboratory performance meeting, exceeding or below the Panel's expectations. Test acceptability criteria were used as the "thresholds" (referred to as meeting expectations) for the biological metrics. Relative variability and uncertainty metrics (CV and PMSD) were derived based on analysis of the data generated in this project and compared to those in the EPA (2000a and 2001a, b) and Fox et al. (2019) studies. Toxicity potency metrics were assessed using a percentile approach (10th, 25th, 75th and 90th percentile). In addition to the threshold values, the Panel conducted preliminary investigations of acceptable frequency of occurrences that do not penalize laboratories occasionally producing marginal data due to documented sub-optimal conditions (e.g., culture or brood board health issues, sample condition, etc.).

These metrics were applied to the datasets as an example of laboratory performance evaluation; but the Panel does not present them as final recommendations. Additional work and vetting of the metrics including defining thresholds for acceptable performance and the importance of each metric is recommended. This would help refine some values, the desired frequency of attainment and investigate other metrics (e.g., IC25, ratio IC25:IC50) to ensure that the approach described in this report is valuable to all parties involved.

Appendix E Table E1 provides further details on the performance metrics described above and applied to historical data and empirical data from two interlaboratory comparison studies (ILS). There is no approved precedent or standard for frequency of occurrence and the values utilized in this example are largely based on best professional judgement.

Analysis of historical data

Most California-accredited laboratories for the *C. dubia* test consistently met the TAC for survival and reproduction. Only two laboratories had relatively low mean reproduction (~15 neonates per surviving females). Laboratory-specific mean of control reproduction, however, varied among laboratories from 18.7 to 37.5 neonates per female, and test control reproduction ranged from <5 to >50 neonates per female. Consequently, CV for neonate production varied appreciably for 9 out of 17 laboratories who reported a mean CV >0.2 in over 50% of their tests. Analysis of the PMSD showed that 8 of the 17 laboratories reported PMSD ≤25 in over 75% of their historical reference toxicant tests, and 7 laboratories had PMSD ≤25 in over 50% of their tests. The observed

range of estimated IC50s varied by more than two-fold among laboratories for sodium chloride (NaCl) or copper chloride (CuCl). Eleven laboratories had over 50% of their IC50s within the 25th and 75th percentile (calculated using historical dataset for valid tests). It should be noted that two of them had over 75% of their IC50s within the 25th and 75th percentile.

In conclusion, over half of laboratories (11 laboratories) consistently met the proposed performance metrics, with four of them exceeding expectations set by the Panel. This suggests that the discharge community has access to qualified testing laboratories capable of performing the *C. dubia* test. However, the range of variability observed in CV of reproduction and IC50s suggested that "standard" practices are not being applied consistently across or within laboratories. This is surprising given the availability of relevant guidance documents, and a decades-old State test accreditation program. Under these conditions, one might reasonably expect that most, if not all, of the accredited laboratories would exhibit performance metrics consistent with regulatory guidance.

Given the range of variability in test performance metrics, the potential linkages between test performance and specific laboratory practices were evaluated. To accomplish this, a questionnaire was sent to each laboratory to request specific information on testing and culturing practices that might affect organism condition and test results. The questionnaires were followed up by phone or video interviews to collect additional information. All data were then analyzed to identify potential variables that might help explain variability in the test metrics. Ultimately, no single or combination of laboratory practices appeared related to variability in control neonate production, variation in control neonate production, or reference toxicant endpoints. Descriptive statistics, multiple linear regression techniques, nor multivariable classification techniques (e.g., random forest) could identify or suggest plausible explanations for differences between laboratories. One possible reason is that unique combinations of laboratory techniques that change over time are responsible for intra-and inter-variability. It should also be noted that each laboratory had a one-of-a-kind profile of experimental practice limiting the ability of the analysis to identify factors driving the result. Finally, it is reasonable to assume that these laboratories' characteristics are highly related and that identifying a single driving factor, or a few factors was not likely to occur.

Baseline interlaboratory comparison study

Because the historical data analyses did not reveal specific laboratory practices that could explain intra- and inter-variability among laboratories, an empirical approach was taken whereby selected sources of variability were controlled during an interlaboratory comparison study (ILS). This baseline ILS consisted of 11 participating laboratories that tested two types of moderately hard dilution water (MHW) and a NaCl-spiked dilution series. The study was comprised of three

separate testing events. The laboratories were asked to utilize their ongoing standard operating procedures (as used for the tests submitted during the historical data compilation), but test procedures were documented in greater detail.

The baseline ILS generated empirical data from 178 unspiked samples and 60 NaCl-spiked samples. The analysis provided similar results as the historical data analysis. Laboratory control mean reproduction ranged from 14.9 to 40.2 neonates per female in valid tests. Five of 11 laboratories met or exceeded the proposed metrics for biological response, test variability and uncertainty, and potency endpoint. Two laboratories met most metrics but exhibited relatively high variability in reproduction and IC50 estimate. The remaining four laboratories had low reproduction CV and PMSD. For these laboratories, the IC50s varied and were not comparable to the rest of the participating laboratories.

Further investigations on the test procedures used by the different laboratories showed that no single (or combination of) factor(s) appeared to be related to the variability observed among test results. However, statistical analysis identified a range of factors that did not seem responsible for the differences between laboratories (including water chemistry parameters such as ion composition, hardness, alkalinity). Thus, these parameters were not prioritized in subsequent tasks. The only factor that appeared related to neonate production in unspiked samples was the age of the female used to generate neonates for test set up; analyses revealed that test organisms obtained from older females tended to produce fewer neonates. However, these test organisms still produced sufficient neonates to pass test acceptability criteria, and this variable did not appear to affect the CV associated with neonate production. Consequently, both the experts and stakeholders expressed caution in over-interpreting these results.

Laboratory visits and roundtable workshop

The lack of identifiable factors that would explain the variability observed within and among laboratories suggested that underlying factors might be diverse, laboratory-specific, intermittent or inconsistent, and not readily captured by data typically collected. Consequently, two members of the Expert Science Panel visited a subset of the laboratories representing a range of variability in test performance metrics to identify potential contributing factors. During laboratory visits, Panel members identified a number of practices that were inconsistent with the method and general laboratory quality control practices. This reinforced the assumption that variability was primarily driven by laboratory-specific factors. Laboratory visits were followed by a workshop attended by the Panel and 12 California accredited laboratories who participated in the baseline ILS study. The two-day roundtable workshop covered 20 topics in four areas of testing: culturing, food and feeding, testing, and documentation and recordkeeping. The laboratories actively

engaged with the Panel and both parties agreed on a list of 16 items to standardize and apply in a second interlaboratory comparison study.

Second interlaboratory comparison study

The findings from the laboratory visits and the recommendations from the roundtable workshop were incorporated into a second interlaboratory study to determine if controlling identified potential causal factors would result in a commensurate reduction in variability of test performance metrics. Ten laboratories participated in the second ILS (1 public and 9 private laboratories), 9 of them had participated in the baseline ILS. Laboratories that could not participate in the second ILS cited staffing issues.

This study was designed identically to the baseline ILS (i.e., same number and types of samples) with the exceptions that fewer water chemistry parameters were measured (e.g., no ion composition analysis of the dilution waters) and participating laboratories followed a common set of laboratory techniques defined in the roundtable workshop. This included use of neonates produced by 6- to 10-day old females to start the test, renewal, or termination of test boards daily at 24 h within a 2-h window, independent quantification of food density by the testing laboratories, food holding times of ≤ 7 days for Yeast-Cerophyll®-Trout Chow (YCT) and ≤ 21 days for green algae, and documentation of split broods on bench sheets at the time of observation. Testing was preceded with a training session for all participating laboratories for these standardized laboratory techniques.

Results from the second ILS indicated that two laboratories did not meet test acceptability criteria in at least one of the three testing rounds. Laboratory control mean reproduction ranged from 25 to 45 neonates per female in valid tests. This was an improvement from the baseline study where three laboratories did not meet test acceptability criteria in at least one of the three testing rounds. For half of the laboratories, the CV for mean neonate production in their controls remained similar or improved compared to the baseline ILS. Three laboratories, however, exhibited a wider distribution of CVs in the second ILS compared to the baseline study. Samples tested as dilution series showed good intra-laboratory agreement, but differences in potency estimates for NaCl remained among laboratories. A noticeable improvement was also observed for 6 laboratories.

Water quality, brood board health, food and age of females showed no obvious correlation with test outcome. Overall, one laboratory showed improvements that may be tied to the standardization of select test methods. However, the inconsistent laboratories remained inconsistent and some of the low-quality data were due to poor organism health or technical

issues. Laboratories that exceeded expectations produced high quality and comparable data in both ILS.

MAIN FINDINGS

This project resulted in several findings related to the current implementation of the *C. dubia* reproduction toxicity test method by California accredited laboratories and the magnitude of variability within and among laboratories.

- Most laboratories can perform the *C. dubia* survival and reproduction test on a routine basis. This assessment was based on evaluation of a set of performance metrics, i.e. frequency of meeting TAC, CV for reproduction. But a more consistent implementation and documentation of the required and suggested laboratory techniques as described in the EPA manual would help improve intra- and inter-laboratory variability.
    - Six of the 12 participating laboratories performed within the Panel's expected level of consistency and comparability during both the historical data analysis and the ILS.
    - Four laboratories performed consistently based on historical data analysis but exhibited increased variability during at least one of the ILS. Two of these laboratories experienced culture or brood board health issues which may have impacted test results. Therefore, these laboratories can likely perform the test well, but will likely benefit from the recommendations provided in this report.
    - Two of the 12 laboratories performed below the Panel's expectations in both the historical data and one of the ILS. These laboratories indicated challenges associated with cultures and technical errors during testing. The recommendations provided will provide opportunities for improvement.

- Variability and uncertainty of test results evaluated using CV reproduction and PMSD revealed that were relatively high for a subset of laboratories.

- In general, no single aspect of culture or test procedures was identified that accounted for a large portion of the variability observed.
    - However, statistical analysis suggested that use of older adults in the brood boards may be associated with poorer performance of test organisms.

- Based on four onsite laboratory assessments and the subsequent workshop, it appeared that variability was most likely a function of multiple laboratory-specific sources that may occur on an intermittent basis, as opposed to any single or consistently-occurring factor.

- Contributing factors may have included preparation of water and diet and associated storage protocols, undesirable laboratory practices (e.g., conducting testing and culturing operations in same area), and poor husbandry procedures resulting in contamination of tests and cultures by micro-organisms.
- Variability within individual laboratories may have been associated with lack of consistency in the application of specific procedures, potentially reflecting a lack of training and sufficient oversight by more experienced staff.

- Overall, standardization of select laboratory techniques produced modest improvements for laboratories with inconsistent or low historical performance.
  - Results suggest that at least one laboratory benefited from greater standardization of laboratory techniques.

- The level of variability within and among laboratories for the *C. dubia* survival and reproduction test was inconsistent with their accreditation status, suggesting that the existing accreditation program including the proficiency testing, are not sufficient to achieve a uniform standard of quality across laboratories that ensures awareness of proper procedures, as well as their implementation.

RECOMMENDATIONS

The recommended guidance resulting from this two-and-a-half-year study falls into one of three categories: (1) Laboratory Best Practices, (2) Accreditation, and (3) Training.

Laboratory Best Practices

These recommendations are based on a review of the standard operating procedures, phone interviews, site visits and the roundtable discussion among laboratories and the ESP. The recommendations are directed largely at the laboratories. The recommendations are divided into "Must do" described in the promulgated method and EPA guideline for freshwater WET testing, and the "Should do" proposed to minimize variation but not required in the EPA method. It is important to note that EPA has already provided definitions of must/shall and may/should for WET toxicity test methods: "Words of obligation" (EPA, 2000a Method guidance). WET test methods often state the procedure without a must or should, and in those instances, it is considered a directive. When WET method manuals use discretionary terms such as "may" or "should" the manual provide flexibility so that the laboratory analyst can optimize successful test completion.

Must do recommended guidance include:

- **Terminate the test when 60% of surviving females in the controls have had three broods, within a 2-h window (i.e., + or - 1 h) of test initiation time.**
- **Independently quantify food concentrations in stock bottles and record amounts added to each test container.**
- **Use source water produced according to the requirements of EPA freshwater WET test methods.**
- **Use known parentage with young from one adult for each concentration and use stratified random or complete randomization of all test cups.**

Should do recommended guidance include:

- **Conduct a detailed quantitative assessment of brood board health prior to testing.**
- **Document split broods on bench sheets daily at the time of the observations.**
- **Renew test solutions daily within + or - 2 h (i.e., 4-h window) of test initiation time.**
- **Update laboratory documentation.**
- **Store reagents to prepare the dilution waters and the reference toxicant appropriately.**

Accreditation

The ESP made recommendations on expanding the goals and implementation of the accreditation process (including proficiency testing) to ensure interlaboratory comparability. Since laboratory accreditation is the responsibility of the State, these recommendations are largely directed at the State accreditation program. These recommendations may increase operating costs for the laboratories, which will likely lead to increased costs to their testing clients. The recommendations are:

- **Increase the number and/or frequency of testing to assess comparability among laboratories.**
- **Collect and evaluate additional data associated comparability testing.**
- **Optimize laboratory audits to ensure effective and consistent implementation of best practices.**

Training Curriculum

Communication through regular WET testing training could ensure that the permittee or permitting authority has the information that they need to make informed decisions. Roundtable discussions and public meetings with stakeholders highlighted the need to provide training materials for an improved understanding of method requirements and data quality objectives. Training recommendations are directed at the State, the laboratories, and at the regulated parties responsible for toxicity testing as a compliance requirement:

- **Implement auditors' training program.**
- **Implement training program with defined performance goals for all personnel involved in performing or reviewing the *C. dubia* test.**
- **Provide guidance to regulated parties to evaluate WET toxicity test data and understand the results.**

CONSTRAINTS

While this study has produced more information on *C. dubia* inter-laboratory variation than any other study in the last 15 years, there are still a number of limitations to the conclusions and recommendations provided. These limitations fall into five categories:

- Constraints associated with the number of laboratories and the timing of the testing, which may not exhibit all the sources of variation possible in the *C. dubia* survival and reproduction test.
- Constraints in the outcomes of the ILS led to difficulties in quantifying the individual variability for each of the nine standardized laboratory best practices.
- Constraints quantifying intra- and inter-laboratory variability associated with testing *C. dubia* survival and reproduction in dilution water of varying hardness.
- Constraints on the number of toxicants evaluated to quantify intra- and inter-laboratory variability for concentration response in the *C. dubia* survival and reproduction test.
- Constraints imposed by the study timeline and due dates, which impeded the Science Panel's opportunity to refine laboratory performance metrics.
- Constraints associated with guidance to implement the recommendations.

TABLE OF CONTENTS

List of Abbreviations and Definitions

ANOVA – Analysis Of Variance

C. dubia – Ceriodaphnia dubia

CV – Coefficient of Variation

DMW – Diluted Mineral Water; EPA recipe for moderately hard water using Perrier® or Evian®

ELAP – State of California Environmental Laboratory Accreditation Program

HW – Hard water is the EPA-developed recipe for reconstituted hard water to use for toxicity testing.

IC – Inhibition Concentration is the toxicant concentration that would cause a given percent reduction in a non-quantal biological measurement for the test population.

IC25 – 25% Inhibitory Concentration

IC50 – 50% Inhibitory Concentration

ILS – Interlaboratory Comparison Study refers to split-sample exercises among California accredited laboratories.

Inter-laboratory variability – The variability between laboratories, measured by comparing results from different laboratories using the same test method and the same test material.

Intra-laboratory variability – The variability within a laboratory, measured when tests are conducted using specific methods under constant conditions in the same laboratory. This variability includes within-test variability.

LC – Lethal Concentration is the toxicant concentration that would cause death in a given percent of the test population.

LC50 – 50% Lethal Concentration

LOEC – Lowest Observed Effect Concentration

MHW – Moderately Hard water is the EPA-developed recipe for reconstituted moderately hard water to use for toxicity testing.

NOEC – No Observed Effect Concentration

PMSD – Percent Minimum Significant Difference is the smallest significant difference from the control expressed as a percentage of the control mean.

PT – Proficiency Testing is a study conducted annually to assess laboratory performance in comparison to the other laboratories across the nation.

QAP – Quality Assurance Plan

SCCWRP – Southern California Coastal Water Research Project

SD – Standard Deviation

SOP – Standard Operating Procedures

SWRCB – State Water Resources Control Board

TAC – Test Acceptability Criteria

TNI – The NELAC Institute, a non-profit organization that operates the National Environmental Laboratory Accreditation Program (NELAP) and the National Environmental Proficiency Testing Program (NEPTP)

TST – Test of Significant Toxicity

U.S. EPA – United States Environmental Protection Agency

YCT – Yeast-Cerophyll-Trout chow

WET – Whole Effluent Toxicity

2. INTRODUCTION

The California State Water Board recently adopted Toxicity Provisions, which include numeric effluent limitations to protect California's enclosed bays, estuaries, and inland water bodies from contaminated discharges. The Toxicity Provisions also include a requirement to use the Test of Significant Toxicity (TST) statistical approach, which controls for both false positive and false negative error rates (Denton et al. 2011). This approach also "restated" the null and alternative hypotheses compared to traditional hypothesis tests; the null hypothesis of the TST being that the sample is toxic. Because of the controls on error rates and the restating of the null hypothesis, the TST approach is more likely to find a sample to be toxic if within-test variability is high. In using this approach, the State Water Resources Control Board (SWRCB) aims to incentivize dischargers to generate high-quality data (i.e., data with low within-test variability). However, dischargers have expressed concerns about the inherent variability in some of the Whole Effluent Toxicity (WET) tests included in the Toxicity Provisions such as the WET test for *Ceriodaphnia dubia* (*C. dubia*) survival and reproduction.

The *C. dubia* survival and reproduction test is a well-established and validated method and was first promulgated in October 1995 and finalized in 2002, over 20 years ago (U.S. EPA 2002a, b, c; U.S. EPA 2016). While the State Water Board has full confidence in the use of *C. dubia* for regulatory programs, they recognized that some laboratories may need to improve their implementation of the *C. dubia* method. For this reason, implementation of the median monthly effluent toxicity limitation for the *C. dubia* test was delayed until January 1, 2024, for some dischargers as specified in the Toxicity Provisions. During this time, the State Water Board has committed to a study, in collaboration with stakeholders and laboratories, to evaluate laboratory performance, investigate factors that can lead to excessive test variability, and provide additional laboratory technique guidance to improve laboratory performance.

The WET test methods (EPA 2002a, b, c) allow laboratories some flexibility when implementing certain laboratory techniques. For example, for the *C. dubia* test method, different types of dilution waters (one made with salts, and one made with diluting a commercial mineral water) can be used. The appendix in the acute manual (EPA 2002a) describes the procedures for culturing and obtaining test organisms of *C. dubia*. In some instances, the promulgated method provides directions for what is to be done and includes non-prescriptive test techniques, leaving laboratories to use their best professional judgement. In 2021, EPA Region 9 and California SWRCB held a virtual *C. dubia* Workshop to help California testing laboratories review and discuss the procedures. This 2-day workshop conducted prior to the start of the current project, consisted of a review of the EPA methods and group discussion of laboratory techniques.

The State of California Environmental Laboratory Accreditation Program (ELAP) accredits all laboratories conducting analysis for regulatory compliance purposes, including the *C. dubia* test. At the start of this study, there were 18 laboratories ELAP accredited to conduct the *C. dubia* test for California and 17 of them participated in the study (**Table 2-1**). Accreditation is based on the demonstration that laboratories are following the testing protocols, properly training their staff, keeping accurate records, demonstrating they can meet data quality objectives for internal reference toxicant samples and nationally distributed proficiency test (PT) samples. While the ELAP process demonstrates that a laboratory capably performs a test, it does not address test variability between laboratories or differences in laboratory techniques that are allowed by the protocols.

It has been hypothesized that the small methodological differences between laboratories may lead to intra- or inter-laboratory variability, which could influence test results. Previous studies have assessed the variability of the *C. dubia* test results within and among laboratories. In the early 2000s, an interlaboratory comparison exercise performed by the EPA found that 22 out of 122 *C. dubia* chronic tests did not meet TAC for survival or reproduction (EPA 2001a, b). The invalid tests were confined to 10 out of 34 participating laboratories. The study reported intra- and inter-laboratory coefficients of variation (CVs) for the IC25 values of effluent and receiving water split samples at 17% and 28%, respectively for the reproduction endpoint. More recently, a smaller interlaboratory comparison exercise was conducted in California to evaluate the reliability of *C. dubia* chronic test for stormwater toxicity (Schiff and Greenstein 2016). Of the nine laboratories that tested split samples of dilution water, three were considered “low comparability” based on three factors including test acceptability, intra-laboratory precision, and inter-laboratory precision. Lack of comparability among a minority of laboratories testing split samples of dilution water was also identified by others (Moore et al. 2000; Diamond et al. 2008). In Fox et al. (2019), NPDES data was examined from routine *C. dubia* survival and reproduction testing data from 2012 to 2015 that had been generated by eight California-accredited laboratories with tests being conducted using moderately hard, hard, and very hard water. The study compared two statistical approaches to determine the influence of laboratory test performance on the false-positive error rate. The study showed the need for laboratories to track their control CV and adopt measures to decrease within test variability (without enumerating how) and found that one laboratory that modified laboratory practices after 2012 showed their CV decreases from 0.31 to 0.17 (at the 75th percentile).

Various studies have focused on the causes of *C. dubia* test variability or ways to optimize the test. The main thrust of these studies has been the water used and organism feeding. Elphick et al. (2011) found that water hardness influenced the sensitivity of the organisms to chloride, with a decrease in toxicity observed as hardness increased. Other studies found acute toxicity associated with major ions (Na+, K+, Ca2+, Mg2+, Cl-, SO42-, and HCO3-/CO32-); salts can be a

confounding factor both in natural waters and anthropogenically influenced waters (Mount et al. 2016, Erickson et al. 2017). Additionally, Mount et al. (2016) found that natural waters with low major ion concentrations caused *C. dubia* to be more sensitive to solutions of some salts. Chronic toxicity tests with *C. dubia* of major ion salts (Mount et al. 2019) showed a similar result to the acute study. Tests were conducted with 6 replicates and 9 closely spaced treatment concentrations were used to better support regression analysis of fairly steep response curves for modelling the data. For both the acute and chronic tests, the variability seemed to increase below an equivalent hardness of about 10-15 mg/L as CaCO3.

One California laboratory conducted multiple studies to reduce sources of variability in their own *C. dubia* tests. The laboratory tested multiple dilution water types and sources and found that synthetic versus natural water had the most impact on reproductive variability, but it was small compared to feeding related aspects (Briden et al. 2017). In a study on the effects of water hardness, the California laboratory also found it had no impact on long-term culture performance (Clark and Briden 2018). However, the laboratory found that organism source and control/dilution water hardness might have an impact on test results. In two of six samples where both a soft and moderately hard water control was used, the interpretation of toxicity differed depending on which control the sample was compared to. The laboratory also conducted two studies looking at the effects of food quality. In the first study, the laboratory found that quality of the food had an impact on the test performance even if inferior quality food was only fed to culture animals, but higher quality was used during the test period (Jorgenson et al. 2017). The study also found that the quality of the algae was the most important factor influencing control variability and was greater than control/culture water parameters, feeding density, food component, culture line, or analyst training. Source of the YCT did not appear to affect test control precision. In their second food study, they found that vendor sourced food was not necessarily of consistent quality (Prosser et al. 2018). The laboratory concluded that it was important to run quality control (QC) tests before using the food in cultures or tests. Little difference was found in reproduction based on variable food components, but when larger volumes of trout chow were digested a negative impact on test performance was observed. The laboratory also noted that the EPA recommendation of a 2-week shelf life for a *Selenastrum* batch may be too restrictive. Visual and olfactory observation of each batch were important to determine shelf life.

Table 2-1. Toxicity testing laboratories accredited for the *C. dubia* test in the state of California at the start of the study in 2021.

Lab Name	Lab Type
ELAP accredited laboratories in California	
49er Water Laboratory	Private
Aqua-Science	Private
Aquatic Bioassay & Consulting Laboratories, Inc.	Private
Aquatic Testing Laboratories	Private
Aquatic Toxicology Laboratory, Aquatic Health Program, UC Davis	Academic
Enthalpy Analytical, LLC	Private
Environmental Monitoring Division (EMD) at Hyperion Treatment Plant	Public
Inland Empire Utilities Agency Laboratory	Public
MBC Aquatic Sciences	Private
McCampbell Analytical, Inc.	Private
Pacific EcoRisk	Private
San Jose Creek Water Quality Laboratory	Public
Wood Environment & Infrastructure Solutions, Inc.	Private
ELAP accredited laboratories outside of California	
EcoAnalysts, Inc.	Private
Eurofins TestAmerica - Corvallis (ASL)	Private
GEI Consultants, Inc.	Private
Tetra Tech's Ecological Testing Facility	Private

3. Study Objectives

The objective of this study was to build on previous efforts and investigate all possible sources of variability in the *C. dubia* reproduction test conducted by California-accredited laboratories. The goal was to provide laboratory technique guidance to: (a) improve the consistency of the execution of the *C. dubia* test method to achieve improved precision within each testing laboratory; and (b) improve the consistency and comparability of *C. dubia* test results among testing laboratories, while retaining the necessary flexibility for environmental relevance (SCCWRP 2021).

The study aimed to answer the following questions:

1. What are the *C. dubia* chronic survival and reproduction toxicity test laboratory techniques used by Environmental Laboratory Accreditation Program (ELAP) accredited laboratories in the state of California?

2. How does variability in control reproduction and/or reference toxicant response in the *C. dubia* chronic survival and reproduction toxicity test compare amongst intra- and inter-laboratory technique differences used by ELAP accredited laboratories?

3. Does standardizing differences in the *C. dubia* chronic toxicity test laboratory techniques reduce intra- and inter-laboratory variability in control reproduction and/or reference toxicant response?

Based on the results of this study, a list of suggested best practices for the *C. dubia* reproduction test laboratory techniques were developed.

Note that this study was not designed to address or quantify false negative or false positive rates for detecting toxicity from known or unknown samples. It was also not expected to eliminate all variability from the test method. Finally, it should be noted that this study was not designed to address aspects of testing that may be more effectively dealt with by appropriate study design: e.g., ion, hardness, or conductivity controls in cases where those variables have the potential to affect test outcomes, but do not represent environmental risks.

4. General Approach

Six tasks were used to address the study objectives. These tasks were sequential with each one informing the details of the next.

1. Create a governance structure to oversee the study
2. Analyze historical data and existing laboratory techniques to identify sources of variability
3. Conduct a baseline interlaboratory comparison study to quantify variability within and among laboratories
4. Agree on a standardized list of laboratory procedures
5. Evaluate the efficacy of standardized laboratory techniques in reducing intra- and inter-variability via a second interlaboratory comparison study
6. Provide final recommended guidance in a Final Report

The first Task created a two-tiered governance structure to ensure transparency and technical rigor. One tier was a Stakeholder Committee comprised of representatives from sectors potentially impacted by the study results. The second tier was an independent Expert Science Panel comprised of scientists experienced in the *C. dubia* test method, biostatistics, and data quality measures, and with no potential conflict with study results. The Expert Science Panel was the final decision-making body. The Stakeholder Committee provided valuable input and context for recommended guidance implementation.

The second Task was comprised of two subtasks. The first subtask was compiling an inventory of laboratory techniques used by ELAP accredited laboratories. The inventory elucidated the level of comparability and differences in test implementation. The second subtask focused on compiling historical testing data from the ELAP accredited laboratories to quantify the level of variability within and among laboratories. Approximately 1,000 tests were compiled from all but one of the ELAP accredited laboratories for this subtask. The inter- and intra-laboratory variability was assessed based on the reproductive endpoints of the test method (e.g., average number of neonates per female). The differences in laboratory techniques were compared to the laboratory test results to attempt relating which laboratory techniques might account for the observed variability in the test outcomes.

The third Task collected new data using a baseline interlaboratory comparison study using well-homogenized, split samples to assess intra- and inter-laboratory variability. The split sample analysis for this subtask supplements the historical data analyses and was intended to confirm possible sources of test variation. The basic study design for the baseline interlaboratory

comparison was to remove the variability associated with samples (which was not possible in the historical data) and quantify the variability introduced by the individual laboratories. Laboratories utilized their existing protocols for all baseline interlaboratory comparison testing.

The fourth Task focused on identifying laboratory practices to standardize to reduce the interlaboratory variability. This task was accomplished using two sub-tasks. The first subtask consisted of on-site laboratory visits by a subset of the Expert Science Panel. These Panel members observed each laboratory's culturing, dilution water preparation, food and feeding, test implementation, quality assurance, amongst other activities. Four labs were visited that comprised a range of size and consistency of quality from the baseline interlaboratory comparison study. The second subtask was a roundtable workshop convening of all the intercalibration-participating laboratories. Based on the differences in laboratory procedures identified during the historical data inventory, the baseline interlaboratory comparison, and the laboratory visits, the goal of the roundtable workshop was to achieve consensus on what laboratory procedures to standardize for Task five.

The fifth Task conducted a second interlaboratory comparison study, which mirrored the baseline interlaboratory comparison, except for the laboratories standardized the list of laboratory procedures agreed to during the roundtable workshop. These laboratory procedures were well-documented, and laboratories trained on how to implement them. The goal was to assess if the standardized laboratory procedures in Task four improved laboratory intra- and interlaboratory variability.

The sixth Task documents the study and lists the final recommended guidance on improving intra- and inter-laboratory variability. This report is the culmination of Task six.

5. Evaluation of Intra-Laboratory and Inter-Laboratory Variability

Several types of information and datasets were evaluated during the two-year project to assess laboratory performance and determine the potential sources of variability.

5.1. Inventory of standard operating procedures, quality assurance plan and historical testing data from California's accredited laboratories

To investigate factors that can lead to test variability, an inventory of laboratory techniques and historical data was created focusing on culture and test conditions, and performance data for control samples and reference toxicant. **Table 5-1** presents the parameters collected. Out of the 18 accredited laboratories, one did not participate due to lack of *C. dubia* data available (i.e., fewer than 15 tests over a 10-year period), and two laboratories provided incomplete information. To compile culture and test condition parameters, a review of the laboratory documentation such as Standard Operating Procedures (SOPs) and Quality Assurance Plans (QAP) was conducted. The performance data collected focused on the last 30 tests or up to 3 years for control samples associated with environmental test samples as well as reference toxicant concentration-response data. Control samples data were used to assess the lab's ability to perform the test, while reference toxicant data were used to assess reproducibility of test organism response. For each test, raw data from individual replicates was collected (i.e., daily neonate counts, survival) along with daily water quality data (i.e., hardness, alkalinity, pH, temperature) and other relevant metadata (e.g., brood board health). Environmental sample toxic response data were not used because there were no expectations of performance and data could not be compared among laboratories. It should be noted that the compiled tests included all samples regardless of whether the tests met TAC or not. The extracted data were hand-entered in a custom database and two independent audits were performed to assess completeness, accuracy, and variability. To verify the information compiled and collect additional data, phone interviews were conducted with key personnel from each laboratory. A survey questionnaire was submitted to the laboratories prior to the phone call and used during the discussion.

Notable differences were observed in all key parts of the test method including dilution water, test termination trigger and feeding techniques among the laboratories (**Appendix A, Tables A2 and A3**). The test termination trigger is an important laboratory technique described in the promulgated method (Section 13.10.9.1), and it is specified that test termination "must be completed when 60% of the females or more have produced three broods". While most

laboratories followed the requirement, the method to determine when the reproduction threshold is met varied greatly. Some laboratories used a strict time window daily while others checked periodically throughout the day to determine if the 60% threshold had been reached. Other laboratories documented using a higher percentage of females having produced three broods to ensure sufficient neonate production (e.g., 70% or 80% of females having three broods). One laboratory implemented a 7-day test consistently independently of the reproduction threshold. Control charts were also different as laboratories used either sodium chloride (NaCl) or copper chloride (CuCl) in different concentrations to prepare their serial dilutions. One laboratory reported using zinc sulfate as the reference toxicant.

Table 5-2. Laboratory techniques and performance variables compiled from California accredited laboratories.

Laboratory practices	Testing and Performance Variables Recorded
Origin of brood stock	Age window at test initiation
Age of culture	Time to reproduction
Culture renewal frequency	Test termination trigger
Dilution water recipe	Test termination window
Source water	Test duration (days)
Dilution water shelf-time	Number of neonates per female per replicate
Reference toxicant name and source	Number of replicate test chambers
YCT vendor, shelf-time	Survival of control females per replicate
YCT concentration in culture and test chamber	Neonate production in control samples (mean, CV)
Algal species	Reference toxicant, LC50
Algae vendor or recipe, shelf time	Reference toxicant, IC50, IC25
Algae concentration in culture and test chamber	PMSD
Feeding frequency	Water hardness
Lab air temperature	Water conductivity
Photoperiod	Water dissolved oxygen
Light source	Water temperature
Sample volume in test chamber	Water pH
Test chamber material, volume, diameter	Water alkalinity

The EPA manual allows some flexibility in source water and dilution water recipes and, as a result, most California accredited laboratories have reported using modified dilution water recipes. Eight laboratories appeared to use one of the dilution water recipes specified in the EPA manual, either DMW or MHW with or without selenium. However, further investigations into the preparation of source water showed that some of them may not be using high quality source water (with a resistance ≥ 18 megaohm-cm). Other laboratories who used the MHW recipe often added different amounts of vitamins and/or selenium or adjusted the salts ratio (referred to herein as modified EPA recipe). Only one laboratory used EPA hard water (HW) for their culture and laboratory controls because the hardness of their test samples is usually outside of the range of MHW targeted in the EPA manual. It should be noted that two of the accredited laboratories did not use any of the dilution water recipes described in the promulgated method (Labs I and K). Food source, preparation and distribution were also different among laboratories. The feeding regime was also vastly different, as laboratories used different vendors and laboratory techniques (purchased or in-house) to produce their YCT and green algae *Raphidocelis subcapitata* stocks and feed them to *C. dubia* cultures. Previous research showed that the quality of food can affect both the number of neonates produced and the variability between tests within a single laboratory (Jorgenson et al. 2017). Other notable differences included material and size of test chambers as well as sample volume used in the test chambers. Finally, differences in methods for water quality measurement were observed. Most laboratories did not use true surrogates, defined as test chambers of the same size and with the same volume of test solution. Instead, laboratories reported the use of larger test chambers or pooling of samples from different cups during the renewal process to ensure a sufficient volume of test solution to submerge the probes.

A total of 551 sets of control data and 452 reference toxicants tests (**Table 5-2**) were entered in the database. Note that a 'set' is comprised of the reproductive output of 10 or 20 replicate test cups and supporting water quality data (temperature, pH, alkalinity, conductivity, and hardness). Only two laboratories did not provide a complete set of control data, Lab H and Lab J. These two laboratories did not participate in the subsequent tasks of this project. Test data collected were typically from 10 replicate chambers, except two laboratories that occasionally conducted their tests using 20 replicates for the submitted control data. All reference toxicant data consisted of five dilutions minimum tested in 10 replicates. Approximately half of the laboratories had ~30 or more tests available within a 1.5 to 5-year period. The other half of the laboratories reported conducting less frequently and submitted between 6 and 25 sets of control or reference toxicant test data. One laboratory, Lab B, provided 7 years of data. Raw data were used to calculate test endpoints such as mean neonate production per surviving female, mean survival of females in controls, IC25/50 for reference toxicant at test termination (**Appendix A**).

Table 5-3. Inventory of historical data compiled from 17 California accredited laboratories.

Labs	Total number of laboratory control tests	Number of laboratory control tests with 10 replicates	Number of laboratory control tests with 20 replicates	Number of reference toxicant tests
A	48	48	0	31
B	48	48	0	47
C	28	28	0	28
D	19	19	0	6
E	49	24	25	30
F	45	37	8	30
G	7	7	0	22
H*	0	0	0	17
I	30	30	0	30
J	7	7	0	21
K	19	19	0	15
L	27	27	0	30
M	59	59	0	34
N	30	30	0	30
O	30	30	0	30
P	80	1	79	28
Q	25	25	0	23
Total	**551**	**439**	**112**	**452**

* Lab H did not respond to request for laboratory control data.

5.2. Evaluating laboratory performance.

Evaluating test performance and consistency is an important component of any monitoring program to ensure that all testing laboratories have a similar level of competency. Initially, a statistical approach was applied (e.g., Analysis of Variance, ANOVA). However, the Panel noted that high variability within and among laboratories limited the ability to detect any significant changes. Therefore, the Panel used a combination of metrics on biological, variability/uncertainty and potency endpoints to assess laboratory performance. The promulgated test method already suggests the use of performance metrics (e.g., CV) to document ongoing laboratory performance and investigate ways to reduce variability (see Section 4.16.3 of the EPA manual). The proposed criteria and metrics used in this study are described below. The Panel is not including the metrics as part of their recommendations, as they require further refinement to ensure that they are reliable and valuable for all parties involved.

- Biological metrics (i.e., TAC) provide information on control test organisms, and reflect culture health and good laboratory practices. Meeting TAC is a requirement of the method.
    - Survival ≥ 80% in laboratory controls.
    - Mean number of neonates per surviving female ≥ 15 in laboratory controls.

- Metrics of variability and uncertainty provide information on consistency of test results within a laboratory. These were evaluated using both laboratory controls and reference toxicant samples.
    - CV for mean number of neonates per female in controls ≤ 0.2, i.e., the standard deviation in the number of neonates is less than or equal to 20% of the mean number of neonates in the group. This value is consistent with observed variability described in EPA (2001b) and Fox et al. (2019) studies. Such a level would be appropriate for confirming the presence of toxicity in environmental and other test samples.
    - PMSD for individual reference toxicant tests ≤25 which corresponds to the 50th percentile using data from this study. Note that this is consistent with the 50th percentile of ≤23 derived from laboratories across the nation (EPA 2001). This metric describes intra-test variability and represents the smallest percent difference that can be statistically detected when a test of mean differences between two concentration groups is conducted.

- Toxicity metrics focused on potency estimates for the reference toxicant and comparability among laboratories. This metric can be evaluated in different ways

depending on the distribution of the data and the desired confidence interval. The metric below is used in the present study as an example of how it can be applied to assess interlaboratory comparability.

- IC50s within the 25th and 75th percentile of all laboratories. There is currently no guideline for this metric in the EPA manual. The only guidance available is to reevaluate IC50 data that fall outside of two standard deviations within a laboratory.

Individual laboratory data was evaluated to assess the frequency of meeting these benchmarks. More details on the rationale and approach are included in **Appendix E**. The use of performance metrics will benefit all parties involved to describe, monitor, and clearly communicate the acceptable level of variability for this test. Testing laboratories can use the data as indicators of test organism condition and implementation of good laboratory practices. Regulators and regulated dischargers can use the information to increase confidence in test results, facilitate data interpretation and improve compliance with water quality monitoring objectives. It is important to note that the metrics used in the analysis of the *C. dubia* datasets are not inclusive or intended to be used as definitive guidelines for laboratory assessment. The State accreditation program could further refine the acceptance metrics and potentially include additional criteria (e.g., IC25, ratio of IC25 and IC50, LOEC, NOEC) based on the goals and objectives of the accreditation program.

5.3. Analysis of historical data

Historical data was evaluated in two ways. Key biological, variability/uncertainty metrics and potency estimates were compared within and among laboratories to assess overall laboratory performance. The potential relationships between laboratory techniques, test factors (e.g., water quality) and test outcomes were then investigated using a variety of linear and non-linear modeling approaches detailed below.

Laboratory performance

Test acceptability criteria for the *C. dubia* chronic toxicity test focus on the performance of laboratory controls and require ≥80% survival and mean production of 15 neonates per surviving female. Historical data analysis showed that all California-accredited laboratories were able to meet the TAC for survival consistently, with 15 out of 17 current and formerly accredited laboratories reporting ≥ 80% in over 90% of their tests. Neonate production, however, was more variable. While most laboratories did achieve the TAC for reproduction, a 2- to 3-fold difference in mean control reproduction was observed among tests within a laboratory (**Figure 5-1**). Out of the 17 laboratories, two (Lab H and J) had an average mean reproduction close to the TAC and did not meet the criterion in more than 1 in 10 (i.e., ~12%) of their tests. The others typically had means of ≥ 20 young per surviving female in over 90% of their tests.

Meeting TAC alone is not sufficient to assess laboratory performance. Thus, the Panel critically evaluated metrics of variability, starting with the CV. Mean CVs for individual tests ranged from 0.14 to 0.32 (**Figure 5-2**). Analysis performed using a CV ≤ 0.2 as a metric of good laboratory performance. indicated that two laboratories (Lab A and Lab P) met this criterion in 8 of 10 of their tests and one laboratory, Lab F, met this criterion in 7 out of 10 tests. It should be noted that Lab A was the only laboratory with a 75th percentile of CV ~ 0.15 indicating that this laboratory had one of the most consistent control reproduction datasets. Five of 17 laboratories were able to achieve a CV of ≤ 0.2 in 50-70% of their tests, meeting the Panel's proposed threshold. These data suggest that a "long-term" average CV of ≤ 0.2 is achievable. The remaining laboratories (9 out of 17) exhibited higher variability in control reproduction suggesting that these laboratories could benefit from greater level of standardization and guidance to improve data consistency.

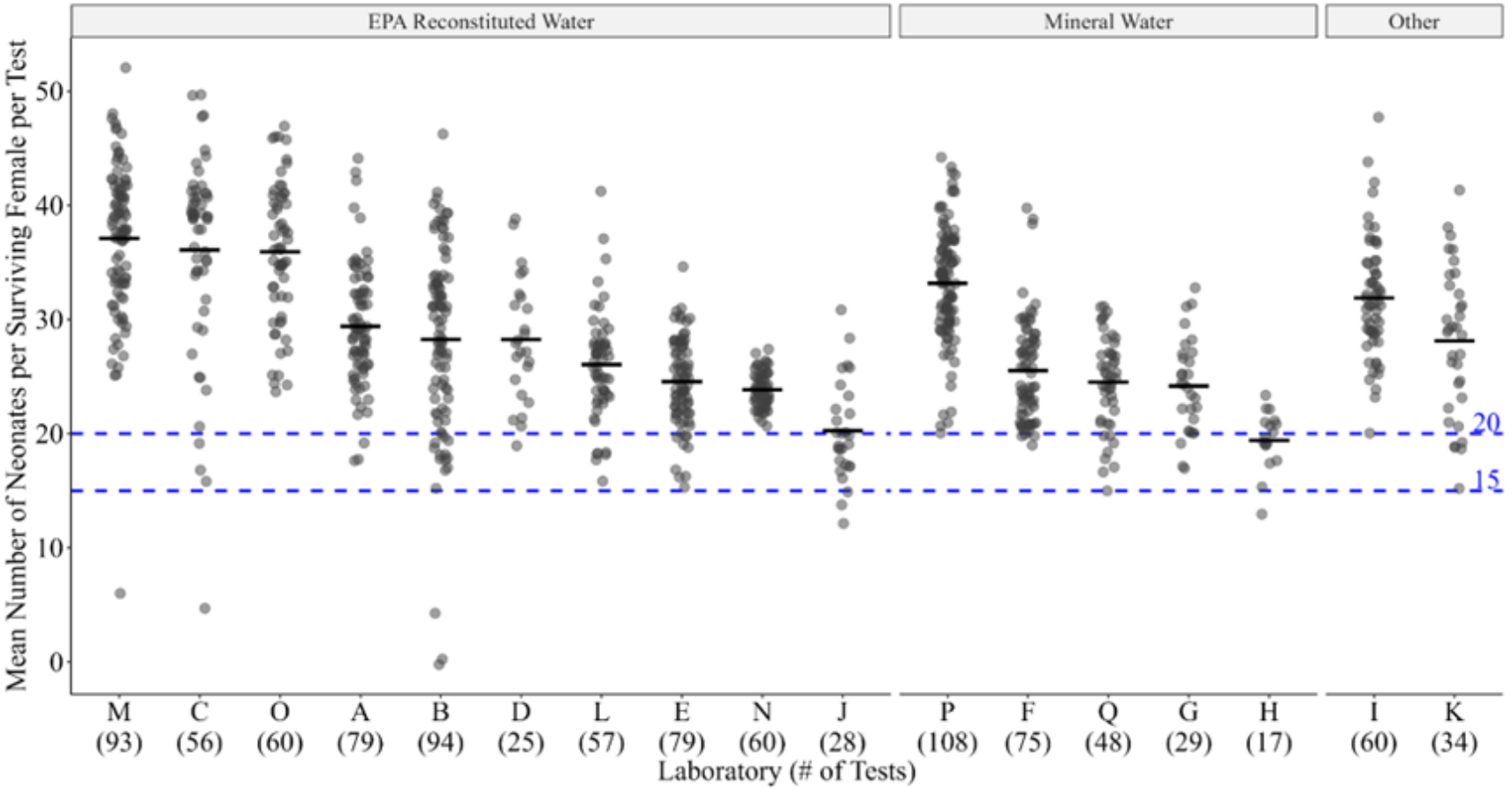
One statistical criterion used in hypothesis testing to assess toxicity test performance is PMSD (Denton et al. 2003). The EPA method guidance (2000a) suggests that PMSD be monitored as part of a testing laboratory's ongoing QA program. While there is no target PMSD value for compliance, the EPA has suggested a PMSD of 37% as the minimum acceptable level, which was the 90th percentile calculated using data from a national interlaboratory comparison study using a reference toxicant approach (EPA 2001a, 2001b). Consistent with the EPA study, the 90th percentile for the data compiled in the current project was 36%. When comparing the mean

PMSD per laboratory to the "upper bound" of ≤ 37%, all laboratories were able to meet this threshold in at least 6 of 10 of tests (**Figure 5-3**). However, the actual PMSD values ranged between 6 and 159 among laboratories (**Appendix A Table A6**). This demonstrates that a guideline of ≤ 37% is not representative of the performance of a competent and consistent laboratory.

When setting the PMSD target value at ≤ 25%, which corresponds to the 50th percentile for the EPA study (EPA 2001a), seven laboratories met the target in over 75% of their reference toxicant tests and five additional laboratories met the target in 50 to 70% of their data. It should be noted that laboratories consistently meeting TAC and a CV of ≤ 0.2 for reproduction (Labs A, F, P) also had the greatest level of consistency in PMSD values. Only two laboratories had < 50% of PMSD values below 25. Overall, these data showed that 7 out of 17 laboratories were able to detect a 25% inhibition level relative to the control on a consistent basis (i.e., over 70% of the time), whereas the remaining laboratories tended to be characterized by higher variability and would likely benefit from a rigorous review of culture conditions and testing procedures.

Ten laboratories used NaCl as their reference toxicant, six laboratories used CuCl and one laboratory used zinc. Therefore, reference toxicant data could not be used to assess interlaboratory agreements among all California accredited laboratories. Mean IC50s ranged between 15 and 65 µg/L for CuCl and between 1000 and 2400 mg/L for NaCl (**Figure 5-4**). For laboratories using NaCl, 9 of 10 laboratories had > 50% of estimated IC50s within the 25th and 75th percentiles of all reference toxicant tests. One of them, Lab Q, exceeded expectations and recorded >75% of IC50s within the 25th and 75th percentiles. This suggests a reasonable level of comparability among most laboratories using NaCl. One outlier was identified, Lab K had a mean IC50 one order of magnitude higher than the other laboratories. For this laboratory, over 50% of estimated IC50s fell outside of the 25th and 75th percentiles. The distribution of IC50s for copper was slightly more variable. Four of the 6 laboratories had > 50% of estimated IC50s within the 25th and 75th percentile of all CuCl reference toxicant tests, and one laboratory was considered an outlier with <50% of IC50 data within the 25th and 10th percentiles.

Figure 5-1. Mean number of neonates per female for each submitted test from the historical dataset. Data is organized by the type of dilution water used by the laboratory for their test controls. Laboratories are in order of high to low mean values within each water type.



Note that most laboratories using the EPA reconstituted water add selenium and/or vitamins (see Appendix A, Table A2).

Figure 5-2. Coefficient of variation for the mean number of neonates per female for each submitted test from the historical dataset. Data is organized by the type of dilution water the laboratory uses in their controls. Laboratories are ordered from high to low mean values within each water type.

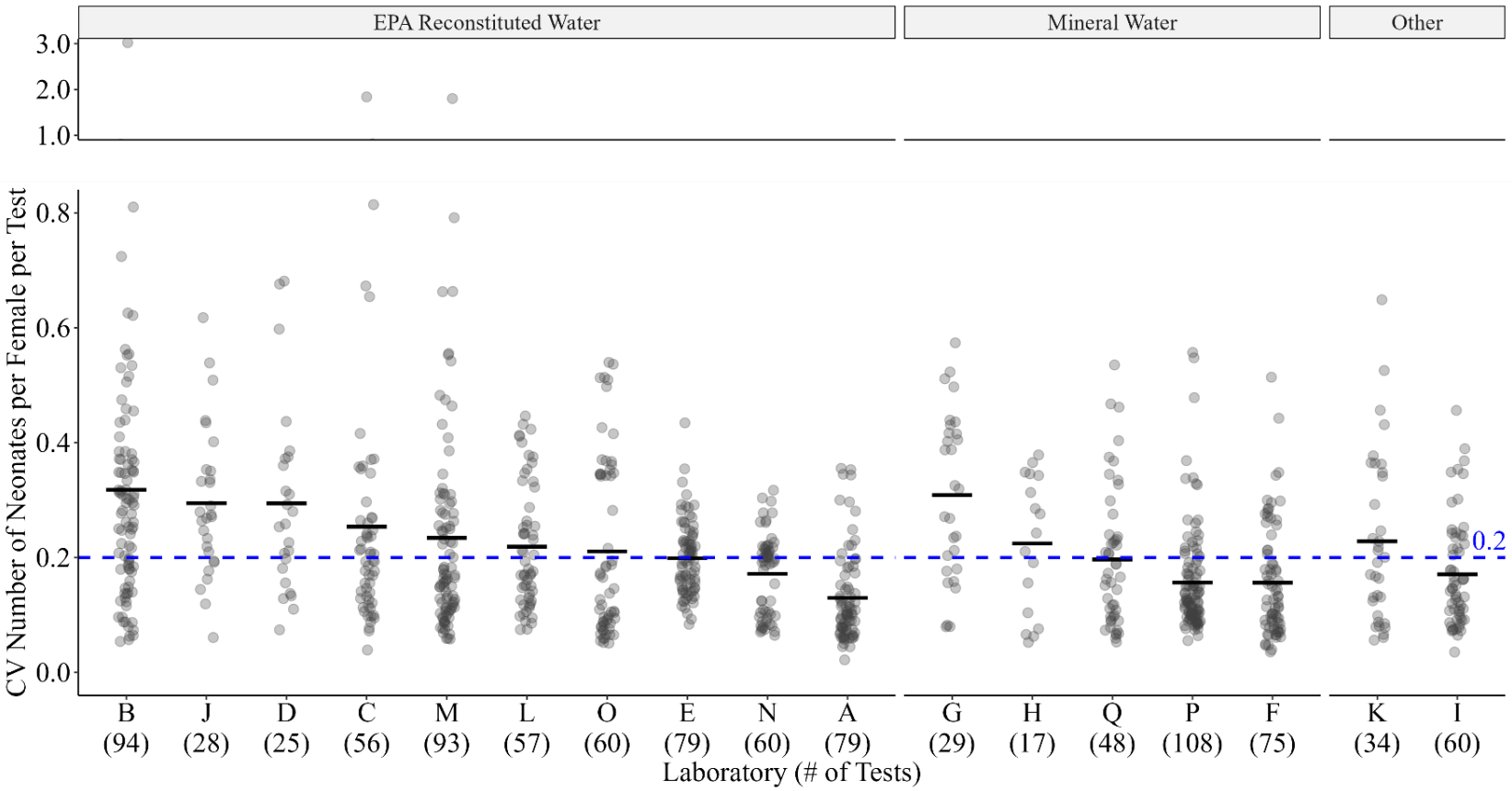


Figure 5-3. Percent minimum significant difference (PMSD) values from the reference toxicant data in the historical dataset. Reference lines indicate proposed laboratory performance criteria. Line at 25 is based on the Science Panel's suggested guideline and is the 50th percentile from EPA (2001a). Line at 37 is based on 90th percentile of data from an EPA (2001a) study. Laboratories are ordered based on their mean IC50 for each reference toxicant type.

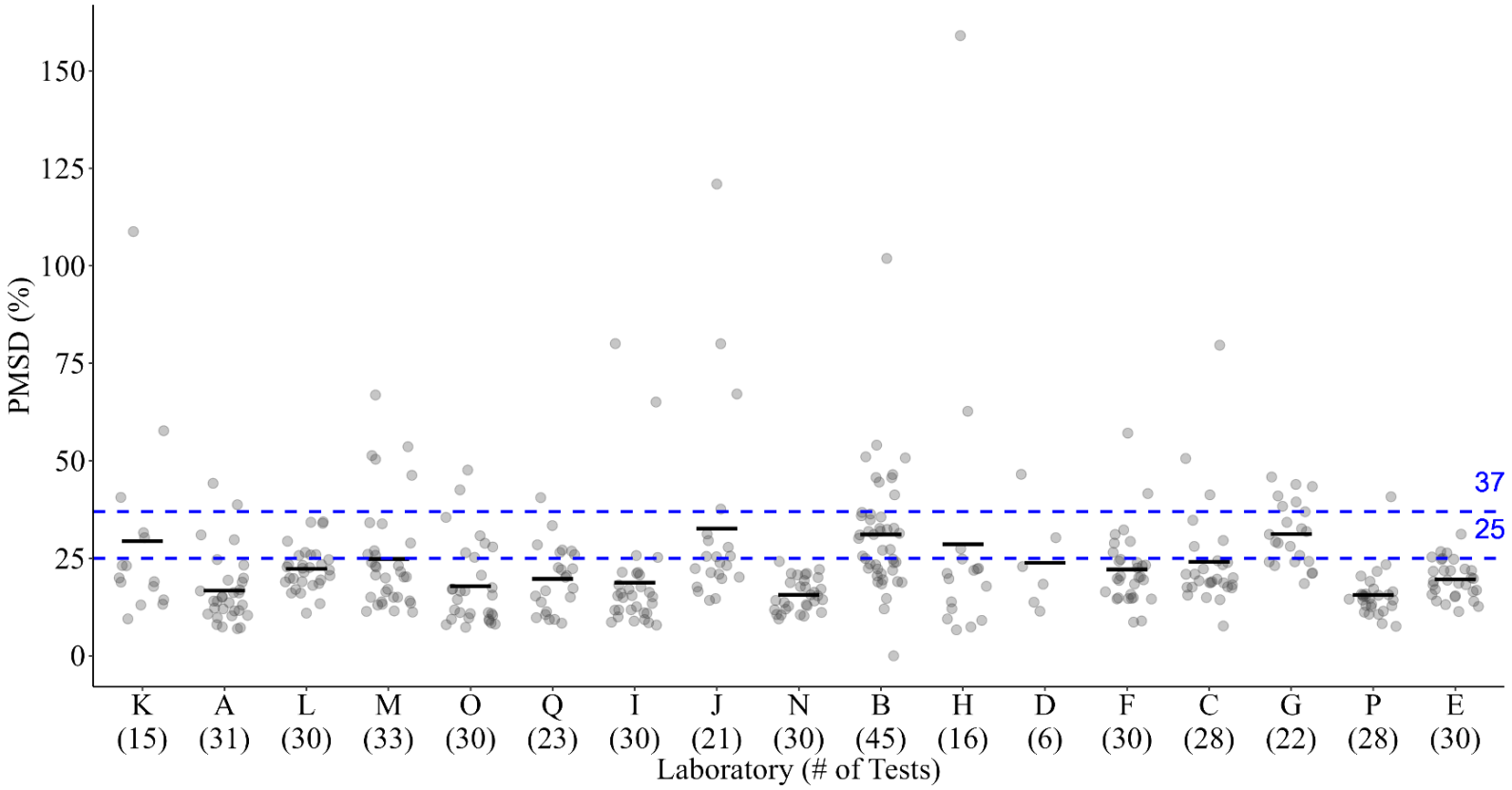
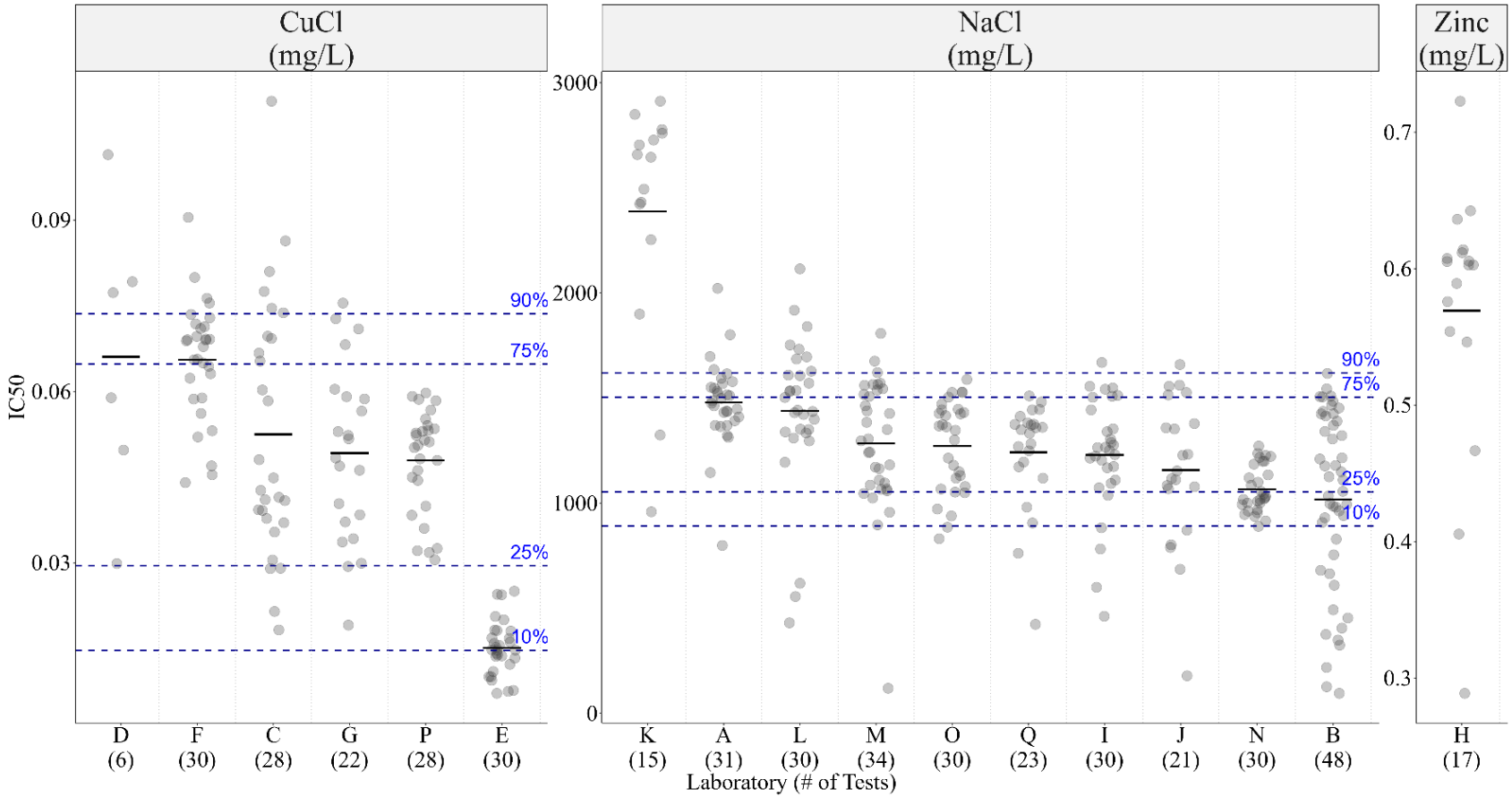


Figure 5-4. IC50 data from the historical dataset. The reference lines are percentiles of the data from each reference toxicant type. Larger, colored dots represent the mean value for each laboratory. The laboratories are ordered by highest to lowest IC50.



Investigating sources of variability in test outcomes.

The relative influence of laboratory techniques and test conditions (listed in **Table 5-1**) on inter- and intra-laboratory patterns of the means and CVs of control neonate production per test were investigated using univariate and multivariate linear modelling. The results indicated that there were no strong relationships between any one of the different laboratory techniques / test conditions with either mean or CV of control neonate production across the different laboratories.

To further explore important factors in neonate production across the historical data, random forest regression models (consensus-based, non-linear modelling) (Breiman 2001; Biau and Scornet 2017) of the mean and CV of control neonate production as a function of all the laboratory techniques / test conditions concurrently and interactively were created. Specifically, regression models were created for each laboratory with either mean or CV of neonate production as the response variable and all metrics **in Table 5-1** used as predictor variables. Variable importance measures extracted from each laboratory's random forest model were ranked by importance and compiled. Measures of test water quality (e.g., temperature, pH) and composition (e.g., conductivity, and dissolved oxygen) were highly ranked for more than 8 out of the 17 laboratories, suggesting that they could potentially be important drivers of both the mean and CV of control neonate production (**Appendix A Tables A10 and A11**). However, there was no agreement in the models to identify common factors among all laboratories. The variable importance results also suggested that the age of the females used in the brood board used to set up the test could be an important factor in mean control neonate production.

The results of the random forest models were used to build more structured, and potentially more diagnostic, multivariate general linear models of water quality, water composition, and female age on mean and CV of neonate production. However, due to the incomplete data record and high degree of heterogeneity in test water and female brood stock within and between laboratories, no conclusive results could be identified as to the sources of the observed variability in control neonate production among the laboratories in their historical test data.

Despite the lack of relationships between laboratory techniques / test conditions and mean or CV of test neonate production, the results did allow the Panel and the Stakeholder Committee to come to a consensus that controlling for water composition, as well as better reporting of test condition data (brood board characteristics, feeding regimes, and testing conditions) could produce a more complete and less "noisy" data set on which more diagnostic analyses could be conducted.

5.4. Split-sample interlaboratory comparison study among California's accredited laboratories

A three-step approach was used to assess the effects of standardizing laboratory practices on laboratory consistency and comparability. First, the laboratories participated in an ILS (referred to as baseline ILS) and tested split samples using their own protocols and provided more detailed data that may not be routinely collected and reported (Robertson-Bryan, Inc/CASA & SCCWRP 2022). **Appendix B** includes an excerpt of the QAP developed for this study. Based on the results of the baseline ILS, the Panel conducted site visits in four of the laboratories and developed a list of topics to be discussed with the participating laboratories. A roundtable workshop was hosted, where laboratories shared their practices with the Panel and the group agreed on a set of practices to standardize in a second ILS. In the second ILS, laboratories tested split samples following the standardized *C. dubia* toxicity testing techniques agreed upon in the Workshop (SCCWRP 2023). **Appendix C** provides an overview of the key study elements for the second ILS.

Eleven California-accredited laboratories (2 public and 9 private) participated in the baseline ILS and nine (1 public and 8 private) in the second ILS. Those who could not participate in one of the two ILS cited issues with their cultures (one laboratory) or lack of staff and time (two laboratories). The baseline and second ILS followed the same testing design, with three rounds of testing per study. For each round, SCCWRP provided four types of split samples to the laboratories.

- Sample 1: MHW water recipe tested at full strength (i.e., 100%, no serial dilution needed). This sample was tested along with one laboratory control consisting of their own dilution water recipe.

- Sample 2A: DMW with Perrier® water tested at full strength (i.e., 100%). This sample was tested along with one laboratory control consisting of their own dilution water.

- Sample 2B-F: Five concentrations of NaCl diluted in DMW with Perrier®. The five samples were prepared at SCCWRP according to the procedure described in the study QAP. These samples were tested as is with no additional sample dilution allowed, along with one laboratory control consisting of their own dilution water.

- Sample 3: NaCl was provided as a solid to each laboratory with detailed instructions to prepare five dilutions using their own dilution water. The laboratory-prepared serial dilution was tested along with one laboratory control consisting of their own dilution water.

All samples were tested in 10 replicate test chambers and the tests were conducted for 8 days with neonate counts recorded daily. Data were submitted electronically to SCCWRP along with

the bench sheets, and data quality (i.e., entry error, accuracy, and completeness) was evaluated by the project team.

5.5. Data analysis for the baseline interlaboratory comparison study

Laboratory performance

A total of 11 laboratories (9 private and 2 public) participated; Lab I could not participate due to microbial contamination causing their culture to crash a few weeks before the start of the ILS. Lab B and M participated in two out of three rounds because the first set of samples was lost during shipping. Lab N reported high mortality in the brood board for one testing round and could only test 3 out of the four samples provided in round 2 (see **Appendix B**). Laboratory performance was considered for individual tests across the three rounds, with an emphasis on laboratory control performance and reference toxicant response.

Most laboratories met TAC target values for survival and reproduction per surviving female (**Figure 5-5**), except for two laboratories. Lab B recorded no survival (i.e., 0%) in all their laboratory controls in round 3 whilst > 90% survival was recorded in all the split-samples provided by SCCWRP and tested during that same round. Analysis of the dilution water pointed to errors during the preparation of the dilution water as the calcium to magnesium ratio was inverted (**Appendix B Table B25**). Lab E only met the TAC for reproduction in 50% of their laboratory controls and did not report any culture issues. Lab M and N also had variable neonate production among individual replicates (**Appendix B Table B3**) but met reproduction TAC in ~80% of their laboratory control samples. Lab N indicated that low reproduction was most likely due to unusually high morality in their test brood board. Lab N also reported a crash of their culture at the beginning of the baseline ILS.

Half of the laboratories had an average CV for neonates per female ≤ 0.20 indicating that they can produce good quality data above the established or suggested guidelines (**Figure 5-6**). Lab A, F and Q reported CVs ≤ 0.15 in at least 10 of their 12 laboratory controls, and Lab G, O and P had CVs ≤ 0.2 in 6 to 9 out of 12 laboratory controls. The remaining 5 laboratories had CVs for reproduction ≥ 0.2 in over 50% of their laboratory controls. Lab B and L achieved a CV ≤ 0.2 in 4 out of 10 tests, while Lab E, M and N exhibited much more variable results (CV ≤ 0.2 in less than 1 in 10 tests).

Spiked samples (samples 2B-F and 3) provided further insight into laboratory performance and comparability. These two sample types were used to (1) assess assay precision (PMSD, Appendix B Table B17) and (2) compare test organism sensitivity to a common reference toxicant (IC25 and

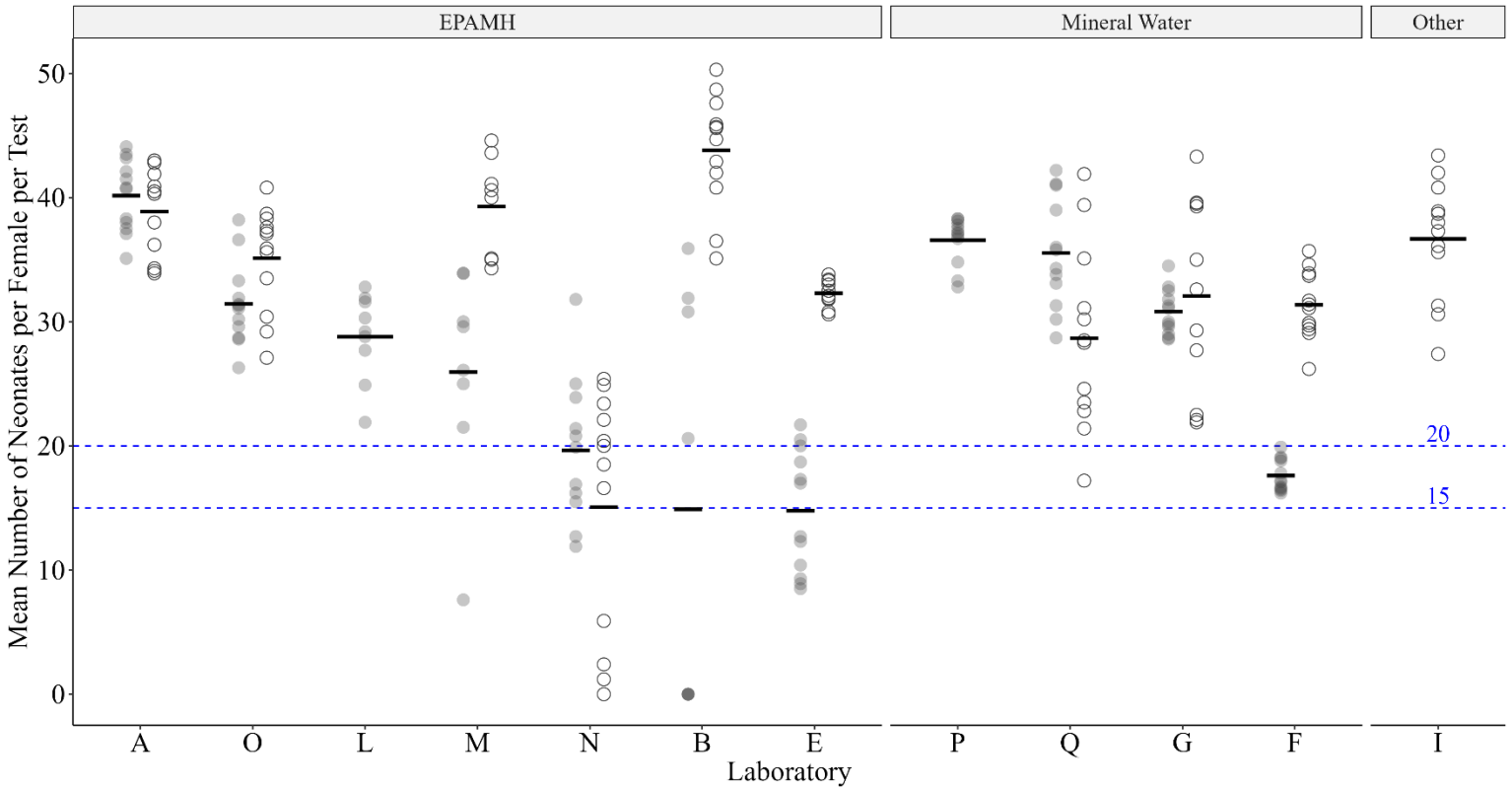
IC50, **Appendix B Tables B15 and B16**) among laboratories. Five laboratories had PMSD values ≤ 25 in over 75% of the tests, with no evidence of bias relative to water type (**Figure 5-7**). This indicates that half of the laboratories participating in the ILS can detect a 25% difference. Lab F and O had between 50 and 70% of their data within the proposed threshold, indicating a good performance. It should be noted that all these laboratories use different dilution water recipes for their own cultures and test controls (**Appendix A Table A2**). The remaining four laboratories had high PMSD above 37 in most samples. Of note, Lab N had PMSD > 100 for the SCCWRP-prepared serial dilutions tested in rounds 1 and 2. This laboratory reported culture issues. Lab B and M had PMSD values between 44 and 53, and 15 and 64. respectively. These same laboratories had high CVs for reproduction. Because the baseline ILS data represents a small sample size and a short period of time, the implication of these results should not be over-interpreted.

To address stakeholders' concerns and evaluate inter-laboratory agreement, statistical analysis of toxicity potency estimates was conducted using the percentile ranking method. The Panel applied this approach for the IC50s and used the interquartile range, the 25th to 75th percentile from the baseline ILS, as lower and upper bounds to characterize data deemed comparable (**Figure 5-8**). IC50 values for 3 out of 11 laboratories fell within that range over 75% of time, including historically high performing laboratories, Lab A and Q. There were no obvious differences in the calculated IC50s by water type. Lab F, G, L and P met expectations with over 50% of the IC50 values falling within the 25th and 75th percentile. Lab B, E, N and O produced variable (up to 2 order of magnitude difference) IC50s among testing rounds with at least 4 out of 6 tests falling outside of the 25/50th percentile range. Lab B, N and O exhibited data characteristics that differ from the level of consistency and comparability characterized in the historical datasets, while Lab E performance remained consistent as their historical IC50s were the least comparable to other historical IC50s for laboratories that used CuCl as their reference toxicant. The percentile approach and proposed frequency of attainment seemed appropriate to identify exceptional, good, and subpar laboratory performances. However, larger datasets normally distributed may benefit from tighter upper and lower bounds. While the Panel focused on IC50s that tended to be less variable, data quality programs may consider applying the percentile ranking approach to assess comparability of IC25s or LC50s.

Unspiked samples (Samples 1 and 2A) tested during the baseline ILS showed variability in neonate reproduction that was often consistent with the variability in the laboratory's own control samples (**Figures C2-C5**). Laboratories with poor performance also had documented issues with cultures, test brood boards or other technical issues described previously. Further evaluation of the data was conducted using the control adjusted mean reproduction data, calculated as follows: sample mean/control mean x 100 (**Appendix C, Tables C7 and C8**). Using a conservative estimate, control adjusted values greater than 90% were considered not different from the control. The analysis revealed that 4 out of the 11 laboratories participating in the baseline ILS accounted for

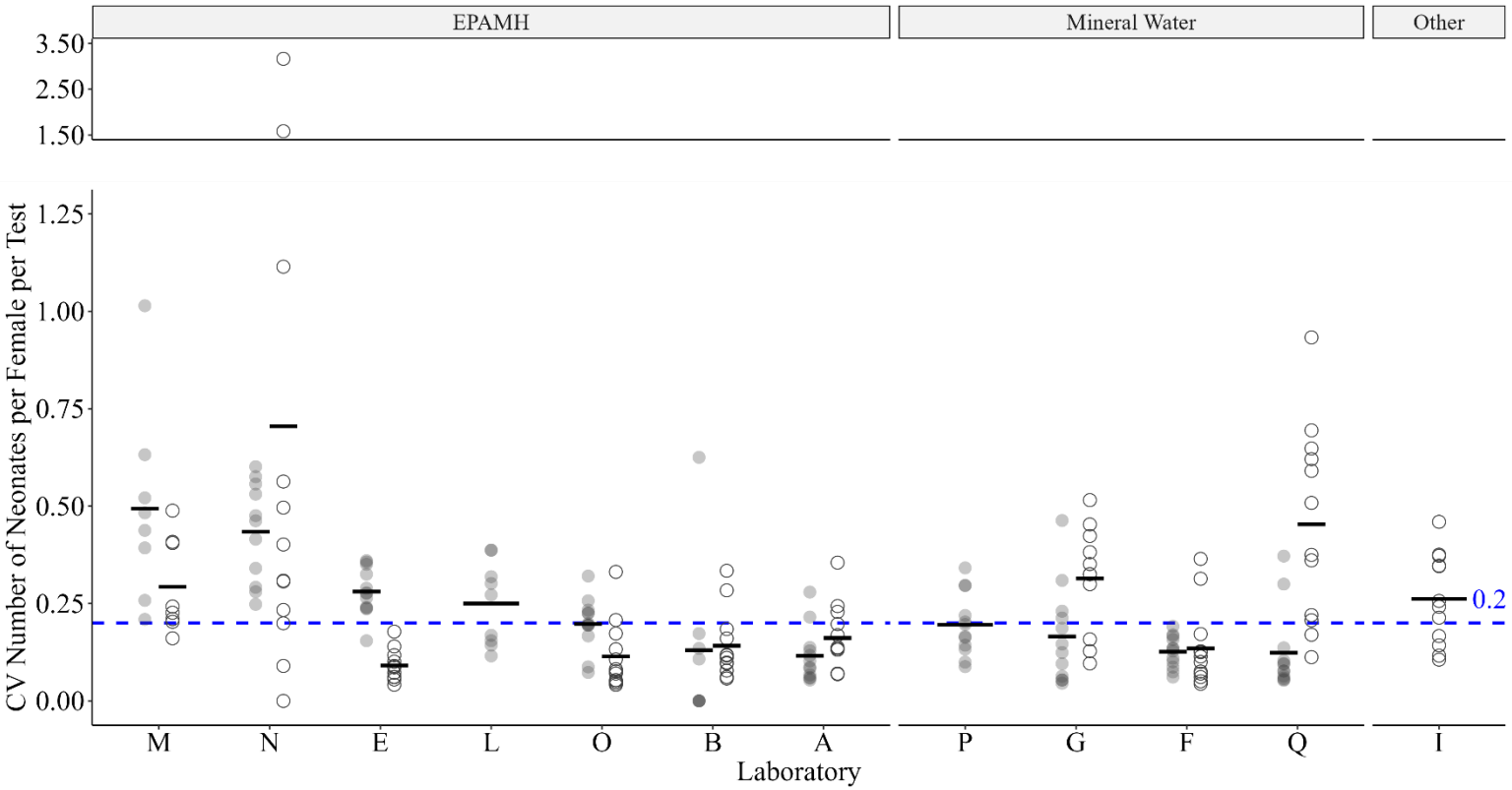
more than 75% of the instances where the adjusted mean was less than 90%. The adjusted control means were between 14 and 142% for the two samples. The only laboratory reporting toxicity consistently was Lab N, which did not perform well in this exercise. Lab N had high CV for control reproduction, high PMSD and low IC50 comparability (See analysis of baseline data above). Lab L had an average adjusted control mean of 90.3% with a range of responses between 76 and 105%. The results suggest that the differences among water types are a function of laboratory performance and not the test method.

Figure 5-5. Mean number of neonates per surviving female in laboratory controls for the baseline and second split-sample exercises. Data is organized by the type of dilution water the laboratory uses in their controls. Laboratories are ordered highest to lowest mean value of the baseline ILS for each laboratory control water type. Closed symbols are for the Baseline and Open are for the second ILS. The black bars are the mean values for each ILS.



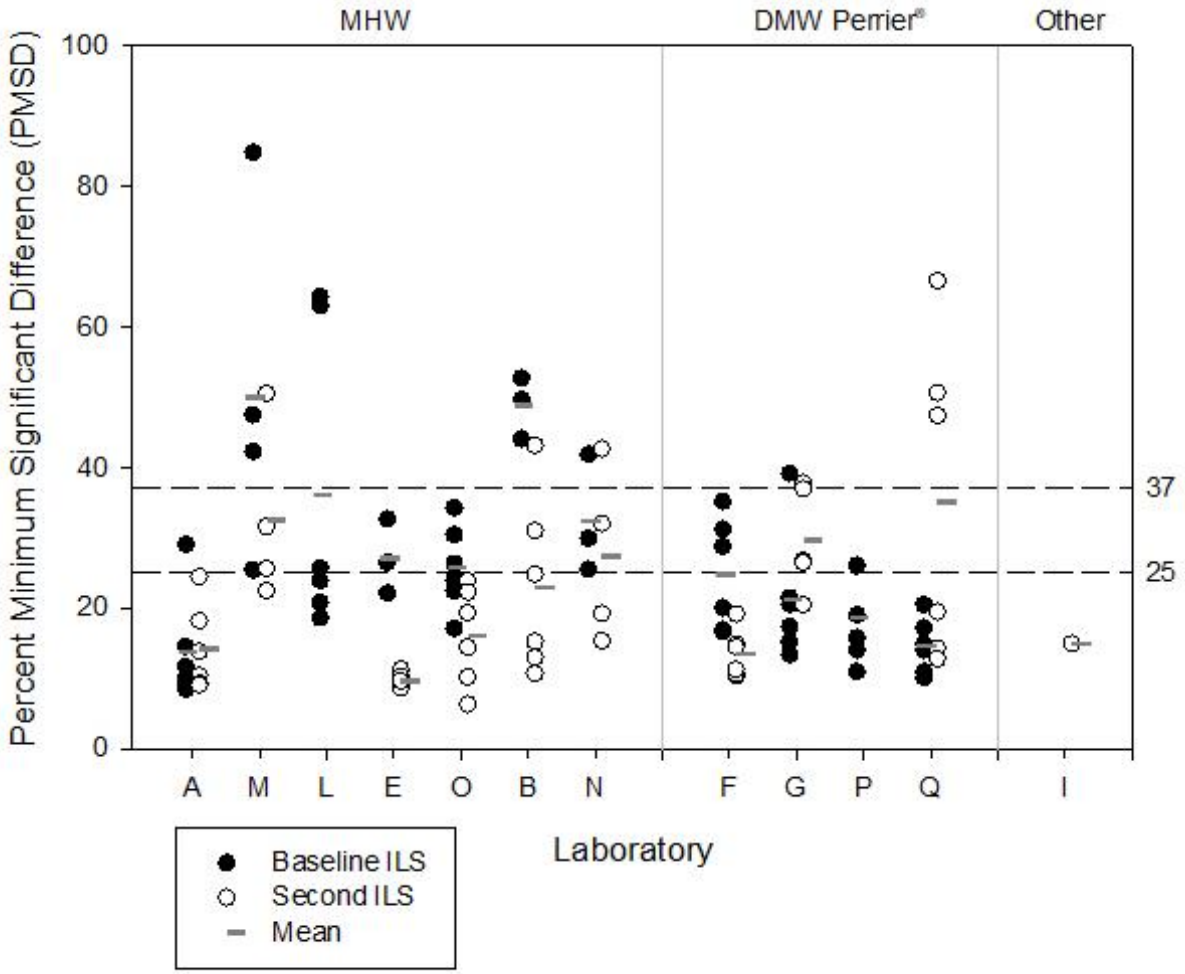
Note: During the baseline ILS Lab B experienced culture crash and Lab M reported high mortality in the brood boards. During the second ILS, Lab Q cited culture issues.

Figure 5-6. Coefficient of variation for the mean number of neonates per female in laboratory controls for the baseline and second split-sample exercises. Data is organized by the type of dilution water the laboratory uses in their controls. Labs are in the same order as mean neonate plot. Closed symbols are for the Baseline and Open are for the second ILS. The black bars are the mean values for each ILS.



Note: During the baseline ILS Lab B experienced culture crash and Lab M reported high mortality in the brood boards. During the second ILS, Lab Q cited culture issues.

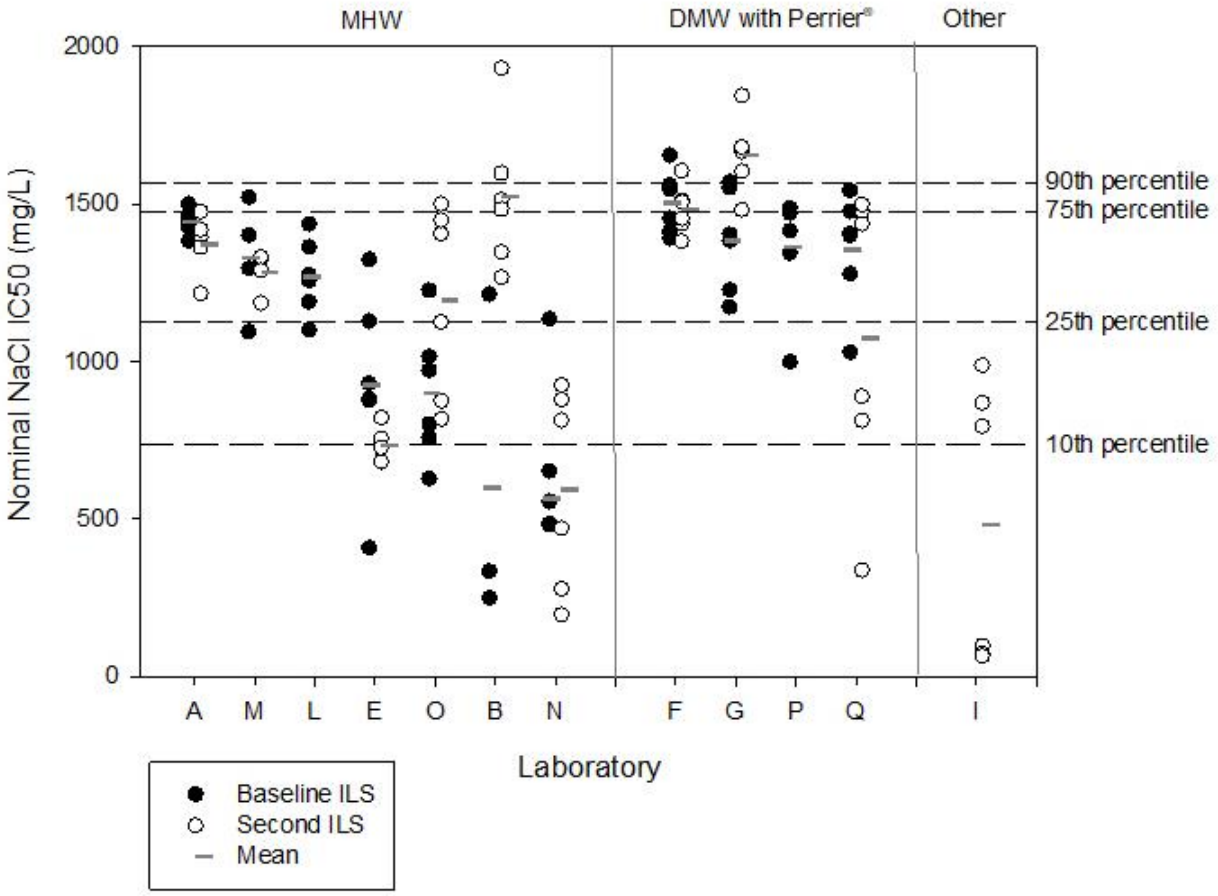
Figure 5-7. Percent minimum significant difference (PMSD) values from both concentration series and both ILS. Reference lines indicate proposed laboratory performance criteria. The dotted lines at 25 and 37 are the estimated 50th and 90th percentile for the current study.



Note: During the baseline ILS Lab B experienced culture crash and Lab M reported high mortality in the brood boards. During the second ILS, Lab Q cited culture issues.

During the second ILs, Lab I had high mortality with all SCCWRP-provided tests. Thus PMSD could not calculated.

Figure 5-8. Nominal IC50 values from both Sample 2 and Sample 3 concentration series from both ILS. The reference lines are based on percentiles of the entire data set.



Note: During the baseline ILS Lab B experienced culture crash and Lab M reported high mortality in the brood boards. During the second ILS, Lab Q cited culture issues and Lab I had high mortality with all SCCWRP-provided tests. Thus, IC50s could not calculated.

Investigating sources of variability in test outcomes

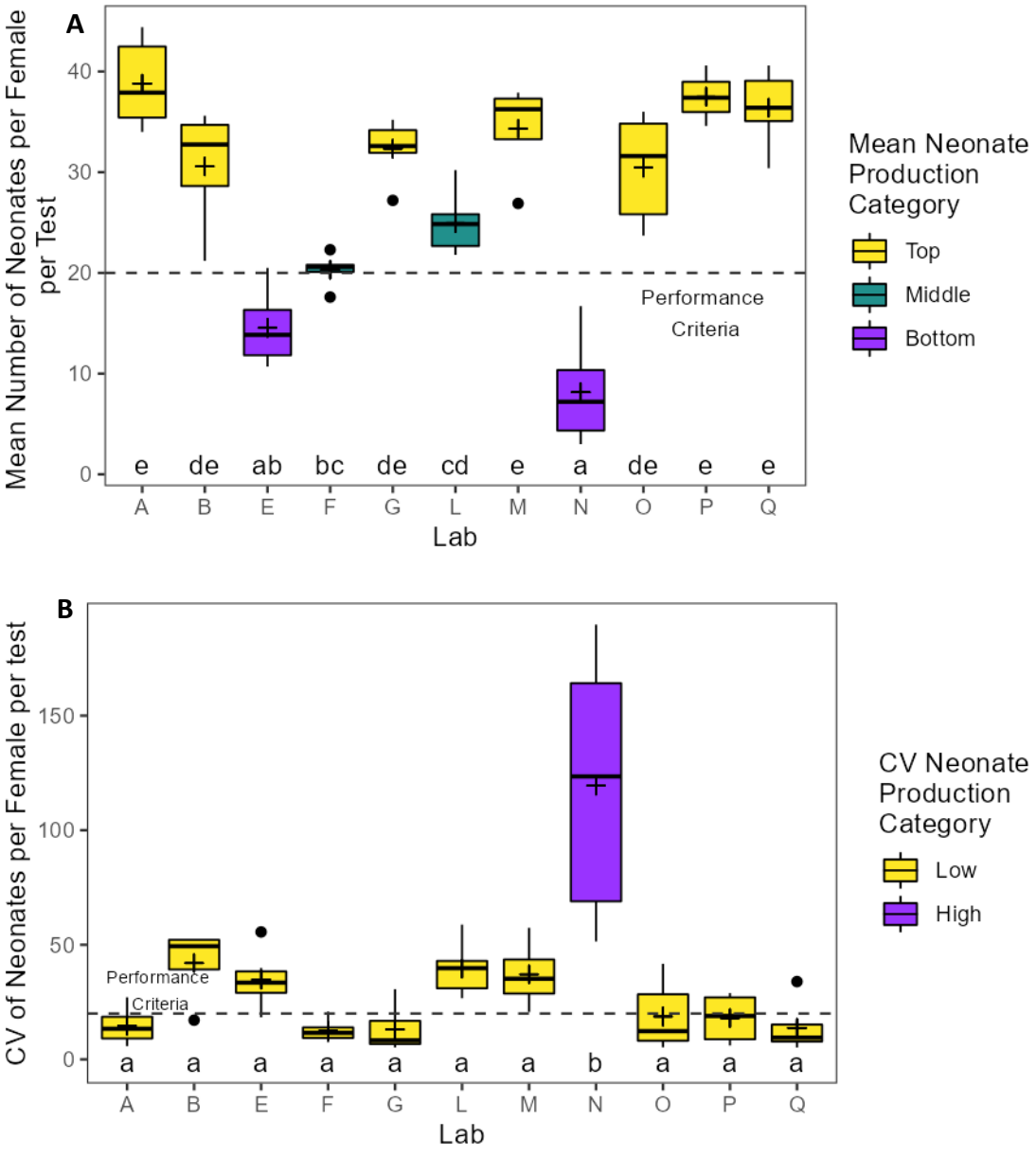
Patterns in control neonate production of Samples 1 and 2A indicated relatively clear groupings of high performing (i.e., high means and low CVs) and low performing (i.e., low means and high CVs) laboratories (**Figure 5-9**). These groupings were quantified using 2-way ANOVAs using laboratory identity and water type (MHW and DMW) as explanatory variables. Water type was never a significant factor (α=0.05), whereas laboratory identity consistently was a meaningful predictor of both the mean and CV of control neonate production.

Based on the outcomes of the ANOVA analyses, measurements that best captured the techniques and test/husbandry conditions of each laboratory – water composition, test water quality, brood board biological characteristics, brood board water quality, and feeding regimes were investigated as potential influential factors in the differences of neonate production among the laboratories. Variable influence was estimated using random forest regressions, with either mean neonate production or CV of neonate production from Samples 1 and 2A as the response variable. The random forest models indicated that the age of the females in the brood board, brood board water quality (e.g., temperature, pH, conductivity), and test water quality (dissolved oxygen and pH) were important factors in mean neonate production for Samples 1 and 2A (**Appendix B Table B29**).

The outputs of the random forest models supported the preliminary conclusions that could be drawn from the lab-to-lab ANOVA analyses, i.e., that practices and conditions inherent to each laboratory were the most likely influences on laboratory performance. Follow-on linear and logistic regression analyses indicated that the younger the age of the females used to initiate a given test the greater the likelihood of test controls to have higher neonate production with lower variance (**Appendix B Figures B33-36**). There were also some indications that food quality – specifically the age of the YCT – might be influencing test performance. The linear and logistic models did not indicate that test water quality and brood board water quality were important in explaining test performance.

Taken together, the ANOVA, random forest, and regression model results did not provide a clear-cut indication as to the cause(s) of lab-to-lab variability but helped the Panel and the Stakeholder Committee to refine the laboratory techniques that could be standardized in the second ILS, including age of the females producing the neonates to start the test and feeding regime.

Figure 5-9. Schematic box plots showing mean (A) and CV (B) of control neonate production from Samples 1 and 2A (i.e., MHW and DMW with Perrier®) among the different labs in the baseline interlaboratory comparison study. The plus symbol indicates the mean value. The letters across the bottom indicate the post-hoc comparisons of each laboratory based upon least-square means Tukey comparisons of the 2-way ANOVA (response=lab|water type). The colors of the bars indicate groupings based upon visual interpretation and the post-hoc comparisons. Colors are not an indication of laboratory performance.



5.6. Data analysis for the second interlaboratory comparison study

A total of nine laboratories (8 private and 1 public) participated in the second ILS. Two laboratories, Lab L and P, that participated in the baseline ILS did not participate in this second exercise due to scheduling and staffing issues. Lab N could only complete two out of three rounds, also due to scheduling and staffing issues. Data were analyzed to evaluate laboratory performance and compare the findings to the baseline ILS and historical data. The goal of this second ILS was to determine whether laboratories producing inconsistent data would show an additional improvement relative to increased control of test procedures. Rather than strict multiple pairwise statistical testing, the Panel evaluated the magnitude of change as a preferable means of determining improvement, if any, between the baseline and second ILS.

All but one laboratory met the TAC for reproduction and recorded ≥ 20 neonates per surviving female. Similarly, all participating laboratories met the TAC for survival. Nonetheless, this is an improvement from the results of the baseline ILS where three laboratories did not meet one or both TAC (**Figure 5-5**). The calculated CV of neonates per female also showed a higher frequency of laboratories achieving a CV ≤ 0.2, and Lab B, E and O showed some improvement compared to the baseline ILS (**Figure 5-6**). Lab G, Q and N exhibited a level of variability that was atypical of their performance in the first ILS and/or in their historical data. Among them, two laboratories reported culture issues prior or during testing or change in personnel. Consistent with previous data analyzed in this project (i.e., baseline and first ILS), Lab A maintained the same reliable level of performance.

Four out of 10 laboratories had PMSD values ≤ 25 in 5 or 6 out of 6 serial dilutions tested, Lab A, F, O and E (*Figure 5-7*). These four laboratories also met TAC consistently and had average CVs ≤ 0.2. This was an improvement in both PMSD values and frequency of attainment for three of these laboratories compared to the baseline ILS. Conversely, PMSD could not be calculated for Lab I due to high mortality in sample 2 and sample 3 testing, despite their laboratory controls meeting the TAC for survival and reproduction. Lab I had a culture crash during the first ILS and did not participate. Lab I also reported high mortality in the brood boards used to set the test in the second ILS. Lab M and B showed improvement in test sensitivity as more than 50% of the PMSD were ≤25. Lab G underperformed for this metric despite good performance according to the historical and baseline ILS data analysis, while Lab N results remained similar to those from the baseline ILS.

Comparisons of the IC50s showed that Lab A and M exceeded expectation with over 75% of tests within the 25/75 percentile. Lab F continued to perform well, with > 50% estimated IC50s within the 25/75th percentile (Figure 5-8). Lab B did show a modest improvement in IC50s distribution

as well as mean CV for reproduction and PMSD relative to the baseline ILS. Among the remaining laboratories with high PMSD, i.e., >25% in ≥ 50% of the tests, Lab E, O, and N did not show any improvement for this metric compared to the baseline ILS while lab Q underperformed.

Analysis of the unspiked samples (Samples 1 and 2A) showed modest improvement in inter-laboratory comparability, although some laboratories remain variable as discussed above. Evaluation of the control adjusted means showed that only two laboratories (B and I) accounted for 75% of the means below 90% (**Appendix C Tables C7 and C8**). These laboratories also had documented issues during testing. Similar to the findings from the baseline sample analyses, the variability in response to the unspiked samples is likely due to individual laboratory performance.

Unfortunately, there were no specific standardized techniques (feeding techniques, age of females producing neonates to start the test, that showed a statistically meaningful correlation with the observed improvement. However, implementation of the recommended practices, including consistency in feeding regime through verification of food density in stock bottles and estimation of food in test cup, are suspected to influence test outcomes, and may have contributed to the overall improvements observed in control performance.

5.7. Conclusions

Detailed analysis of the historical data and results of the two ILS showed that six out of the 12 laboratories that participated in the project can perform the *C. dubia* reproduction test and meet test acceptability criteria consistently. Lab A stood out as one of the most consistent laboratories throughout. Lab F, P and Q were also considered high performing laboratories and exceeded expectations for over half of the datasets evaluated. Standardization of laboratory techniques may not improve the performance of these laboratories, but some of the Panel recommendations in the Accreditation and Training categories will improve tracking and documentation of their performance over time. Lab L and O also met the Panel's level of expectations. The other six laboratories exhibited greater variability as evidenced by their frequency of meeting the proposed metrics for CV for neonate production, potency endpoints and/or PMSD. Their overall performance was characterized as acceptable. It should be noted that two of these laboratories underperformed in one of the ILS because of brood board health issues, while the other three reported technical errors impacting dilution water or test maintenance during the ILS. It is understandable that laboratory performance can vary with time, thus the findings are not unusual or alarming. However, this project shows that frequent assessment of laboratory performance is important to understand and remediate any issues that may arise. One laboratory, Lab B, showed clear improvements in performance over the course of this project. This is evidenced by their ability to meet the proposed performance metrics more consistently in the second ILS compared to their historical data and the first ILS data. Overall, implementation of standardized practices (e.g., detailed brood board health assessment, high quality source water, quantification of food) and performance metrics put forward by the Panel should help improve consistency and ensure that laboratory performance is comparable across California accredited laboratories.

6. Panel's Recommendations to Improve Laboratory Performance and Comparability

The recommended guidance falls into three categories: Best Practices, Accreditation, and Training. These were formulated based on assessment of the overall performance of the laboratories (using both historical and ILS data) and observations during site visits and the roundtable workshop.

6.1. Best Practices

The *C. dubia* EPA promulgated method allows for flexibility in select laboratory techniques to facilitate animal handling and retain necessary flexibility in testing local environmental samples using nationwide guidance. Laboratories can choose the water recipe used to culture the organisms and dilute test samples, food source, age of neonates to start the test (within a 24 h window), etc. This was reflected in the evaluation of the laboratory SOPs and roundtable discussions which revealed that no two California accredited laboratories conduct tests with exactly the same laboratory practices (**Table 5-1** and **Appendix A**). However, there are test parameters that are requirements of the methods which are not flexible, and this study also found that these requirements can be interpreted differently among laboratories. For example, not all California accredited laboratories follow the requirement to end the test after 60% or more females have produced three broods (**Table 6-1**) and some purposely wait for 70% to 80% of the females to produce 3 broods. One laboratory conducts a standard 7-day test regardless of brood status. Such practices could influence toxicity determination when testing environmental samples. To clarify the requirements of the test method and provide additional guidance to improve consistency and comparability, the Panel is providing two sets of guidance. The first set are considered "must do's" and are requirements of the method with the rationale provided in EPA documents. The second set are suggested "should do" recommendations as presented in the EPA manual that may improve laboratory performance.

Below are the "must do" recommendations.

Recommendation #1: Terminate the test when 60% of surviving females in the controls have had three broods, within a 2-h window (i.e., +/- 1 h of test initiation time).

Termination of the *C. dubia* test when 60% of the females have produced three broods is one of the test acceptability criteria but not all laboratories follow this requirement (**Table 6-1**). While

the test duration is not specifically prescribed in the method manual, during the Panel's laboratory visits and data compiled during phone interviews, it was noted that some laboratories are intentionally delaying test termination either to meet test acceptability criteria for control reproduction, to allow delayed reproduction in test concentrations to catch up with reproduction in controls, or to accommodate the test breakdown within the lab's schedule. These practices can mask toxicity that is expressed through delayed reproduction. It also produces a source of variability between laboratories, which this study was intended to reduce. For example, analysis of the NaCl-spiked samples tested during the two ILS shows how test termination trigger can affect the estimated inhibitory concentrations (**Figure 6-1**). As the method states in Section 13.10.9.1, test termination must be completed when 60% of the females or more have produced three broods. It is the Panel recommendation that such a decision must be made daily in 24-h increments, within a 2-h window of test initiation time.

While the time window is not specifically stated in the method, this is assumed throughout the manual. In multiple instances, "daily" renewal and test "days" are also referred to as 24-h periods. Section 8.5.4 states that the use of samples for static renewal tests be at 24 h, 48 h, and/or 72 h after first use. Section 13.10.6.1.2 also states that daily routine water chemistry must be measured at the end of each 24-h exposure period. This infers that daily water changes are performed within a reasonably narrow and fixed window each day. It should also be noted that determination of acute endpoints (e.g., lethal concentration at 24 h or 48 h) requires a precise exposure window (section 13.1.2) which further implies daily checks and test termination at 24-h intervals. Finally, if sufficient toxicity occurs, the 24, 48, and 96 h LC50's could be reported for the test as well.

Table 6-4.Inventory of laboratory techniques extracted from the laboratory SOPs and phone interviews.

Lab	Water recipe	Feeding method	Test termination; % females having 3 broods	*Test termination window*
A	MHW + vitamins + Se	In test solution	≥60%	*Strict window with single check*
B	Modified MHW	In test solution	≥60%	*Strict window with single check*
C	HW + vitamins + Se	Not provided	≥60%	*No specific window with periodic checks*
D	MHW + Se	Not provided	≥60%	*Strict window with single check*
E	MHW	In test cup	None	*Test always runs for seven days*
F	80% DIW: 20% Perrier®	In test cup	≥60%	*No specific window with single check*
G	80% DIW: 20% Perrier®	In test cup	≥80%	*No specific window with periodic checks*
H	80% DIW: 20% Evian®	Not provided	≥70%	*No specific window with periodic checks*
I	Hoheisel* +vitamins + Se	In test cup	≥60%	*Strict window with single check*
J	Not provided	Not provided	Not provided	Not provided
K	L1650% + vitamins + Se	Not provided	≥60%	*Strict window with periodic checks*
L	MHW + vitamins	In test cup	≥60%	*Strict window with periodic checks*
M	Modified MHW + vitamins	In test cup	≥60%	*Strict window with single check*
N	MHW + Se	In test cup	≥60%	*Strict window with single check*
O	MHW + vitamins + Se	In test solution	≥60%	*No specific window with single check*
P	80% DIW: 20% Perrier®	In test solution	≥60%	*No specific window with periodic checks*
Q	80% DIW: 20% Perrier®	In test cup	≥60%	*Strict window with single check*

Abbreviations: MHW = EPA moderately hard water; HW = EPA hard water; Se = selenium; DIW = deionized water. *Hoheisel et al. (2011)

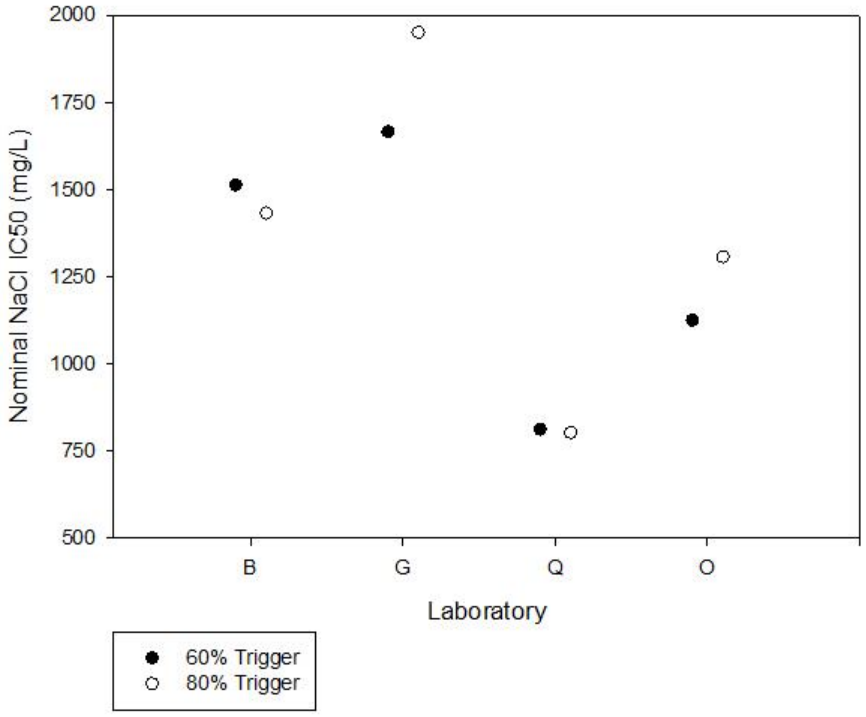


Figure 6-9. Example of differences in IC50s for tests ended at 60% (●) versus 80% (○) of surviving females having produced three broods. Data was collected during the two interlaboratory comparison studies for this project.

Recommendation #2: Independently quantify food concentrations in stock bottles and record amounts added to each test container.

The effects of food quality and quantity on *C. dubia* reproduction and toxicity test results are well documented (Cerda and Olive 1993; Norberg-King and Schimdt 1993; Jorgenson et al. 2017; Prosser et al. 2018). To support healthy cultures and organisms, the EPA manual provides directives describing the density of YCT and algae to feed daily (see Section 13.10.5). In this study, several practices were uncovered that are not consistent with the manual's directives. Four laboratories indicated not quantifying the density of the stocks used during the test. Interestingly, three of these laboratories (B, E and M) had a significant increase in neonate production in the controls in the second ILS (**Figure 5-5**) which required them to independently quantify the food and estimate the food concentrations in test cups. Other documented practices that are not compatible with the test requirements include filtration of the YCT solution to remove suspended solids. Labs using such technique reported not re-adjusting the density before feeding. To ensure

that the density of YCT and algae is within range, laboratories must verify the density of each batch purchased or produced before use and record the volume dispensed to confirm the targeted concentrations in test cups. This is particularly important for laboratories that provide food for the organisms in newly-prepared test solutions instead of adding the food to individual test cups.

Recommendation #3: Use source water produced according to the requirements of the EPA freshwater WET test methods and measure resistance to confirm ongoing water quality.

C. dubia is sensitive to water quality, making it a good indicator species to assess toxicity. To maintain good water quality in the cultures, EPA 2002a provides guidance on the purity of the source water. Section 7.2.2.2 recommends the use of four cartridges to produce deionized source water: (1) ion exchange, (2) ion exchange, (3) carbon, and (4) organic cleanup. This should be followed by a final filtration step. During this study, requests for information on source water treatment systems revealed that laboratories use a variety of water treatment approaches (**Table 6-2**), and few of them are as recommended in the manual. It was noted that most laboratories do not include the filtration step at the end of the treatment train to remove bacteria. Discussions with the laboratories also revealed that most laboratories do not conduct and/or document routine maintenance of their system. Section 5.4.2.1 of the manual requires that laboratories have high quality deionized water providing a resistance of 18 megaohm-cm, also known as Type 1 water (ASTM 1999). However, most of the laboratories visited did not have continuous monitoring of the water resistance to verify that source water consistently meets high standards and did not have a record of the values to refer to. Laboratories must continually or routinely (daily) measure and record the resistance of source water to confirm that it is > 18 megaohm-cm and suitable for use in preparing dilution waters.

Table 6-5. Water treatment methods for laboratories that participated in the study. This was collected during the second ILS and data is not available for the other laboratories.

Lab	Source Water	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Final Water Type
A	Municipal	Carbon cartridge	Ion exchange	Ion exchange	Filtration	UV disinfection	N/A	1
B	Municipal	Carbon cartridge	Ion exchange	Ion exchange	Other organic clean-up	Other organic clean-up	N/A	1
E	Municipal	Filtration	Ion exchange	Carbon cartridge	N/A	N/A	N/A	1
F	Municipal	Filtration	Carbon cartridge	Reverse osmosis	Ion exchange	N/A	N/A	2
G	Municipal	Filtration	Carbon cartridge	Reverse osmosis	Ion exchange	N/A	N/A	2
I	Municipal	Carbon cartridge	Reverse osmosis	Ion exchange	Filtration	UV disinfection	N/A	1
M	Municipal	Ion exchange	Carbon cartridge	Filtration	UV disinfection	Disinfection: 0.2 µm filter	N/A	1
N	Municipal	Filtration	Carbon cartridge x3	Filtration	Reverse osmosis	Ion exchange	UV disinfection	2+
O	Municipal	Carbon cartridge	Water softener	Filtration	Reverse osmosis	Ion exchange	N/A	1
Q	Arrowhead Distilled Water	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Recommendation #4: Randomize test cups in the test chamber.

Randomization of test cups on the test boards is a requirement of the test method (section 13.10.2.2) and laboratories must document the set up used for each test. The EPA manual requires the known parentage and provides a procedure for placing test concentration on the board using a stratified random procedure. While most laboratories randomly assign organisms in test cups using blocking by known parentage (section 13.10.2.4), the Science Panel observed that few laboratories randomized the placement of the cups in the exposure chamber. The impact of randomization on test outcome cannot be statistically demonstrated, but randomization is a requirement of the statistical approaches used to analyze test data. This laboratory technique is part of good laboratory practices to avoid any bias due to light, temperature, or other gradients in the exposure chamber. Generating a number randomization templates for the test concentrations can make setting up and transferring the animals easier (cf., EPA 2002a).

In addition to clarifications of the method requirements, the Panel recommends the following "should do's" best practices. These were developed based on findings of the present study.

Recommendation #5: Conduct a detailed quantitative assessment of brood board health prior to testing.

Documenting brood board health prior to testing is a requirement of the method (13.6.16.11.1). However, the method does not specify the level of detail needed. Initial assessment of the historical data indicated that four out of 17 laboratories consistently documented the health of their brood board in detail, including counts of unhealthy adults and neonates, and the presence of males. Observations recorded by the other laboratories were typically limited to the presence or number of dead organisms. Due to the lack of data collected on brood board health, no meaningful statistical analysis of related historical data could be performed during this study. The Panel noted that three of the five laboratories who did not participate in the ILS cited issues with their culture prior to testing as a rationale. Therefore, the Panel recommends that laboratories track daily neonate production per female, mortality, number of males, and the size, appearance and movement of adults and neonates in the cultures. Such information provides an ongoing assessment of culture health and serves as an indicator of poor organism quality prior to starting the test. Bench sheets and training materials produced for the second ILS can be used for this purpose (**Appendix D**).

Recommendation #6: Document split brood on bench sheets daily at the time of the observations.

The *C. dubia* test relies on the identification of three separate broods for 60% of the females as an indicator to terminate the test. *C. dubia* has a rapid reproduction rate and can develop and release their brood daily. In some cases, females may release the neonates from a single brood on two consecutive days, which is known as a split brood. The determination of a brood is an important laboratory technique as it directly impacts when the test can be ended and the total number of neonates in three broods. Review of the laboratories' bench sheets revealed that some of them do not add any observations of the females or neonates during daily maintenance. The determination is conducted post hoc by the data analyst, largely based on numbers of neonates. This practice can lead to misidentification of a brood and affect mean neonate production. This was observed firsthand by the Panel during their site visit. In **Figure 6-2**, the top graphic shows a bench sheet of daily neonate counts without any observations, suggesting that each day constitutes a single brood and could underestimate neonate production (see total count for replicate #5 and #6). The bottom graphic shows the same bench sheet with observations of the females' appearance and neonates' size. These observations can be used to identify split broods and better estimate the size of a given brood. Therefore, it is strongly recommended that laboratories make notes of possible split broods on the datasheets at the time of the observations. Bench sheets and training materials produced for the second ILS can be used for this purpose **(Appendix D)**.

Recommendation #7: Renew test solutions daily within +/- 2 h (i.e., 4-h window) of test initiation time.

As described in Recommendation #1, the *C. dubia* test method is written with the expectation of daily observations and measurements in 24 h increments. Regular renewals can help maintain the integrity of the test and ensure comparable exposures of test organisms each day. Therefore, daily renewal of test solutions should be performed within a consistent window. While this study did not collect data to evaluate these impacts on test outcomes, the practice is consistent with other requirements of the test method, including water chemistry measurements. Further, water chemistries should be measured and recorded on day 0 and on the additional test days as provided in EPA's test methods errata (2016).

		#Neonates/Replicate									
	Day	1	2	3	4	5	6	7	8	9	10
Lab Control	1	0	0	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	3	0	0	0
	4	0	0	0	0	2	1	0	1	0	2
	5	4	3	5	0	2	4	3	2	2	2
	6	6	7	6	7	6	7	13	7	12	10
	7	9	10	13	14	14	15	6	10	7	6
	8	7	6	6	5	8	10	3	12	10	12
	Total	19	20	24	21	10	12	22	10	21	14
										Mean	**17.3**

		#Neonates/Replicate									
	Day	1	2	3	4	5	6	7	8	9	10
Lab Control	1	0	0	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	3	0	0	0
	4	0	0	0	0	2*	1*	0	1*	0	2*
	5	4	3	5	0	2*	4*	3	2*	2	2*
	6	6	7	6	7	6	7	13	7	12	10
	7	9	10	13	14	14	15	6	10	7	6
	8	7	6	6	5	8	10	3	12	10	12
	Total	19	20	24	21	24	27	22	20	21	20
										Mean	**21.8**

Day 4: Control replicate 6 and 10, females have neonates in pouches.
Day 5: Control replicate 5, 6, 8 and 10, neonates are larger than in the other replicates.

Figure 6-10. Reproduction bench sheet collected from a participating laboratory.

Recommendation #8: Upgrade laboratory documentation

Information from historical data collection, laboratory standard operating procedures, and Panel laboratory visits suggested that routine laboratory operations could be improved with more complete and readily available documentation. This includes documentation of reagent holding times and expiration dates, dilution water preparation and holding times, food preparation and holding times, and randomization templates. Balance logs should be maintained along with calibration records for the balance, and weights that are used to verify the balance is reading correctly. The record book should have a preparation date, the hardness, alkalinity, pH, along with an expiration date. The source of food that is purchased (individual components or prepared foods) should be recorded in a record book. Sometimes, such information was not properly documented, but more often, information was documented in a format or location that was not readily accessible to laboratory staff that were routinely conducting tests. It is recommended that information about reagent, dilution water, or food holding times be documented directly on the container as well as on data forms that are easily accessible to the laboratory staff. The

preparation and use of randomization templates can also make test randomization easier and less error prone.

Recommendation #9: Store reagents to prepare the dilution waters and the reference toxicant appropriately.

The salts used to prepare the EPA synthetic reconstituted water and the reference toxicant NaCl are moisture sensitive. To avoid moisture absorption and changes in the weight of the salts, it is recommended to store them in a desiccator as soon as they arrive in the laboratory. During the laboratory's visits, the Panel noted that these reagents were kept on the bench and not protected from humidity. The recipe for reconstituted water was developed to meet a specific range for water quality parameters, including hardness and alkalinity (section 7.2.3.1 of the EPA manual), and moist salts can impact the ionic composition of the reconstituted water. Similarly, the target ionic balance of the refence toxicant solution can be compromised if the salts used have a higher water content than expected.

6.2. Accreditation

California's Environmental Laboratory Accreditation Program (ELAP) has many aspects, but two stand out. The first is Proficiency Testing (PT) where split samples are sent to laboratories. The second is in-person laboratory audits where State staff visit each laboratory and observe lab activities against a checklist of requirements in the promulgated method.

The State of California currently uses the national PT program developed by EPA Discharge Monitoring Report–Quality Assurance (DMR-QA) study program, which consists of laboratories testing a single PT sample annually for the species and method required by the NPDES permit. An unknown sample is purchased from a third-party vendor and tested as a dilution series. Laboratories must ensure that WET test methods/procedures follow instructions from the PT Provider and EPA's promulgated WET test manuals. Laboratories report one test endpoint (concentration) for each DMR-QA WET test code required. Further, for laboratory performance quality assurance purposes only, the point estimation techniques that produce test endpoints such as IC25 are the preferred statistical approach used for calculating test endpoints for effluent chronic toxicity tests. However, laboratories choose the statistical approach that allows calculation of the test endpoint(s). All PT results are analyzed by the third-party vendor based on one of the toxicity endpoints required by the method: LC50, IC25, NOEC (NOEC-survival, NOEC-growth or NOEC-reproduction). The third-party vendor is also responsible for comparing the data to the other laboratories across the country for the same test method. This approach, however, is limited due to the lack of the supporting data generated to report the test result and is a missed opportunity to evaluate other metrics required in EPA method or by the State, and wide

acceptance criteria preventing the detection of important differences in performance across laboratories.

Laboratory audits ensure that test methods are conducted in a consistent manner among laboratories using a defined quality system. For some test procedures, the State may use third party assessors that may not have a strong understanding of the toxicity test methods. For other procedures, the State relies on a single assessor with decades of auditing experience. Evaluation of the accredited laboratory's SOPs during the current study indicated that some deviations from the promulgated method are documented (**Table 6-1**), and site visits of select labs by a subset of the Panel members observed additional inconsistencies with the promulgated method. These findings suggest that audit checklists may not be applied consistently across laboratories and/or that follow-up visits to ensure compliance are not conducted. It should be noted that no Panel member accompanied assessors on an audit, so the inconsistencies are not first-hand observations.

One key concern brought forward by the Stakeholder Advisory Committee is the interlaboratory agreement among ELAP accredited laboratories. Indeed, it is problematic to estimate the level of comparability under the current ELAP program as laboratories can use different dilution waters and reference toxicants in their own QA programs, and can test PT samples from different PT providers. Results of the baseline and second ILS showed that approximately half of the laboratories can produce consistent and comparable results, while the others experienced greater variability than desirable (**Figure 6-3**).

The following recommendations are largely directed at the State but may also have impact on the laboratories who may need to complete additional testing and reporting, as well as the clients who would potentially absorb some additional costs for the increased accreditation requirements.

Recommendation #10: Increase the number and/or frequency of testing to assess comparability among laboratories.

Currently, the only mechanism to assess interlaboratory comparability among laboratories seeking accreditation for the *C. dubia* chronic test, is through the national PT study conducted once per year. The study consists of testing one sample, with similar test conditions as the NPDES permit, generating one endpoint for each test and then comparing the results to the other laboratories across the nation. Because evaluation of PT results is based on data from the participating laboratories, falling within two standard deviations of the national average allows for more interlaboratory variability than may be acceptable to the State. Moreover, the PT study

provides a brief snapshot on laboratory performance but is not sufficient to address intra-laboratory consistency. The current project showed that data quality can vary from year to year and good performing laboratories can have uncharacteristically difficult periods (e.g., culture crash) (**Figure 6-3**).

To demonstrate that ELAP-accredited laboratories can produce comparable results, the Panel recommends analyses of split-samples for all laboratories seeking State accreditation. This could be achieved by participating in a bi-annual interlaboratory comparison study or increasing the frequency of PT sample testing. Another option for consideration would be to require all laboratories to use the same reference toxicant, which are typically tested by accredited laboratories to assess intra-laboratory precision. This approach has unique advantages because laboratories already run monthly reference toxicant tests, the data will be immediately available for the laboratory's own use in terms of QA attainment and will offer similar points of comparison across different laboratories for accreditation and procurement purposes. However, it is important to note that this may not constitute a true split-sample, like the national PT sample, since not all laboratories use the same dilution water. Regardless of the approach used, the Panel recommends that several non-compliance test metrics be evaluated and compared among laboratories such as control neonate production (mean and CV), IC25 and IC50, concentration-response evaluations, and the calculated PSMD.

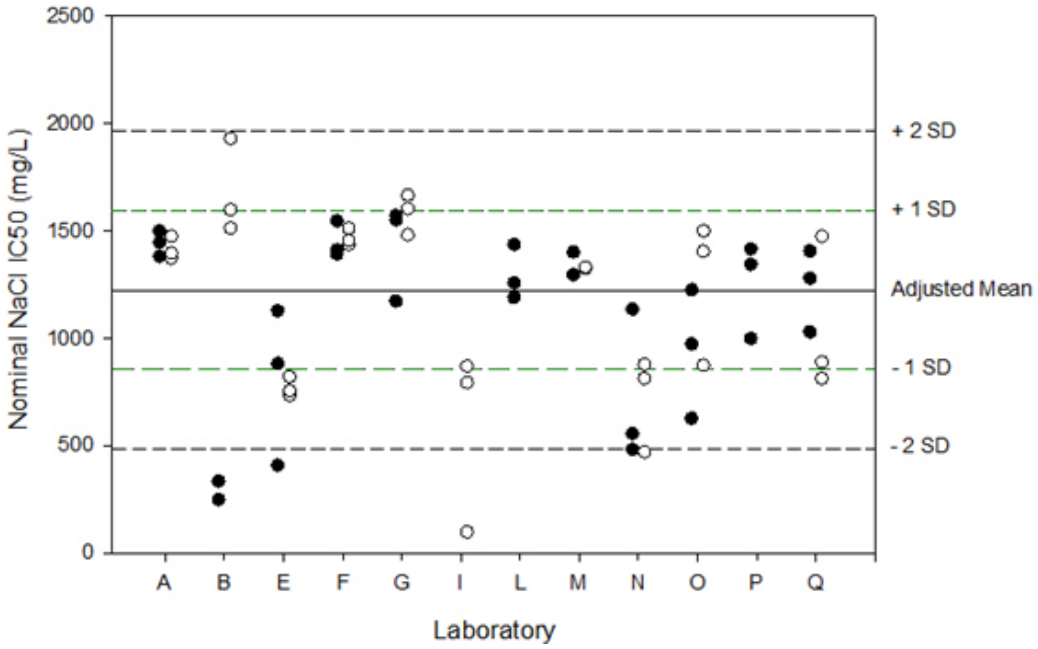


Figure 6-11. Comparisons of IC50s for NaCl-spiked split samples tested during the two interlaboratory comparison studies conducted in Fall 2022 (●) and Spring 2023 (○). All data are plotted (n=59) but only those that met test acceptability criteria (n= 48) were included in the calculations of the mean and standard deviation.

Recommendation #11: Collect and evaluate additional data associated with the PT sample.

Currently, there are three test results evaluated during a PT study: LC50, IC25, NOEC reproduction, NOEC (NOEC-Growth, NOEC-reproduction, NOEC-survival) (**Table 6-3**). These metrics represent only a small fraction of the valuable information that can be collected during PT sample testing. Submission of a data packet with the results of the comparison testing exercise will allow the evaluation of the mean and CV of neonate production in the controls and treatment concentrations, as well as the required evaluation of the concentration response, IC25, PMSD and water chemistry. Reviewing these test parameters proved useful during the two ILS. The data were used by the Panel to evaluate overall laboratory performance based on their historical data and results of the two ILS. Additional information should also be collected to assess replicate level control and treatment responses, brood board health and water quality (**Table 6-3**). A data framework has already been developed to accommodate these data types through the ILS during the current project and could be valuable in the future. Data submittal templates, data submittal portal, data quality assurance checkers, and pre-programmed data analysis routines have already been designed and utilized and are available to the State. The State will need to host these data submittal requirements and dedicate staff to troubleshoot data submittals and analyze/synthesize the data received.

Table 6-6. *C. dubia* chronic test endpoints to evaluate as part of proficiency testing.

	Endpoints
Current proficiency testing data	• LC50 • IC25 survival and reproduction • NOEC-survival and NOEC-reproduction
Additional data recommended by the Panel	• Water chemistry daily measurements (initial and final, daily) • Concentration response evaluation • Mean and CV of control neonate production. • IC50 survival and reproduction • PMSD • Ratio IC25/IC50 (reproduction)

Recommendation #12: Optimize laboratory audits to ensure effective and consistent implementation of best practices.

Laboratory audits aim to ensure that accredited laboratories are performing toxicity tests according to the promulgated method requirements and that data quality systems are implemented correctly. Thus, effective audits should be developed to assess critical laboratory techniques (e.g., dilution water preparation, feeding, test termination), and to review data quality procedures in place to maintain laboratory performance goals. This task requires significant time and effort by accreditation officials and is best performed by a trained team of auditors (see Recommendation #13). In the current study, the Panel identified some violations of the test method protocol that can be identified during laboratory audits. To facilitate laboratory audits, the Panel recommends that the State's accreditation checklist be updated to include acceptable practices (e.g., source water treatment, feeding regime, test maintenance) and a list of responses for each type of deviation. There may be other options for optimizing laboratory audits. For example, if there are some issues which are a common thread among laboratories, this could lead to identification of additional "best practices". Another example could be strengthening the review of the corrective actions performed by the laboratories in response to audit findings. Finally, if there are a subset of laboratories which are struggling to maintain their quality standards, these laboratories could be audited more frequently. Formalized approaches will benefit both the State and the laboratories by ensuring that laboratories are equitably and effectively evaluated.

6.3. Training

Although the test method has been in use since 1984, training on the method continues to be needed. EPA has training modules (http://www.epa.gov/water - WET training) and TNI has training on WET as well, but additional training may be needed with hands-on type of events. A workshop (EPA Region 9 and California SWRCB) was held just prior to the Expert Panel being convened to provide opportunities for the labs in California to have an open discussion about the *C. dubia* survival and reproduction test method to be able to discuss water conditions, food conditions, ending the test, split broods, data analysis and more. This was held in hopes that labs would be performing the test following the procedures discussed in the workshop. However, as staff changes occur within laboratories, regulators, and regulated entities, continuous access to training is needed to communicate to laboratories, regulators, and the regulated to make appropriate and informed decisions. Roundtable discussions and public meetings with stakeholders highlighted the need to provide training materials for an improved understanding of method requirements and data quality objectives. Training recommendations are directed at the

State, the testing laboratories, and at the regulated parties responsible for toxicity testing as a compliance requirement in the following areas:

- Implement training program with defined performance goals for all personnel involved in performing and/or reviewing the *C. dubia* test. This could include how the cultures are managed, the test procedure is performed, collection of effluent samples, reference toxicant testing, data review and analysis, reports that are filed, and PT testing requirements.

- Develop training documentation for standard testing and provide it to the testing laboratories.

- Implement auditors' training program. This type of training would augment auditors' knowledge of the method requirements. The training would aid the development and implementation of training performance goals for laboratory personnel.

- Regulated parties may not have the experience and knowledge to review the data or know how to select laboratories. Training is needed to provide guidance on how to evaluate and review WET test data.

Recommendation #13: Implement auditors' training program.

For effective auditing, State auditors must have a good understanding of the test method including test-specific requirements and general toxicology principles. The State currently has one official auditor responsible for visiting over a dozen labs for accreditation on half a dozen methods. As more auditors become involved, it will be important to ensure that their audits are conducted with similar rigor and attention to detail. A training program would serve to reinforce the key toxicology principles, review and optimize checklists, and develop a forum for ongoing discussion.

Recommendation #14: Implement training program with defined performance goals for all personnel involved in the *C. dubia* test.

As part of the historical data compilation and the two ILS, the Panel wanted to assess expertise and experience as potential variables leading to intra- and inter-laboratory variability. The assumption was knowledgeable and experienced laboratory staff will produce more consistent and reliable test results. Amongst the California accredited laboratories, anecdotal data compiled during the laboratory interviews illustrated some labs had long-term staff with decades of

experience while others had frequent staff turnover. However, not all labs had a pre-defined training program, few expressed any pre-defined criteria for performance goals, and no labs identified any formal re-testing or continuing education requirements. The greatest challenge, however, was that few of the laboratories documented their training and those that did so did not conduct or record their training in the same fashion. Ultimately, the Panel could not evaluate the effectiveness of expertise and/or experience for reducing intra- and inter-laboratory variability.

The Panel recommends a training program be created to support the laboratories. The primary objective is to create a training program for each level of activity in the laboratory from bench staff to laboratory managers, and that performance goals are associated with a demonstration of training (and re-training) competence. Moreover, the training should have appropriate documentation which is similar among laboratories so that regulated dischargers can compare training success across potential clients and auditors can utilize the information for troubleshooting accreditation challenges. The training program can be created by the State with assistance of experts in this area who can coordinate the training with a consortium of the testing laboratories, by the regulated dischargers that require well-trained laboratories for compliance assessment, regional trade associations (i.e., Society of Environmental Toxicology and Chemistry), the National Environmental Laboratory Accreditation Program Institute (TNI), or a combination of these groups.

Recommendation #15: Provide guidance to regulated parties to evaluate WET toxicity test data.

The regulated parties on the Stakeholder Committee repeatedly expressed concern about the variability of results within and among California accredited laboratories. However, the Panel noticed that some regulated parties had much more experience and expertise about toxicity testing than others. Some regulated dischargers have their own laboratories and are intimately familiar with the metrics that identify laboratories who produce consistent, high-quality data. In contrast, some regulated parties only have a single compliance officer who is addressing the myriad of facility compliance issues including water, air, and biosolids, amongst others. These compliance officers are typically not aquatic toxicologists, and aquatic toxicity is a very small part of their compliance requirements.

The Panel recommends providing guidance to regulated parties on fundamental toxicological concepts, WET testing in specific, and educating compliance assessment staff on the performance metrics that constitute a consistently well-performing laboratory. The goal of this guidance is to empower regulated parties to assess the quality of the laboratories they might wish to retain, and to more fully understand the output of the toxicity tests that determine their compliance. This

type of guidance, which can be conducted through a variety of media (written, video, in-person) need not be long or highly detailed but should be understandable and easily digested by non-toxicologists. Ideally, this guidance should come from a partnership of the regulated dischargers for whom the guidance is directed but can be done in collaboration with laboratories and/or the State.

7. CONSTRAINTS

While this study has produced more information on *C. dubia* inter-laboratory variation in more than 15 years, there are still a number of constraints to the conclusions and recommendations provided. These constraints fall into six categories. The first category is the constraints associated with the number of laboratories and the timing of the testing. While the study attempted to generate data for every accredited laboratory in California, 12 out of 17 laboratories participated in the ILS. During this two-year study, at least two laboratories let their accreditation lapse, and at least two laboratories did not provide complete historical datasets. Moreover, for the 12 laboratories who did participate, the magnitude and sources of variability observed during this two-year study may not be similar to variability (or lack of variability) observed prior to or following this study. Implementing the recommendation for increasing the number and/or frequency of proficiency testing samples could help address this limitation in the future.

The second category is the limited capability to quantify the individual variability for each of the nine standardized laboratory best practices. One important finding was the improvement for some laboratories and subsequent reduction in variability from the baseline ILS to the second ILS. The primary difference between the baseline ILS and the second ILS was standardizing the nine laboratory best practices in the recommendations. However, since all nine best practices were changed at the same time, we cannot quantify which best practice provided the most benefit and which provided the least, only the improvement cumulatively across all best practices. To provide this information on the variability associated with each individual best practice, a new study would need to be designed and implemented separating each best practice one at a time.

The third category is the limited capability to quantify variability associated with testing *C. dubia* in dilution water of varying hardness. The focus of the study was on implementation of test methods to improve consistency and comparability of test results. Therefore, this study was designed to quantify the variability using the default dilution water in the test method (section 4.4.1); MHW or DMW. However, stakeholders within the regulated community (CASA, CASQA, Agricultural) and public laboratories wanted to assess the intra- and inter-laboratory variability associated with testing *C. dubia* in HW which was outside of scope. The levels of variability may or may not be different in this atypical dilution water hardness. Thus, no recommendation is

provided for minimizing variability in the *C. dubia* reproduction test using HW. To address this limitation, a new study would need to be designed and implemented. However, it should be noted the test method does provide guidance for how to control variability when testing *C. dubia* in HW (section 7.1.1.).

The fourth category is the study's constraint regarding the reference toxicant tested. Quantifying intra- and inter-laboratory variability for the *C. dubia* reproduction test utilized NaCl for concentration-response in both ILS. This concentration-response data was critical in evaluating laboratory performance and consistency. However, the intra- and inter-laboratory variability may differ using other toxicants, especially for those that do not routinely use NaCl. While this element of the study plan was rigorously discussed and evaluated by the Science Panel and the Stakeholders, it was clear insufficient resources and time were available to test additional toxicants.

The fifth category was the timeline. While several extensions were granted, overall timeline and due dates impeded the Science Panel's process. For example, additional time would have provided the opportunity to refine laboratory performance metrics, including developing additional guidance to implement these metrics. These activities are encouraged to continue at the conclusion of the study by the State, the regulated dischargers, and the laboratories.

The last constraint of the study was guidance on implementing the recommendations. The Panel acknowledged that stakeholders would benefit from additional guidance on who and how, and the recommendations should be implemented. However, many of these implementation decisions have multiple options and different timelines that should be discussed with the laboratories and/or the regulated community in a thoughtful, effective, and efficient process facilitated by California-based decision makers, and not by a short-term Panel of experts who will not be in California beyond the measure of this study.

8. References

American Society for Testing and Materials (ASTM). 1999. Standard practice for conducting an interlaboratory study to determine precision of a test method, E691-20. In: Annual Book of ASTM Standards, Vol. 14.05. Philadelphia, PA.

Biau, G. and E. Scornet. 2017. A random forest guided tour. Test 25: 197-227.

Breiman, L. 2001. Random forests. Machine Learning 45: 5-32.

Briden, A.M., S.L. Clark, B.C. Jorgenson, R.S. Ogle, and J. Cotsifas. 2017. *Ceriodaphnia dubia* chronic toxicity test variability: Evaluation of water as a source of test variability. Society of Environmental Toxicology and Chemistry North America 38th Annual Meeting. Minneapolis, MN.

Clark, S.L. and A.M. Briden. 2018. Effect of Ceriodaphnia dubia culture hardness on laboratory control test performance. Society of Environmental Toxicology and Chemistry North America 39th Annual Meeting. Sacramento, CA.

Denton, D.L., J. Diamond, and L. Zheng. 2011. Test of significant toxicity: a statistical application for assessing whether an effluent or site water is truly toxic. Environmental Toxicology and Chemistry 30: 1117-1126.

Denton, D.L., J.F. Fox, J.F., and F.A. Fulk. 2003. Enhancing toxicity test performance by using a statistical criterion. Environmental Toxicology and Chemistry 22: 2323-2328.

Diamond, J., P. Stribling, M. Bowersox, and H. Latimer. 2008. Evaluation of effluent toxicity as an indicator of aquatic life condition in effluent dominated streams: A pilot study. Integrated Environmental Assessment and Management 4:456-470.

Elphick, J.R.F., K.D. Bergh, and H.C. Bailey. 2011. Chronic toxicity of chloride to freshwater species: Effects of harness and implication for water quality guidelines. Environmental Toxicology and Chemistry 30:239-246.

EPA. 2000a. Understanding and accounting for method variability in whole effluent toxicity applications undert the National Pollutant Discharge Eliminatin System Program. Office of Wastewater Management. EPA 833-R-00-003 .

EPA. 2000b. Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136). Office of Water, Washington, DC. EPA 821-B-00-004. July 2000

EPA. 2001a. Final report: Interlaboratory variability study of EPA short-term chronic and acute whole effluent toxicity test methods. Volume 1. Office of Water, Office of Science and Technology. EPA 821-B-01-004.

EPA. 2001b. Final report: Interlaboratory variability study of EPA short-term chronic and acute whole effluent toxicity test methods. Volume 2. Office of Water, Office of Science and Technology. EPA 821-B-01-005.

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.

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

EPA. 2002c. Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms. 5th edition. Office of Water, Washington, DC. EPA-821-R-02-012.

EPA. 2016. Whole effluent toxicity methods errata sheet. Office of Water. EPA821-R-02-012-ES.

Erickson, R.J., D.R. Mount, T.L. Highland, J.R. Hockett, D.J. Hoff, C.T. Jenson, T.J. Norberg-King, and K.N. Peterson. 2017. The acute toxicity of major ion salts to *Ceriodaphnia dubia*: II. Empirical relationships in binary salt mixtures. Environmental Toxicology and Chemistry 36:1525-1537.

Fox, J.F., D.L. Denton, J. Diamond, and R. Stuber. 2019. Comparison of false-positive rates of 2 hypothesis-test approaches in relation to laboratory toxicity test performance. Environmental Toxicology and Chemistry 38:511-523.

Hoheisel, S.M., R.J. Erickson, T.L. Highland, J.R. Hockett, D.J. Hoff, T.J. Norberg-King, T.W. Valenti, and D.R. Mount. 2011. Development and initial evaluation of a reconstituted water formulation that better represents natural waters. Society of Environmental Toxicology and Chemistry North America 32nd Annual Meeting. Boston, MA.

Jorgenson, B.C., S.L. Clark, A.M. Briden, R.S. Ogle, and J. Cotsifas. 2017. *Ceriodaphnia dubia* chronic toxicity variability: Evaluation of food as a source of test variability. Society of Environmental Toxicology and Chemistry North America 38th Annual Meeting. Minneapolis, MN.

Moore, T.F., S.P. Canton, and M. Grimes. 2000. Investigating the incidence of type I errors for chronic whole effluent toxicity testing using Ceriodaphnia dubia. Environmental Toxicology and Chemistry 19:118-122.

Mount, D.R., R.J. Erickson, T.L. Highland, J.R. Hockett, D.J. Hoff, C.T. Jenson, T.J. Norberg-King, K.N. Peterson, Z.M. Polaske, and S. Wisniewski. 2016. The acute toxicity of major ion salts to *Ceriodaphnia dubia*: I. Influence of background water chemistry. Environmental Toxicology and Chemistry 35:3039-3057.

Mount, D.R., R.J. Erickson, B.B. Forsman, T.L. Highland, J.R. Hockett, D.J. Hoff, C.T. Jenson, and T.J. Norberg-King. 2019. Chronic Toxicity of Major Ion Salts and Their Mixtures to *Ceriodaphnia dubia.* Environmental Toxicology and Chemistry 38(4): 769–783.

Norberg-King, T.J., and S. Schmidt. 1993. Comparison of effluent toxicity results using *Ceriodaphnia dubia* cultured on several diets. Environmental Toxicology and Chemistry 12:1945-1955.

Prosser, K.N., A.M. Briden, S.L. Clark, and B.C. Jorgenson. 2018. Assessing food preparation and storage as sources of *Ceriodaphnia dubia* culture quality and chronic test variability. Society of Environmental Toxicology and Chemistry North America 39th Annual Meeting. Sacramento, CA.

Robertson-Bryan, Inc/CASA & SCCWRP. 2022. Quality Assurance Project Plan (QAPP) for the State Water Resources Control Board *Ceriodaphnia dubia* Study - Baseline Interlaboratory Testing. Approved 8/11/2022. http://ftp.sccwrp.org/pub/download/DOCUMENTS/CeriodaphniaQA/August2022SciencePanelMaterials2.pdf

Southern California Coastal Water Research Project (SCCWRP). 2021. Conceptual Workplan: Quality Assurance Evaluation of the *Ceriodaphnia dubia* Reproduction Test. http://ftp.sccwrp.org/pub/download/DOCUMENTS/CeriodaphniaQA/May2021DraftWorkplan.pdf

Southern California Coastal Water Research Project (SCCWRP). 2023. Quality Assurance Project Plan (QAPP) for the State Water Resources Control Board Ceriodaphnia dubia Study - Second Interlaboratory Comparison Study. Approved 4/18/23. http://ftp.sccwrp.org/pub/download/DOCUMENTS/CeriodaphniaQA/April2023Materials2.pdf

9. APPENDICES

Appendix A – Summary of historical data and laboratory-specific techniques.

The appendix is divided into six sections:

1) Inventory of available data
2) Biological response data
3) Water quality data
4) Reference toxicant data
5) Correlations among data variables
6) Potential sources of variance in test performance

Each section and/or graphic is appended with a caption to summarize important methodological information. No assessment or conclusion is provided.

Inventory of available data

Inventory updates include additional requests for lab specific information as well as additional quality assurance reviews for seemingly unusual data. The inventory includes some summary statistics requested by the Panel.

Table A1. Summary of the inventory of available test control and reference toxicant test data provided to SCCWRP from the accredited laboratories. Number of test controls are divided into how many replicates were utilized in each test.

Lab	# of Test Controls			# of Ref Tox tests (incl. control)*
	Total # Test Controls	# of tests with 10 reps	# of tests with 20 Reps	
A	48	48	0	31
B	48	48	0	47
C	28	28	0	28
D	19	19	0	6
E	49	24	25	30
F	45	37	8	30
G	7	7	0	22
H	0	0	0	17
I	30	30	0	30
J	7	7	0	21
K	19	19	0	15
L	27	27	0	30
M	59	59	0	34
N	30	30	0	30
O	30	30	0	30
P	80	1	79	28
Q	25	25	0	23
Total	551	439	112	452

* All reference toxicant controls had 10 replicates

Table A2. Inventory of lab techniques that was available to SCCWRP from the accredited laboratories at the time of historical data compilation.

Lab	Light Intensity	Brood Board WQ	Chamber Material	Chamber Volume (ml)	Chamber Diameter (cm)	Sample Volume (ml)	Dilution Water Formula
A	Not measured	Not measured	glass	20	2.54	20	MHW + Vitamins and Selenium
B	Sent	Not measured	glass	30	5.5	15	MHW
C	Not measured	Not measured	polystyrene	29.57	4	15	MH + Vitamins and Selenium
D	Sent Ranges	Not measured	polystyrene	29.5	4	15	MHW + Selenium
E	Measure Quarterly	Not measured	polypropylene	29.5	3.8	30	MHW
F	Measure every 6 months	Not measured	polypropylene	29.57	4.29	15	DMW; 80% DIW: 20% Perrier®
G	Not sent	Not measured	glass	26	2.9	15	DMW; 80% DIW: 20% Perrier®
H	Unknown	Unknown	polystyrene	26	2.54	Not sent	DMW; 80% DIW: 20% Evian®
I	Measure annually	Sent	plastic	36.9	4.5	15	Hoheisel
J	Unknown	Unknown	Unknown	Unknown	Unknown	15	MHW
K	Sent Average	Sent	glass	20	2.8	15	L1650%, proprietary formula
L	Not measured	Not measured	plastic	29.57	4	15	MHW + Selenium
M	Have by test	Sent	polystyrene	Not sent	3.8	15	MHW
N	Sent Ranges	Not measured	polystyrene	29.57	4	15	MHW + Selenium
O	Measure every 6 months	Sent	polystyrene	36.97	4.45	15	MHW + Selenium

Lab	Light Intensity	Brood Board WQ	Chamber Material	Chamber Volume (ml)	Chamber Diameter (cm)	Sample Volume (ml)	Dilution Water Formula
P	Have by test	Sent	polystyrene	Not sent	6	30	DMW; 80% DIW: 20% Perrier®
Q	Sent graphs	Not measured	polystyrene	30	4.32	15	DMW; 80% DIW: 20% Perrier®

MHW= EPA moderately hard water recipe using salts

MH= EPA hard water recipe using salts

DMW= EPA recipe for moderately hard water using diluted mineral water

DIW= Deionized water

Unclear= More than one recipe listed in SOP

Table A3. Additional lab techniques that were available to SCCWRP from the accredited laboratories at the time of historical data compilation.

Lab	YCT Source	Cerophyl	Trout Chow	Algae Source	Termination Trigger	Termination Window
A	Purchased	*NA*	*NA*	In-house	≥60% having 3 broods	*Strict window with single check*
B	Purchased	*NA*	*NA*	Purchased	≥60% having 3 broods	*Strict window with single check*
C	In-house	*Fisher Science Cereal Grass Media*	*Aquamax 100*	In-house	≥60% having 3 broods	*No specific window with periodic checks*
D	Purchased	*NA*	*NA*	Purchased	≥60% having 3 broods	*Strict window with single check*
E	Purchased	*NA*	*NA*	Purchased	NA	*Test always runs seven days*
F	In-house	*Wheatgrass from ABS*	*Thomas Fish*	Purchased	≥60% having 3 broods	*No specific window with single check*
G	In-house	*Wheatgrass from ?*	*TetraMin Flakes*	Purchased	≥80% having 3 broods	*No specific window with periodic checks*
H	In-house	*Starwest Botanicals Organic Alfalfa Leaf*	*TetraMin Flakes*	In-house	≥70% having 3 broods	*No specific window with periodic checks*
I	In-house	*Frontier Powdered Wheat Grass*	*Silver Cup Trout Chow by Skretting*	In-house	≥60% having 3 broods	*Strict window with single check*

Lab	YCT Source	Cerophyl	Trout Chow	Algae Source	Termination Trigger	Termination Window
K	In-house	*Alfalfa from Co-Op*	*Skreeting Trout Chow*	In-house	≥60% having 3 broods	*Strict window with periodic checks*
L	Purchased	*NA*	*NA*	Purchased	≥60% having 3 broods	*Strict window with periodic checks*
M	Purchased	*NA*	*NA*	Purchased	≥60% having 3 broods	*Strict window with single check*
N	In-house	*Carolina Cereal*	*Purina Trout Chow + Carolina Daphnia Food (50:50)*	In-house	≥60% having 3 broods	*Strict window with single check*
O	In-house	*Pines Wheat Grass*	*Zinfer Bros. Fin Fish Starter*	In-house	≥60% having 3 broods	*No specific window with single check*
P	In-house and ARO	*Cereal Grass Media from Ward's*	*Zeigler Finfish Starter*	Purchased	≥60% having 3 broods	*No specific window with periodic checks*
Q	Purchased	*NA*	*NA*	Purchased	≥60% having 3 broods	*Strict window with single check*

NA= Not applicable

Table A4. Summary status of biological data that is available to SCCWRP from the accredited laboratories. The data is divided into test controls vs reference toxicant test per laboratory.

		Neonate/Female			Test Length (d)			Time to First Brood (d)			Number of Males per Test			Age at Start of Test (h)		
Lab	Test Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
A	Sample Control	29.9	48	18.9 - 44	6.5	48	6 - 7	3.6	48	3.0 - 4.2	0	48	0	NA	0	NA
A	Ref Tox Control	28.7	31	18.9 - 40.2	6.5	31	6 - 7	3.7	31	3 - 4.5	0	31	0	NA	0	NA
B	Sample Control	27.6	48	3.8 - 41.4	7.0	48	6 - 8	3.8	48	2.9 - 4.4	0	48	0	NA	0	NA
B	Ref Tox Control	26.7	47	0 - 46.1	6.6	47	2 - 8	3.5	47	0 - 5.1	0	47	0	NA	0	NA
C	Sample Control	35.8	28	5.3 - 50.2	7.0	28	6 - 7	4.1	28	3.6 - 5.4	0	28	0	5.5	16	3.2 - 7.1
C	Ref Tox Control	35.8	28	19.4 - 50.0	7.0	28	7 - 7	4.0	28	3.6 - 4.5	0.04	28	0 - 1	5.5	15	3.2 - 7.0
D	Sample Control	27.2	19	13.4 - 39.6	6.7	19	6 - 7	3.9	19	3.2 - 4.3	0	19	0	7.3	19	5.5 - 9.1
D	Ref Tox Control	28.4	6	17.2 - 35.1	7.0	6	6 - 7	3.9	6	3.0 - 4.3	0	6	0	6.3	6	4.8 - 7.8
E	Sample Control	24.9	49	15 - 31.3	7.0	49	7 - 7	3.9	49	3.4 - 4.9	0	49	0	NA	0	NA
E	Ref Tox Control	25.1	30	17.6 - 35.3	7.0	30	7 - 7	3.9	30	3.6 - 4.3	0	30	0	NA	0	NA
F	Sample Control	26.0	45	19.2 - 40.3	6.5	45	6 - 7	3.8	45	3.0 - 4.2	0	45	0	12.2	44	6.2 - 23.8
F	Ref Tox Control	25.5	30	20.1 - 39.3	6.4	30	6 - 7	3.9	30	2.7 - 4.0	0	30	0	11.5	30	6.0 - 23.8
G	Sample Control	26.2	7	19.6 - 33.9	6.7	7	6 - 7	3.4	7	2.9 - 4.2	0	7	0	17.6	5	8.0 - 24.0
G	Ref Tox Control	23.1	22	15.7 - 28.4	6.7	22	6 - 8	3.8	22	2.7 - 5.1	0.09	22	0 - 1	16	21	8.0 - 24.0
H	Sample Control	NA	0	NA	NA	0	NA	NA	0	NA	NA	0	NA	NA	0	NA
H	Ref Tox Control	18.7	17	12.5 - 23.4	6.2	17	6 - 7	4.0	17	3.6 - 5.0	0	17	0	NA	0	NA
I	Sample Control	32.0	30	20.7 - 43.3	6.2	30	6 - 8	3.4	30	2.9 - 4.1	0.03	30	0 - 1	19.5	30	8.0 - 24.0
I	Ref Tox Control	31.4	30	23.1 - 44.0	6.0	30	5 - 6	3.3	30	2.7 - 4.0	0.03	30	0 - 1	19.7	30	8.0 - 24.0
J	Sample Control	17.2	7	11.2 - 22.1	6.9	7	6 - 8	4.1	7	3.3 - 4.7	0	7	0	NA	0	NA
J	Ref Tox Control	20.1	21	13.4 - 31.6	6.5	21	6 - 8	3.7	21	3.0 - 4.5	0.05	21	0 - 1	NA	0	NA
K	Sample Control	27.2	19	12.7 - 38.4	6.5	19	5 - 8	3.9	19	2.7 - 5.3	0	19	0	19.0	19	10 - 23.9
K	Ref Tox Control	27.6	15	15.6 - 41.9	6.5	15	6 - 7	3.5	15	1.8 - 4.2	0	15	0	20.0	14	14.2 - 23.9
L	Sample Control	25.5	27	16.6 - 35.7	7.1	27	7 - 8	4.0	27	3.1 - 4.6	0.04	27	0 - 1	3.3	25	2.8 - 7.5
L	Ref Tox Control	25.8	30	17 - 41.3	7.1	30	7 - 8	4.0	30	3.3 - 4.7	0.07	30	0 - 1	3.3	28	2.8 - 7.5
M	Sample Control	37.5	59	25.4 - 52.1	6.5	59	6 - 8	3.4	59	2.7 - 4.6	0	59	0	20.6	58	7.4 - 24.0
M	Ref Tox Control	35.0	34	3.4 -43.6	6.3	34	5 - 8	3.3	34	1.1 - 4.4	0	34	0	21.8	31	7.9 - 24.0
N	Sample Control	24.3	30	21.6 - 27	7.0	30	7 - 7	3.6	30	3.1 - 4.2	0	30	0	7.8	30	7.0 - 8.5

		Neonate/Female			Test Length (d)			Time to First Brood (d)			Number of Males per Test			Age at Start of Test (h)		
Lab	Test Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
N	Ref Tox Control	23.8	30	21.1 - 26.9	7.0	30	7 - 7	3.6	30	3.0 - 3.9	0	30	0	7.8	30	7.0 - 8.5

Table A4 Continued. Summary status of biological data that is available to SCCWRP from the accredited laboratories.

		Neonate/Female			Test Length (d)			Time to First Brood (d)			Number of Males per Test			Age at Start of Test (h)		
Lab	Test Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
O	Sample Control	35.3	30	21.9 - 47.5	6.2	30	6 - 7	3.3	30	2.4 - 4.1	0.33	30	0 - 2	18.5	30	7.5 - 24.0
O	Ref Tox Control	36.4	30	21.6 - 46.9	6.2	30	6 - 7	3.4	30	2.4 - 4.0	0.17	30	0 - 2	18.8	30	9.0 - 24.0
P	Sample Control	33.1	80	21.1 - 41.8	6.3	80	6 - 8	3.4	80	3.0 - 4.0	0	80	0	18.9	79	3.3 - 23.9
P	Ref Tox Control	33.1	28	20.7 - 43.8	6.1	28	6 - 8	3.1	28	3.0 - 4.1	0	28	0	20.6	28	2.6 - 24.6
Q	Sample Control	24.5	25	16.0 - 31.9	6.0	25	5 - 7	3.4	25	2.9 - 4.2	0	25	0	6.1	18	2.9 - 8.9
Q	Ref Tox Control	24.4	23	12.8 - 31.0	6.1	23	5 - 8	3.4	23	3.0 - 4.5	0	23	0	6.2	16	2.9 - 8.9

Table A5. Summary status of water quality data that is available to SCCWRP from the accredited laboratories. The data is divided into test sample controls and reference toxicant test controls per laboratory.

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)			Conductivity (µS/cm)			pH			Dissolved Oxygen (mg/L)			Temperature (°C)		
Lab	Test Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
A	Sample Control	89.6	48	81.9 - 96	62.1	48	57 - 64	315	48	291 - 331	7.9	48	7.8 - 8.1	8.0	48	7.7 - 8.2	24.0	48	24.0 - 24.0
A	Ref Tox Control	90.0	31	81.9 - 96	62.1	31	57 - 64	316	31	291 - 330	8.0	31	7.8 - 8.1	8.0	31	7.7 - 8.2	24.0	31	24.0 - 24.1
B	Sample Control	97.3	35	88.9 - 107.7	68.1	35	56 - 90.5	358	48	310 - 420	7.6	48	7.1 - 8.1	8.1	48	6.4 - 9.4	24.8	48	24.2 - 25.6
B	Ref Tox Control	95.0	35	89 - 106.3	65.9	35	60 - 75.2	368	47	330 - 415	7.5	47	7.1 - 8.1	7.8	47	6.0 - 9.4	24.9	47	24.0 - 26.1
C	Sample Control	152	26	80 - 190	110	28	60 - 148	557	28	330 - 711	8.1	28	7.9 - 8.6	7.7	28	7.1 - 8.3	25.1	28	24.5 - 25.9
C	Ref Tox Control	171	26	105 - 190	119	28	112 - 126	616	28	588 - 713	8.2	28	7.8 - 8.6	7.7	28	7.1 - 8.2	25.0	28	24.6 - 25.6
D	Sample Control	87.3	19	80 - 96	55.7	19	37 - 64	335	19	284 - 399	8.0	19	7.7 - 8.2	8.0	19	7.8 - 8.2	25.0	19	24.3 - 25.6
D	Ref Tox Control	87.8	6	80 - 94	57.7	6	52 - 63	333	6	322 - 346	7.8	6	7.4 - 8.0	8.1	6	7.8 - 8.4	24.8	6	24.1 - 25.4
E	Sample Control	95.0	49	60 - 220	67.2	49	60 - 198	348	49	262 - 574	7.9	49	7.5 - 8.2	7.4	49	6.9 - 8.2	24.0	49	24.0 - 24.1
E	Ref Tox Control	91.7	30	60 - 100	61.2	30	60 - 65	343	30	332 - 356	7.9	30	7.5 - 8.2	7.5	30	6.9 - 7.9	24.0	30	24.0 - 24.0
F	Sample Control	87.9	45	81 -99	81.5	45	64 - 98	377	45	185 - 1072	8.2	45	7.9 - 8.5	8.1	45	7.7 - 8.4	24.7	45	24.0 - 25.5
F	Ref Tox Control	88.5	30	81 - 99	83.3	30	67 - 98	190	30	186 - 195	8.2	30	8.0 - 8.3	8.1	30	7.9 - 8.4	24.7	30	23.8 - 25.3
G	Sample Control	87.4	5	81 - 94	NA	0	NA	198	7	174 - 204	7.9	7	7.7 - 8.1	7.9	7	7.7 - 8.2	24.8	7	24.2 - 25.1
G	Ref Tox Control	86.1	21	81 - 98	NA	0	NA	195	22	171 - 214	8.0	22	7.8 - 8.2	8.1	22	7.8 - 8.4	24.8	22	24.3 - 25.4
H	Sample Control	NA	0	NA	NA	0	NA	NA	0	NA	NA	0	NA	NA	0	NA	NA	0	NA
H	Ref Tox Control	82.4	9	51.8 - 103.6	83.7	14	72 - 98	NA	0	NA	8.1	17	7.9 - 8.2	8.1	17	8.0 - 8.3	25.6	17	24.0 -26.6
I	Sample Control	90.9	30	85 - 95	73.3	30	68 - 76	358	30	269 - 415	7.9	30	7.7 - 8.1	7.7	30	6.6 - 8.5	24.8	30	24.3 - 25.5
I	Ref Tox Control	90.9	30	85 - 95	73.3	30	68 - 76	411	30	362 - 743	8.0	30	7.8 - 8.1	7.7	30	6.1 - 8.5	24.8	30	24.1 - 25.5
J	Sample Control	NA	0	NA	NA	0	NA	328	7	311 - 338	7.9	7	7.8 - 8.1	7.7	7	7.2 - 8.0	25.2	7	24.6 - 25.7
J	Ref Tox Control	NA	0	NA	NA	0	NA	324	21	308 - 333	7.9	21	7.5 - 8.1	7.7	21	7.2 - 8.0	25.0	21	24.1 - 25.6
K	Sample Control	81.9	17	8 - 108	59.9	17	54 - 84	299	19	258 - 442	7.8	19	7.3 - 8.1	7.8	19	7.4 - 8.0	24.4	19	23.5 - 25.1
K	Ref Tox Control	92	1	NA	56.0	1	NA	299	15	258 - 319	7.7	15	6.7 - 8.0	7.8	15	7.2 - 8.3	24.4	15	23.4 - 25.6
L	Sample Control	98.8	27	86 - 100	61.5	27	57 - 64	413	27	383 - 442	8.0	27	7.4 - 8.3	8.1	27	7.3 - 8.7	24.4	27	24.0 - 25.1
L	Ref Tox Control	98.9	30	86 - 100	61.5	30	57 - 64	416	30	381 - 447	7.8	30	7.4 - 8.2	8.2	30	7.3 - 8.7	24.3	30	24.0 - 25.0
M	Sample Control	96.7	58	80 - 170	69.3	58	60 - 116	351	59	299 - 582	8.3	59	8.0 - 8.5	6.8	59	6.2 - 7.2	24.9	59	23.0 - 27.6
M	Ref Tox Control	NA	0	NA	NA	0	NA	346	34	332 - 372	8.2	34	8.0 - 8.6	6.8	34	6.4 - 7.2	24.9	34	23.8 - 26.1
N	Sample Control	87.7	30	81 - 91	58.3	30	24 - 61	299	30	212 - 330	8.2	30	8.1 - 8.5	8.6	30	8.3 - 8.8	24.7	30	24.3 - 25.1

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)			Conductivity (µS/cm)			pH			Dissolved Oxygen (mg/L)			Temperature (°C)		
Lab	Test Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
N	Ref Tox Control	87.5	30	81 - 91	59.5	30	57 - 61	302	30	271 - 327	8.1	30	8.0 - 8.2	8.6	30	8.4 - 8.8	24.7	30	24.3 - 25.3

Table A5 continued. Summary status of water quality data that is available to SCCWRP from the accredited laboratories.

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)			Conductivity (µS/cm)			pH			Dissolved Oxygen (mg/L)			Temperature (°C)		
Lab	Test Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
O	Sample Control	94.9	30	89 - 99	61.9	30	60 - 40	365	30	356 - 374	7.9	30	7.7 - 8.1	8.3	30	8.0 - 8.8	24.9	30	24.3 - 25.2
O	Ref Tox Control	94.9	30	89 - 99	61.9	30	60 - 40	365	30	357 - 372	7.9	30	7.6 - 8.0	8.3	30	7.8 - 8.7	24.9	30	24.4 - 25.4
P	Sample Control	94.6	80	87 - 99	88.7	75	75 - 84	205	80	193 - 212	8.1	80	7.8 - 8.4	8.4	80	8.2 - 8.7	24.9	80	24.3 - 25.4
P	Ref Tox Control	93.9	28	87 - 99	89.0	27	84 - 97	205	28	199 - 210	8.1	28	7.7 - 8.4	8.4	28	8.2 - 8.7	24.8	28	24.4 - 25.3
Q	Sample Control	87.7	19	84 - 112	86.9	19	78 - 101	203	25	191 - 275	8.2	25	7.9 - 8.5	8.0	25	7.9 - 8.3	25.1	25	24.1 - 26.4
Q	Ref Tox Control	87.7	17	83 - 111	87.7	17	78 - 101	206	23	194 - 269	8.2	23	7.8 - 8.4	8.0	23	7.7 - 8.3	25.1	23	24.0 - 26.3

Biological response data

This section summarizes the basic biological response data from the accredited laboratories. The left half of each lab's distribution is a box plot with box hinges equivalent to the 75th,50th (median), and 25th percentiles. The whiskers are 1.5x the quartile ranges. The right half of each labs distribution is a violin plot. The circle symbol is the lab mean.

Figure A1. Mean number of neonates per female per test for controls for each laboratory. This was created by averaging the number neonates per female across all replicates in each test. The test acceptability criterion for controls is a mean of 15 neonates per female per test.

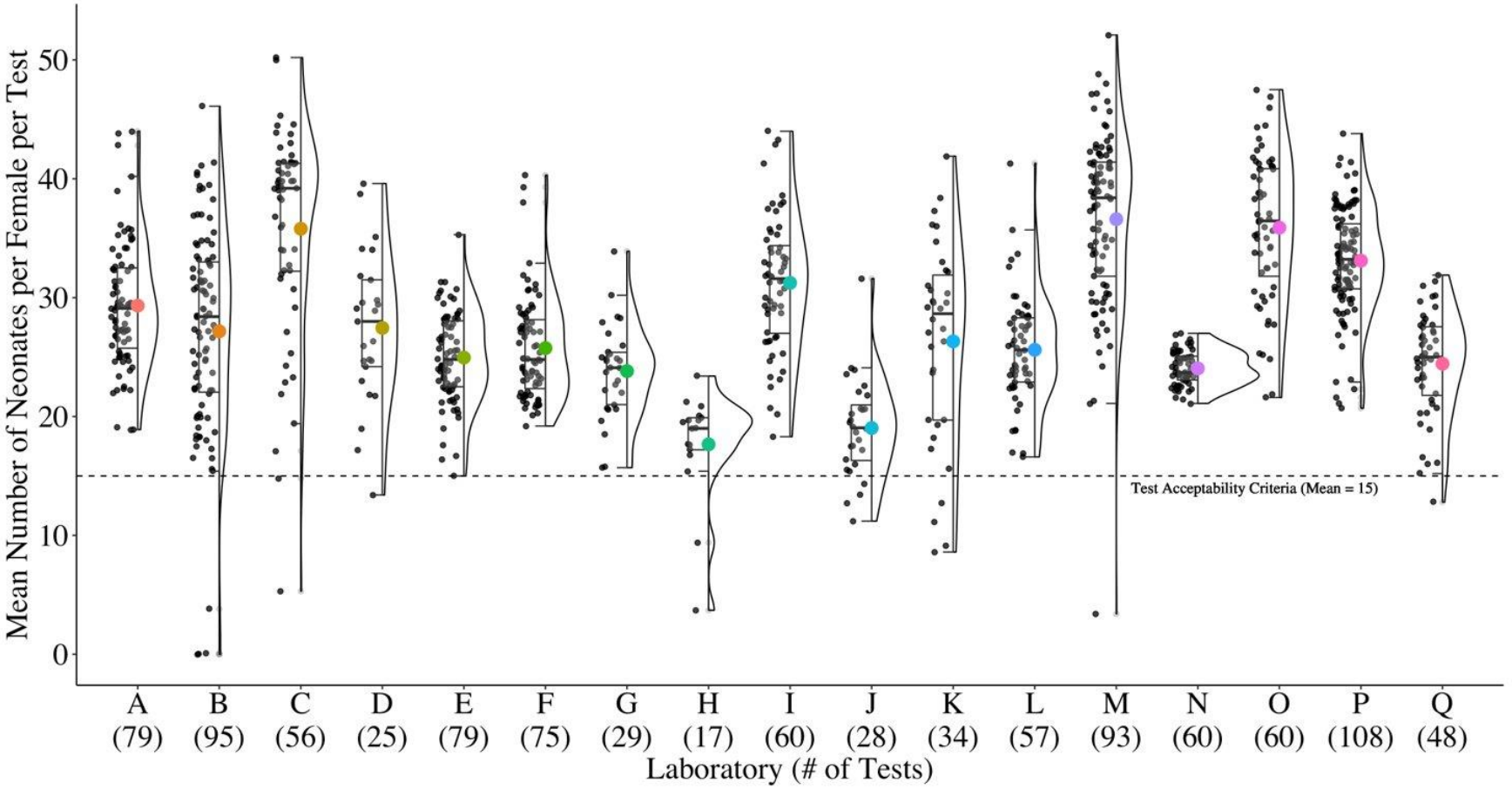


Figure A2. Coefficient of variation (CV) for number of neonates per female per test for controls for each laboratory. CVs are calculated as the SD/mean for each test. See the next plot for a zoom in on a reduced scale from 0.0 – 1.0.

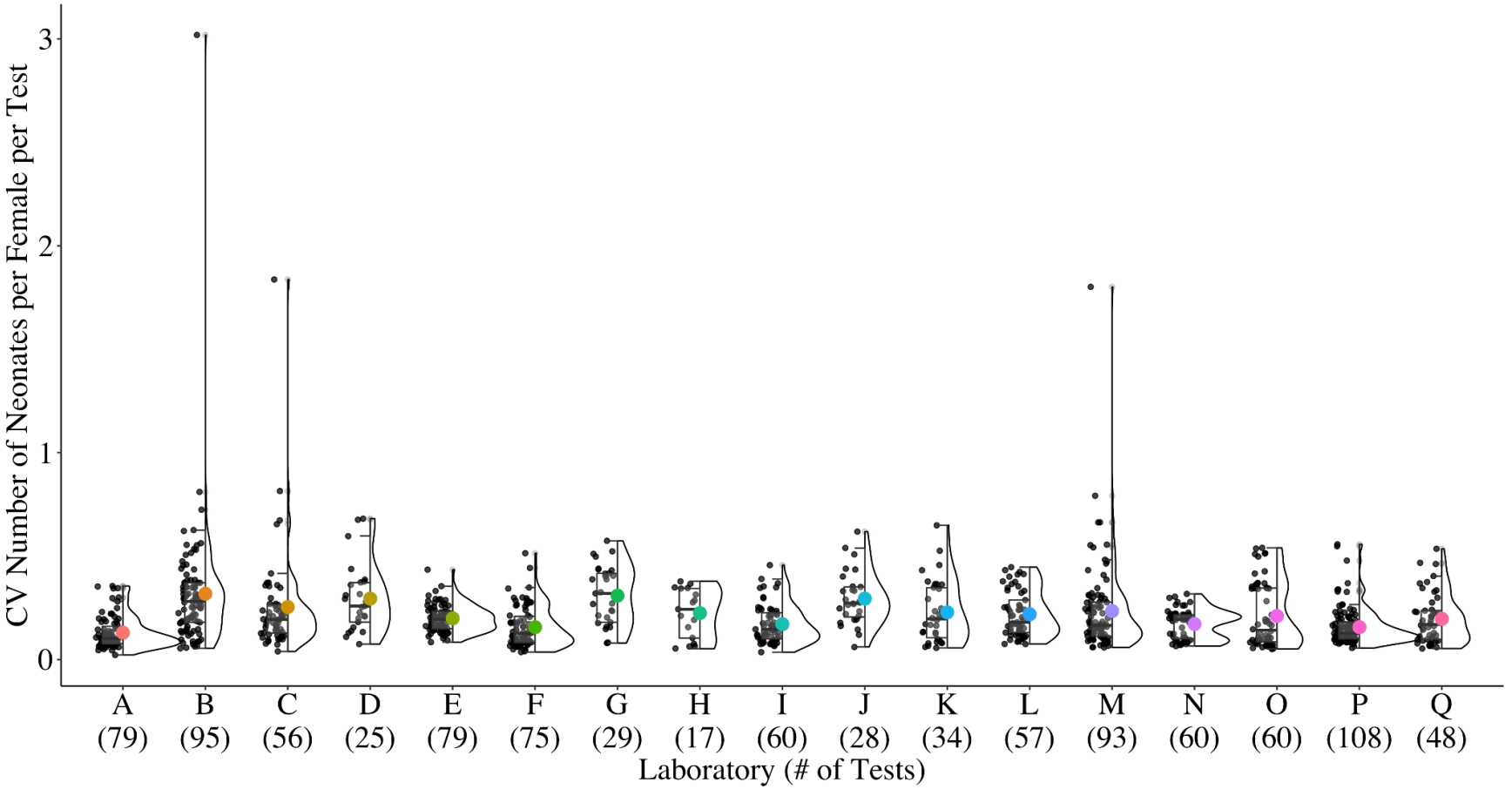


Figure A3. Coefficient of variation (CV) for number of neonates per female per test for controls for each laboratory. CVs are calculated as the SD/mean for each test. This plot is a zoom in on a reduced scale from 0.0 – 1.0. See the previous graph for the full scale from 0 – 8. Current unofficial EPA guidance is that long-term average CVs (the circle symbol) should be equal to or less than 0.15 for each laboratory.

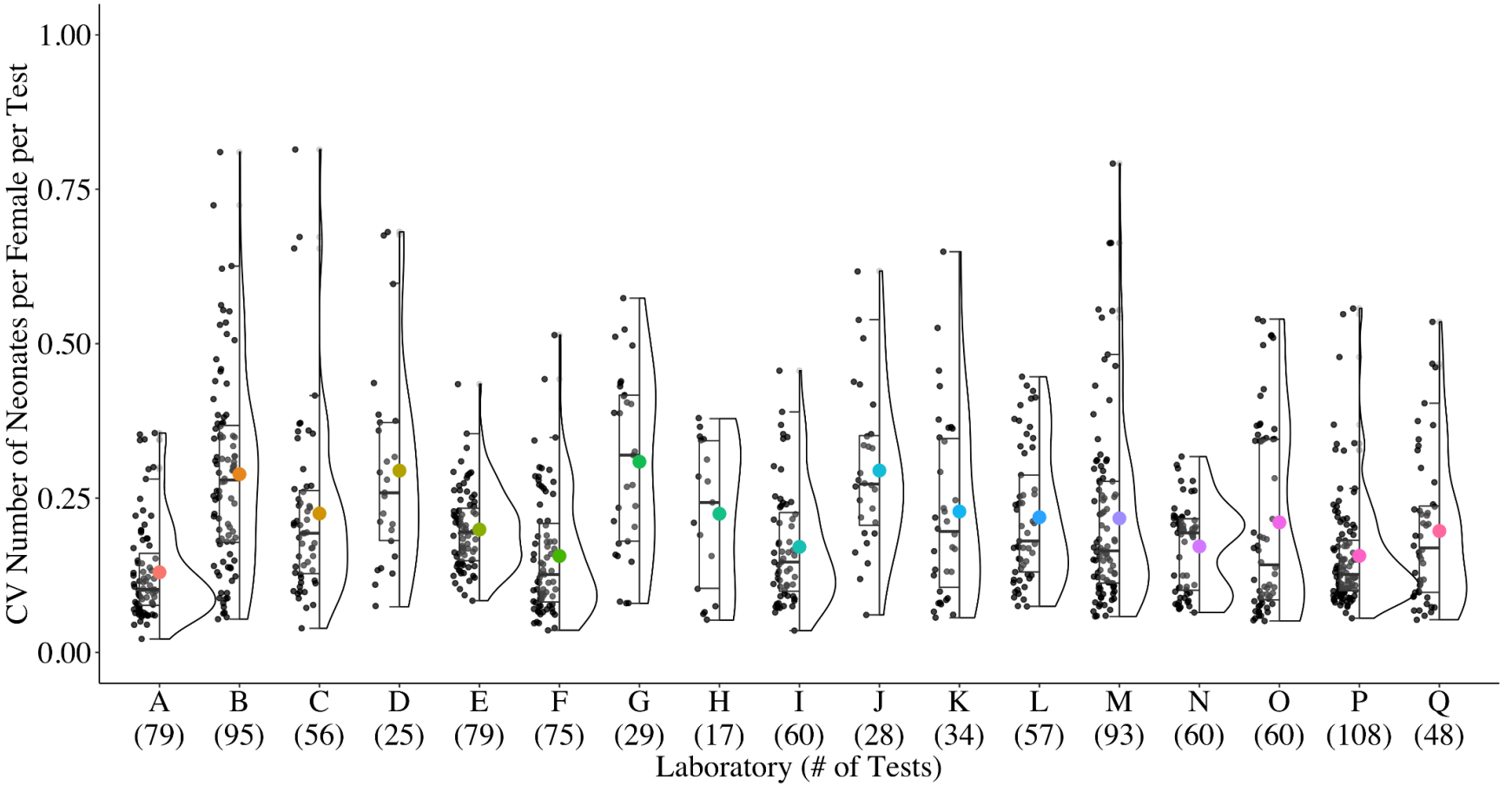


Figure A4. Example daily production – Lab C. Neonates per female per control replicate per day for three tests with different reproduction rates.

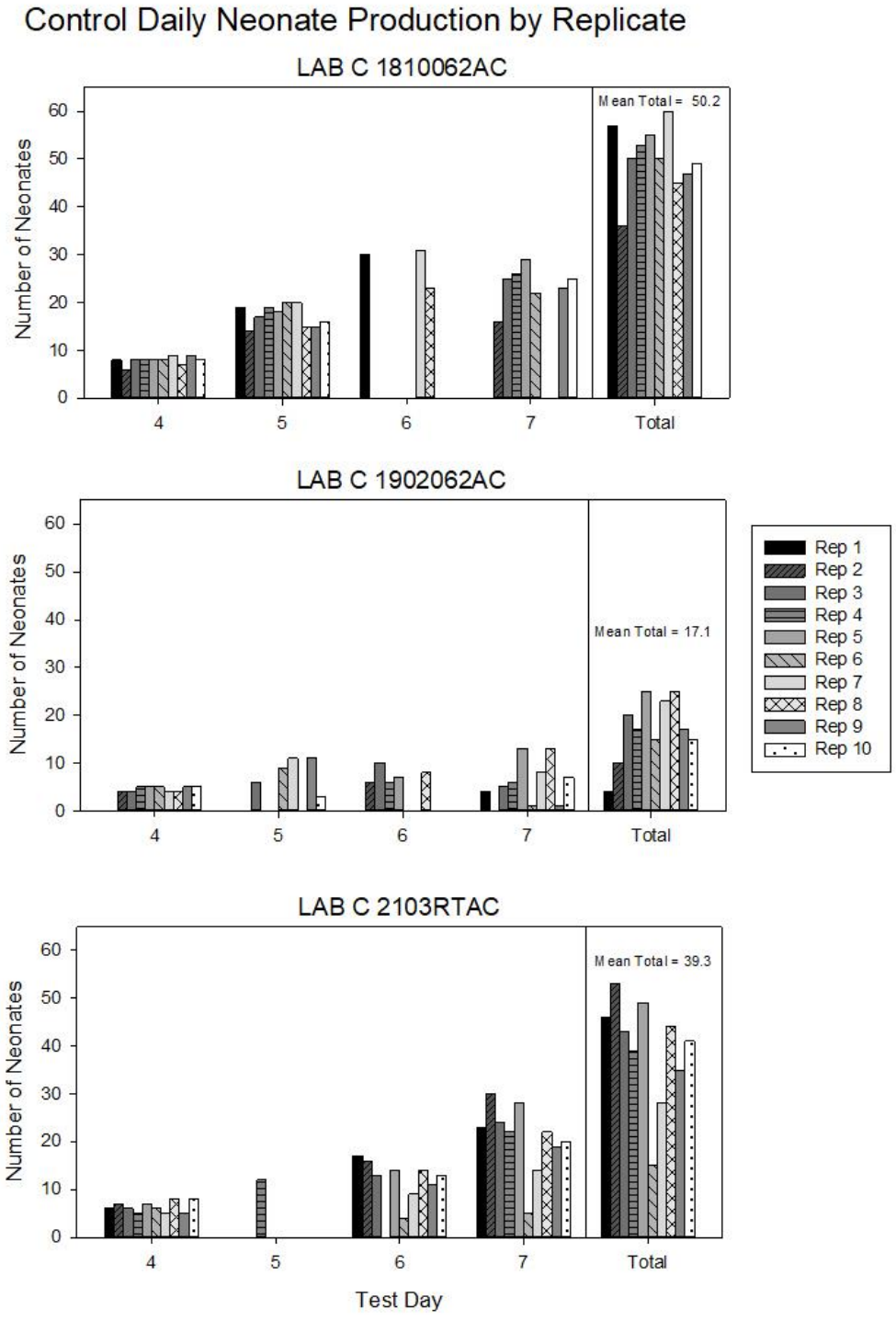


Figure A5. Example daily production – Lab E. Neonates per female per control replicate per day for three tests with different reproduction rates.

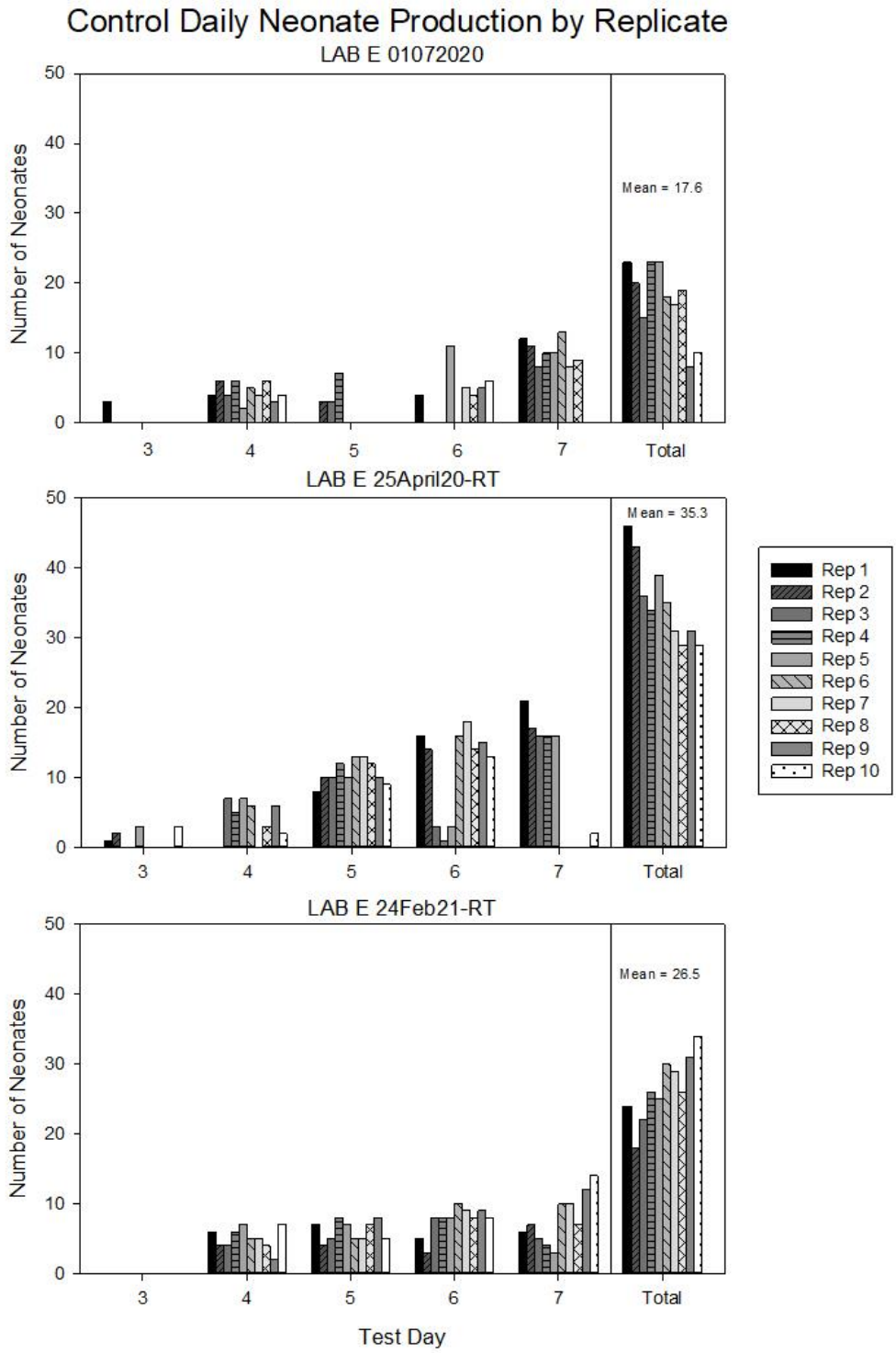


Figure A6. Example daily production – Lab M. Neonates per female per control replicate per day for three tests with different reproduction rates.

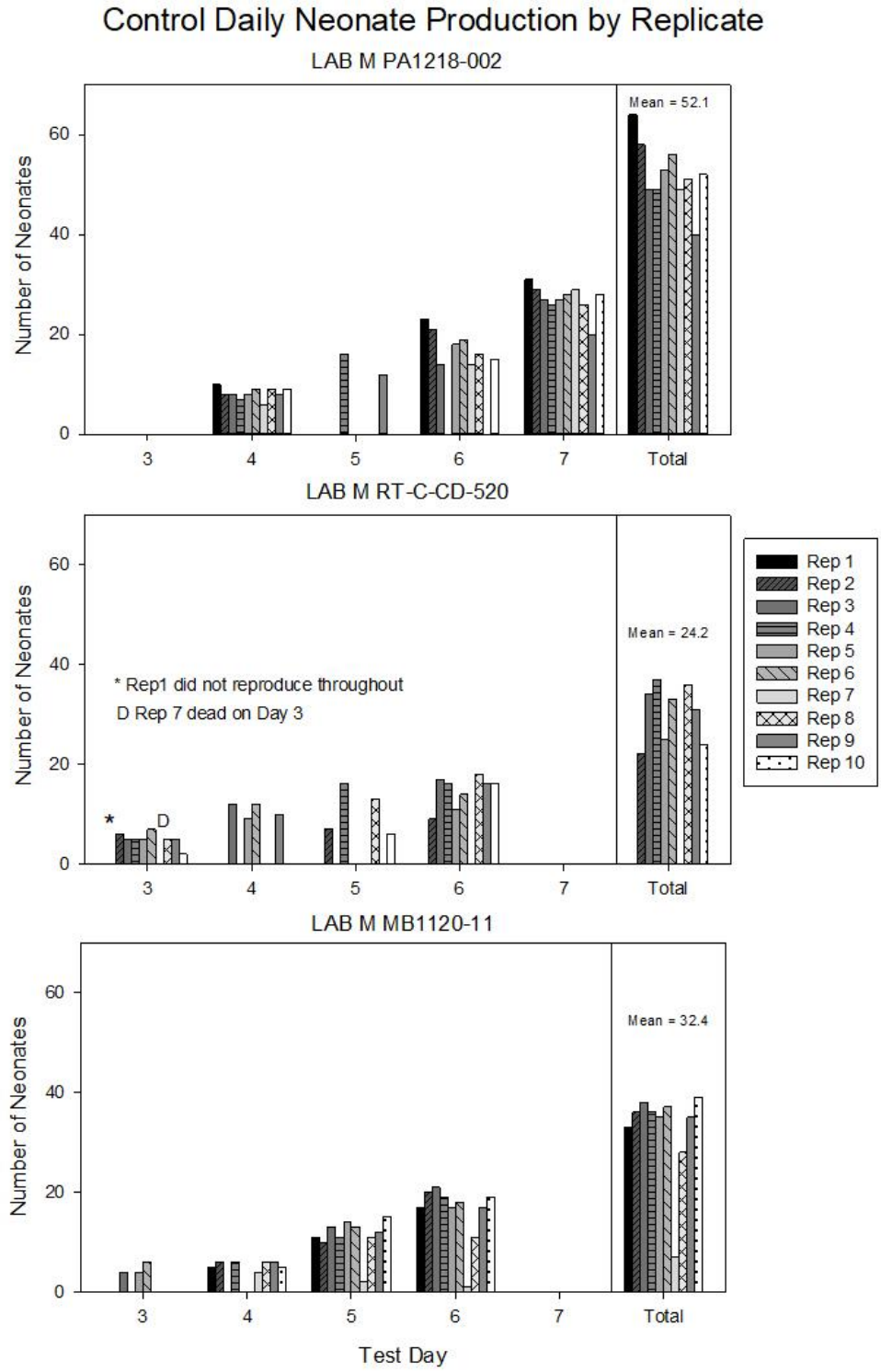


Figure A7. Test length by laboratory. The test termination was identified by each lab for each test.

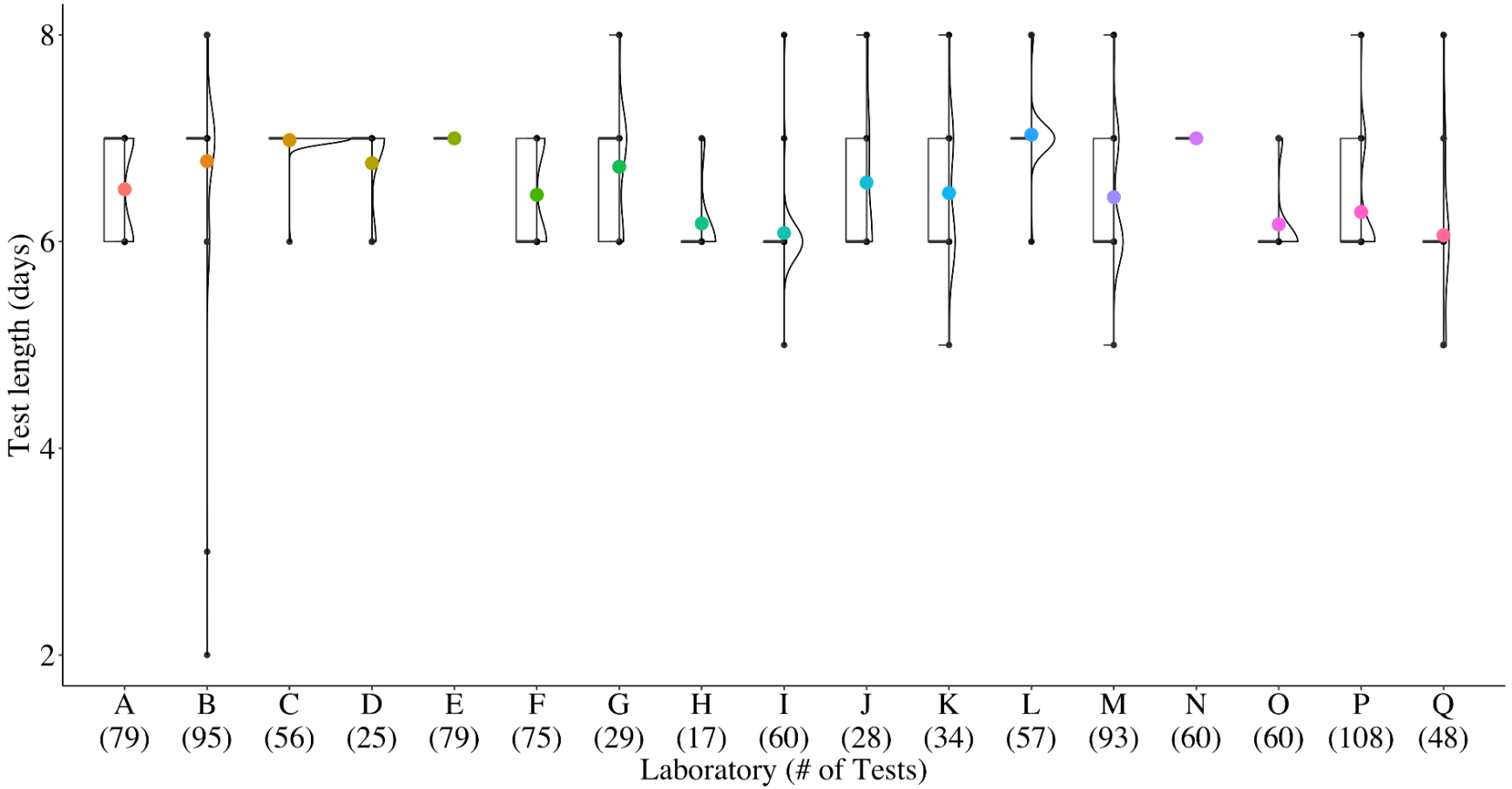


Figure A8. Mean survival of control replicates per test per laboratory. Each black dot is the percent survival of in the control for each test. The colored dots are the mean survival for all tests within each laboratory. Test acceptability criteria is 80% mean survival.

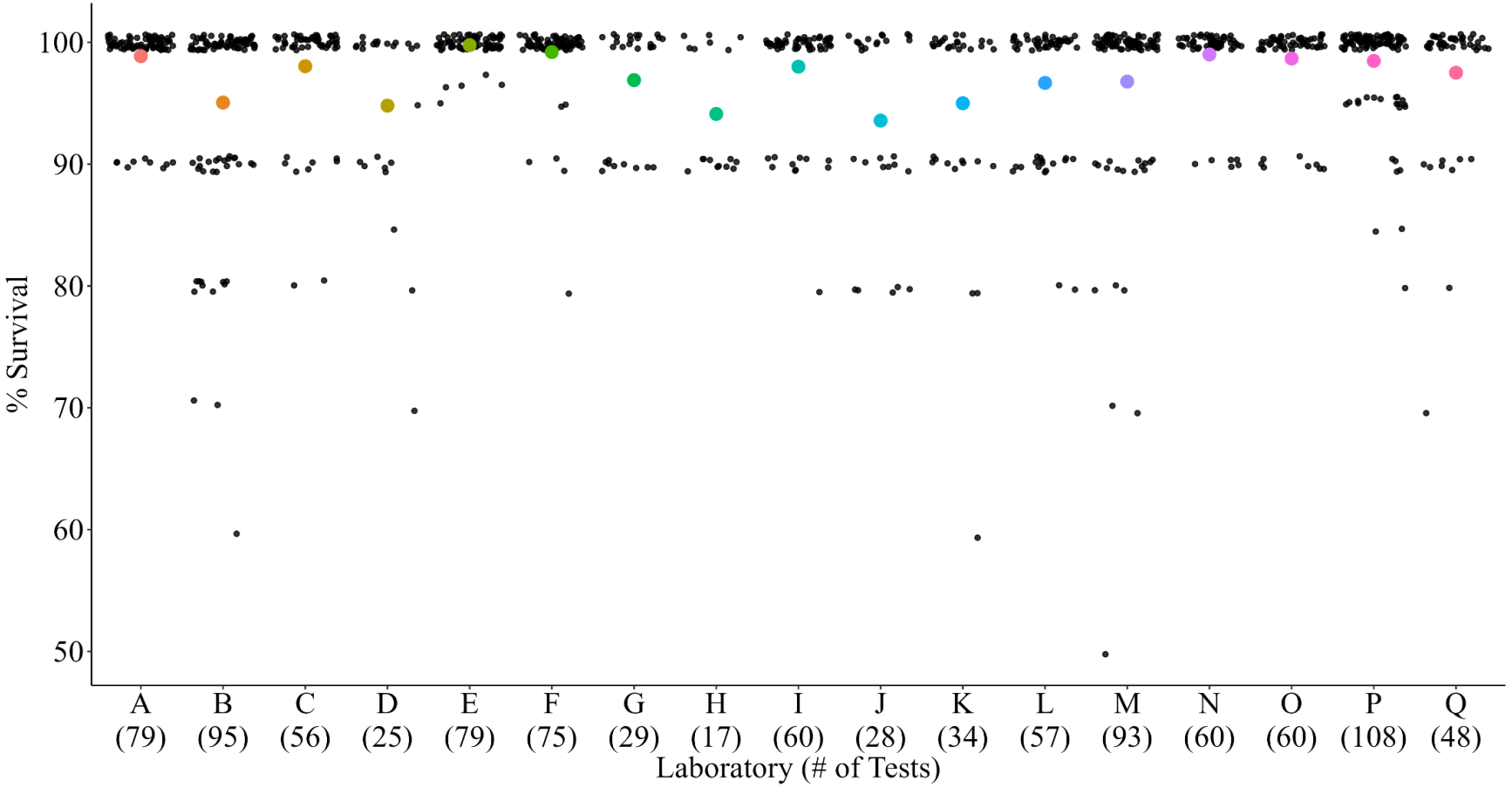
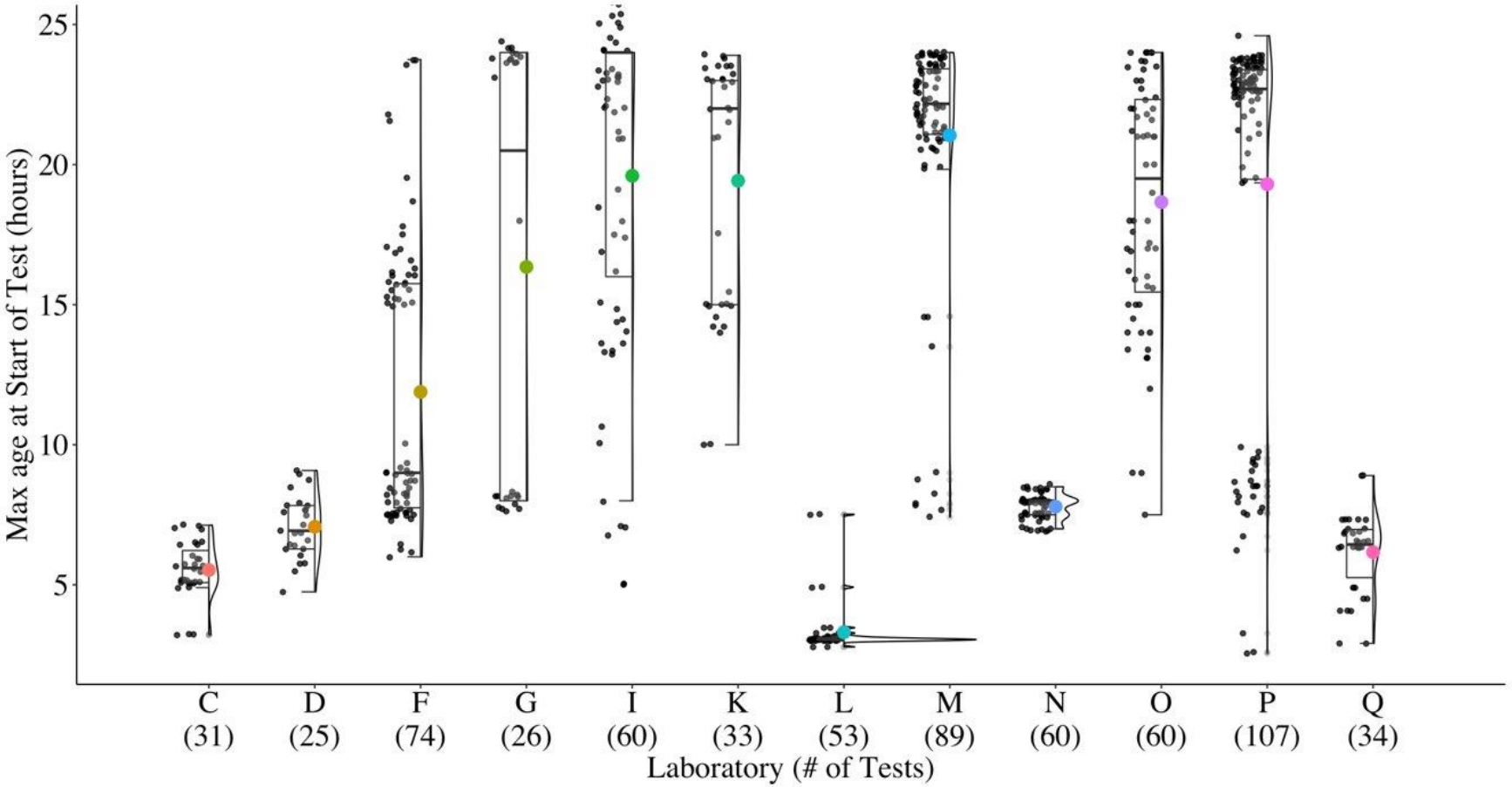


Figure A9. Age of females at test initiation per test per laboratory. Only 12 of the 17 labs provided age at test initiation. Max age was used because not all labs track the range of times of individual replicates. Test guidance requires replicates be within 8 hours of each other and no replicate greater than 24 hours old.



Water quality data

Water quality is a requirement of every test.

However, not all labs measured or reported water quality in the same fashion. For example, pH, DO, and temperature were measured daily while hardness, alkalinity, and conductivity were measured between once per test and everyday depending on the lab. Some labs measured from test chambers while others utilized surrogates. For the following analyses, we averaged data to once per test, which is the most common time step across the most labs. All units were standardized for the following analyses.

Figure A10. Mean Hardness for each test per laboratory.

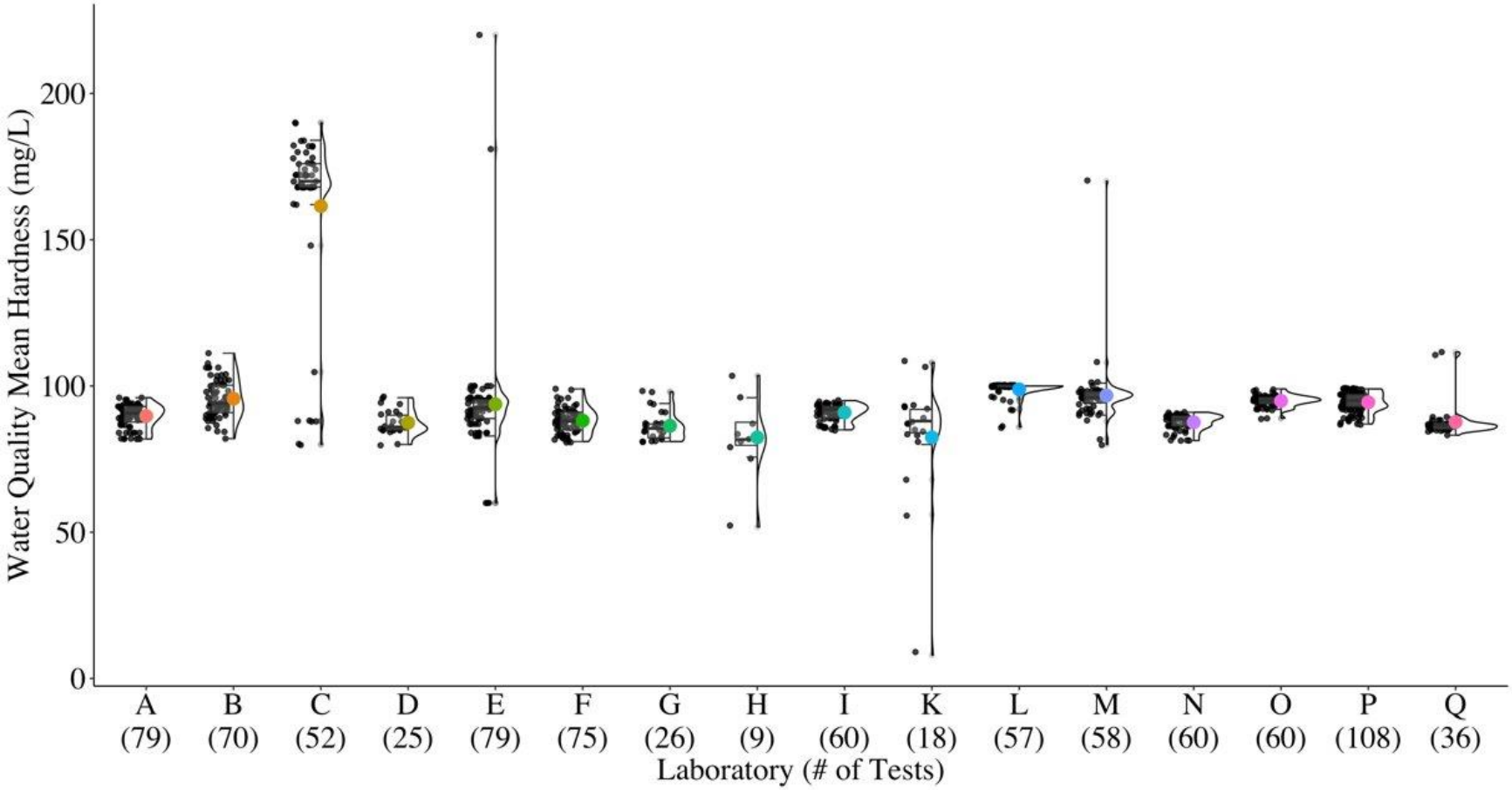


Figure A11. Mean conductivity for each test per laboratory.

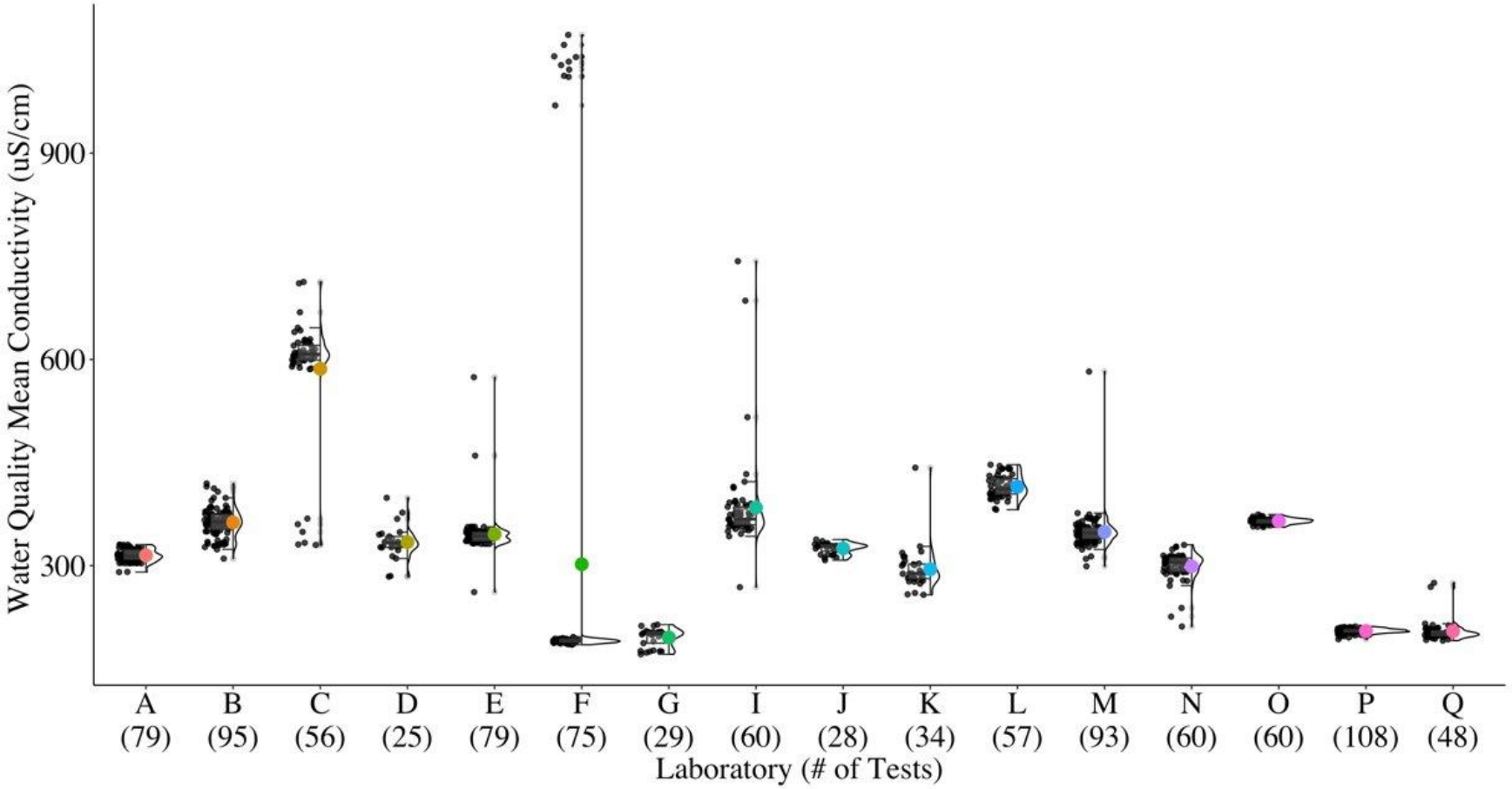


Figure A12. Mean alkalinity for each test per laboratory.

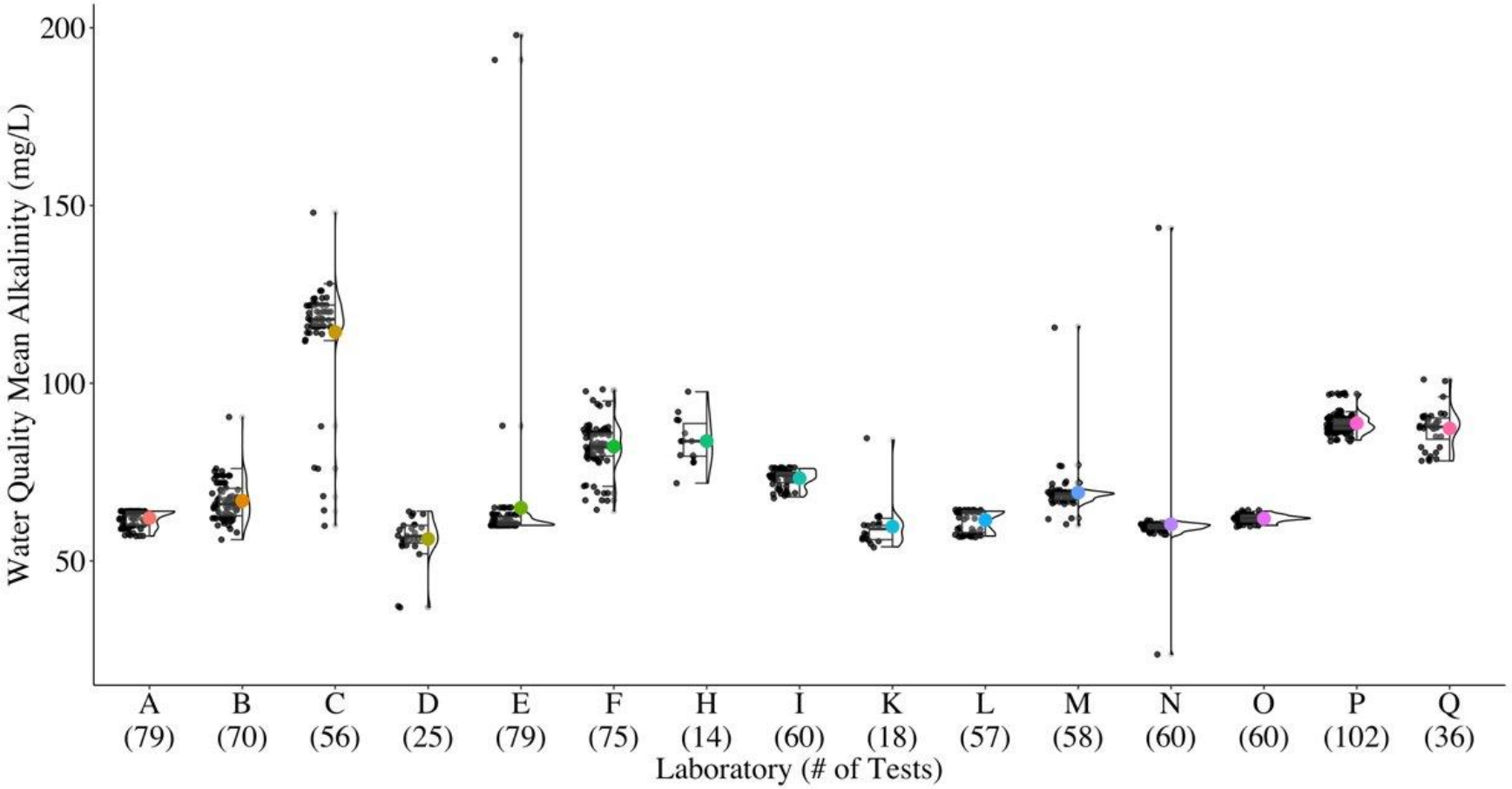


Figure A13. Mean Dissolved oxygen for each test per laboratory.

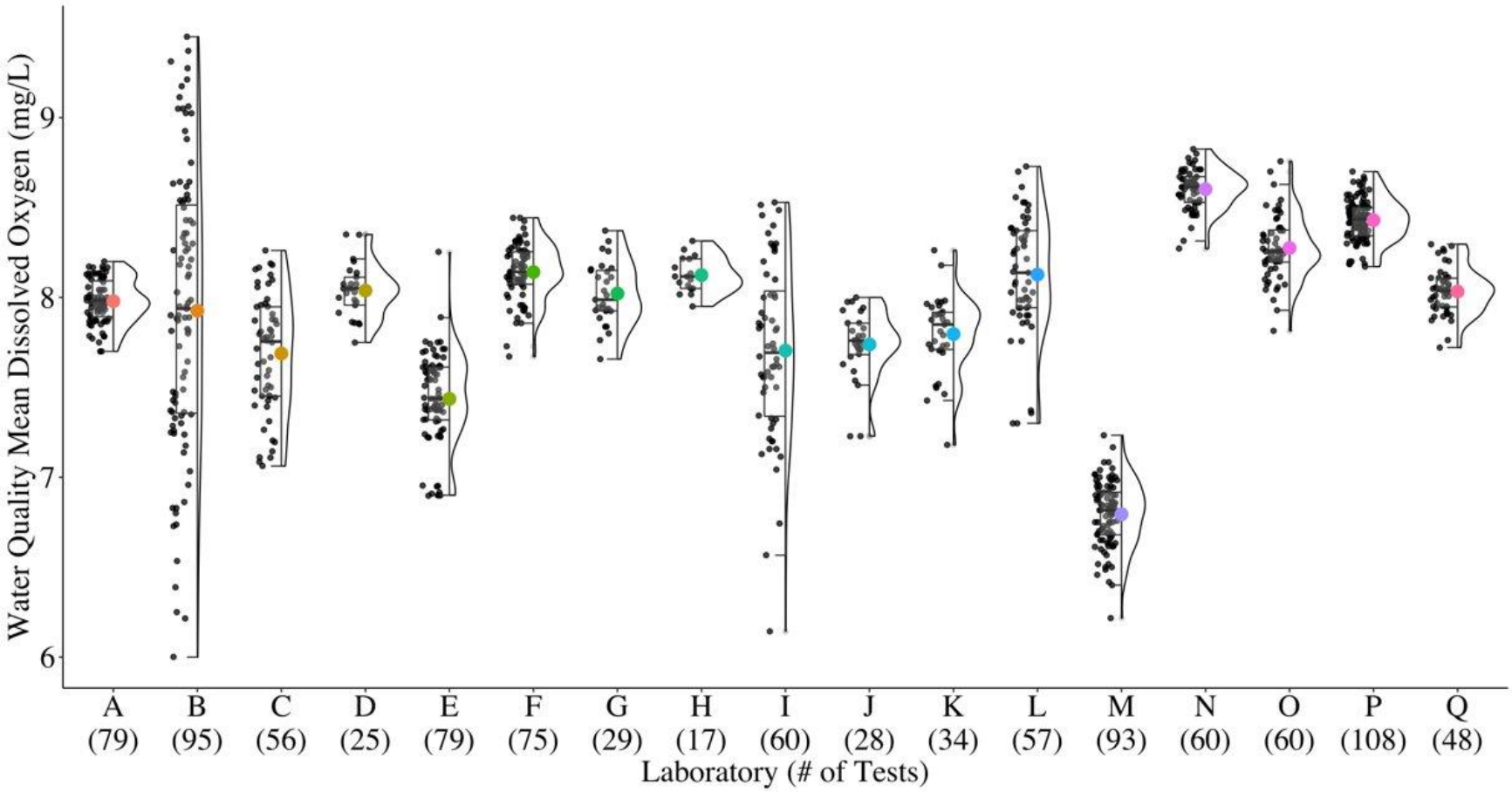


Figure A14. Mean pH for each test per laboratory.

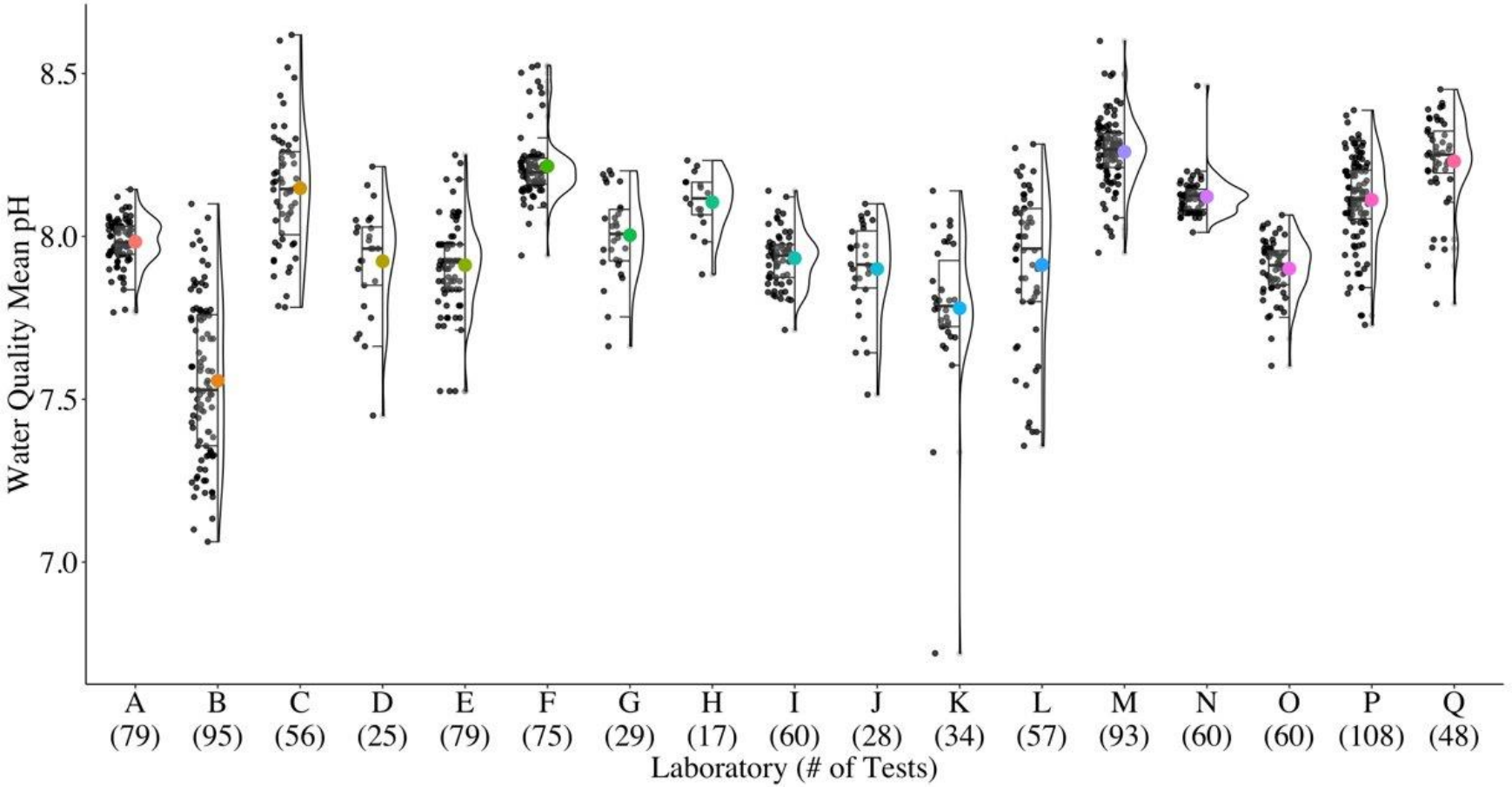
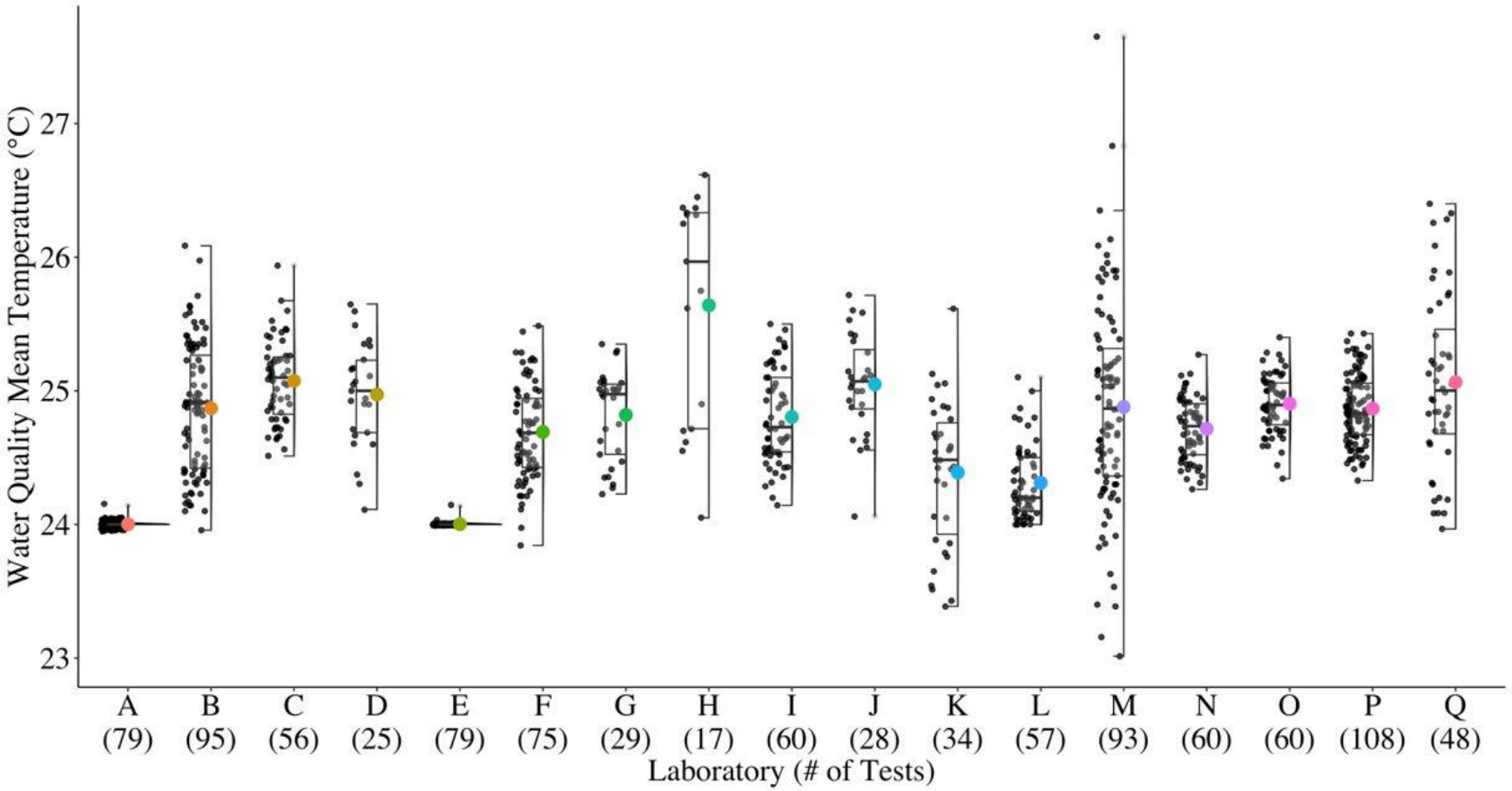


Figure A15. Mean temperature for each test per laboratory.



Reference toxicant data

All of the laboratories provided reference toxicant data, but not all laboratories use the same reference toxicant. Here we present both survival and reproduction endpoints.

Figure A16. Reference toxicant LC50 by test per laboratory. Not all laboratories used the same reference toxicant. LC50 values were provided by each laboratory.

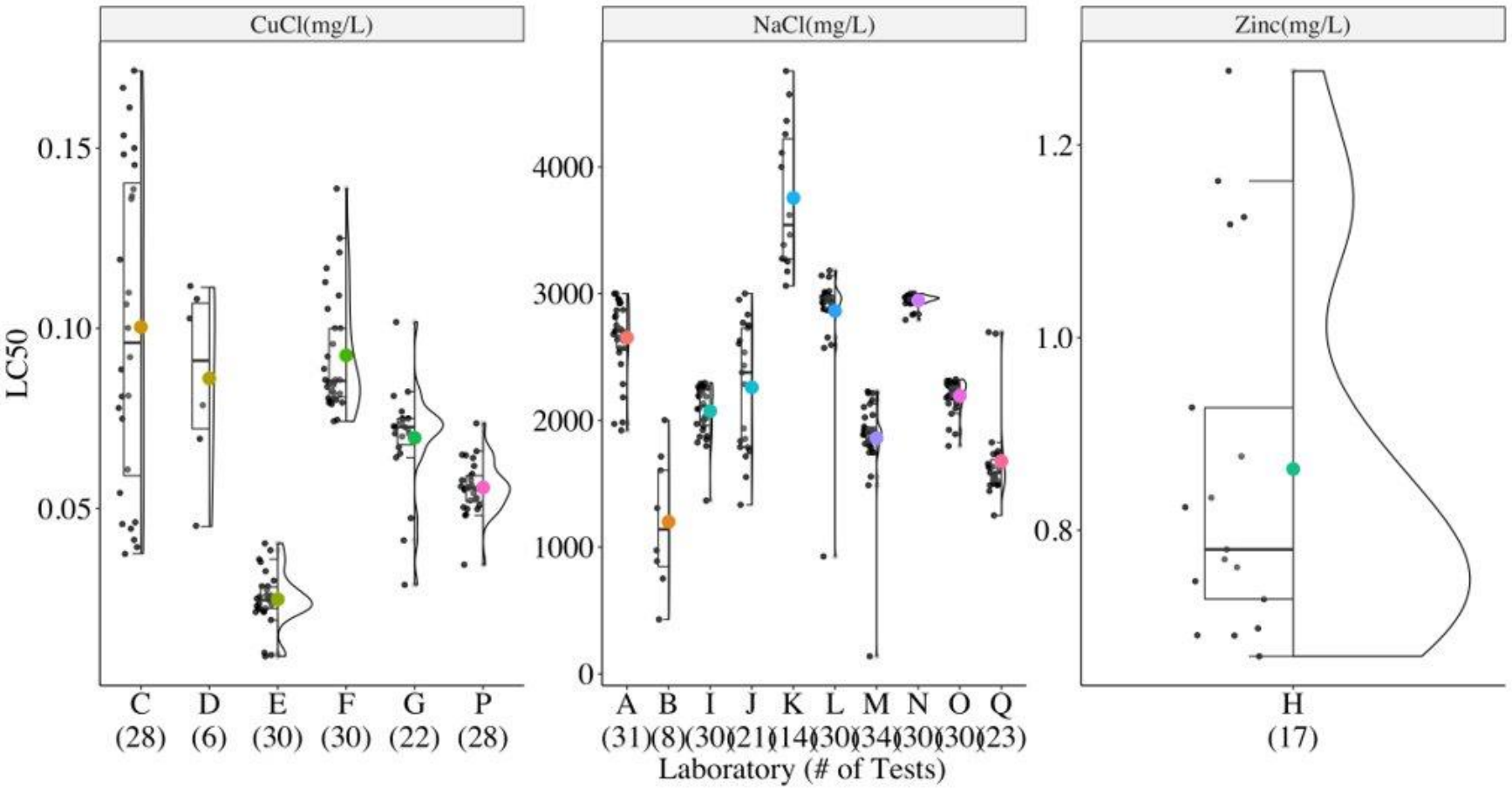


Figure A17. Reference toxicant IC50 by test per laboratory. Not all laboratories used the same reference toxicant. IC50 values were provided by each laboratory.

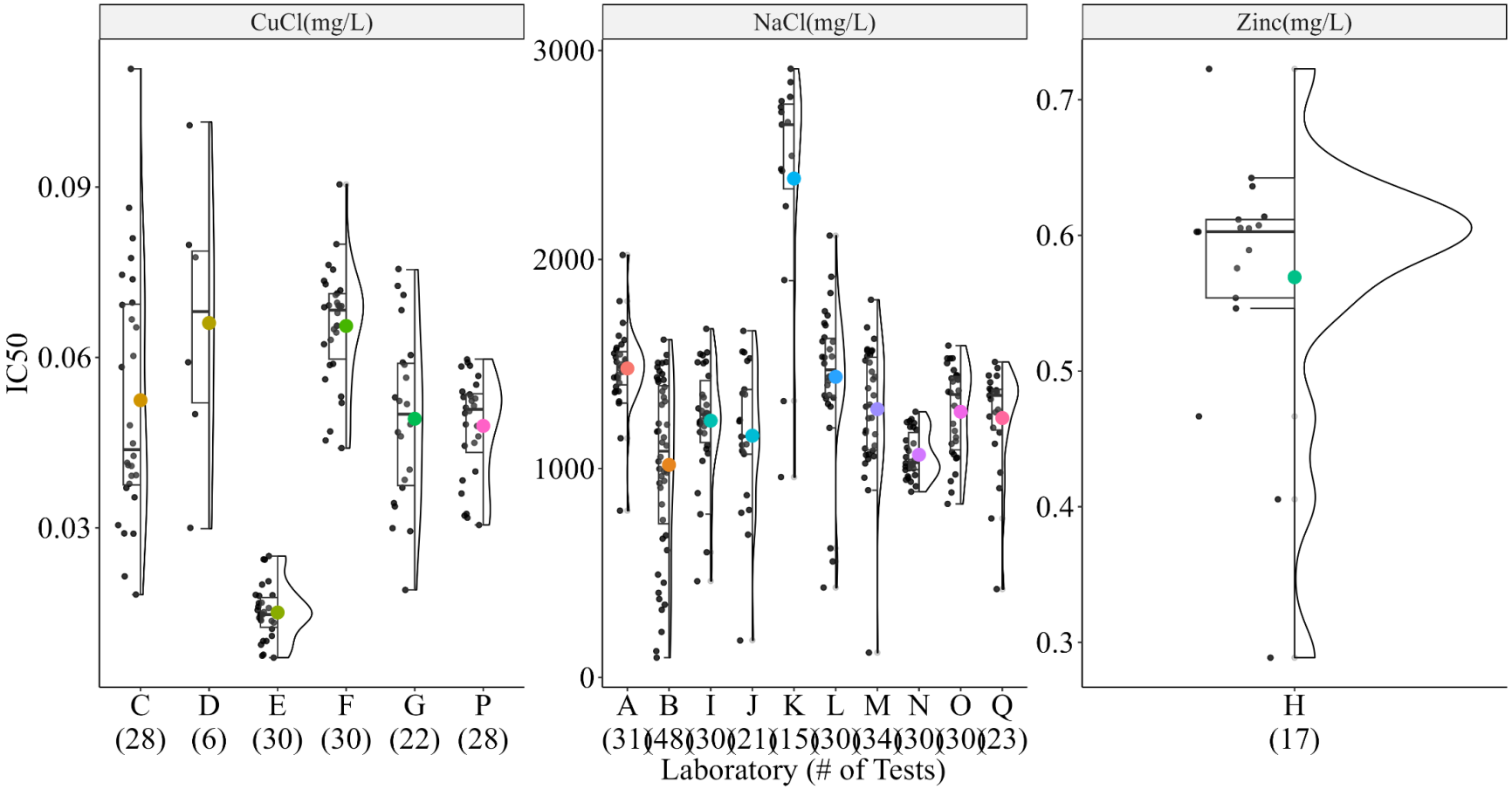


Table A6. Point estimate data for historical sodium chloride-based reference toxicant tests. Data are expressed in nominal mg/L.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
A	1041	543	1190	1369	1229	1468	2182
A	1049	816	1155	1369	1213	1441	1921
A	851	112	1199	1364	806	1524	2704
A	1134	959	1246	1532	1394	1658	2704
A	1067	383	1202	1437	1235	1520	2923
A	1243	887	1422	1697	1481	1903	2955
A	1106	928	1191	1432	1331	1489	2824
A	381	231	774	1145	478	1300	2636
A	1193	1028	1254	1486	1379	1532	2723
A	514	320	682	798	473	969	2444
A	1067	776	1212	1463	1178	1559	2957
A	1254	1156	1319	1568	1483	1640	2926
A	1140	493	1207	1447	1349	1493	2627
A	1145	1081	1196	1435	1391	1470	2286
A	1170	1064	1232	1515	1431	1588	2929
A	1186	1112	1248	1498	1436	1564	2537
A	923	603	1054	1321	1221	1386	1972
A	912	781	1053	1313	1215	1386	2824
A	1226	1112	1286	1522	1435	1574	2696
A	1212	1107	1259	1551	1446	1617	2571
A	1274	1211	1316	1593	1527	1662	3000
A	1244	908	1370	1615	1413	1832	2871
A	1219	1160	1249	1511	1467	1548	2679
A	1508	1377	1675	2021	1760	2413	2868
A	1370	1248	1472	1801	1653	1951	3000

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
A	1064	1005	1110	1410	1360	1455	2750
A	980	765	1174	1391	1170	1535	1984
A	1184	1009	1257	1474	1360	1520	2717
A	1151	949	1243	1546	1408	1662	2571

Table A6 continued. Historical point estimate data for sodium chloride-based reference toxicant tests (in nominal mg/L).

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
A	1240	1177	1313	1578	1506	1677	2654
A	1319	1267	1344	1635	1566	1683	2814
B	861	462	1200	1425	858	1697	-
B	474	363	754	1210	971	1334	-
B	589	237	900	1184	825	1349	-
B	375	103	1037	993	407	1431	-
B	288	89	504	609	376	892	-
B	111	90	554	930	640	1249	-
B	1312	625	1454	1781	1535	1977	-
B	1057	811	1211	1452	1254	1586	-
B	576	180	722	998	851	1224	-
B	1126	949	1260	1505	1391	1602	-
B	1147	1031	1285	1544	1422	1659	-
B	802	200	891	985	868	1310	-
B	219	133	586	680	407	1170	2000
B	250	105	516	492	211	792	-
B	109	87	164	218	173	608	-
B	458	374	771	941	718	1209	-
B	159	86	335	324	172	484	740
B	88	52	267	348	105	625	981
B	475	109	731	828	442	1181	-
B	269	82	415	453	212	694	-
B	210	171	336	375	230	624	883
B	723	580	831	963	840	1211	-
B	429	260	559	754	590	860	-

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
B	89	58	304	405	117	595	-
B	1004	672	1201	1424	1239	1581	-
B	1002	416	1149	1506	1375	1609	-

Table A6 continued. Historical point estimate data for sodium chloride-based reference toxicant tests (in nominal mg/L).

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
B	395	198	700	1109	488	1328	-
B	1166	498	1295	1509	1111	1594	-
B	672	82	814	908	612	1250	-
B	90	63	224	1056	186	1263	-
B	163	87	485	1125	454	1646	-
B	831	291	1208	1485	1242	1664	-
B	671	410	844	1034	807	1352	-
B	779	433	1029	1226	952	1390	-
B	1006	849	1093	1342	1261	1401	-
B	740	390	1099	1321	1110	1478	-
B	323	105	591	664	395	905	1581
B	63	45	137	126	90	207	1300
B	47	38	82	95	77	265	430
B	929	618	1192	1390	1194	1565	-
B	731	116	811	980	737	1242	-
B	1112	1019	1174	1436	1364	1487	-
B	101	74	1142	1305	1018	1474	-
B	1154	400	1300	1616	1405	1751	-
I	787	638	964	1174	1051	1260	1798
I	978	767	1141	1542	1477	1579	2025
I	845	662	1084	1250	973	1390	2266
I	657	439	871	1223	990	1352	2197

Table A6 continued. Historical point estimate data for sodium chloride-based reference toxicant tests (in nominal mg/L).

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
I	822	773	924	1186	1104	1300	2269
I	841	564	1031	1214	1093	1342	2277
I	705	559	821	1203	921	1565	2188
I	1070	729	1303	1555	1389	1697	2290
I	938	838	997	1443	1295	1548	1924
I	656	475	843	1527	1391	1583	2002
I	848	639	1001	1264	1038	1377	2238
I	1135	975	1252	1547	1450	1608	2132
I	660	553	753	1037	880	1331	1875
I	271	227	386	1072	455	1207	2188
I	493	324	678	882	730	882	1955
I	452	379	539	781	725	847	2216
I	831	717	984	1230	1167	1288	2089
I	920	680	1112	1508	1309	1592	2093
I	231	188	345	461	376	746	1367
I	536	385	761	1093	952	1206	2256
I	987	767	1110	1512	1437	1554	1826
I	283	193	528	599	387	888	1850
I	1131	882	1259	1503	1348	1599	1973
I	892	660	1118	1276	948	1487	1872
I	919	735	1048	1353	1224	1495	2250
I	738	594	970	1304	1187	1438	2238
I	725	314	1088	1168	861	1274	2017
I	843	686	1028	1288	1075	1546	2296
I	519	426	617	1110	1005	1174	2000

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
I	654	365	1143	1266	868	1532	2244
J	877	564	1167	1549	1358	1636	2727
J	1019	654	1109	1227	1099	1306	1554

Table A6 continued. Historical point estimate data for sodium chloride-based reference toxicant tests (in nominal mg/L).

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
J	1150	443	1466	1591	1382	1698	2630
J	826	634	826	1232	924	1232	1790
J	772	231	1104	1556	1353	1628	2286
J	567	201	1135	1353	888	1582	1936
J	89	77	149	178	154	367	1333
J	521	259	626	685	494	750	2377
J	452	338	661	871	641	1043	2537
J	707	186	1083	1083	709	1346	1717
J	640	189	785	1078	851	1198	2771
J	855	750	1020	1378	1190	1535	3000
J	557	412	730	1068	843	1525	2750
J	1173	447	1474	1560	1340	1689	2835
J	438	211	655	801	667	1086	1855
J	559	352	866	1116	796	1116	2434
J	1235	1122	1363	1514	1355	1639	2952
J	694	297	694	1111	877	1394	1837
J	192	150	287	788	589	1053	1788
J	824	491	1034	1153	975	1261	1759
J	775	400	1162	1358	976	1601	2605
K	2115	956	2246	2743	2603	2831	3620
K	2166	924	2252	2777	2602	2835	3252
K	1728	472	2030	2645	2570	2705	4000
K	2131	1501	2475	2759	2402	2994	4111
K	1297	796	1691	2253	1672	2616	3061
K	960	820	2109	2704	2400	2825	4257

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
K	1524	1172	1938	2494	2181	2658	3277
K	2090	1675	2293	2727	2571	2862	3384
K	131	104	484	1324	208	2541	4757

Table A6 continued. Historical point estimate data for sodium chloride-based reference toxicant tests (in nominal mg/L).

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
K	1393	406	2070	2423	1970	2713	3176
K	2104	1530	2406	2736	2277	2937	3274
K	180	130	468	958	468	1694	4364
K	1659	473	2394	2431	1681	2929	-
K	2273	2090	2413	2849	2726	2942	3463
K	1667	219	2293	2658	2141	2875	4571
L	1032	644	1316	1685	1352	2078	2889
L	1173	1027	1293	1534	1452	1657	3020
L	1189	481	1416	1840	1535	2252	2923
L	372	188	477	618	527	678	926
L	500	195	1121	1398	1022	1552	2939
L	1127	240	1321	1627	1230	1901	2871
L	566	226	1097	1338	775	1783	2879
L	551	406	976	1194	811	1695	3000
L	335	211	432	555	433	732	2654
L	1267	863	1382	1732	1533	1877	3132
L	831	700	1051	1309	955	1501	2978
L	1168	242	1394	1918	1742	2223	2965
L	239	168	1316	1606	1226	2082	2936
L	811	599	1121	1352	943	1573	2931
L	1145	418	1324	1609	1377	1794	2871
L	838	696	1042	1570	977	1853	2970
L	423	264	1067	1333	741	1484	2931
L	936	410	1168	1421	1081	1561	2879
L	472	287	1183	1502	1323	1746	2968

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
L	864	733	1112	1296	964	1476	2854
L	1091	752	1293	1535	1356	1687	2972
L	1077	385	1252	1430	1235	1548	2596

Table A6 continued. Historical point estimate data for sodium chloride-based reference toxicant tests (in nominal mg/L).

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
L	1216	465	1451	1752	1290	2235	3143
L	1219	1062	1437	1695	1533	1978	3012
L	304	197	553	430	369	1151	2571
L	1328	912	1517	2114	1781	2310	2988
L	1000	470	1114	1441	1313	1532	2971
L	862	623	1153	1436	1016	1612	3182
L	1159	495	1278	1531	1417	1634	2989
L	710	233	1130	1350	1103	1536	2955
M	856	564	1093	1296	1045	1550	2043
M	667	78	1198	1067	280	1602	1556
M	695	124	917	1133	851	1516	1886
M	910	255	1231	1438	1093	1621	1816
M	985	813	1239	1507	1293	1626	1846
M	751	685	802	956	890	1043	1488
M	807	449	1048	1350	1107	1535	1935
M	1042	733	1348	1537	1276	1698	1905
M	549	242	903	1164	649	1511	1746
M	837	495	986	1384	1241	1487	1920
M	1122	843	1342	1564	1358	1695	2229
M	59	47	261	119	94	462	139
M	329	124	626	896	520	1186	1745
M	805	701	937	1182	1008	1475	2215
M	1132	936	1281	1564	1433	1654	2130
M	1231	941	1320	1621	1528	1680	1836
M	784	458	1046	1240	1034	1517	2106

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
M	1009	515	1245	1544	1347	1630	1904
M	601	445	859	1086	680	1354	1754
M	724	237	909	1305	1007	1503	1909

Table A6 continued. Historical point estimate data for sodium chloride-based reference toxicant tests (in nominal mg/L).

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
M	941	744	1221	1463	1139	1614	1920
M	1095	855	1271	1560	1468	1647	1920
M	847	546	984	1095	914	1457	1887
M	1511	1321	1516	1807	1681	1810	1929
M	988	760	1223	1517	1329	1615	1900
M	1101	869	1253	1570	1498	1635	1906
M	785	657	904	1065	920	1371	2160
M	670	274	866	1171	889	1465	1788
M	683	549	742	1046	948	1145	1800
M	1313	508	1408	1675	1443	1739	2215
M	522	217	967	1293	792	1487	2024
M	767	669	838	1024	957	1097	1950
M	696	536	779	1058	960	1149	1841
M	586	251	1193	1426	1177	1605	1884
N	731	675	781	962	889	1134	3000
N	806	630	986	1271	1043	1393	2833
N	813	745	891	1222	993	1350	2971
N	754	698	805	1017	937	1195	2971
N	717	440	857	1025	879	1282	3000
N	812	680	882	1218	989	1329	2939
N	821	631	1031	1236	986	1401	2955
N	747	665	800	1016	923	1180	2955
N	754	688	800	1020	918	1222	2971
N	751	677	854	1001	865	1278	2971
N	776	512	878	1120	920	1320	2955

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
N	814	699	933	1198	968	1346	2938
N	749	570	805	1055	931	1226	2939
N	705	630	754	937	882	1038	2955

Table A6 continued. Historical point estimate data.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
N	728	673	752	968	918	1017	2923
N	695	649	726	889	829	953	2922
N	724	621	777	946	867	1125	2923
N	711	573	796	1085	966	1203	2955
N	832	631	900	1227	1092	1329	2970
N	708	644	745	915	856	991	2971
N	720	631	784	993	913	1163	2793
N	725	577	836	1099	936	1289	2841
N	787	581	878	1134	939	1289	2971
N	815	642	871	1199	1049	1282	2986
N	748	541	796	1043	925	1171	2986
N	729	585	778	1000	913	1129	2939
N	728	654	789	954	854	1154	2971
N	732	576	789	1032	937	1159	2955
N	752	690	808	1009	918	1191	2955
N	804	713	875	1185	1019	1288	2971
O	782	643	868	1149	1081	1244	2299
O	721	567	840	1180	1007	1351	2275
O	1000	752	1152	1424	1307	1497	2275
O	1020	864	1143	1471	1380	1534	2190
O	699	220	1291	1373	619	1542	2296
O	1116	803	1286	1588	1536	1659	2249
O	773	623	955	1214	1055	1430	2133
O	893	557	1146	1443	1281	1592	2263
O	465	339	814	1367	640	1495	2293

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
O	822	484	1100	1431	1362	1498	2176
O	736	575	899	1346	1206	1442	2181
O	909	605	1150	1504	1435	1574	2215

Table A6 continued. Historical point estimate data.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
O	1132	289	1446	1526	1165	1638	2133
O	722	587	874	1416	1334	1488	2094
O	874	684	1040	1426	1350	1492	2244
O	1008	891	1131	1527	1475	1584	2275
O	974	279	1446	1446	771	1721	2313
O	526	263	772	1067	882	1239	1798
O	966	512	1333	1527	1382	1635	2324
O	747	590	891	1363	1266	1433	2275
O	621	535	691	1079	966	1249	2287
O	542	256	825	1118	598	1388	2229
O	469	230	685	1131	940	1380	2179
O	427	361	489	885	701	987	2314
O	519	385	671	1049	602	1348	1925
O	745	540	885	1300	1119	1416	2199
O	269	165	752	830	396	1028	1891
O	539	451	674	1052	797	1173	2242
O	708	246	858	971	866	1091	2186
O	578	188	900	939	587	1216	2053
Q	528	218	1008	1195	1014	1339	1492
Q	1054	871	1153	1371	1268	1437	2684
Q	1073	754	1252	1382	1012	1502	1443
Q	1257	994	1264	1510	1338	1522	1631
Q	1038	765	1166	1359	1184	1444	1500
Q	1168	1026	1252	1445	1351	1501	1735
Q	1111	1029	1178	1413	1358	1459	1588

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
Q	1062	1000	1135	1374	1334	1424	1826
Q	1165	1078	1232	1443	1386	1488	1250
Q	787	651	864	1118	941	1239	1577

Table A6 continued. Historical point estimate data.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
Q	1060	717	1252	1375	1196	1502	1634
Q	900	729	1102	1282	1025	1402	1526
Q	791	687	1045	1172	899	1363	1758
Q	1040	932	1095	1360	1309	1397	1727
Q	992	233	1144	1332	1072	1429	1556
Q	886	767	1018	1269	1154	1345	1556
Q	1191	1121	1252	1481	1431	1535	1632
Q	870	706	1136	1249	971	1424	1489
Q	1024	756	1211	1349	1034	1474	1526
Q	725	234	872	980	837	1263	1622
Q	483	215	605	761	666	882	2696
Q	521	231	682	905	703	1102	1489
Q	176	104	296	423	209	634	1657

Table A7. Point estimate data for historical copper-based reference toxicant tests. Data are expressed in nominal mg/L.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
C	0.062	0.056	0.066	0.081	0.076	0.085	0.15
C	0.058	0.024	0.064	0.078	0.072	0.086	0.136
C	0.037	0.027	0.050	0.06	0.045	0.070	0.089
C	0.045	0.036	0.058	0.067	0.054	0.076	0.081
C	0.053	0.041	0.060	0.07	0.063	0.075	0.081
C	0.031	0.027	0.040	0.045	0.038	0.061	0.061
C	0.062	0.051	0.063	0.075	0.067	0.075	0.075
C	0.019	0.014	0.025	0.029	0.022	0.033	0.038
C	0.040	0.022	0.054	0.065	0.049	0.074	0.110
C	0.028	0.026	0.030	0.038	0.035	0.043	0.044
C	0.027	0.021	0.033	0.041	0.037	0.047	0.107
C	0.066	0.046	0.071	0.086	0.073	0.094	0.161
C	0.029	0.025	0.033	0.039	0.035	0.044	0.046
C	0.028	0.009	0.044	0.048	0.036	0.126	0.171
C	0.018	0.006	0.033	0.033	0.011	0.049	0.092
C	0.033	0.008	0.035	0.041	0.032	0.045	0.100
C	0.030	0.017	0.036	0.043	0.035	0.056	0.119
C	0.030	0.024	0.035	0.039	0.035	0.047	0.046
C	0.029	0.025	0.033	0.037	0.034	0.041	0.041
C	0.020	0.016	0.026	0.030	0.024	0.036	0.039
C	0.028	0.015	0.033	0.041	0.036	0.054	0.054
C	0.052	0.043	0.057	0.074	0.070	0.076	0.145
C	0.039	0.024	0.053	0.058	0.046	0.069	0.078
C	0.023	0.019	0.039	0.069	0.044	0.089	0.148
C	0.074	0.059	0.087	0.111	0.090	0.126	0.167

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
C	0.008	0.005	0.015	0.018	0.011	0.025	0.154
C	0.011	0.007	0.018	0.021	0.015	0.025	0.139
C	0.024	0.016	0.032	0.039	0.030	0.055	0.137

Table A7 continued. Historical point estimate data.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
D	0.046	0.043	0.049	0.059	0.056	0.062	0.079
D	0.032	0.009	0.047	0.050	0.032	0.060	0.070
D	0.055	0.045	0.071	0.079	0.065	0.096	0.103
D	0.080	0.075	0.087	0.101	0.098	0.106	0.111
D	0.036	0.025	0.059	0.077	0.054	0.091	0.108
D	0.019	0.008	0.028	0.030	0.023	0.044	0.045
E	0.006	0.006	0.006	0.008	0.007	0.008	0.009
E	0.007	0.003	0.008	0.010	0.009	0.015	0.026
E	0.010	0.004	0.013	0.017	0.013	0.018	0.022
E	0.008	0.007	0.010	0.014	0.009	0.018	0.022
E	0.006	0.004	0.006	0.007	0.006	0.008	0.009
E	0.006	0.004	0.007	0.007	0.006	0.008	0.010
E	0.009	0.007	0.012	0.015	0.010	0.018	0.024
E	0.008	0.007	0.009	0.014	0.011	0.016	0.028
E	0.009	0.008	0.010	0.016	0.015	0.018	0.033
E	0.007	0.006	0.008	0.009	0.009	0.011	0.028
E	0.012	0.005	0.015	0.018	0.015	0.020	0.021
E	0.009	0.008	0.012	0.016	0.012	0.018	0.026
E	0.012	0.010	0.014	0.018	0.017	0.020	0.025
E	0.008	0.007	0.011	0.011	0.008	0.017	0.019
E	0.009	0.007	0.012	0.016	0.013	0.018	0.028
E	0.016	0.013	0.018	0.025	0.023	0.027	0.040
E	0.008	0.007	0.012	0.014	0.009	0.019	0.025
E	0.009	0.008	0.010	0.017	0.011	0.020	0.036
E	0.014	0.009	0.017	0.021	0.017	0.025	0.026

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
E	0.005	0.004	0.006	0.010	0.008	0.013	0.030
E	0.008	0.007	0.011	0.014	0.010	0.017	0.023
E	0.004	0.002	0.011	0.015	0.011	0.018	0.023

Table A7 continued. Historical point estimate data.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
E	0.015	0.015	0.015	0.020	0.020	0.020	0.021
E	0.017	0.013	0.021	0.024	0.021	0.031	0.035
E	0.008	0.006	0.009	0.012	0.009	0.015	0.026
E	0.009	0.008	0.010	0.015	0.012	0.017	0.022
E	0.014	0.010	0.019	0.024	0.019	0.032	0.038
E	0.008	0.006	0.011	0.015	0.011	0.017	0.025
E	0.012	0.009	0.015	0.018	0.016	0.020	0.022
E	0.008	0.007	0.009	0.013	0.011	0.015	0.021
F	0.036	0.031	0.049	0.047	0.040	0.073	0.125
F	0.068	0.054	0.076	0.090	0.078	0.110	0.139
F	0.035	0.012	0.047	0.045	0.038	0.068	0.117
F	0.038	0.033	0.051	0.053	0.043	0.069	0.084
F	0.053	0.037	0.060	0.069	0.056	0.074	0.082
F	0.038	0.031	0.052	0.059	0.045	0.071	0.096
F	0.048	0.031	0.062	0.069	0.047	0.079	0.092
F	0.040	0.028	0.057	0.062	0.044	0.072	0.080
F	0.053	0.041	0.060	0.069	0.059	0.073	0.085
F	0.057	0.042	0.062	0.071	0.061	0.075	0.080
F	0.048	0.037	0.060	0.066	0.051	0.074	0.079
F	0.061	0.053	0.065	0.076	0.070	0.083	0.100
F	0.052	0.022	0.062	0.069	0.048	0.076	0.100
F	0.062	0.058	0.065	0.080	0.077	0.082	0.121
F	0.037	0.031	0.055	0.056	0.041	0.070	0.086
F	0.034	0.029	0.037	0.044	0.040	0.049	0.084
F	0.040	0.034	0.056	0.059	0.045	0.071	0.086

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
F	0.058	0.047	0.064	0.075	0.068	0.085	0.105
F	0.045	0.039	0.053	0.064	0.059	0.069	0.086
F	0.054	0.037	0.061	0.070	0.057	0.074	0.079

Table A7 continued. Historical point estimate data.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
F	0.059	0.051	0.063	0.074	0.069	0.078	0.113
F	0.059	0.045	0.063	0.073	0.064	0.075	0.082
F	0.036	0.028	0.046	0.052	0.041	0.067	0.089
F	0.044	0.035	0.060	0.063	0.046	0.073	0.079
F	0.048	0.037	0.059	0.066	0.050	0.072	0.081
F	0.059	0.046	0.063	0.073	0.066	0.077	0.085
F	0.052	0.037	0.063	0.069	0.049	0.075	0.109
F	0.057	0.046	0.061	0.071	0.065	0.074	0.075
F	0.045	0.034	0.056	0.065	0.057	0.070	0.074
F	0.057	0.046	0.062	0.072	0.066	0.075	0.083
G	0.059	0.034	0.063	0.073	0.062	0.075	0.075
G	0.052	0.038	0.061	0.068	0.054	0.074	0.073
G	0.038	0.027	0.053	0.053	0.042	0.068	0.071
G	0.031	0.012	0.037	0.040	0.027	0.050	0.073
G	0.033	0.010	0.044	0.046	0.035	0.064	0.070
G	0.016	0.007	0.033	0.037	0.020	0.061	0.074
G	0.014	0.007	0.026	0.029	0.016	0.040	0.102
G	0.036	0.011	0.057	0.060	0.042	0.072	0.067
G	0.037	0.030	0.051	0.052	0.043	0.067	0.075
G	0.015	0.009	0.017	0.019	0.016	0.021	0.029
G	0.034	0.021	0.049	0.048	0.038	0.066	0.065
G	0.021	0.011	0.031	0.030	0.019	0.037	0.075
G	0.008	0.006	0.058	0.059	0.011	0.072	0.071
G	0.056	0.038	0.063	0.071	0.063	0.075	0.072
G	0.025	0.007	0.031	0.034	0.022	0.038	0.041

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
G	0.041	0.032	0.053	0.059	0.043	0.069	0.077
G	0.026	0.009	0.045	0.047	0.036	0.065	0.073
G	0.019	0.017	0.023	0.034	0.022	0.057	0.064

Table A7 continued. Historical point estimate data.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
G	0.027	0.008	0.054	0.052	0.028	0.069	0.082
G	0.030	0.020	0.032	0.038	0.033	0.041	0.047
G	0.063	0.032	0.063	0.075	0.065	0.076	0.081
G	0.026	0.009	0.053	0.057	0.026	0.069	0.075
P	0.047	0.044	0.049	0.058	0.056	0.062	0.066
P	0.030	0.026	0.042	0.045	0.035	0.054	0.054
P	0.028	0.024	0.039	0.040	0.032	0.051	0.050
P	0.026	0.025	0.028	0.032	0.030	0.036	0.034
P	0.040	0.035	0.045	0.053	0.050	0.056	0.059
P	0.042	0.040	0.044	0.053	0.052	0.054	0.056
P	0.040	0.032	0.045	0.051	0.046	0.055	0.055
P	0.027	0.025	0.031	0.036	0.031	0.045	0.050
P	0.039	0.032	0.046	0.053	0.046	0.057	0.058
P	0.042	0.035	0.045	0.054	0.050	0.056	0.064
P	0.040	0.027	0.046	0.051	0.041	0.056	0.056
P	0.038	0.026	0.043	0.051	0.042	0.054	0.056
P	0.024	0.022	0.026	0.031	0.029	0.034	0.050
P	0.030	0.021	0.036	0.044	0.037	0.049	0.055
P	0.035	0.030	0.042	0.050	0.046	0.054	0.060
P	0.036	0.029	0.041	0.048	0.045	0.051	0.052
P	0.048	0.044	0.049	0.059	0.055	0.063	0.065
P	0.041	0.018	0.043	0.052	0.049	0.054	0.055
P	0.048	0.044	0.050	0.059	0.056	0.062	0.074
P	0.025	0.022	0.029	0.038	0.033	0.044	0.051
P	0.045	0.040	0.047	0.055	0.052	0.057	0.058

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
P	0.037	0.030	0.041	0.048	0.043	0.051	0.053
P	0.042	0.038	0.044	0.053	0.050	0.054	0.056
P	0.025	0.021	0.027	0.032	0.029	0.036	0.048

Table A7 continued. Historical point estimate data.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
P	0.022	0.008	0.026	0.033	0.028	0.040	0.048
P	0.044	0.037	0.048	0.057	0.054	0.059	0.065
P	0.033	0.029	0.037	0.046	0.043	0.048	0.053
P	0.049	0.044	0.051	0.060	0.056	0.063	0.062

Table A8. Point estimate data for historical zinc-based reference toxicant tests. Data are expressed in nominal mg/L.

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
H	0.179	0.090	0.257	0.289	0.183	0.337	0.669
H	0.481	0.437	0.529	0.607	0.576	0.638	1.13
H	0.296	0.267	0.334	0.467	0.401	0.532	0.927
H	0.496	0.420	0.533	0.614	0.565	0.639	0.747
H	0.554	0.510	0.624	0.723	0.659	0.844	1.28
H	0.479	0.320	0.529	0.603	0.540	0.636	0.770
H	0.271	0.213	0.323	0.405	0.347	0.501	0.691
H	0.393	0.315	0.465	0.554	0.435	0.595	0.834
H	0.537	0.483	0.538	0.642	0.605	0.645	1.12
H	0.474	0.359	0.541	0.612	0.524	0.661	0.877
H	0.514	0.439	0.546	0.637	0.588	0.661	1.16
H	0.439	0.359	0.511	0.576	0.519	0.624	0.691
H	0.399	0.318	0.501	0.554	0.469	0.617	0.728
H	0.478	0.450	0.503	0.603	0.584	0.619	0.698
H	0.529	0.486	0.536	0.636	0.607	0.641	0.762
H	0.479	0.452	0.501	0.605	0.587	0.621	0.780

Lab	IC25	Lower 95% CI Bound	Upper 95% CI Bound	IC50	Lower 95% CI Bound	Upper 95% CI Bound	LC50
H	0.347	0.295	0.432	0.546	0.501	0.577	0.824

Table A9. Percent minimum significant difference (PMSD) for reference toxicants in the historical dataset.

Lab	Mean PMSD	CV of Mean PMSD	Range of Mean PMSD
A	16.73	0.534	6.95 - 44.23
B	31.83	0.468	12.02 - 102
C	24.11	0.574	7.65 – 79.66
D	23.88	0.544	11.44 - 46.54
E	19.62	0.246	11.36 – 31.19
F	22.16	0.437	8.64 – 57.11
G	31.24	0.262	18.52 – 45.86
H	28.62	1.299	6.66 – 159.05
I	18.78	0.828	7.88 – 80.06
J	32.65	0.795	14.22 - 121
K	29.40	0.858	9.47 - 109
L	22.33	0.254	10.95 – 34.35
M	22.68	0.493	11.22 – 53.63
N	15.64	0.268	9.46 – 24.21
O	17.86	0.599	7.31 – 47.63
P	15.61	0.392	7.56 – 40.79
Q	18.77	0.394	8.37 – 33.39

Correlations

The following plots are heat maps of Spearmen Rank Correlations among variables. Biological variables, lab technique, and water quality variables are all combined to give a first impression picture of relationships. Data are analyzed on a per test basis (so test means or single values, whichever was available). The correlation coefficient values that created the heat map can be found at https://data.sccwrp.org/owncloud/apps/files_sharing/get.php?token=e346721d85d30d687ab75143706ea49606b7d47a&path=/2021-12-06-19:00:08. As noted in the data summary tables, not all labs provided all of the data, so the individual parameters evaluated will vary among different labs.

Figure A18. Heat map of spearman rank correlations between all variables with all labs and all tests combined. Not the best way to do this, but a big picture start.

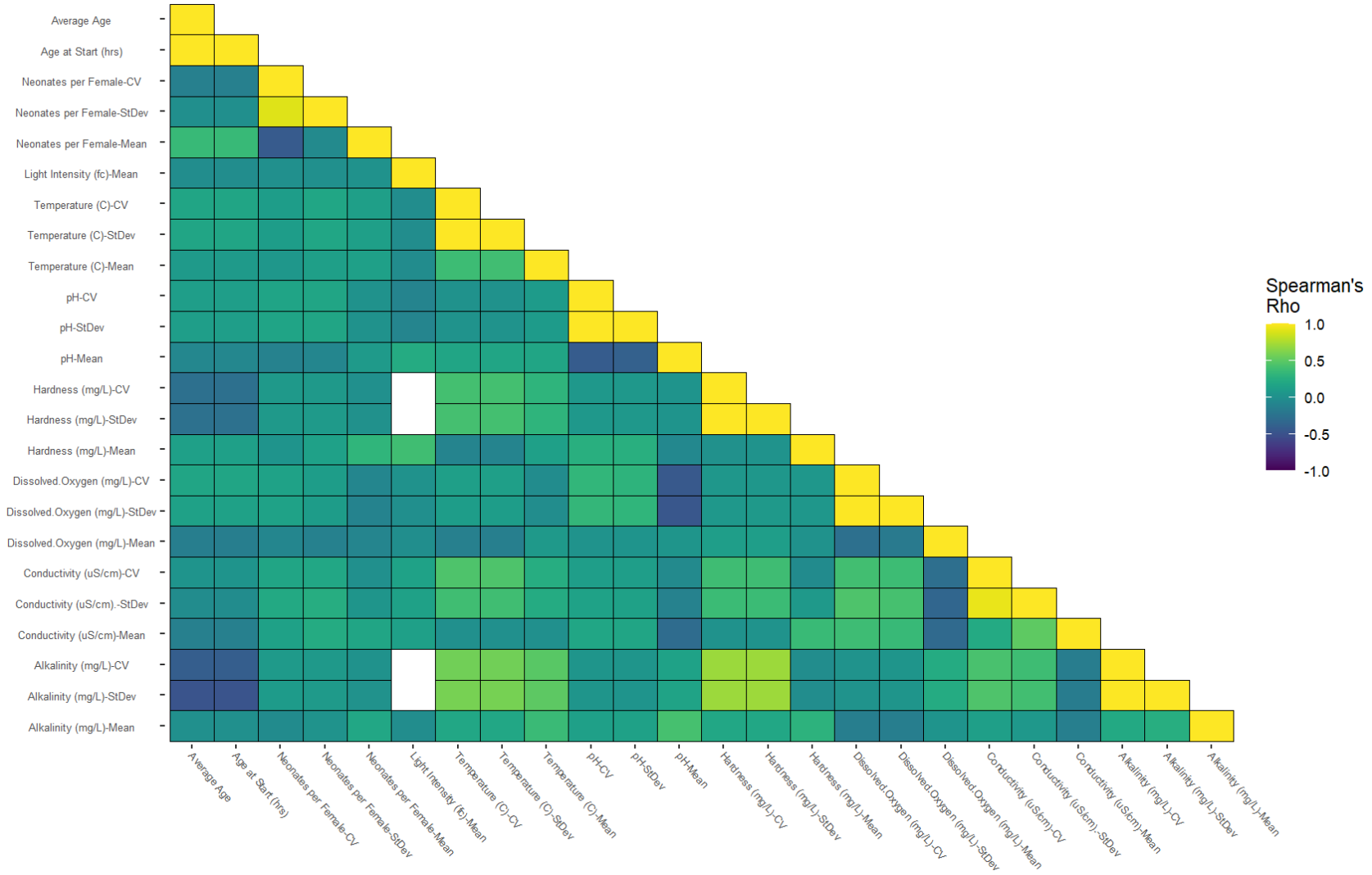
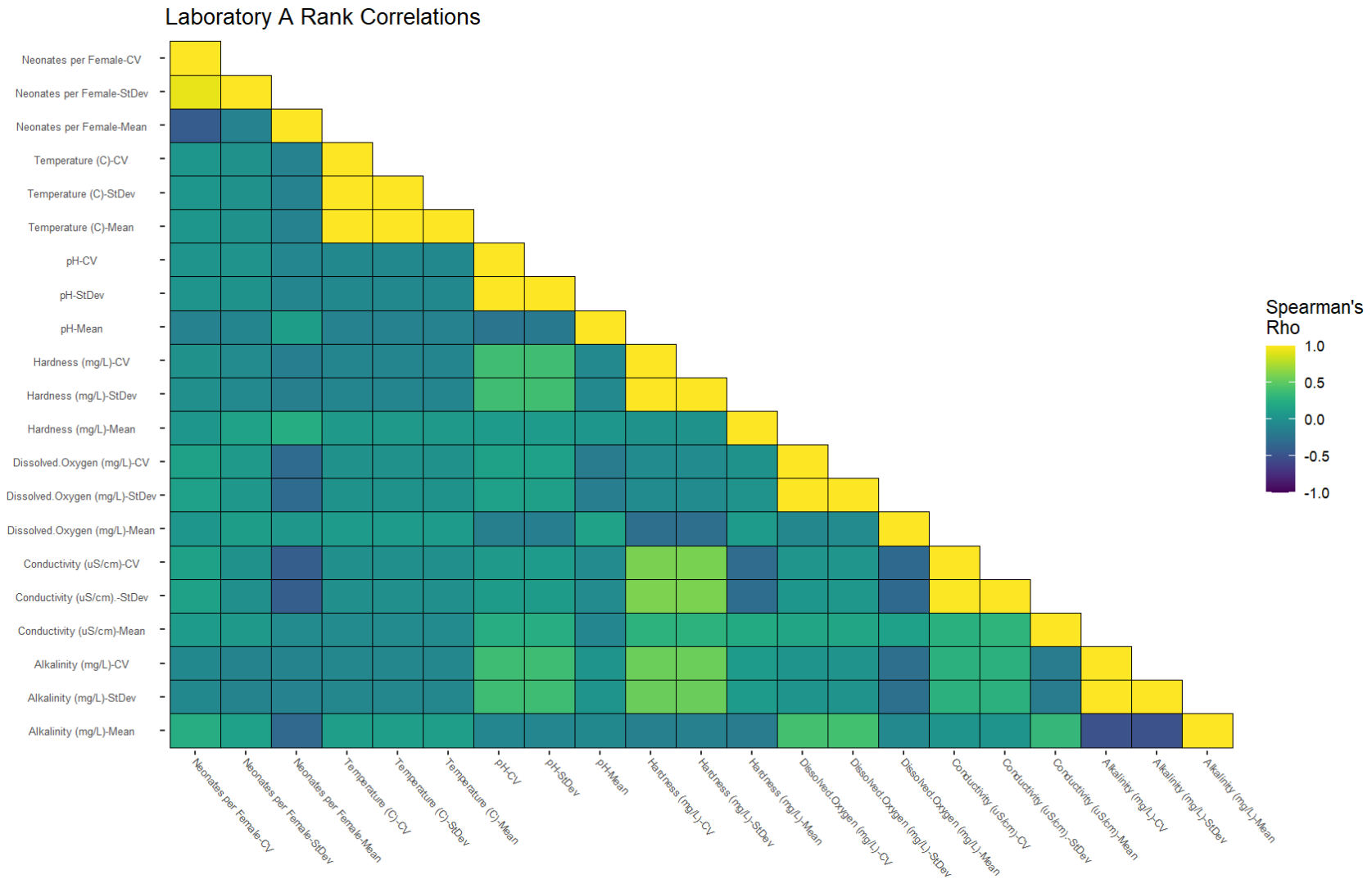
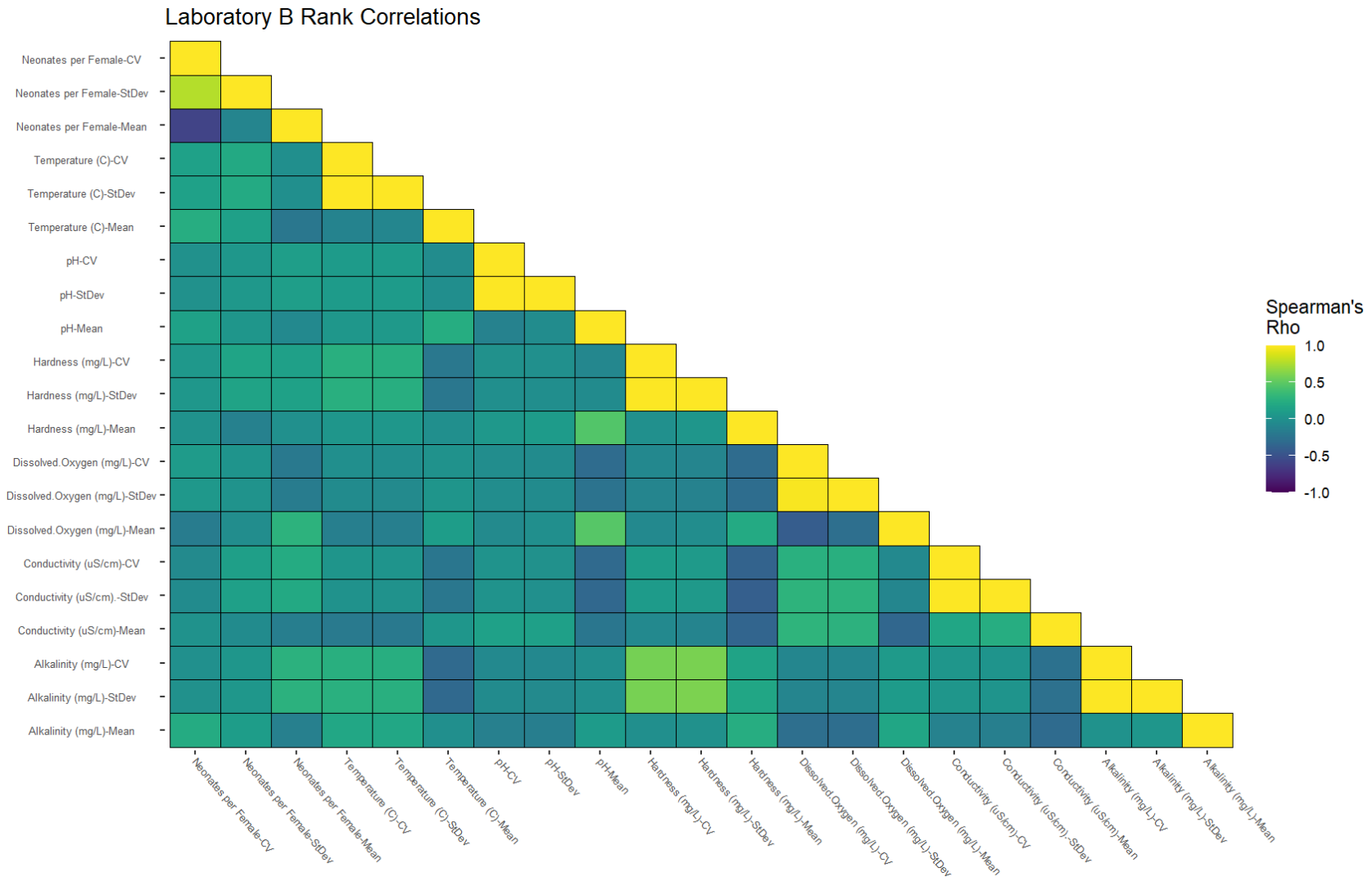
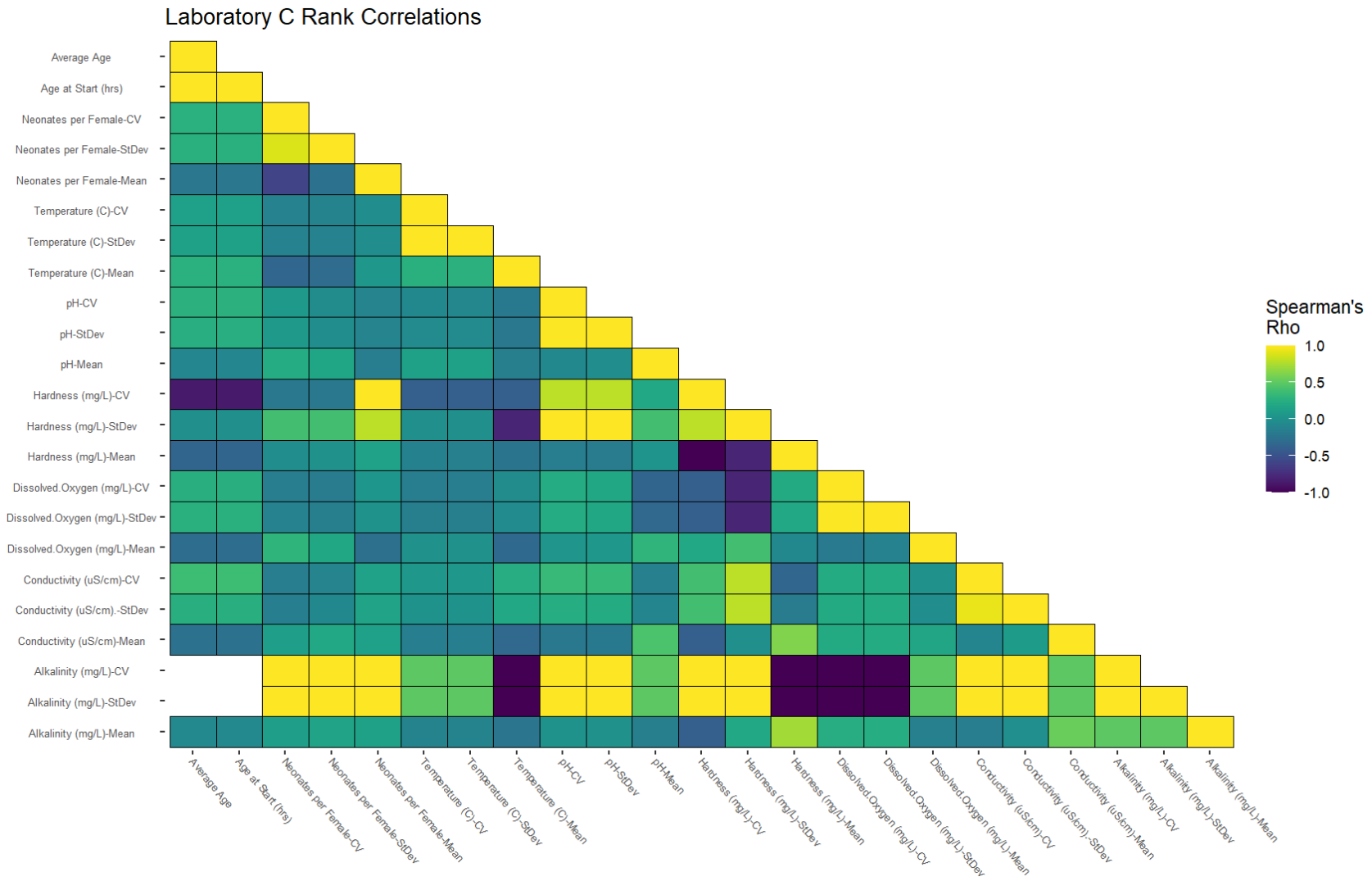
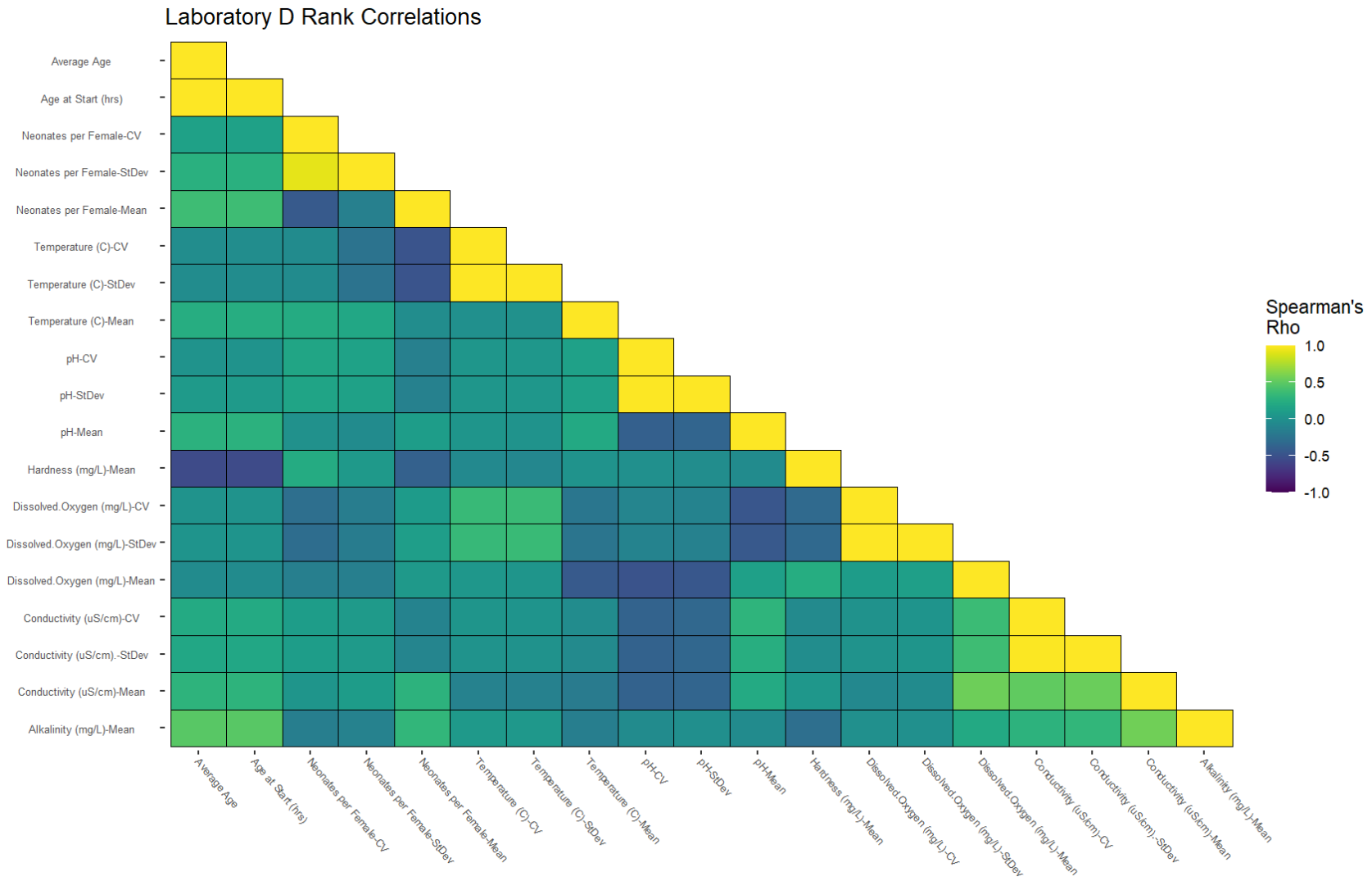


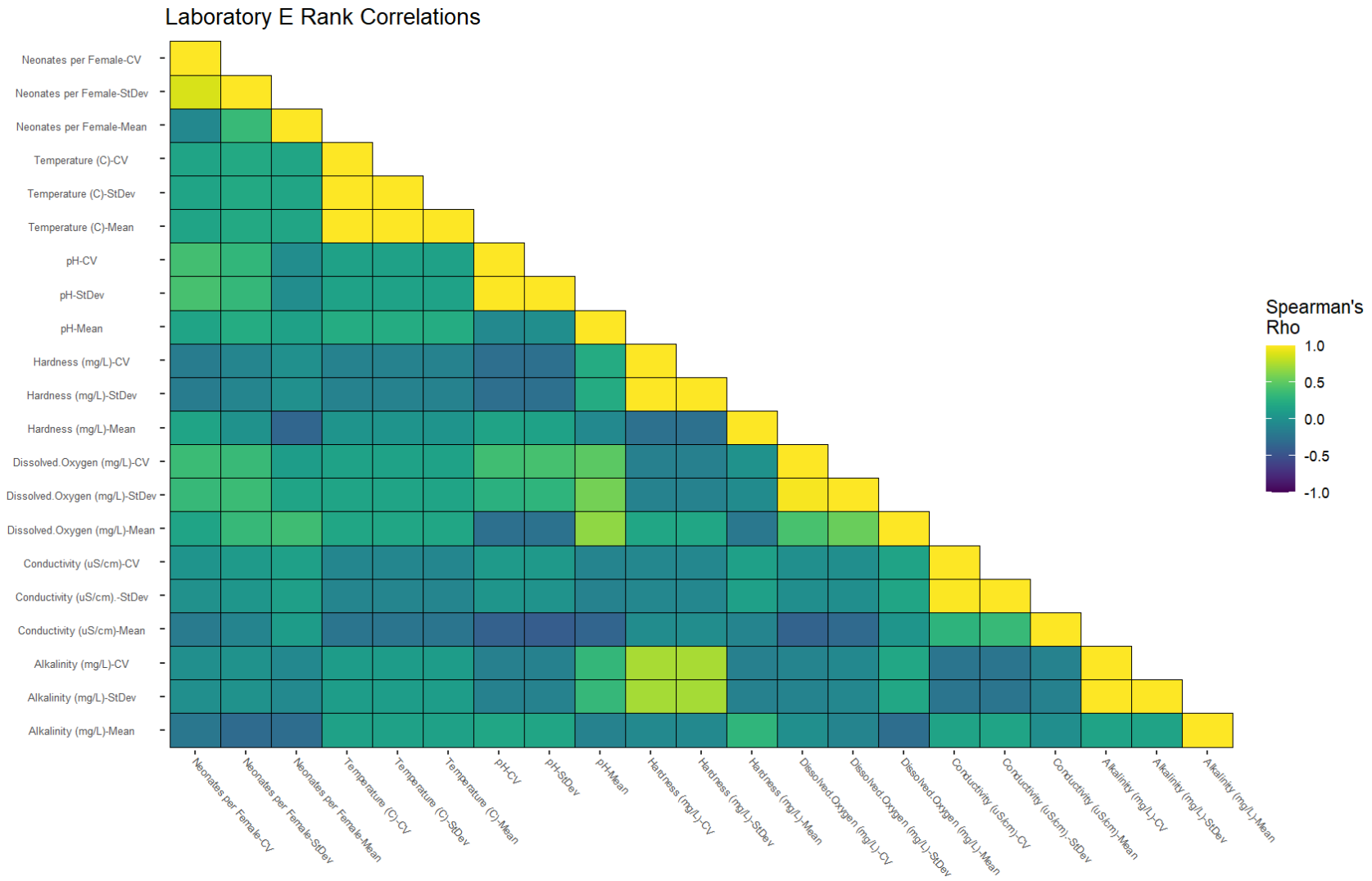
Figure A19. Spearman Rank Correlations by laboratory

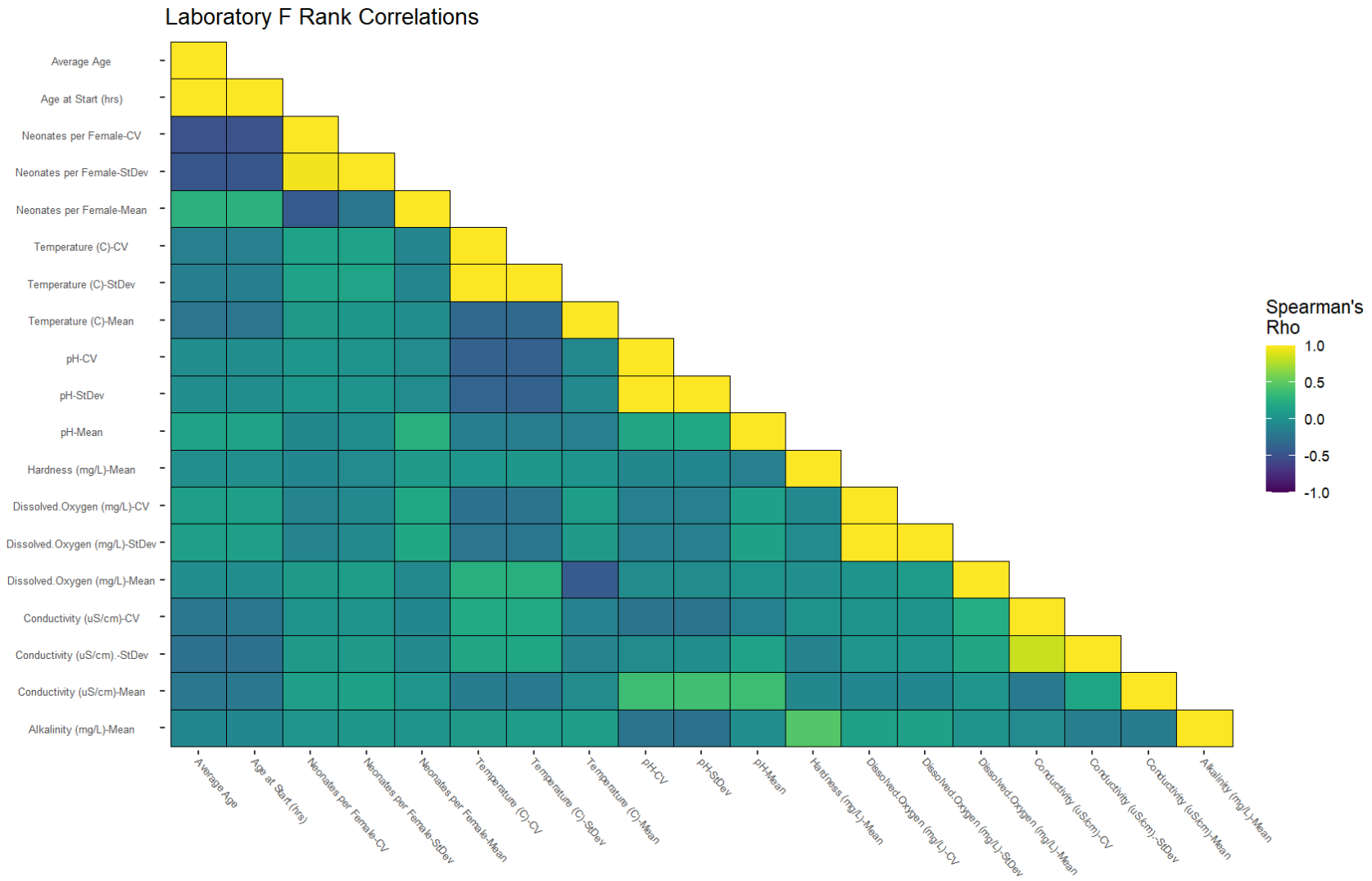


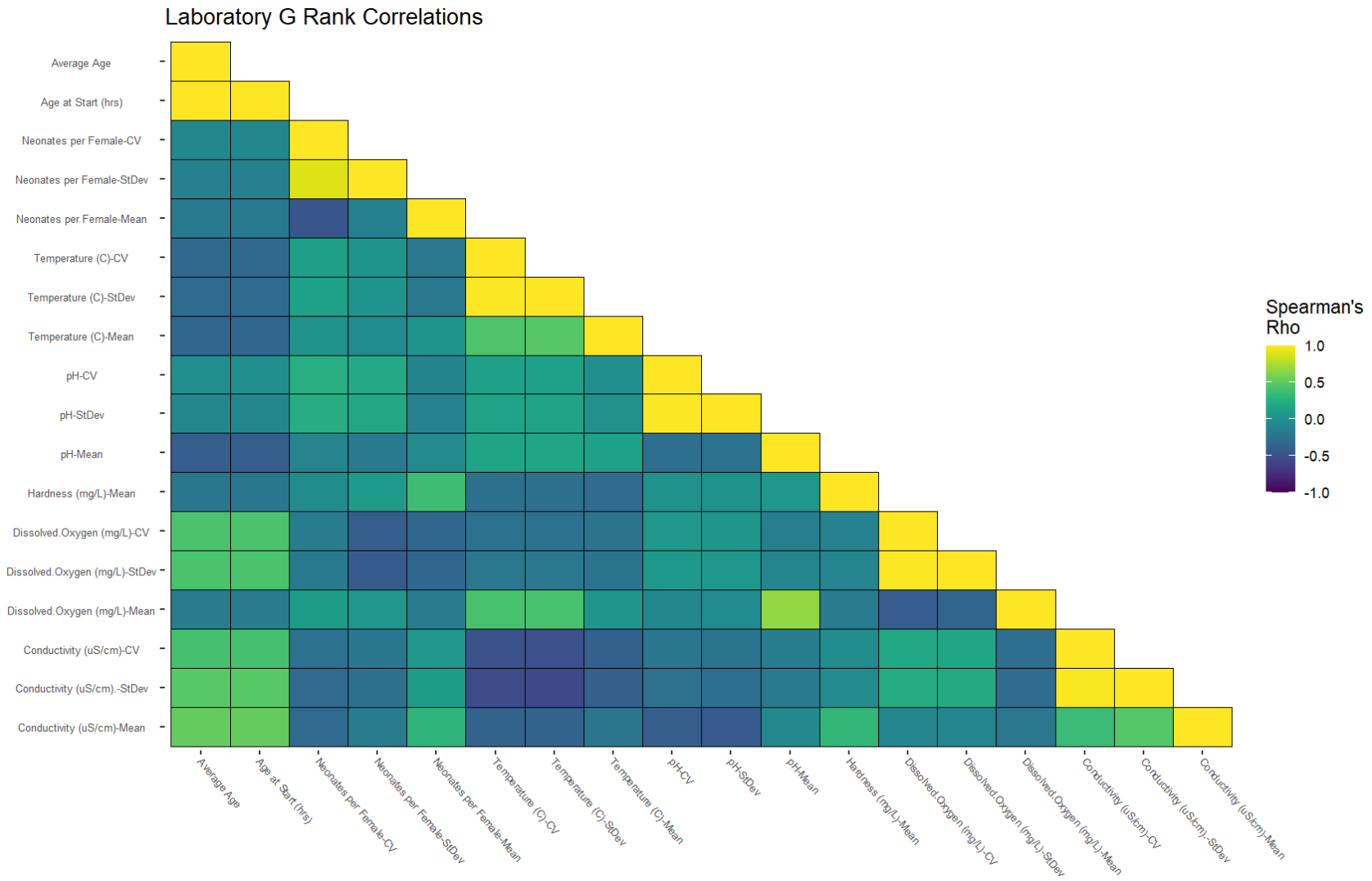

Laboratory B Rank Correlations
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-CV
Hardness (mg/L)-StDev
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-CV
Alkalinity (mg/L)-StDev
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

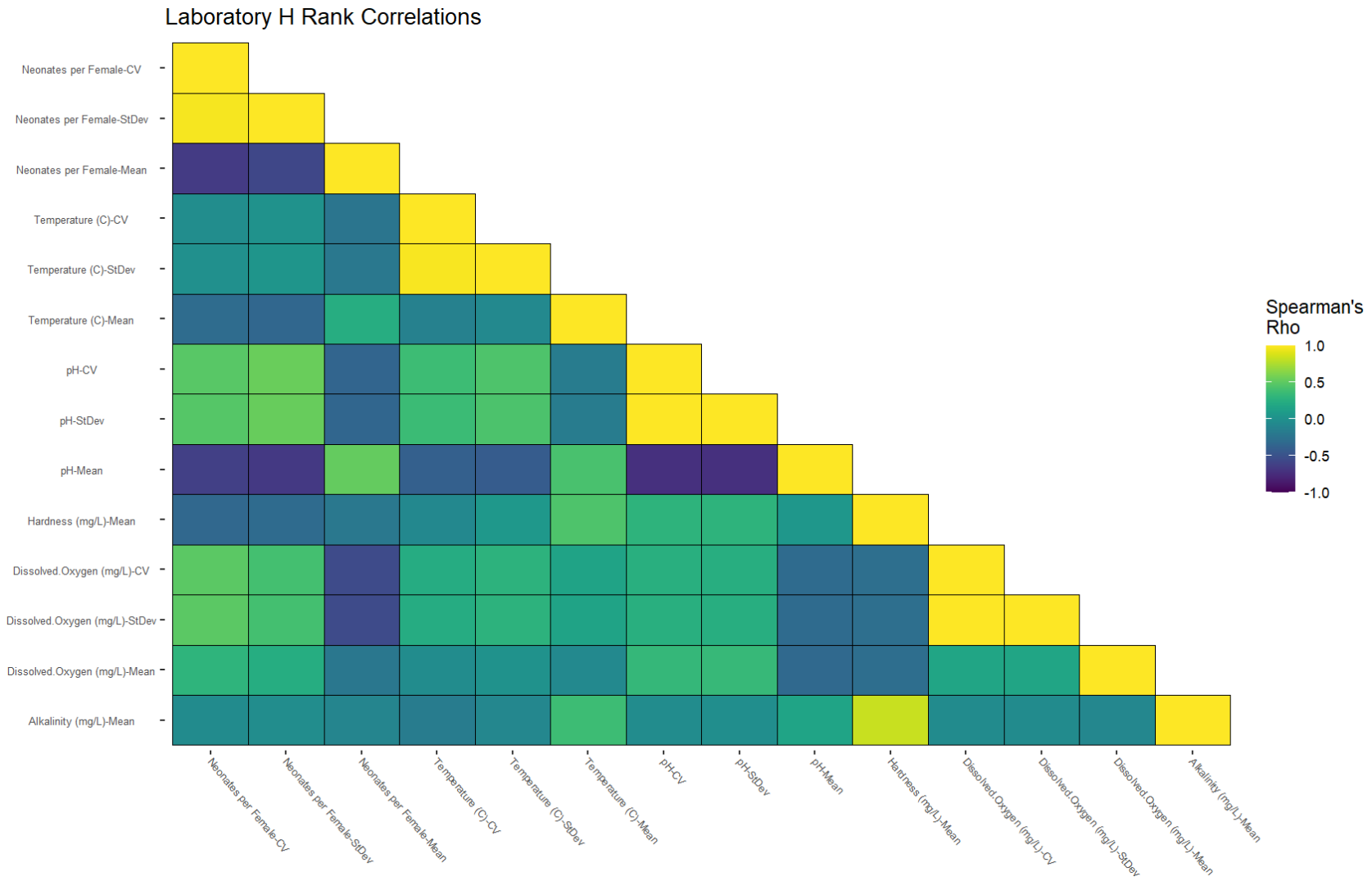

Laboratory C Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-CV
Hardness (mg/L)-StDev
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-CV
Alkalinity (mg/L)-StDev
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

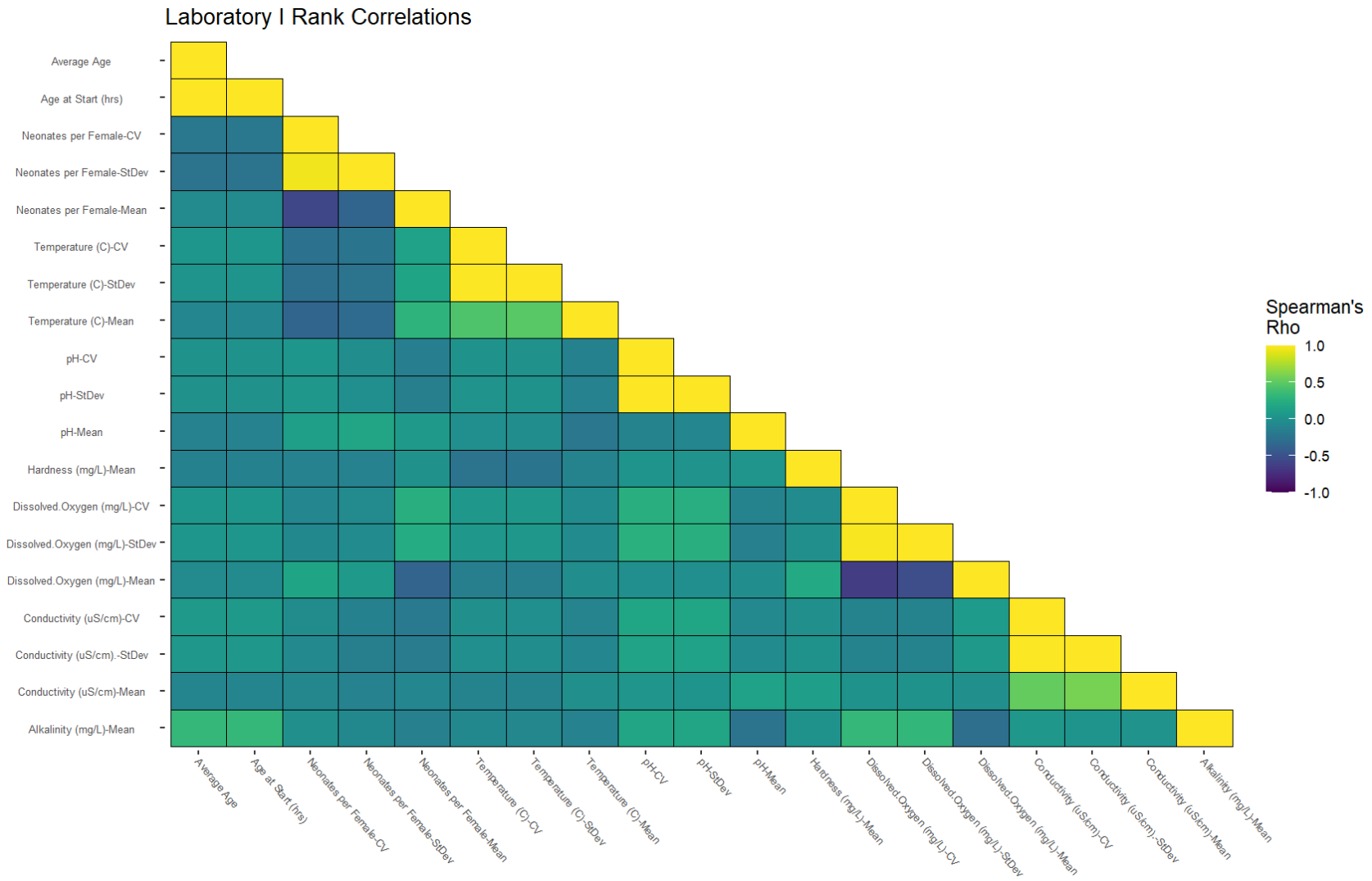

Laboratory D Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

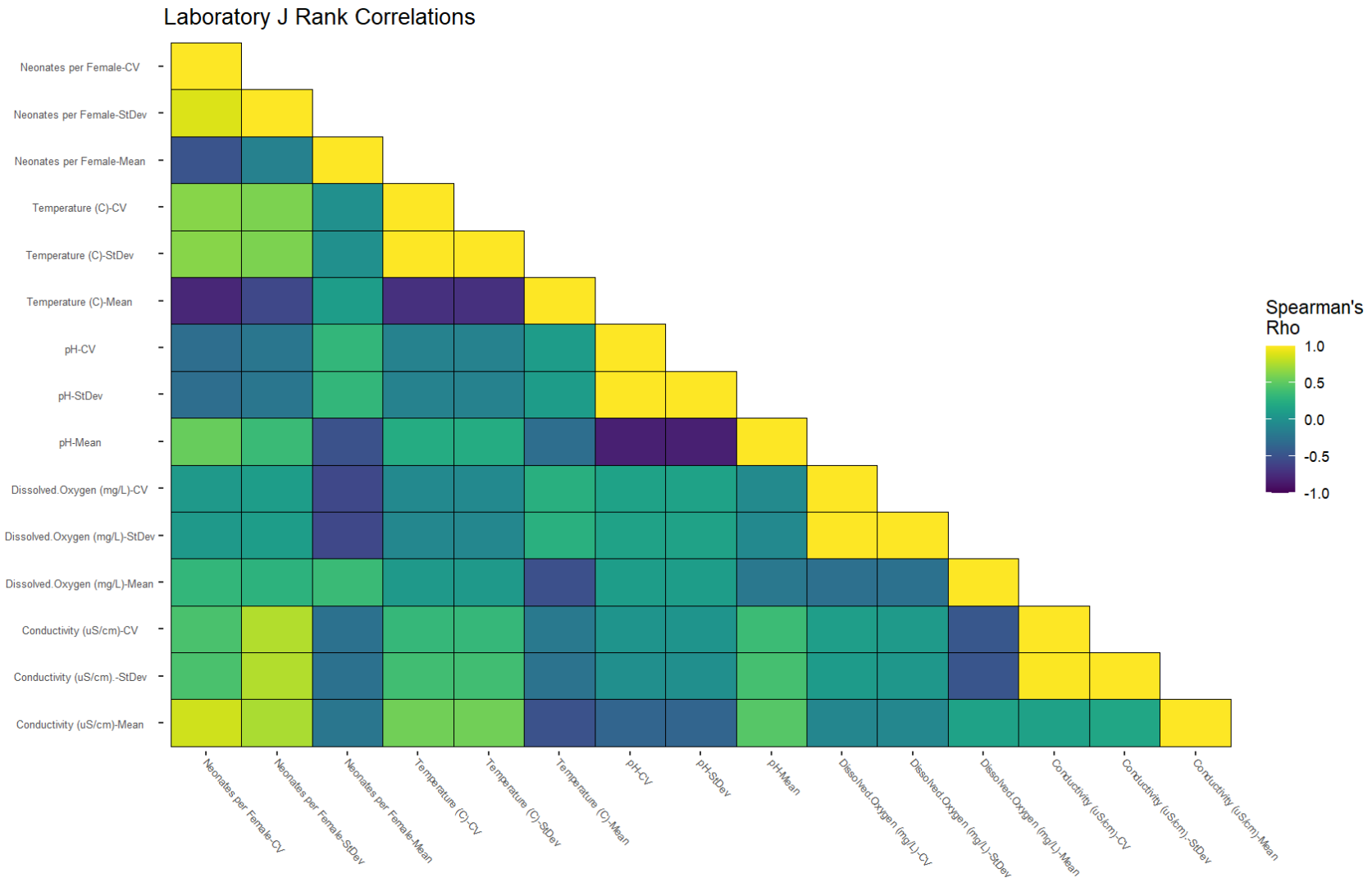

Laboratory E Rank Correlations
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-CV
Hardness (mg/L)-StDev
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-CV
Alkalinity (mg/L)-StDev
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

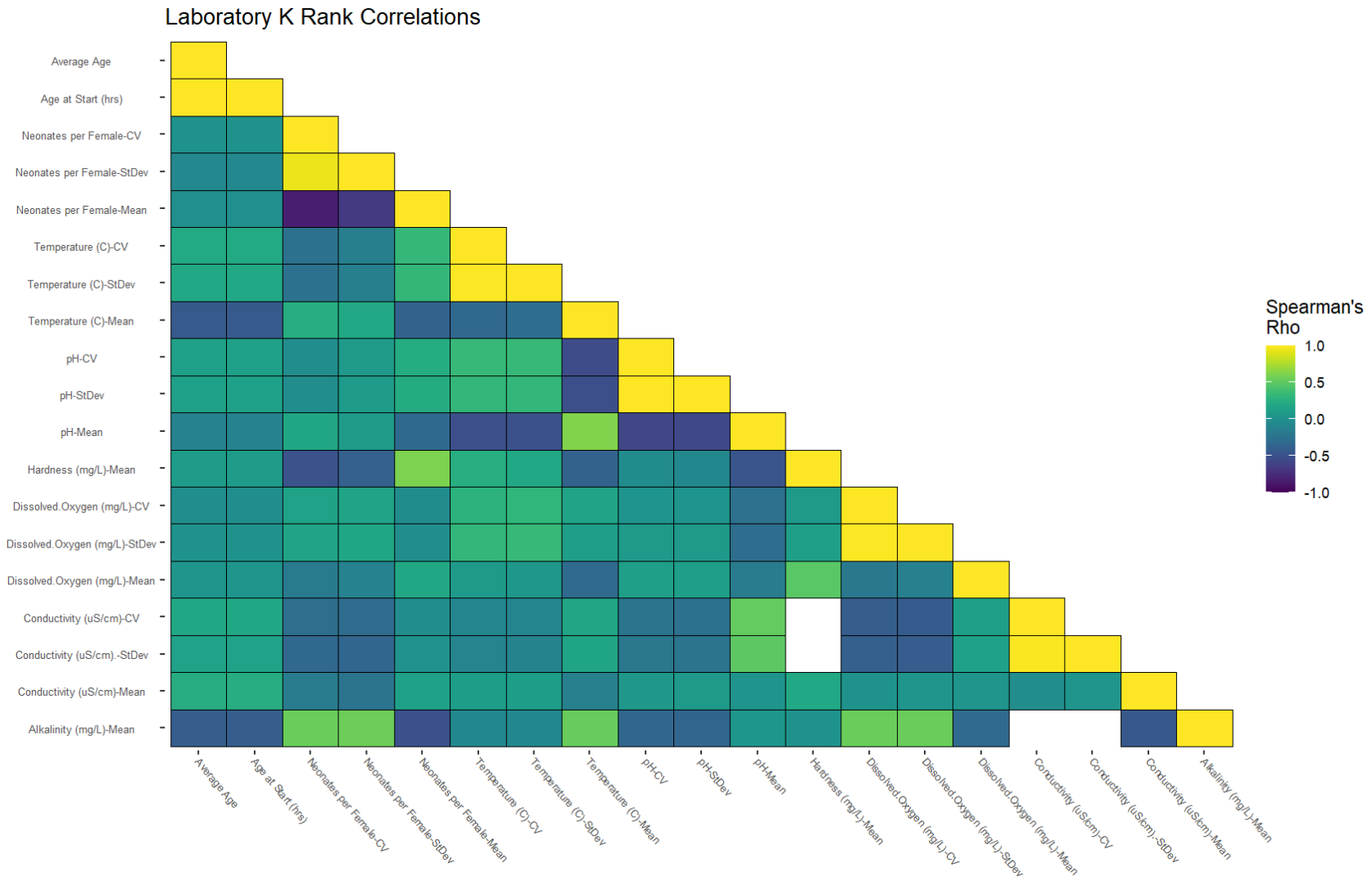

Laboratory F Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

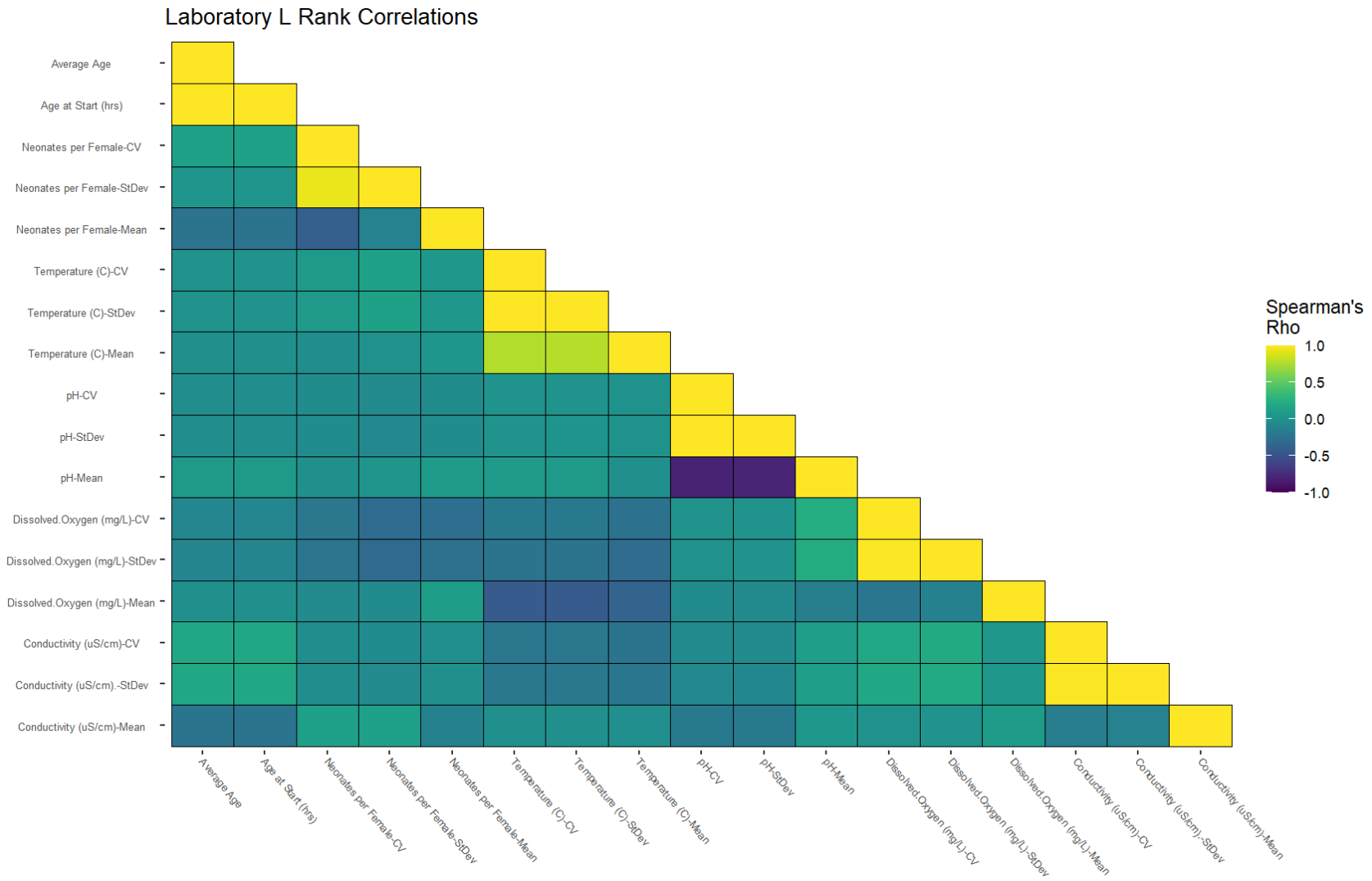

Laboratory G Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

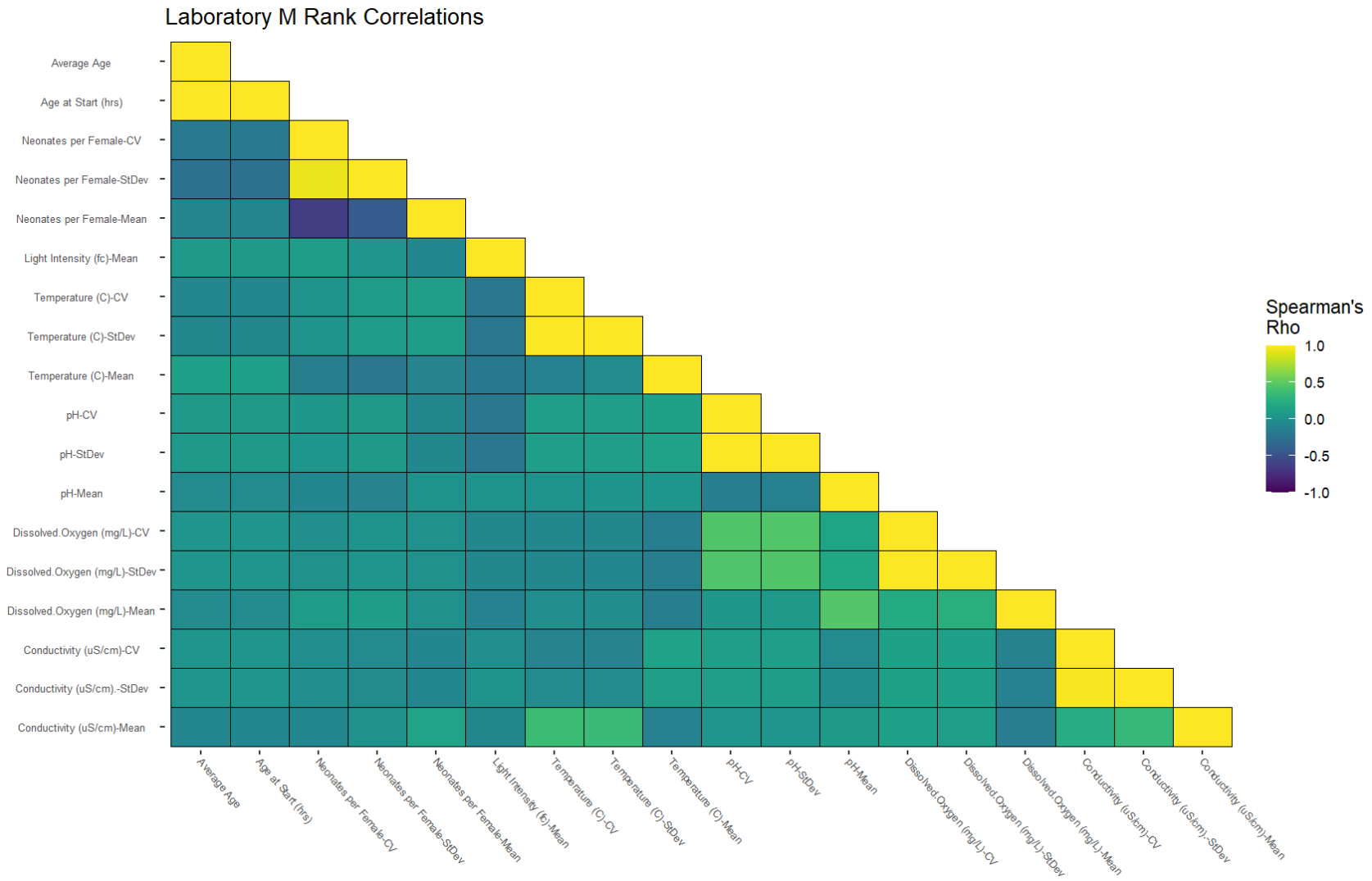

Laboratory H Rank Correlations
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

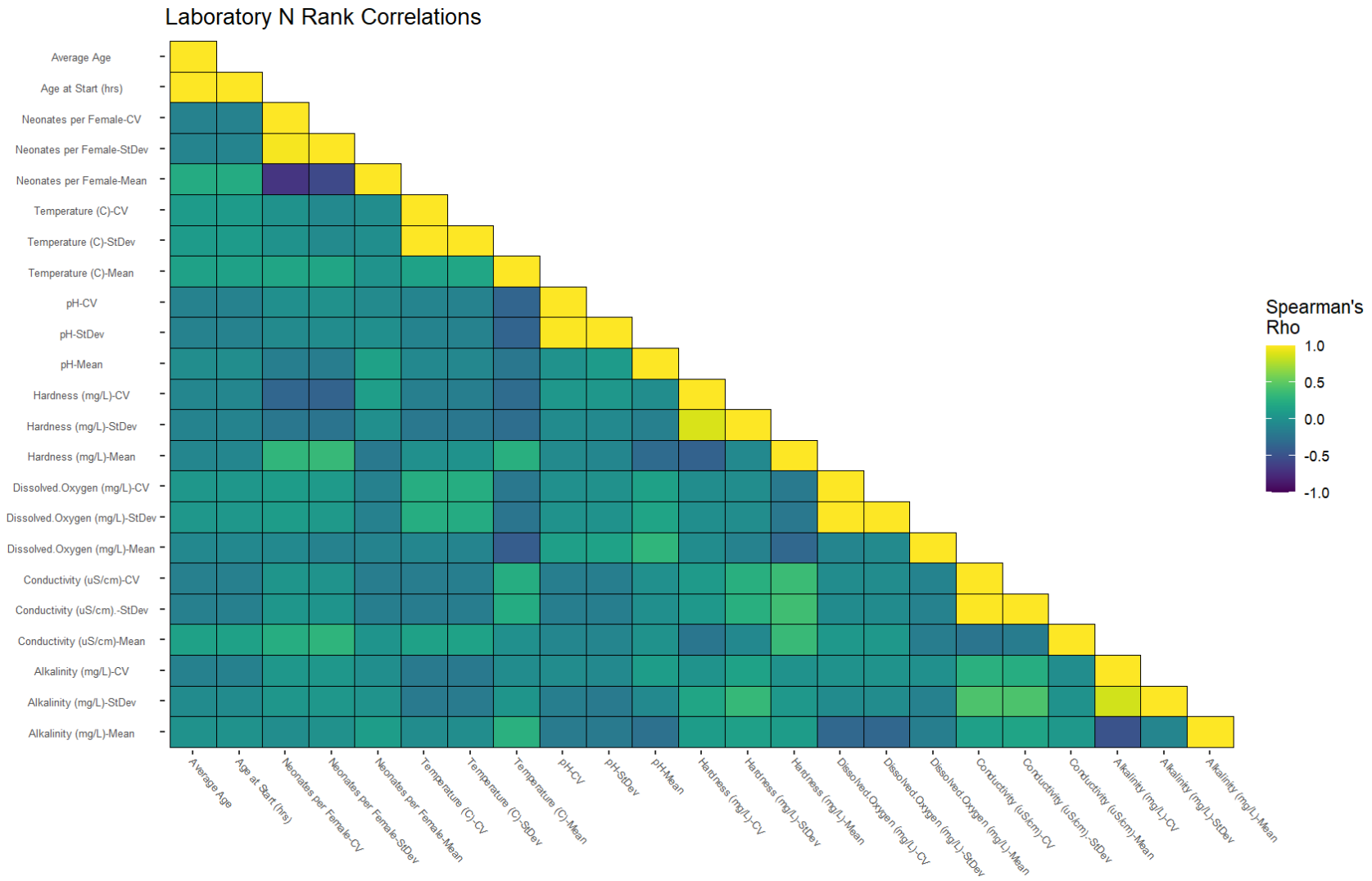

Laboratory I Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

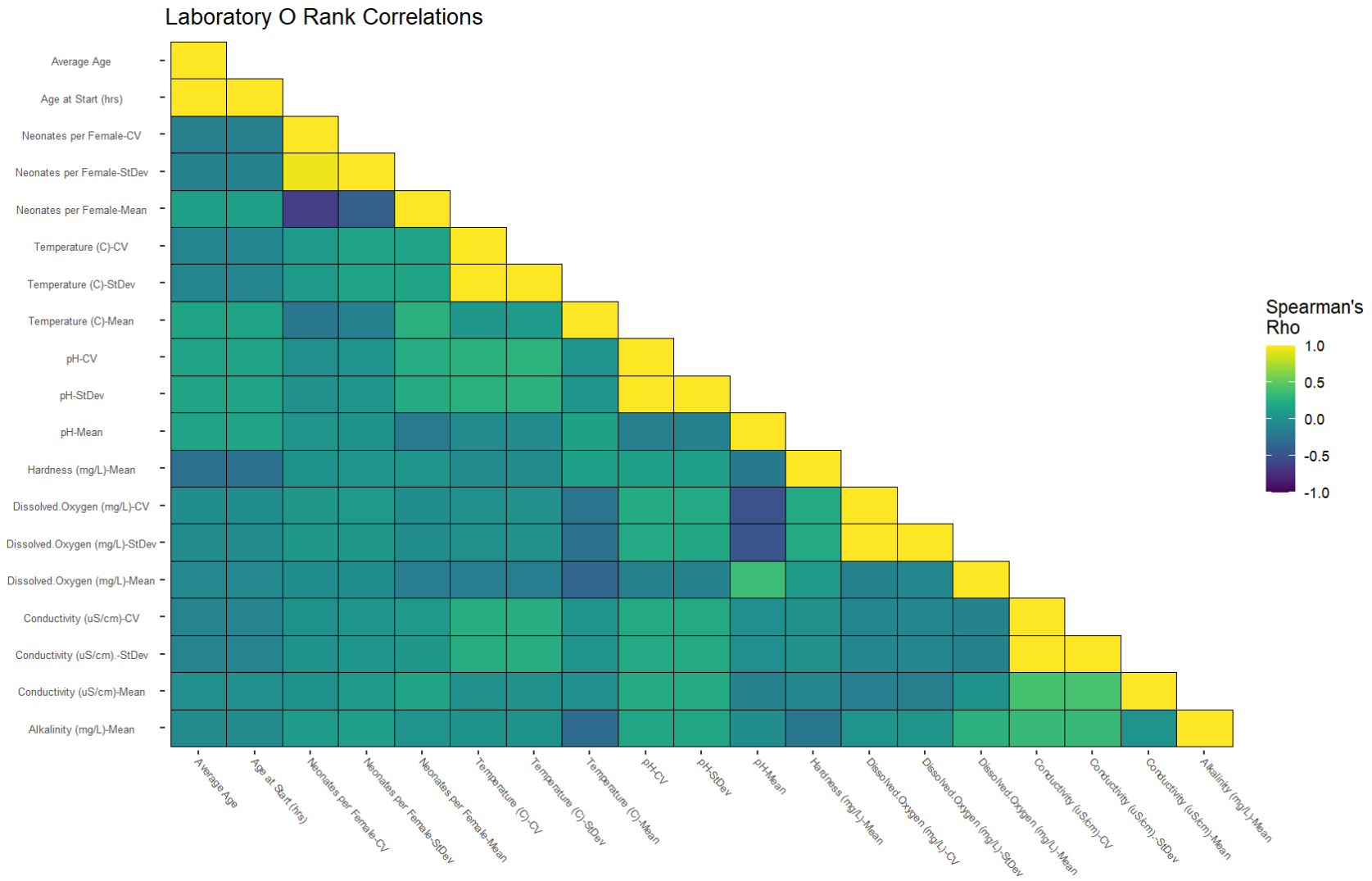

Laboratory J Rank Correlations
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

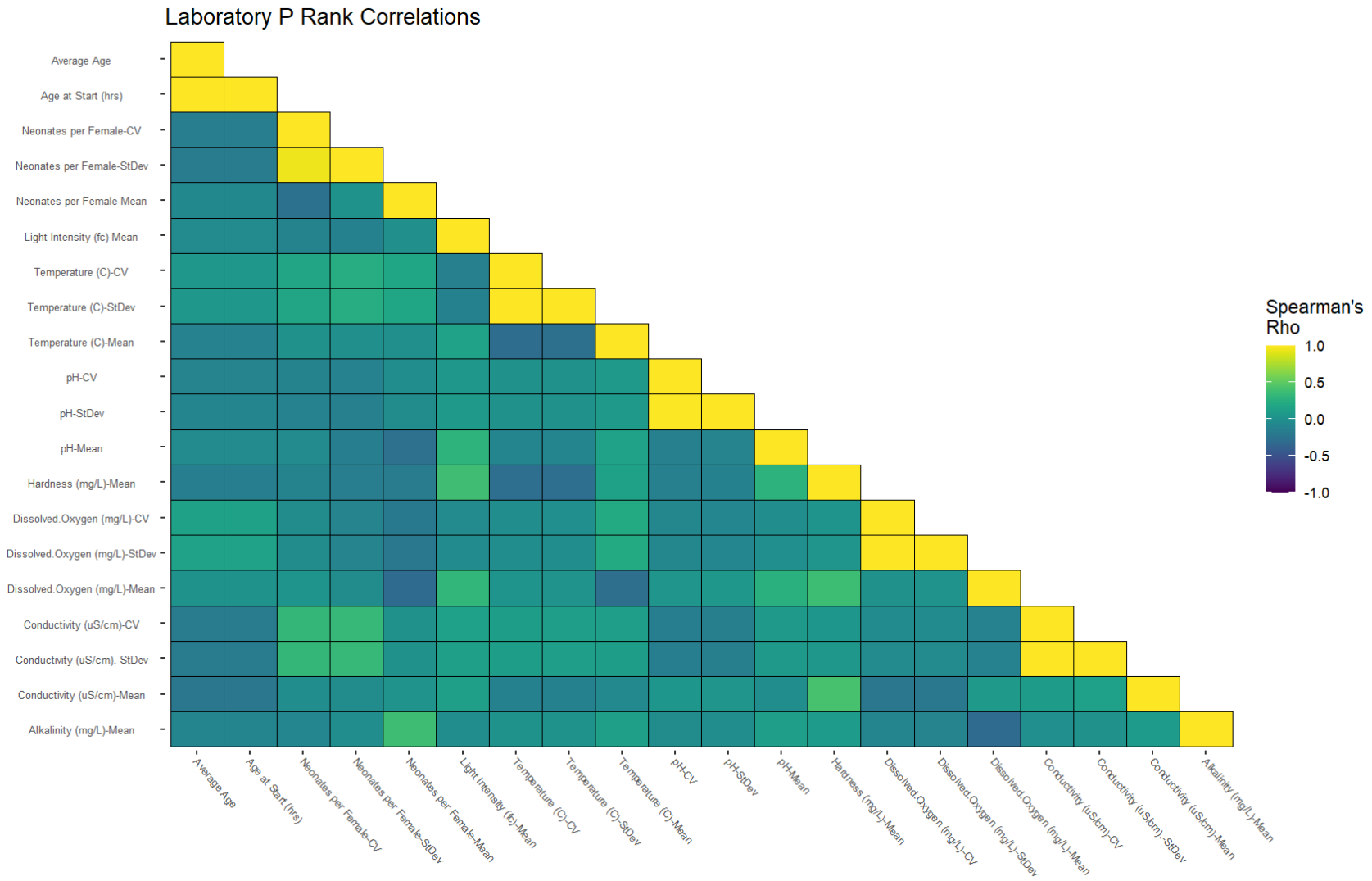

Laboratory K Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0


Laboratory L Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0


Laboratory M Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Light Intensity (fc)-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0


Laboratory N Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-CV
Hardness (mg/L)-StDev
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-CV
Alkalinity (mg/L)-StDev
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0


Laboratory O Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0


Laboratory P Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Light Intensity (fc)-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

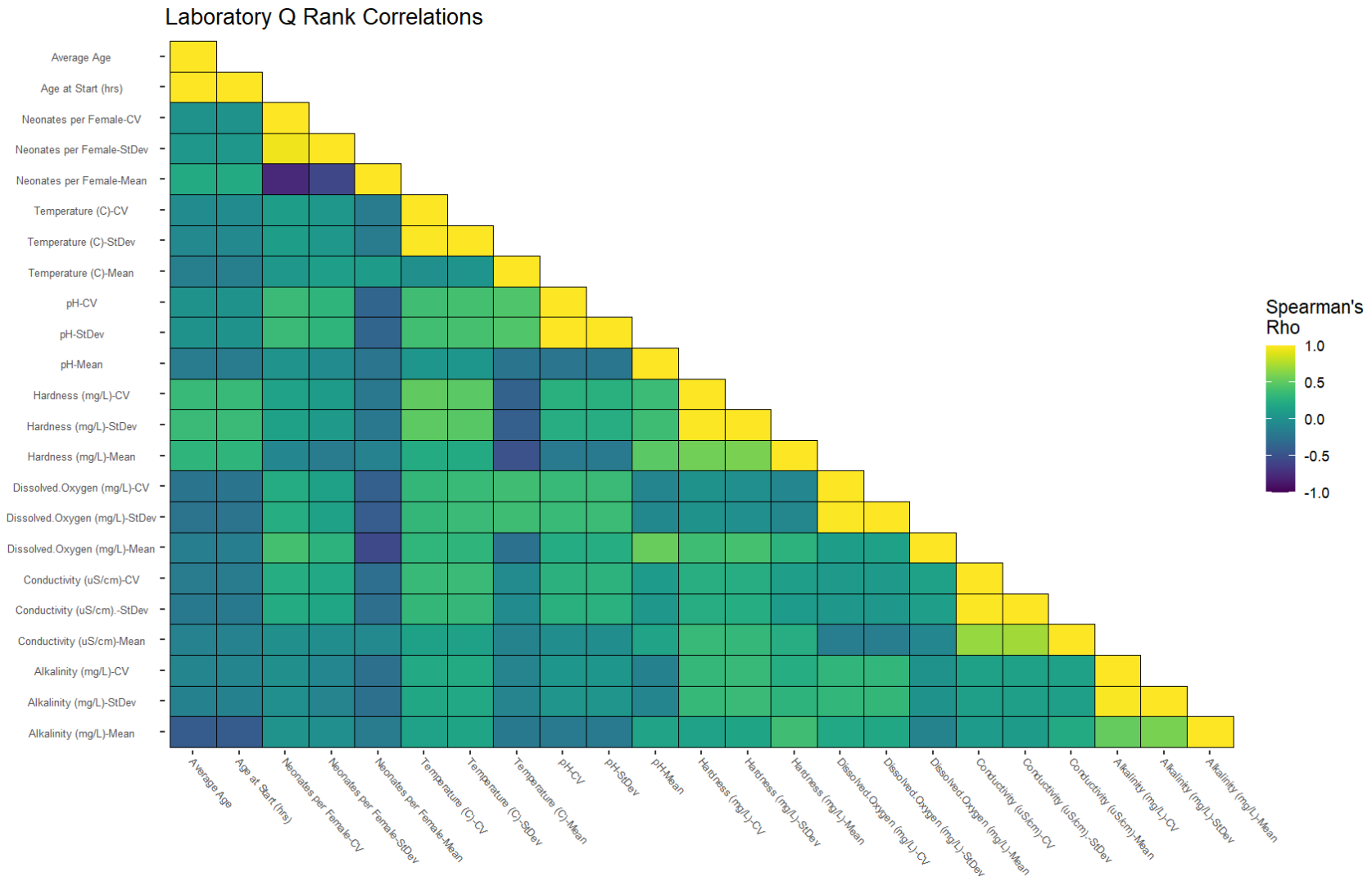

Laboratory Q Rank Correlations
Average Age
Age at Start (hrs)
Neonates per Female-CV
Neonates per Female-StDev
Neonates per Female-Mean
Temperature (C)-CV
Temperature (C)-StDev
Temperature (C)-Mean
pH-CV
pH-StDev
pH-Mean
Hardness (mg/L)-CV
Hardness (mg/L)-StDev
Hardness (mg/L)-Mean
Dissolved.Oxygen (mg/L)-CV
Dissolved.Oxygen (mg/L)-StDev
Dissolved.Oxygen (mg/L)-Mean
Conductivity (uS/cm)-CV
Conductivity (uS/cm).-StDev
Conductivity (uS/cm)-Mean
Alkalinity (mg/L)-CV
Alkalinity (mg/L)-StDev
Alkalinity (mg/L)-Mean
Spearman's Rho
1.0
0.5
0.0
-0.5
-1.0

Table A10. Variable importance values from random forest regression models of mean of control neonate production as predicted by the reported lab techniques and test conditions in the historical test data. Predictor variables are ranked by their importance within a given lab. Importance is measured by the % increase in Mean Square Error when the variable was omitted from model runs. N is the number of tests used to create the model.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
A	Mean Neonates	Alkalinity (mg/L)-Mean	95.9	78	1
A	Mean Neonates	Conductivity (uS/cm)-CV	48.0	78	2
A	Mean Neonates	Conductivity (uS/cm)-StDev	47.7	78	3
A	Mean Neonates	Conductivity (uS/cm)-Mean	36.7	78	4
A	Mean Neonates	pH-Mean	30.5	78	5
A	Mean Neonates	Hardness (mg/L)-Mean	29.0	78	6
A	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	23.9	78	7
A	Mean Neonates	Hardness (mg/L)-CV	20.3	78	8
A	Mean Neonates	Hardness (mg/L)-StDev	18.9	78	9
A	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	14.7	78	10
A	Mean Neonates	Dissolved Oxygen (mg/L)-CV	14.4	78	11
A	Mean Neonates	Year	13.9	78	12
A	Mean Neonates	pH-CV	13.3	78	13
A	Mean Neonates	Alkalinity (mg/L)-StDev	12.0	78	14
A	Mean Neonates	Alkalinity (mg/L)-CV	11.7	78	15
A	Mean Neonates	pH-StDev	9.1	78	16

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
A	Mean Neonates	Temperature (C)-CV	0.0	78	18.5
A	Mean Neonates	Test Replicates	0.0	78	18.5
A	Mean Neonates	Temperature (C)-StDev	0.0	78	18.5
A	Mean Neonates	Temperature (C)-Mean	0.0	78	18.5
B	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	51.2	95	1
B	Mean Neonates	Temperature (C)-Mean	23.6	95	2

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
B	Mean Neonates	Dissolved Oxygen (mg/L)-CV	23.4	95	3
B	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	21.2	95	4
B	Mean Neonates	Year	20.9	95	5
B	Mean Neonates	Conductivity (uS/cm)-CV	20.3	95	6
B	Mean Neonates	Conductivity (uS/cm)-StDev	17.7	95	7
B	Mean Neonates	pH-Mean	10.6	95	8
B	Mean Neonates	pH-StDev	7.7	95	9
B	Mean Neonates	pH-CV	6.2	95	10
B	Mean Neonates	Conductivity (uS/cm)-Mean	4.3	95	11

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
B	Mean Neonates	Temperature (C)-StDev	2.2	95	12
B	Mean Neonates	Temperature (C)-CV	2.0	95	13
B	Mean Neonates	Test Replicates	0.0	95	14
C	Mean Neonates	Temperature (C)-Mean	25.6	56	1
C	Mean Neonates	Year	24.9	56	2
C	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	23.8	56	3
C	Mean Neonates	Conductivity (uS/cm)-Mean	17.3	56	4
C	Mean Neonates	Hardness (mg/L)-Mean	16.3	56	5
C	Mean Neonates	pH-Mean	16.1	56	6
C	Mean Neonates	Alkalinity (mg/L)-Mean	12.0	56	7
C	Mean Neonates	pH-StDev	11.9	56	8
C	Mean Neonates	pH-CV	11.9	56	9
C	Mean Neonates	Conductivity (uS/cm)-CV	10.8	56	10
C	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	9.7	56	11
C	Mean Neonates	Dissolved Oxygen (mg/L)-CV	6.6	56	12
C	Mean Neonates	Temperature (C)-CV	1.8	56	13
C	Mean Neonates	Conductivity (uS/cm)-StDev	1.3	56	14
C	Mean Neonates	Temperature (C)-StDev	0.5	56	15
C	Mean Neonates	Test Replicates	0.0	56	16

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
D	Mean Neonates	Age at Start (hrs)	30.5	25	1
D	Mean Neonates	Average Age	30.4	25	2
D	Mean Neonates	Temperature (C)-StDev	20.3	25	3
D	Mean Neonates	Temperature (C)-CV	19.3	25	4
D	Mean Neonates	Hardness (mg/L)-Mean	13.0	25	5
D	Mean Neonates	Conductivity (uS/cm)-Mean	7.6	25	6
D	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	6.1	25	7
D	Mean Neonates	Dissolved Oxygen (mg/L)-CV	2.7	25	8
D	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	0.8	25	9
D	Mean Neonates	Alkalinity (mg/L)-Mean	0.6	25	10
D	Mean Neonates	Temperature (C)-Mean	0.2	25	11
D	Mean Neonates	pH-StDev	0.1	25	12
D	Mean Neonates	Test Replicates	0.0	25	13
D	Mean Neonates	pH-CV	0.0	25	14
D	Mean Neonates	Year	-0.8	25	15
D	Mean Neonates	pH-Mean	-3.7	25	16
D	Mean Neonates	Conductivity (uS/cm)-StDev	-4.4	25	17
D	Mean Neonates	Conductivity (uS/cm)-CV	-6.3	25	18

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
E	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	95.4	139	1
E	Mean Neonates	pH-StDev	75.9	139	2
E	Mean Neonates	pH-CV	73.3	139	3
E	Mean Neonates	Hardness (mg/L)-Mean	65.6	139	4
E	Mean Neonates	Dissolved Oxygen (mg/L)-CV	64.6	139	5
E	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	63.4	139	6
E	Mean Neonates	pH-Mean	62.2	139	7
E	Mean Neonates	Alkalinity (mg/L)-Mean	62.0	139	8

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
E	Mean Neonates	Alkalinity (mg/L)-CV	60.4	139	9
E	Mean Neonates	Conductivity (uS/cm)-CV	58.5	139	10
E	Mean Neonates	Conductivity (uS/cm)-Mean	58.3	139	11
E	Mean Neonates	Alkalinity (mg/L)-StDev	56.0	139	12
E	Mean Neonates	Conductivity (uS/cm)-StDev	54.3	139	13
E	Mean Neonates	Hardness (mg/L)-CV	46.0	139	14
E	Mean Neonates	Hardness (mg/L)-StDev	44.0	139	15
E	Mean Neonates	Year	43.4	139	16
E	Mean Neonates	Test Replicates	30.6	139	17

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
E	Mean Neonates	Temperature (C)-Mean	14.6	139	18
E	Mean Neonates	Temperature (C)-CV	14.4	139	19
E	Mean Neonates	Temperature (C)-StDev	14.3	139	20
F	Mean Neonates	Age at Start (hrs)	30.8	75	1
F	Mean Neonates	Average Age	30.2	75	2
F	Mean Neonates	pH-CV	24.0	75	3
F	Mean Neonates	pH-StDev	22.3	75	4
F	Mean Neonates	Alkalinity (mg/L)-Mean	17.9	75	5
F	Mean Neonates	Hardness (mg/L)-Mean	14.5	75	6
F	Mean Neonates	Year	10.6	75	7
F	Mean Neonates	pH-Mean	7.0	75	8
F	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	7.0	75	9
F	Mean Neonates	Conductivity (uS/cm)-Mean	2.6	75	10
F	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	2.6	75	11
F	Mean Neonates	Dissolved Oxygen (mg/L)-CV	2.4	75	12
F	Mean Neonates	Temperature (C)-StDev	0.4	75	13
F	Mean Neonates	Conductivity (uS/cm)-StDev	0.4	75	14
F	Mean Neonates	Test Replicates	-1.0	75	15
F	Mean Neonates	Temperature (C)-CV	-1.1	75	16

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
F	Mean Neonates	Conductivity (uS/cm)-CV	-6.0	75	17
F	Mean Neonates	Temperature (C)-Mean	-7.4	75	18
G	Mean Neonates	Conductivity (uS/cm)-Mean	36.9	29	1
G	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	30.6	29	2
G	Mean Neonates	Dissolved Oxygen (mg/L)-CV	26.6	29	3
G	Mean Neonates	Average Age	7.5	29	4
G	Mean Neonates	Age at Start (hrs)	5.5	29	5
G	Mean Neonates	Temperature (C)-Mean	3.5	29	6
G	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	0.1	29	7
G	Mean Neonates	Test Replicates	0.0	29	8
G	Mean Neonates	Hardness (mg/L)-Mean	-0.5	29	9
G	Mean Neonates	Year	-0.7	29	10
G	Mean Neonates	Conductivity (uS/cm)-CV	-2.5	29	11
G	Mean Neonates	pH-CV	-3.1	29	12
G	Mean Neonates	Temperature (C)-CV	-3.3	29	13
G	Mean Neonates	Temperature (C)-StDev	-3.4	29	14
G	Mean Neonates	pH-StDev	-3.5	29	15
G	Mean Neonates	Conductivity (uS/cm)-StDev	-3.6	29	16
G	Mean Neonates	pH-Mean	-9.5	29	17
H	Mean Neonates	Dissolved Oxygen (mg/L)-CV	17.0	17	1

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
H	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	16.7	17	2
H	Mean Neonates	pH-CV	13.7	17	3
H	Mean Neonates	pH-StDev	11.3	17	4
H	Mean Neonates	pH-Mean	11.2	17	5
H	Mean Neonates	Test Replicates	0.0	17	6
H	Mean Neonates	Temperature (C)-StDev	-3.0	17	7
H	Mean Neonates	Temperature (C)-CV	-3.0	17	8
H	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	-3.5	17	9

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
H	Mean Neonates	Temperature (C)-Mean	-5.9	17	10
H	Mean Neonates	Alkalinity (mg/L)-Mean	-6.0	17	11
H	Mean Neonates	Year	-7.7	17	12
I	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	32.7	60	1
I	Mean Neonates	Alkalinity (mg/L)-Mean	26.5	60	2
I	Mean Neonates	Temperature (C)-Mean	24.6	60	3
I	Mean Neonates	Hardness (mg/L)-Mean	22.1	60	4
I	Mean Neonates	Temperature (C)-StDev	16.4	60	5
I	Mean Neonates	Dissolved Oxygen (mg/L)-CV	16.3	60	6

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
I	Mean Neonates	Temperature (C)-CV	12.8	60	7
I	Mean Neonates	pH-StDev	10.4	60	8
I	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	10.3	60	9
I	Mean Neonates	pH-CV	9.1	60	10
I	Mean Neonates	Year	3.8	60	11
I	Mean Neonates	Conductivity (uS/cm)-CV	2.3	60	12
I	Mean Neonates	Age at Start (hrs)	1.8	60	13
I	Mean Neonates	Conductivity (uS/cm)-Mean	1.5	60	14
I	Mean Neonates	Conductivity (uS/cm)-StDev	1.4	60	15
I	Mean Neonates	Test Replicates	0.0	60	16
I	Mean Neonates	Average Age	-0.4	60	17
I	Mean Neonates	pH-Mean	-4.5	60	18
K	Mean Neonates	Temperature (C)-Mean	29.1	34	1
K	Mean Neonates	Temperature (C)-CV	23.2	34	2
K	Mean Neonates	Temperature (C)-StDev	23.1	34	3
K	Mean Neonates	pH-Mean	18.6	34	4
K	Mean Neonates	pH-CV	15.4	34	5
K	Mean Neonates	pH-StDev	12.8	34	6
K	Mean Neonates	Year	12.3	34	7

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
K	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	4.8	34	8
K	Mean Neonates	Age at Start (hrs)	3.8	34	9
K	Mean Neonates	Average Age	3.5	34	10
K	Mean Neonates	Dissolved Oxygen (mg/L)-CV	2.8	34	11
K	Mean Neonates	Test Replicates	0.0	34	12
K	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	-2.7	34	13
K	Mean Neonates	Conductivity (uS/cm)-Mean	-5.4	34	14
L	Mean Neonates	Average Age	31.5	57	1
L	Mean Neonates	Age at Start (hrs)	31.5	57	2
L	Mean Neonates	pH-Mean	22.1	57	3
L	Mean Neonates	Year	18.1	57	4
L	Mean Neonates	Dissolved Oxygen (mg/L)-CV	16.3	57	5
L	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	8.3	57	6
L	Mean Neonates	Conductivity (uS/cm)-Mean	7.0	57	7
L	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	1.6	57	8
L	Mean Neonates	Test Replicates	0.0	57	9
L	Mean Neonates	pH-StDev	-1.9	57	10
L	Mean Neonates	pH-CV	-3.2	57	11
L	Mean Neonates	Temperature (C)-StDev	-3.2	57	12

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
L	Mean Neonates	Temperature (C)-CV	-3.7	57	13
L	Mean Neonates	Temperature (C)-Mean	-3.8	57	14
L	Mean Neonates	Conductivity (uS/cm)-StDev	-5.7	57	15
L	Mean Neonates	Conductivity (uS/cm)-CV	-7.0	57	16
M	Mean Neonates	Year	56.3	95	1
M	Mean Neonates	Conductivity (uS/cm)-StDev	18.3	95	2
M	Mean Neonates	Conductivity (uS/cm)-CV	15.6	95	3
M	Mean Neonates	Conductivity (uS/cm)-Mean	7.2	95	4
M	Mean Neonates	Temperature (C)-StDev	6.0	95	5

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
M	Mean Neonates	pH-StDev	5.8	95	6
M	Mean Neonates	pH-CV	5.3	95	7
M	Mean Neonates	Dissolved Oxygen (mg/L)-CV	4.3	95	8
M	Mean Neonates	Temperature (C)-CV	3.1	95	9
M	Mean Neonates	Light Intensity (fc)-Mean	2.7	95	10
M	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	2.0	95	11
M	Mean Neonates	Test Replicates	0.0	95	12

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
M	Mean Neonates	Temperature (C)-Mean	-1.1	95	13
M	Mean Neonates	Age at Start (hrs)	-1.8	95	14
M	Mean Neonates	Average Age	-3.3	95	15
M	Mean Neonates	pH-Mean	-3.4	95	16
M	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	-4.5	95	17
N	Mean Neonates	Hardness (mg/L)-Mean	55.9	60	1
N	Mean Neonates	Year	24.6	60	2
N	Mean Neonates	Hardness (mg/L)-CV	16.1	60	3
N	Mean Neonates	Alkalinity (mg/L)-StDev	14.5	60	4
N	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	12.6	60	5
N	Mean Neonates	Alkalinity (mg/L)-Mean	12.1	60	6
N	Mean Neonates	Conductivity (uS/cm)-CV	8.7	60	7
N	Mean Neonates	Age at Start (hrs)	8.7	60	8
N	Mean Neonates	Conductivity (uS/cm)-StDev	7.8	60	9
N	Mean Neonates	Hardness (mg/L)-StDev	6.7	60	10
N	Mean Neonates	Conductivity (uS/cm)-Mean	6.6	60	11
N	Mean Neonates	Average Age	6.3	60	12
N	Mean Neonates	Alkalinity (mg/L)-CV	4.9	60	13
N	Mean Neonates	Dissolved Oxygen (mg/L)-CV	3.1	60	14
N	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	2.8	60	15
N	Mean Neonates	Test Replicates	0.0	60	16

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
N	Mean Neonates	pH-Mean	0.0	60	17
N	Mean Neonates	pH-CV	-0.1	60	18
N	Mean Neonates	Temperature (C)-Mean	-0.7	60	19
N	Mean Neonates	pH-StDev	-1.0	60	20
N	Mean Neonates	Temperature (C)-StDev	-15.2	60	21
N	Mean Neonates	Temperature (C)-CV	-16.9	60	22
O	Mean Neonates	Hardness (mg/L)-Mean	34.9	60	1
O	Mean Neonates	pH-Mean	30.0	60	2
O	Mean Neonates	Temperature (C)-Mean	29.3	60	3
O	Mean Neonates	Year	25.9	60	4
O	Mean Neonates	Conductivity (uS/cm)-Mean	21.6	60	5
O	Mean Neonates	Alkalinity (mg/L)-Mean	15.6	60	6
O	Mean Neonates	Average Age	12.4	60	7
O	Mean Neonates	Age at Start (hrs)	11.0	60	8
O	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	7.0	60	9
O	Mean Neonates	Temperature (C)-CV	7.0	60	10
O	Mean Neonates	pH-CV	5.0	60	11
O	Mean Neonates	Temperature (C)-StDev	4.3	60	12
O	Mean Neonates	Dissolved Oxygen (mg/L)-CV	3.4	60	13
O	Mean Neonates	pH-StDev	2.2	60	14
O	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	1.8	60	15

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
O	Mean Neonates	Conductivity (uS/cm)-StDev	1.8	60	16
O	Mean Neonates	Conductivity (uS/cm)-CV	0.4	60	17
O	Mean Neonates	Test Replicates	0.0	60	18
P	Mean Neonates	Alkalinity (mg/L)-Mean	69.2	108	1
P	Mean Neonates	pH-StDev	36.4	108	2
P	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	33.2	108	3
P	Mean Neonates	Dissolved Oxygen (mg/L)-CV	32.4	108	4

Table A10 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
P	Mean Neonates	Hardness (mg/L)-Mean	30.3	108	5
P	Mean Neonates	Year	29.6	108	6
P	Mean Neonates	pH-CV	25.2	108	7
P	Mean Neonates	pH-Mean	24.6	108	8
P	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	22.9	108	9
P	Mean Neonates	Temperature (C)-Mean	22.4	108	10
P	Mean Neonates	Light Intensity (fc)-Mean	17.0	108	11
P	Mean Neonates	Conductivity (uS/cm)-Mean	16.2	108	12
P	Mean Neonates	Temperature (C)-StDev	13.2	108	13

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
P	Mean Neonates	Temperature (C)-CV	13.1	108	14
P	Mean Neonates	Age at Start (hrs)	12.0	108	15
P	Mean Neonates	Average Age	11.6	108	16
P	Mean Neonates	Conductivity (uS/cm)-CV	10.5	108	17
P	Mean Neonates	Conductivity (uS/cm)-StDev	9.0	108	18
P	Mean Neonates	Test Replicates	2.6	108	19
Q	Mean Neonates	Dissolved Oxygen (mg/L)-Mean	45.1	48	1
Q	Mean Neonates	Year	33.1	48	2
Q	Mean Neonates	Temperature (C)-Mean	26.2	48	3
Q	Mean Neonates	Dissolved Oxygen (mg/L)-CV	15.3	48	4
Q	Mean Neonates	Dissolved Oxygen (mg/L)-StDev	13.4	48	5
Q	Mean Neonates	Conductivity (uS/cm)-CV	10.1	48	6
Q	Mean Neonates	Conductivity (uS/cm)-StDev	6.0	48	7
Q	Mean Neonates	pH-Mean	3.8	48	8
Q	Mean Neonates	pH-StDev	1.7	48	9
Q	Mean Neonates	Test Replicates	0.0	48	10
Q	Mean Neonates	pH-CV	-1.2	48	11
Q	Mean Neonates	Conductivity (uS/cm)-Mean	-2.5	48	12
Q	Mean Neonates	Temperature (C)-StDev	-7.7	48	13
Q	Mean Neonates	Temperature (C)-CV	-8.0	48	14

Table A11. Variable importance values from random forest regression models of mean of control neonate production as predicted by the reported lab techniques and test conditions in the historical test data. Predictor variables are ranked by their importance within a given lab. Importance is measured by the % increase in Mean Square Error when the variable was omitted from model runs. N is the number of tests used to create the model.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
A	CV of Neonates	Conductivity (uS/cm)-CV	9.6	78	1
A	CV of Neonates	Conductivity (uS/cm)-StDev	9.1	78	2
A	CV of Neonates	pH-StDev	6.6	78	3
A	CV of Neonates	pH-CV	6.3	78	4
A	CV of Neonates	Dissolved Oxygen (mg/L)-CV	5.9	78	5
A	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	5.8	78	6
A	CV of Neonates	Alkalinity (mg/L)-Mean	5.5	78	7
A	CV of Neonates	Hardness (mg/L)-Mean	4.5	78	8
A	CV of Neonates	Alkalinity (mg/L)-CV	3.8	78	9
A	CV of Neonates	Conductivity (uS/cm)-Mean	2.5	78	10
A	CV of Neonates	Alkalinity (mg/L)-StDev	2.1	78	11
A	CV of Neonates	pH-Mean	1.4	78	12
A	CV of Neonates	Temperature (C)-CV	0.0	78	14.5
A	CV of Neonates	Test Replicates	0.0	78	14.5
A	CV of Neonates	Temperature (C)-StDev	0.0	78	14.5
A	CV of Neonates	Temperature (C)-Mean	0.0	78	14.5
A	CV of Neonates	Year	-0.4	78	17
A	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	-3.4	78	18

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
A	CV of Neonates	Hardness (mg/L)-StDev	-5.0	78	19
A	CV of Neonates	Hardness (mg/L)-CV	-5.0	78	20
B	CV of Neonates	Temperature (C)-StDev	4.4	95	1
B	CV of Neonates	Conductivity (uS/cm)-CV	4.3	95	2
B	CV of Neonates	Conductivity (uS/cm)-StDev	3.9	95	3
B	CV of Neonates	Temperature (C)-CV	3.6	95	4
B	CV of Neonates	pH-StDev	2.6	95	5
B	CV of Neonates	pH-CV	2.6	95	6

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
B	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	2.5	95	7
B	CV of Neonates	Dissolved Oxygen (mg/L)-CV	2.0	95	8
B	CV of Neonates	Temperature (C)-Mean	1.8	95	9
B	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	1.5	95	10
B	CV of Neonates	Year	0.5	95	11
B	CV of Neonates	Test Replicates	0.0	95	12
B	CV of Neonates	pH-Mean	-0.4	95	13
B	CV of Neonates	Conductivity (uS/cm)-Mean	-0.5	95	14

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
C	CV of Neonates	pH-Mean	17.6	56	1
C	CV of Neonates	Conductivity (uS/cm)-Mean	16.0	56	2
C	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	14.1	56	3
C	CV of Neonates	pH-StDev	10.7	56	4
C	CV of Neonates	pH-CV	10.6	56	5
C	CV of Neonates	Year	8.0	56	6
C	CV of Neonates	Hardness (mg/L)-Mean	7.5	56	7
C	CV of Neonates	Temperature (C)-CV	6.6	56	8
C	CV of Neonates	Temperature (C)-StDev	6.2	56	9
C	CV of Neonates	Temperature (C)-Mean	5.9	56	10
C	CV of Neonates	Alkalinity (mg/L)-Mean	5.1	56	11
C	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	4.8	56	12
C	CV of Neonates	Dissolved Oxygen (mg/L)-CV	3.5	56	13
C	CV of Neonates	Conductivity (uS/cm)-CV	0.5	56	14
C	CV of Neonates	Test Replicates	0.0	56	15
C	CV of Neonates	Conductivity (uS/cm)-StDev	-2.0	56	16
D	CV of Neonates	Year	16.0	25	1
D	CV of Neonates	Dissolved Oxygen (mg/L)-CV	2.4	25	2
D	CV of Neonates	Test Replicates	0.0	25	3
D	CV of Neonates	Average Age	-0.3	25	4
D	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	-0.6	25	5

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
D	CV of Neonates	Age at Start (hrs)	-1.2	25	6
D	CV of Neonates	pH-CV	-2.4	25	7
D	CV of Neonates	pH-StDev	-2.4	25	8
D	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	-4.5	25	9
D	CV of Neonates	pH-Mean	-6.2	25	10
D	CV of Neonates	Hardness (mg/L)-Mean	-6.9	25	11
D	CV of Neonates	Alkalinity (mg/L)-Mean	-7.7	25	12
D	CV of Neonates	Temperature (C)-CV	-9.0	25	13
D	CV of Neonates	Temperature (C)-StDev	-9.3	25	14
D	CV of Neonates	Conductivity (uS/cm)-CV	-10.2	25	15
D	CV of Neonates	Conductivity (uS/cm)-StDev	-12.1	25	16
D	CV of Neonates	Conductivity (uS/cm)-Mean	-12.6	25	17
D	CV of Neonates	Temperature (C)-Mean	-12.8	25	18
E	CV of Neonates	Conductivity (uS/cm)-Mean	78.9	139	1
E	CV of Neonates	pH-Mean	75.4	139	2
E	CV of Neonates	pH-CV	75.3	139	3
E	CV of Neonates	Alkalinity (mg/L)-Mean	73.5	139	4
E	CV of Neonates	pH-StDev	72.9	139	5
E	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	70.7	139	6
E	CV of Neonates	Hardness (mg/L)-Mean	69.0	139	7
E	CV of Neonates	Dissolved Oxygen (mg/L)-CV	64.3	139	8

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
E	CV of Neonates	Conductivity (uS/cm)-StDev	63.7	139	9
E	CV of Neonates	Conductivity (uS/cm)-CV	62.4	139	10
E	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	61.6	139	11
E	CV of Neonates	Year	51.9	139	12
E	CV of Neonates	Hardness (mg/L)-CV	46.9	139	13
E	CV of Neonates	Hardness (mg/L)-StDev	44.5	139	14
E	CV of Neonates	Alkalinity (mg/L)-CV	28.4	139	15
E	CV of Neonates	Alkalinity (mg/L)-StDev	25.7	139	16

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
E	CV of Neonates	Test Replicates	19.4	139	17
E	CV of Neonates	Temperature (C)-StDev	8.6	139	18
E	CV of Neonates	Temperature (C)-Mean	6.7	139	19
E	CV of Neonates	Temperature (C)-CV	6.2	139	20
F	CV of Neonates	Age at Start (hrs)	51.2	75	1
F	CV of Neonates	Average Age	50.4	75	2
F	CV of Neonates	Alkalinity (mg/L)-Mean	36.7	75	3
F	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	18.4	75	4

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
F	CV of Neonates	Conductivity (uS/cm)-Mean	7.5	75	5
F	CV of Neonates	Conductivity (uS/cm)-StDev	7.2	75	6
F	CV of Neonates	Conductivity (uS/cm)-CV	7.0	75	7
F	CV of Neonates	pH-Mean	3.4	75	8
F	CV of Neonates	Hardness (mg/L)-Mean	1.9	75	9
F	CV of Neonates	Test Replicates	0.5	75	10
F	CV of Neonates	Temperature (C)-Mean	-2.0	75	11
F	CV of Neonates	pH-StDev	-3.4	75	12
F	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	-3.7	75	13
F	CV of Neonates	pH-CV	-3.8	75	14
F	CV of Neonates	Year	-4.0	75	15
F	CV of Neonates	Dissolved Oxygen (mg/L)-CV	-6.2	75	16
F	CV of Neonates	Temperature (C)-CV	-8.5	75	17
F	CV of Neonates	Temperature (C)-StDev	-13.1	75	18
G	CV of Neonates	Conductivity (uS/cm)-Mean	21.1	29	1
G	CV of Neonates	pH-Mean	9.3	29	2
G	CV of Neonates	pH-CV	7.4	29	3
G	CV of Neonates	pH-StDev	6.1	29	4
G	CV of Neonates	Test Replicates	0.0	29	5
G	CV of Neonates	Temperature (C)-CV	-2.8	29	6
G	CV of Neonates	Year	-3.0	29	7

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
G	CV of Neonates	Temperature (C)-StDev	-3.5	29	8
G	CV of Neonates	Conductivity (uS/cm)-CV	-3.7	29	9
G	CV of Neonates	Hardness (mg/L)-Mean	-4.4	29	10
G	CV of Neonates	Conductivity (uS/cm)-StDev	-5.0	29	11
G	CV of Neonates	Age at Start (hrs)	-5.5	29	12
G	CV of Neonates	Average Age	-5.7	29	13
G	CV of Neonates	Dissolved Oxygen (mg/L)-CV	-6.3	29	14
G	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	-6.3	29	15
G	CV of Neonates	Temperature (C)-Mean	-9.2	29	16
G	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	-11.8	29	17
H	CV of Neonates	pH-Mean	23.9	17	1
H	CV of Neonates	pH-CV	20.0	17	2
H	CV of Neonates	pH-StDev	19.2	17	3
H	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	14.4	17	4
H	CV of Neonates	Dissolved Oxygen (mg/L)-CV	12.2	17	5
H	CV of Neonates	Temperature (C)-StDev	7.1	17	6
H	CV of Neonates	Temperature (C)-CV	5.0	17	7
H	CV of Neonates	Year	2.2	17	8
H	CV of Neonates	Test Replicates	0.0	17	9

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
H	CV of Neonates	Temperature (C)-Mean	-2.2	17	10
H	CV of Neonates	Alkalinity (mg/L)-Mean	-7.2	17	11
H	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	-8.8	17	12
I	CV of Neonates	pH-StDev	7.8	60	1
I	CV of Neonates	Age at Start (hrs)	6.4	60	2
I	CV of Neonates	pH-CV	6.3	60	3
I	CV of Neonates	Alkalinity (mg/L)-Mean	6.2	60	4
I	CV of Neonates	Average Age	5.7	60	5
I	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	5.1	60	6
I	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	2.2	60	7

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
I	CV of Neonates	pH-Mean	2.0	60	8
I	CV of Neonates	Temperature (C)-StDev	1.3	60	9
I	CV of Neonates	Temperature (C)-CV	1.0	60	10
I	CV of Neonates	Test Replicates	0.0	60	11
I	CV of Neonates	Dissolved Oxygen (mg/L)-CV	-1.2	60	12
I	CV of Neonates	Hardness (mg/L)-Mean	-1.6	60	13
I	CV of Neonates	Conductivity (uS/cm)-CV	-2.2	60	14

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
I	CV of Neonates	Conductivity (uS/cm)-StDev	-2.5	60	15
I	CV of Neonates	Year	-5.5	60	16
I	CV of Neonates	Temperature (C)-Mean	-5.7	60	17
I	CV of Neonates	Conductivity (uS/cm)-Mean	-6.7	60	18
K	CV of Neonates	Temperature (C)-CV	18.7	34	1
K	CV of Neonates	Temperature (C)-StDev	16.7	34	2
K	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	10.0	34	3
K	CV of Neonates	Year	5.7	34	4
K	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	5.7	34	5
K	CV of Neonates	pH-Mean	4.2	34	6
K	CV of Neonates	pH-StDev	3.8	34	7
K	CV of Neonates	pH-CV	3.1	34	8
K	CV of Neonates	Temperature (C)-Mean	3.0	34	9
K	CV of Neonates	Dissolved Oxygen (mg/L)-CV	2.6	34	10
K	CV of Neonates	Test Replicates	0.0	34	11
K	CV of Neonates	Average Age	0.0	34	12
K	CV of Neonates	Age at Start (hrs)	-2.3	34	13
K	CV of Neonates	Conductivity (uS/cm)-Mean	-5.0	34	14
L	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	4.8	57	1
L	CV of Neonates	Dissolved Oxygen (mg/L)-CV	1.3	57	2
L	CV of Neonates	Average Age	1.0	57	3
L	CV of Neonates	Age at Start (hrs)	0.7	57	4

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
L	CV of Neonates	Test Replicates	0.0	57	5
L	CV of Neonates	pH-CV	-4.9	57	6
L	CV of Neonates	pH-Mean	-7.1	57	7
L	CV of Neonates	pH-StDev	-7.2	57	8
L	CV of Neonates	Temperature (C)-StDev	-8.9	57	9
L	CV of Neonates	Temperature (C)-Mean	-9.0	57	10
L	CV of Neonates	Conductivity (uS/cm)-Mean	-9.1	57	11
L	CV of Neonates	Year	-9.1	57	12
L	CV of Neonates	Temperature (C)-CV	-10.0	57	13
L	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	-11.5	57	14
L	CV of Neonates	Conductivity (uS/cm)-CV	-14.0	57	15
L	CV of Neonates	Conductivity (uS/cm)-StDev	-15.5	57	16
M	CV of Neonates	Dissolved Oxygen (mg/L)-CV	6.6	95	1
M	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	5.6	95	2
M	CV of Neonates	Temperature (C)-StDev	5.6	95	3
M	CV of Neonates	Temperature (C)-CV	5.0	95	4
M	CV of Neonates	Conductivity (uS/cm)-StDev	3.8	95	5
M	CV of Neonates	Conductivity (uS/cm)-Mean	2.9	95	6
M	CV of Neonates	pH-CV	2.9	95	7
M	CV of Neonates	pH-StDev	2.4	95	8

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
M	CV of Neonates	Conductivity (uS/cm)-CV	1.8	95	9
M	CV of Neonates	Temperature (C)-Mean	1.1	95	10
M	CV of Neonates	pH-Mean	0.6	95	11
M	CV of Neonates	Test Replicates	0.0	95	12
M	CV of Neonates	Year	-0.6	95	13
M	CV of Neonates	Age at Start (hrs)	-1.2	95	14
M	CV of Neonates	Average Age	-1.3	95	15
M	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	-2.3	95	16
M	CV of Neonates	Light Intensity (fc)-Mean	-6.4	95	17

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
N	CV of Neonates	Hardness (mg/L)-CV	39.1	60	1
N	CV of Neonates	Hardness (mg/L)-Mean	31.3	60	2
N	CV of Neonates	Year	23.5	60	3
N	CV of Neonates	Conductivity (uS/cm)-Mean	20.6	60	4
N	CV of Neonates	Hardness (mg/L)-StDev	19.9	60	5
N	CV of Neonates	pH-CV	11.6	60	6
N	CV of Neonates	pH-StDev	10.9	60	7
N	CV of Neonates	Conductivity (uS/cm)-CV	9.4	60	8
N	CV of Neonates	Age at Start (hrs)	9.4	60	9
N	CV of Neonates	Average Age	8.8	60	10
N	CV of Neonates	Alkalinity (mg/L)-Mean	8.2	60	11

N	CV of Neonates	Temperature (C)-Mean	7.3	60	12
N	CV of Neonates	Conductivity (uS/cm)-StDev	6.4	60	13
N	CV of Neonates	Alkalinity (mg/L)-StDev	5.6	60	14
N	CV of Neonates	Alkalinity (mg/L)-CV	2.8	60	15
N	CV of Neonates	pH-Mean	0.6	60	16
N	CV of Neonates	Test Replicates	0.0	60	17
N	CV of Neonates	Temperature (C)-CV	0.0	60	18
N	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	-0.4	60	19
N	CV of Neonates	Temperature (C)-StDev	-1.9	60	20
N	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	-3.2	60	21
N	CV of Neonates	Dissolved Oxygen (mg/L)-CV	-3.7	60	22
O	CV of Neonates	Conductivity (uS/cm)-CV	11.9	60	1
O	CV of Neonates	Conductivity (uS/cm)-StDev	11.0	60	2
O	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	5.8	60	3
O	CV of Neonates	Dissolved Oxygen (mg/L)-CV	5.0	60	4
O	CV of Neonates	Average Age	4.2	60	5
O	CV of Neonates	Age at Start (hrs)	4.2	60	6
O	CV of Neonates	Alkalinity (mg/L)-Mean	4.2	60	7

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
O	CV of Neonates	Temperature (C)-CV	2.9	60	8
O	CV of Neonates	Test Replicates	0.0	60	9
O	CV of Neonates	Temperature (C)-StDev	-0.2	60	10
O	CV of Neonates	Year	-2.5	60	11
O	CV of Neonates	Conductivity (uS/cm)-Mean	-2.6	60	12
O	CV of Neonates	pH-StDev	-3.1	60	13
O	CV of Neonates	Temperature (C)-Mean	-3.5	60	14
O	CV of Neonates	pH-Mean	-4.1	60	15
O	CV of Neonates	Hardness (mg/L)-Mean	-4.6	60	16
O	CV of Neonates	pH-CV	-5.2	60	17
O	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	-14.8	60	18
P	CV of Neonates	pH-StDev	43.4	108	1
P	CV of Neonates	Conductivity (uS/cm)-CV	37.0	108	2
P	CV of Neonates	pH-CV	32.8	108	3
P	CV of Neonates	Conductivity (uS/cm)-StDev	32.5	108	4
P	CV of Neonates	Year	30.8	108	5
P	CV of Neonates	Alkalinity (mg/L)-Mean	22.3	108	6
P	CV of Neonates	Hardness (mg/L)-Mean	13.8	108	7
P	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	10.0	108	8
P	CV of Neonates	Dissolved Oxygen (mg/L)-CV	9.3	108	9
P	CV of Neonates	Age at Start (hrs)	8.9	108	10

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
P	CV of Neonates	Average Age	8.7	108	11
P	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	7.9	108	12
P	CV of Neonates	Temperature (C)-StDev	7.3	108	13
P	CV of Neonates	Temperature (C)-CV	6.9	108	14
P	CV of Neonates	Light Intensity (fc)-Mean	4.8	108	15
P	CV of Neonates	Conductivity (uS/cm)-Mean	2.4	108	16
P	CV of Neonates	Temperature (C)-Mean	2.4	108	17
P	CV of Neonates	Test Replicates	-1.6	108	18

Table A11 continued. Variable importance values from random forest.

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
P	CV of Neonates	pH-Mean	-2.0	108	19
Q	CV of Neonates	Dissolved Oxygen (mg/L)-Mean	27.6	48	1
Q	CV of Neonates	Conductivity (uS/cm)-StDev	15.4	48	2
Q	CV of Neonates	Conductivity (uS/cm)-CV	14.3	48	3
Q	CV of Neonates	pH-CV	4.3	48	4
Q	CV of Neonates	pH-StDev	3.9	48	5
Q	CV of Neonates	Year	1.4	48	6
Q	CV of Neonates	Test Replicates	0.0	48	7

Lab	Response Variable	Predictor Variable	Variable Importance (%MSE Change)	n	Variable Importance Rank
Q	CV of Neonates	Temperature (C)-Mean	-2.9	48	8
Q	CV of Neonates	Conductivity (uS/cm)-Mean	-4.1	48	9
Q	CV of Neonates	pH-Mean	-6.3	48	10
Q	CV of Neonates	Temperature (C)-StDev	-9.6	48	11
Q	CV of Neonates	Dissolved Oxygen (mg/L)-CV	-10.4	48	12
Q	CV of Neonates	Temperature (C)-CV	-11.8	48	13
Q	CV of Neonates	Dissolved Oxygen (mg/L)-StDev	-16.4	48	14

Figure A20. A heat map summarizing the frequency at which different lab technique or test condition variables were selected within the top 5 most important variables in explaining the pattern in mean control neonate production in the historical test data from each lab by the random forest models detailed in Tables B29 and 30. The cooler/darker the color the more frequently that variable was assigned that particular rank. Frequencies are expressed as a percentage of the number of labs that reported a given technique/condition variable to account for differences in reporting rate among labs for different variables.

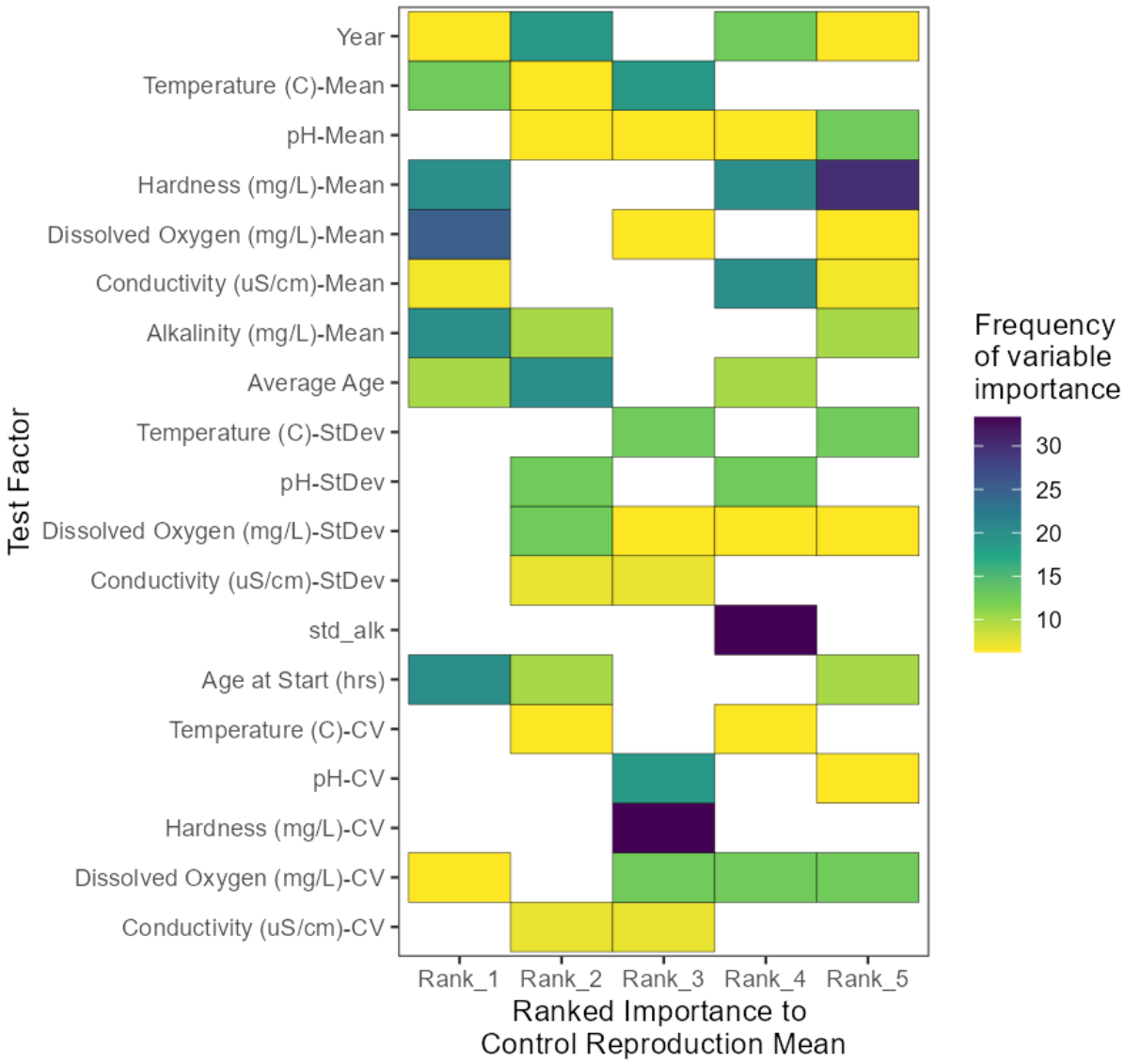
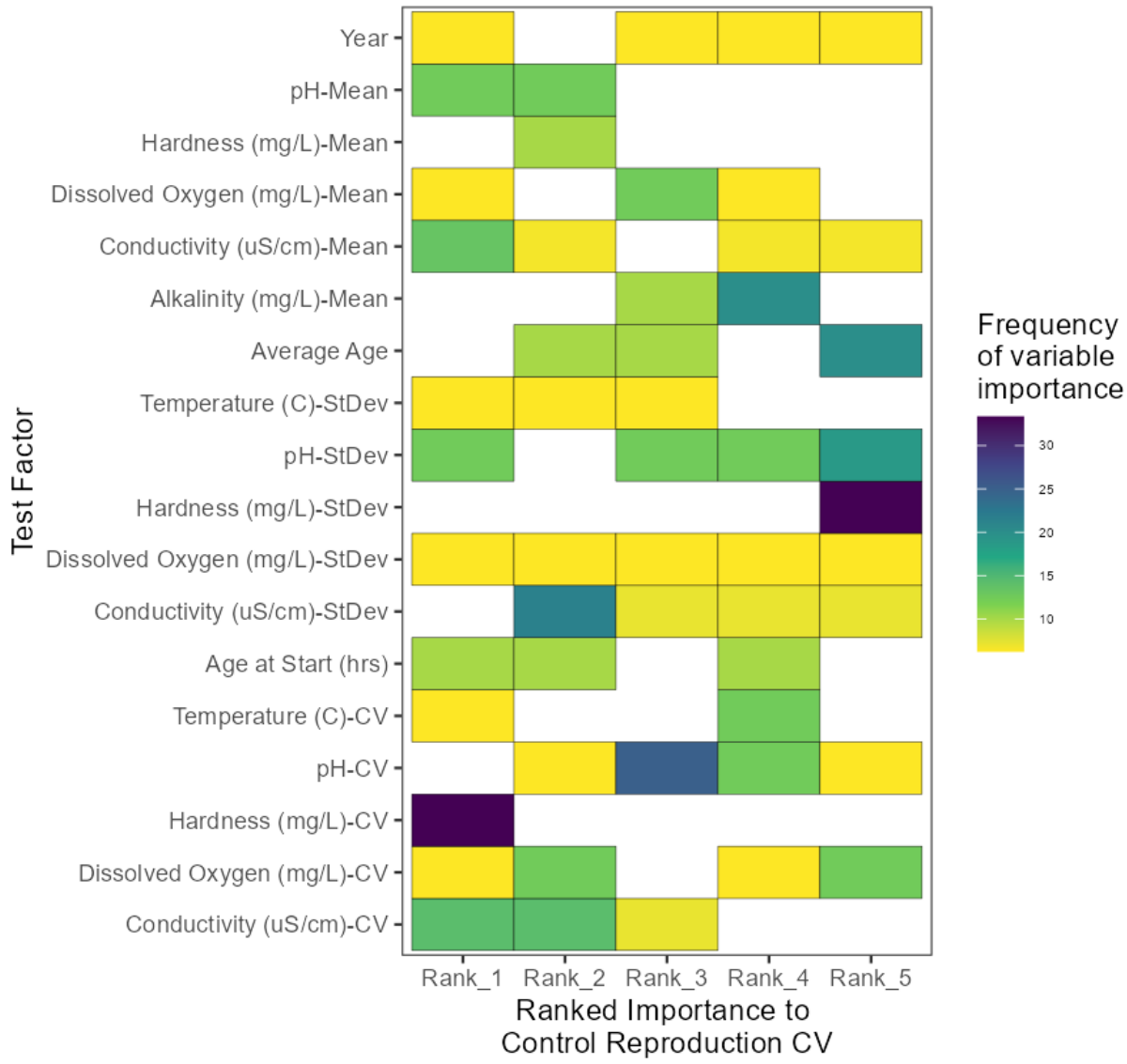


Figure A21. A heat map summarizing the frequency at which different lab technique or test condition variables were selected within the top 5 most important variables in explaining the pattern in CV of control neonate production in the historical test data from each lab by the random forest models detailed in Tables B29 and 30. The cooler/darker the color the more frequently that variable was assigned that particular rank. Frequencies are expressed as a percentage of the number of labs that reported a given technique/condition variable to account for differences in reporting rate among labs for different variables.



Appendix B – Study plan and summary data for the baseline intercalibration study.

Overview of baseline testing procedure

The specific objective of the baseline testing is to collect additional *C. dubia* chronic toxicity data and a more complete/consistent lab technique dataset across California-accredited laboratories. Twelve (12) laboratories participated in an intercomparison exercise consisting of several split samples tested in three separate testing batches. This testing design is proposed to generate a minimum of seven (7) control datasets per participating laboratory. This was statistically determined based on analyses of the width of the confidence interval to assess intra-laboratory precision. Our analyses indicated that the grand mean for control neonate production from 7 separate tests (each test performed with 10 replicates) would increase our confidence that such mean would fall within the historical control grand mean +/- 5 neonates.

Split samples to be tested include:

Sample 1: Moderately hard water recipe #1 (EPA MHW-salts) to be tested at full strength (i.e., 100%). This sample was tested along with one (1) laboratory control consisting of the lab's own dilution water recipe.

Sample 2A: Moderately hard water recipe #2 (EPA DMW); Perrier®) to be tested at full strength (i.e., 100%). This sample was tested along with one (1) laboratory control consisting of the lab's own dilution water recipe.

Sample 2B-F: 5 concentrations of sodium chloride (NaCl) diluted in MHW recipe #2 (i.e., Perrier®). All samples were prepared at SCCWRP according to the procedure described earlier in the QAPP. These samples were tested as is (i.e., no additional sample dilution allowed) along with one (1) laboratory control consisting of the lab's own dilution water recipe.

Sample 3: NaCl was provided (as a solid) to each lab with detailed instructions to prepare 5 dilutions using the lab's own dilution water. This serial dilution was tested along with one (1) laboratory control consisting of the lab's own dilution water recipe. *Note that Sample 3 is now replacing the requirement for each lab to test their routine reference toxicant with each testing batch.*

Summary of standard operating procedures

Participating laboratories (n= 12) analyzed three separate test batches within a ~ 8-week window, using their own standard operating procedures for the *C. dubia* chronic toxicity test. A summary

of standard operating procedures (SOPs), test acceptability criteria (TAC) and measurement expectations are provided in **Table 1** and in the QAPP. However, all laboratories were required to meet the following specifications:

- All tests were carried out to 8 days (i.e., 192 hours).
- All samples, including lab controls, were performed with 10 replicate chambers.
- Assignment of neonates at test set-up must use the randomized blocking by known parentage, using only brood board chambers with a minimum of 8 neonates from the adult on test initiation day. Each test (i.e., sample and associated laboratory control) was treated as independent for blocking and randomization, except for samples 2A and 2B-F and the two associated controls that must be blocked by the same known parentage.
- A 500 mL-sample of their own dilution water was collected at test initiation using the container provided by SCCWRP and shipped back to SCCWRP within 24 hours. This sample was used for analysis of ion composition.
- Test solutions were renewed daily within a 24 +/- 1 hour window to enhance the comparability of neonate counts among laboratories. Specific time of renewal (hours and minutes) were recorded and initialed.

Additionally, participating laboratories were required to report data that may not be currently documented/reported including (note that the specifics for taking these measurements are provided in the QAPP):

- Number of males, unhealthy and dead adults, and dead neonates in the brood board. This data is to be collected for all days from every chamber within any brood boards that are used to initiate the test. The expectation is that this would be about 6 to 10 days of data depending on the age of the brood board at test initiation
- Specific beginning and end time window for age of neonates at test initiation
- Water quality parameters (air and water temperature, pH, DO, conductivity) at test initiation, termination, and before and after daily renewal, to the decimal place specified in the QAPP. If possible, water temperature was also continuously monitored at the test location.
- Light intensity and twice daily air temperature within the testing area at the time of the experiments and reported in the units specified in the QAPP.

Table 1. Summary of test conditions and test acceptability criteria (TAC) for the *Ceriodaphnia dubia* survival and reproduction test.

Parameter[1]	**Description**
Test organism	*Ceriodaphnia dubia*
Protocol(s)	EPA/821/R-02-013; EPA 821-R-02012-ES [1]
Exposure	Static, daily renewal
No. replicate test chambers	10 replicates per sample/dilution
Sample holding time[2]	Up to 48 hours before test initiation
Test duration	8 days, i.e., 192 hours
Endpoints	Survival and reproduction (number of neonates per female)
Laboratory control	One laboratory dilution water control per test sample
Water quality measurements	Daily: air and water temperature in℃, pH and dissolved oxygen in mg/L reported with 0.1 precision; conductivity in µS/cm. Continuous monitoring of water temperature, if possible. 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 (TAC)	80% or greater survival and an average of 15 or more live neonates per surviving female in the controls at test termination (i.e., 8 days)

[1] Parameters and test conditions used in this study are suitable for investigative/non-compliance testing but may be different than those required for NPDES permit testing. [2] This is a deviation from the promulgated method.

[1] USEPA. 2016. Whole Effluent Toxicity Methods Errata Sheet. 28 p. Office of Water EPA 821-R-02012-ES. December 2016 https://www.epa.gov/cwa-methods; USEPA. 2002. Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, Fourth Edition. EPA-821-R-02-013.US. Environmental Protection Agency, Office of Water, Washington, DC.

Overview of split sample preparation and distribution

Bulk test water samples were prepared in the SCCWRP laboratories as described in the QAPP using large sample containers with spigots and thoroughly mixed on a large-capacity stirrer to ensure that the samples are homogeneous. The number of samples to be tested by the participating laboratories are presented in **Table 2**. Bulk samples were allowed to equilibrate for up to 48 hours before preparing the split-samples that were shipped to the laboratories. Subsampling of the bulk test samples were conducted using 3.8 L cubitainers filled to the top. All cubitainers were randomly filled in two steps. First, each cubitainer was filled halfway. Then the cubitainers were filled the rest of the way in no particular order. Each cubitainer was labeled with a unique sample ID and stored in the dark in the walk-in fridge at 4 °C less than 48 hours before shipping them to the participating laboratories.

Table 2. Number of split-samples to be tested by the 12 participating laboratories for each round. Three testing rounds were completed for this study.

Sample ID	Number of samples per lab per round	Number of sample dilutions to test	Number of lab control to include per sample
1	1	-*	1
2A	1	-*	1
2B-G	5	-*	1
3	1¥	5	1

*Water samples DO NOT require further dilution before testing.
¥ Sample 3 was shipped as a powder with instructions to prepare the serial dilution for testing.

To evaluate their preparation method and prevent unexpected toxicity, SCCWRP prepared bulk water samples and sent them to one laboratory for a *C. dubia* chronic toxicity test. If unspiked samples are not toxic and a **Concentration**-response appears normal for the dilution series, the preparation method was deemed suitable for the ILS. SCCWRP prepared fresh bulk samples and split them in individual cubitainers as described above. Since all methods and equipment were the same for subsequent rounds, this preliminary testing was only carried out for round one.

To ensure that all subsamples are representative of the original bulk test samples, two subsamples were collected in separate vessels from each cubitainer before shipment. The first set (50 mL) was used to measure conductivity, alkalinity, and hardness. The subsample was discarded after the measurements are completed. The second set (500 mL) was collected for ion composition analysis. Due to sample volume requirements, SCCWRP collected ion composition

analysis from each cubitainer before shipping. These subsamples were collected in 500 mL HDPE bottles, filled to the top, and shipped to the analytical laboratory (Physis) to measure bicarbonate, carbonate, chloride, fluoride, nitrate, sulfate, selenium, and major cations (calcium, phosphate, magnesium, sodium). The analyses were completed within 21 days of sampling to meet holding time requirements.

Split samples were shipped to each laboratory starting August 22, 2022, according to the schedule presented below. Samples were shipped on wet ice using priority overnight (OnTrac or FedEx) service to the laboratories to the addresses in Appendix A of the QAPP. The shipments included chain-of-custody (COC) forms completed by SCCWRP, and a copy of the study plan and testing instructions.

Upon delivery, temperature, conductivity, pH, dissolved oxygen, hardness, and alkalinity were measured and recorded for each sample to document their stability before testing is initiated. These measurements were made from a small subsample poured into a clean secondary vessel. Probes and any other measuring equipment cannot be used in the cubitainer, and the subsample used for water quality were discarded after use (subsample cannot be used for testing or as a chemistry or archived sample). Additionally, a 125-mL sample was collected from each cubitainer at the time of test initiation and archived. Once all chemistry and water quality samples were collected by both SCCWRP and the laboratories, there was more than 3 L remaining in each cubitainer to conduct the 8-day *C. dubia* test.

For sample 3, 14.00 g of NaCl was weighed and placed in 100 mL HDPE containers. Each laboratory received one container with instructions to prepare the serial dilution using their own lab dilution water (i.e., dilute the supplied NaCl in 7.0 L of their own dilution water and perform a 50% dilution series to generate a total of 5 dilutions). Similar to the split-water samples, once the dilutions were prepared, the laboratories recorded temperature, conductivity, pH, dissolved oxygen, hardness, and alkalinity for each dilution at test initiation. A 125-mL sample was collected and archived from each dilution at test initiation.

Note that approximately one (1) hour prior to test initiation and water changes, the volume of water needed to renew the test solutions should be adjusted/maintained at test temperature.

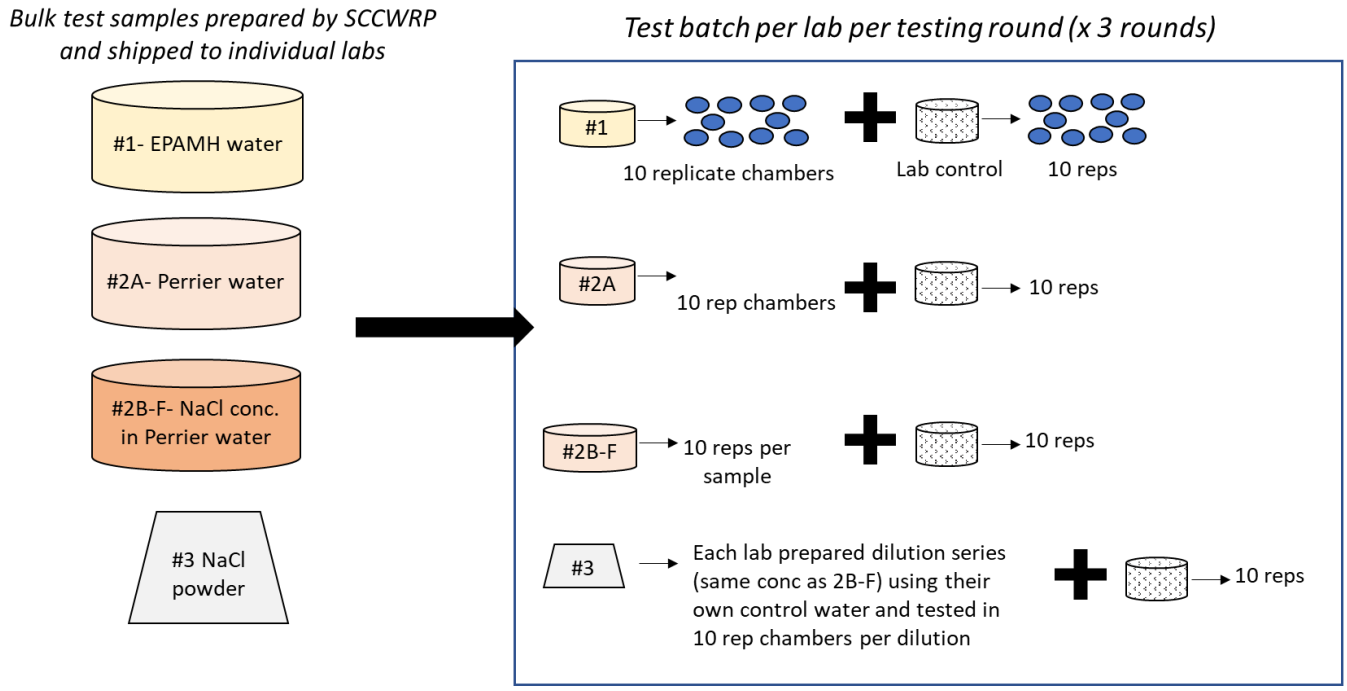
Data submission

SCCWRP provided 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 were submitted to the SCCWRP data portal. Data required include:

- Laboratory information

- Sample information upon receipt (time, temperature, condition, and more as described above)
- Testing conditions including dilution water and food recipe
- Brood board health data
- Bench water quality data for testing, survival and reproduction counts
- Control charts for reference toxicant tests for the last 12 months

Figure B1: Overview of the *C. dubia* baseline study design.



Inventory of available data

Summary of lab participation and data collected:

- Eleven labs participated in testing.
    - A 12th lab (Lab I) suffered a ciliate infection in their culture and test early in Round 1 and had to abort the test. They could not get their cultures healthy to participate in the subsequent rounds
- For Round 1, nine laboratories tested samples.
    - Labs B and M had their samples arrive too late for testing (3 days after shipping)
    - Lab N had culture problems and was unable to test Sample 3
- For Round 2, eleven laboratories tested samples.
    - Samples for the same two labs, B and M, with shipping issues in Round 1 arrived one day late but were tested within holding time.
- For Round 3, eleven laboratories tested samples, all on time.
    - Lab B had a problem with their dilution water, which led to the deaths of all organisms within a day. Therefore, all their lab controls and the Sample 3 series were unsuccessful. They did carry out the Sample 1 and 2A-F series to completion.
- Lab L tested only three sets of controls in each round instead of the requested four (they tested one control for the 2A-F series instead of having a separate control for 2A)
- Out of 132 expected laboratory controls from 11 labs, 117 were tested to completion

Biological response data

Table B1. Summary of biological data for Sample 1 collected from the eleven laboratories participating in the baseline C. dubia interlaboratory study. For each lab, the data is presented as mean of 3 rounds (except for labs B and M who could not participate in round 1). N values refer to the number of tests included in the mean and CV.

		Neonates/Adult Female				Number of Broods (8 days)				Time to First Brood			
Lab	Sample Type	Mean	N	Range	CV	Mean	N	Range	CV	Mean	N	Range	CV
A	EPA MHW (1) (salts)	38.6	3	35-44	0.13	4.1	3	4.0-4.1	0.014	4.0	3	4.0-4.0	0
B	EPA MHW (1) (salts)	35.0	2	34-36	0.024	3.0	2	3.1-3.4	0.065	4.0	2	4.0-4.1	0.017
E	EPA MHW (1) (salts)	16.0	3	13-21	0.25	4.7	3	4.3-5.0	0.075	4.2	3	4.0-4.6	0.076
F	EPA MHW (1) (salts)	20.1	3	18-21	0.072	4.5	3	4.4-4.6	0.022	3.4	3	3.1-3.5	0.069
G	EPA MHW (1) (salts)	31.6	3	27-35	0.13	4.1	3	4.0-4.1	0.014	3.3	3	3.0-3.9	0.16
L	EPA MHW (1) (salts)	24.3	3	22-26	0.092	3.2	3	2.8-3.6	0.12	4.2	3	4.1-4.8	0.032
M	EPA MHW (1) (salts)	32.4	2	27-38	0.24	3.0	2	2.3-3.8	0.35	3.6	2	3.1-4.0	0.18
N	EPA MHW (1) (salts)	7.1	3	3-11	0.59	1.8	3	0.9-2.8	0.51	4.6	3	4.3-5.0	0.075
O	EPA MHW (1) (salts)	31.0	3	24-35	0.19	4.2	3	4.0-4.5	0.069	3.0	3	3.0-3.0	0

		Neonates/Adult Female				Number of Broods (8 days)				Time to First Brood			
Lab	Sample Type	Mean	N	Range	CV	Mean	N	Range	CV	Mean	N	Range	CV
P	EPA MHW (1) (salts)	38.6	3	36-41	0.065	4.0	3	3.9-4.0	0.015	4.0	3	4.0-4.0	0
Q	EPA MHW (1) (salts)	36.2	3	30-41	0.14	3.6	3	3.4-3.7	0.043	3.9	3	3.6-4.0	0.060

Table B2. Summary of biological data for Sample 2A collected from the eleven laboratories participating in the baseline C. dubia interlaboratory study. For each lab, the data is presented as mean of 3 rounds (except for labs B and M who could not participate in round 1). N values refer to the number of tests included in the mean and CV.

		Neonates/Adult Female				Number of Broods (8 days)				Time to First Brood			
Lab	Sample Type	Mean	N	Range	CV	Mean	N	Range	CV	Mean	N	Range	CV
A	EPA MHW (DMW); Perrier® Water (2A)	39.0	3	34-43	0.12	4.0	3	3.9-4.0	0.015	4.0	3	4.0-4.0	0
B	EPA MHW (DMW); Perrier® Water (2A)	26.2	2	21-31	0.27	3.0	2	2.8-3.1	0.072	4.0	2	4.0-4.0	0
E	EPA MHW (DMW); Perrier® Water (2A)	15.4	3	11-21	0.34	4.7	3	4.4-4.8	0.049	4.2	3	4.0-4.4	0.050
F	EPA MHW (DMW); Perrier® Water (2A)	21.2	3	20-22	0.047	4.7	3	4.5-5.0	0.053	3.2	3	3.0-3.5	0.091
G	EPA MHW (DMW); Perrier® Water (2A)	33.1	3	32-35	0.043	4.0	3	4.0-4.1	0.014	3.7	3	3.0-4.1	0.16
L	EPA MHW (DMW); Perrier® Water (2A)	25.6	3	22-30	0.16	3.0	3	2.3-3.4	0.20	4.2	3	4.1-4.3	0.032
M	EPA MHW (DMW); Perrier® Water (2A)	35.2	2	35-35	0.008	3.8	2	3.5-4.0	0.094	3.6	2	3.3-4.0	0.14
N	EPA MHW (DMW); Perrier® Water (2A)	9.2	3	4-17	0.73	1.8	3	1.2-2.9	0.51	5.1	3	4.0-6.6	0.27
O	EPA MHW (DMW); Perrier® Water (2A)	29.9	3	24-36	0.21	4.0	3	3.9-4.1	0.029	3.0	3	3.0-3.0	0
P	EPA MHW (DMW); Perrier® Water (2A)	37.0	3	35-38	0.056	3.8	3	3.7-4.1	0.060	4.0	3	4.0-4.0	0

		Neonates/Adult Female				Number of Broods (8 days)				Time to First Brood			
Lab	Sample Type	Mean	N	Range	CV	Mean	N	Range	CV	Mean	N	Range	CV
Q	Perrier Water (2A)	36.6	3	35-40	0.070	3.9	3	3.9-4.0	0.015	4.0	3	4.0-4.0	0

Table B3. Summary of biological data for laboratory dilution water collected from the eleven laboratories participating in the baseline C. dubia interlaboratory study. For each lab, the data is presented as mean of 3 rounds (except for labs B and M who could not participate in round 1). N values refer to the number of tests included in the mean and CV.

		Neonates/Adult Female				Number of Broods (8 days)				Time to First Brood			
Lab	Sample Type	Mean	N	Range	CV	Mean	N	Range	CV	Mean	N	Range	CV
A	Lab Water	40.2	12	35-44	0.072	4.0	12	3.7-4.1	0.025	4.0	12	4.0-4.0	0
B	Lab Water	29.8	4	21-36	0.22	3.2	4	2.4-3.8	0.19	4.3	4	4.0-5.1	0.13
E	Lab Water	14.9	12	8-22	0.32	4.5	12	4.1-5.0	0.067	4.3	12	3.8-4.8	0.075
F	Lab Water	17.6	12	16-20	0.072	4.4	12	4.0-4.9	0.071	3.4	12	3.0-3.9	0.096
G	Lab Water	30.8	12	29-35	0.060	4.0	12	3.2-4.2	0.070	3.5	12	3.0-4.0	0.13
L	Lab Water	28.9	9	22-33	0.13	3.1	9	2.6-3.8	0.12	4.3	9	4.0-5.0	0.088
M	Lab Water	26.0	8	8-34	0.33	2.7	8	1.1-3.4	0.26	3.7	8	3.0-4.2	0.13
N	Lab Water	19.6	11	12-32	0.30	3.4	11	2.9-4.2	0.12	4.0	11	3.3-4.4	0.094
O	Lab Water	31.4	12	26-38	0.106	4.3	12	4.0-4.7	0.045	3.0	12	3.0-3.2	0.019
P	Lab Water	36.6	12	33-38	0.054	3.9	12	3.7-4.0	0.036	4.0	12	4.0-4.1	0.007
Q	Lab Water	35.5	12	29-42	0.13	3.7	12	3.1-4.0	0.094	4.0	12	4.0-4.0	0

Table B4. Additional biological data collected by the participating laboratories across all samples. N values refer to the number of tests included in the mean and CV.

	Min. Age @ Test Start (h)				Max. Age @ Test Start (h)				Number of Males per Test		Calculated Test Duration (Days to 60% of females having 3 broods)		
Lab	Mean	N	Range	CV	Mean	N	Range	CV	Mean	N	Mean	N	Range
A	6.7	12	6-8	0.098	14.6	12	14-15	0.035	0	12	7.0	12	7-7
B	0	4	0-0	0	6.8	4	6.8-6.8	0	0	4	7.2	4	7-8
E	4.3	12	4-5	0.11	8.0	12	8-8	0	0	12	6.5	12	6-7
F	10.8	12	8-16	0.34	18.5	12	15-24	0.20	0	12	6.1	12	6-7
G	12.2	12	6-20	0.53	18.7	12	14-25	0.26	0	12	6.2	12	6-7
L	4.0	9	1-8	0.53	15.5	9	5-24	0.49	0	9	7.2	9	7-8
M	12.7	8	6-18	0.45	19.5	8	14-24	0.23	0	8	6.6	8	6-8
N	6.4	11	4-13	0.67	14.4	11	12-21	0.30	0	11	6.7	11	6-8
O	14.3	12	7-22	0.36	20.8	12	15-24	0.18	0	12	6.0	12	6-6
P	1.8	12	1-3	0.42	6.3	12	5-8	0.13	0	12	6.7	12	6-7
Q	2.1	12	0.5-5	0.59	6.6	12	5-10	0.20	0	12	7.0	12	7-7

Table B5. Summary of number of neonates per surviving female in EPA Moderately Hard Water (Sample 1) from the eleven laboratories participating in the baseline C. dubia interlaboratory study. The mean was calculated by taking the total number of neonates produced for all females and dividing by the number of surviving females. For each lab, the data is presented as mean of 3 rounds (except for labs B and M who could not participate in round 1). N values refer to the number of tests included in the mean and CV.

		Neonates/ Surviving Female			
Lab	Sample Type	Mean	N	Mean Range	CV
A	EPA MHW (1)	39.9	3	35-44	0.16
B	EPA MHW (1)	39.5	2	34-44	0.12
E	EPA MHW (1)	15.1	3	11-20	0.27
F	EPA MHW (1)	19.5	3	18-21	0.16
G	EPA MHW (1)	31.6	3	27-35	0.15
L	EPA MHW (1)	29.3	3	28-31	0.18
M	EPA MHW (1)	36.0	2	30-42	0.27
N	EPA MHW (1)	13.8	3	10-17	0.62
O	EPA MHW (1)	31.0	3	24-35	0.20
P	EPA MHW (1)	38.5	3	36-41	0.15
Q	EPA MHW (1)	37.3	3	34-41	0.062

Table B6. Summary of number of neonates per surviving female in Perrier based Moderately Hard Water (Sample 2A) from the eleven laboratories participating in the baseline C. dubia interlaboratory study. The mean was calculated by taking the total number of neonates produced for all females and dividing by the number of surviving females. For each lab, the data is presented as mean of 3 rounds (except for labs B and M who could not participate in round 1). N values refer to the number of tests included in the mean and CV.

Lab	Sample Type	Mean	N	Mean Range	CV
A	DMW) Perrier® Water (2A)	39.0	3	34-43	0.13
B	DMW) Perrier® Water (2A)	27.9	2	21-35	0.45
E	DMW) Perrier® Water (2A)	14.0	3	11-17	0.43
F	DMW) Perrier® Water (2A)	21.2	3	20-22	0.089
G	DMW) Perrier® Water (2A)	33.1	3	32-35	0.11
L	DMW) Perrier® Water (2A)	30.8	3	27-34	0.22
M	DMW) Perrier® Water (2A)	38.2	2	37-39	0.28
N	DMW) Perrier® Water (2A)	17.6	3	7-25	0.61
O	DMW) Perrier® Water (2A)	29.9	3	24-36	0.18
P	DMW) Perrier® Water (2A)	39.1	3	38-41	0.082
Q	DMW) Perrier® Water (2A)	36.6	3	35-40	0.12

Table B7. Summary of number of neonates per surviving female in laboratory control water from the eleven laboratories participating in the baseline C. dubia interlaboratory study. The mean was calculated by taking the total number of neonates produced for all females and dividing by the number of surviving females. For each lab, the data is presented as mean of 3 rounds (except for labs B and M who could not participate in round 1). N values refer to the number of tests included in the mean and CV.

		Neonates/ Surviving Female			
Lab	Sample Type	Mean	N	Mean Range	CV
A	Lab Water	40.5	12	37-44	0.099
B	Lab Water	32.1	4	23-36	0.23
E	Lab Water	14.8	12	8-22	0.28
F	Lab Water	17.8	12	16-21	0.13
G	Lab Water	33.0	12	29-42	0.11
L	Lab Water	30.4	9	27-35	0.21
M	Lab Water	29.2	8	11-38	0.37
N	Lab Water	23.1	11	15-35	0.34
O	Lab Water	32.0	12	29-38	0.20
P	Lab Water	38.3	12	35-47	0.14
Q	Lab Water	36.1	12	31-42	0.11

Table B8. Number neonates per female calculated at the protocol trigger (60% of surviving females reaching 3 broods) and at an alternate trigger used by some labs (80% of surviving females reaching 3 broods), in the lab's dilutions water. N values refer to the number of tests included in the mean and CV.

		Neonates/Adult Female (60% Trigger)				Neonates/Adult Female (80% Trigger)			
Lab	Sample Type	Mean	N	Range	CV	Mean	N	Range	CV
A	Lab Water	40.2	12	35-44	0.072	40.2	12	35-44	0.072
B	Lab Water	29.8	4	21-36	0.22	29.8	4	21-36	0.22
E	Lab Water	14.9	12	8-22	0.32	15.6	12	12-22	0.30
F	Lab Water	17.6	12	16-20	0.072	17.6	12	16-20	0.072
G	Lab Water	30.8	12	29-35	0.060	30.8	12	29-34	0.060
L	Lab Water	28.9	9	22-33	0.13	29.5	9	26-33	0.077
M	Lab Water	26.0	8	8-34	0.33	26.2	8	8-34	0.32
N	Lab Water	19.6	11	12-32	0.30	20.7	11	12-32	0.31
O	Lab Water	31.4	12	26-38	0.106	31.4	12	26-38	0.106
P	Lab Water	36.6	12	33-38	0.054	36.6	12	33-38	0.052
Q	Lab Water	35.5	12	29-42	0.13	35.5	12	29-42	0.13

Table B9. Age of females in brood boards on the day their neonates were used to initiate testing during the baseline C. dubia ILS. Note that the N value is variable between laboratories, depending on how many brood boards they used to initiate testing. Some labs do single, large brood boards, while others do multiple smaller boards.

Lab	Mean Age of Female at Test Initiation (Days)	N	Range
A	6.3	3	6 - 7
B	9.0	2	9 - 9
E	10.7	3	10 - 11
F	9.7	9	9 - 11
G	7.7	6	7 - 8
L	12.4	14	12 - 14
M	8.0	3	7 - 9
N	7.6	7	7 -11
O	8.0	4	8 - 8
P	6.8	5	6 - 7
Q	9.3	3	9 - 10

Table B10. Information on the algae used for feeding *Ceriodaphnia* during ILS testing.

Lab	Algae Source	Algae Concentration (cells/mL)	Algae Concentration Measurement
A	In-house	250,000	By lab for each batch
B	ABS	695,600	By the supplier
E	ABS	233,333	By the supplier
F	ARO	233,333	By lab for each batch
G	ABS	200,000	By the supplier
L	ABS	220,000	By the supplier
M	ABS	210,000	By the supplier
N	In-house	213,000	By lab for each batch
O	In-house	245,000	By lab for each batch
P	ABS	300,000	By the supplier
Q	ABS	215,000	By lab for each batch

NA= Information not supplied by lab.

Table B11. Information on the YCT used for feeding *Ceriodaphnia* during ILS testing.

Lab	YCT Source	YCT Recipe	Feeding Method	YCT Concentration in Test Chamber (µl/ml)	Feeding rate (ml
A	ARO	Fleishman's Yeast+Blue Seal Alfalfa+Zeigler #1 Finfish Crumble Trout Chow	In test solution	0.0075	0.113
B	ABS	NA	In test solution	0.0168*	2.52
E	ABS	NA	Direct addition	0.0067	0.101
F	In-house	Fleischmann's Yeast + Pines Wheatgrass + Thomas Fish Co Trout Chow	Direct addition	0.0067	0.101
G	ABS	NA	Direct addition	0.0067	0.101
L	ABS	NA	Direct addition	0.0067	0.101
M	ARO	NA	Direct addition	0.0067	0.101
N	In-house	Trout Chow (Purina Aquamax Fry Starter 100) / Carolina Daphnia Food (4 oz) + Fleischmann's baker's yeast (one pouch 7 grams) + Cerophyl (Wards Cereal Grass Media)	Direct addition	0.005*	0.101
O	In-house	Fleishman's active dry yeast + Pines Wheatgrass + Purina Trout Chow (supplied by ABS)	In test solution	0.007	0.105
P	ARO	NA	In test solution	0.0067	0.101
Q	ABS	NA	Direct addition	0.0067	0.101

NA= Not applicable, lab purchases YCT.

- Note: Lab B was 2.5X more than the amount of YCT described by EPA; and Lab N was 0.75X less than the amount described by EPA.

Table B12. Individual test batches not meeting test acceptability criterion for reproduction (≥ 15 neonates/surviving female).

Lab	Test Round	Mean Neonates/Surviving Female
E	Round 2	12.3
E	Round 2	12.7
E	Round 3	8.9
E	Round 3	10.4
E	Round 3	9.3
E	Round 3	8.5
M	Round 2	10.9

Note: All four Round 3 tests for Lab B did not meet test acceptability due to zero survival.

Table B13. Individual test batches not meeting acceptable brood board mortality (< 20%). Note that all labs had at least 8 neonates per brood board female used to initiate tests and no females were older than 14 days.

Lab	Test Round	Brood Board Percent Mortality	Mean Neonates per Female in Lab Control
N	Round 1	42	15.5
N	Round 1	42	12.7
N	Round 1	42	11.9
L	Round 1	23	28.8
L	Round 2	36*	32.8
L	Round 2	36*	27.7
L	Round 2	36*	31.9

* Average mortality of four broods used to initiate the three test batches.

Table B14. Brood board health parameters recorded in boards used to initiate the tests during the baseline C. dubia interlaboratory study. Data is expressed as percentage of brood board cups exhibiting a health issue category per brood board. N values refer to the number of brood boards used by each lab to initiate all of their tests.

	Unhealthy Adult			Dead Adult			Male			Unhealthy Neonates			Dead Neonates		
Lab	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
A	0	5	NA	0	5	NA	0	5	NA	0	5	0-0	0	5	NA
B	0	2	NA	11.2	2	7.5-15.0	0	2	NA	0	2	0-0	0	2	NA
E	0	3	NA	0	3	NA	0	3	NA	0	3	0-0	0	3	NA
F	1.1	9	0-10	1.5	9	0-6.7	0	9	NA	0.2	9	0-1.7	0.2	9	0-1.7
G	3.3	6	0-20	0.6	6	0-3.3	1.1	6	0-6.7	0	6	0-0	0	6	NA
L	3.3	14	0-10	17.6	14	0-46.7	1.9	14	0-10.0	0	14	0-0	0	14	NA
M	0	3	NA	5.6	3	0-11.7	0	3	NA	9.5	3	1.7-21.7	20.0	3	11.7-33.3
N	0	7	NA	11.0	7	0-41.7	0	7	NA	0	7	0-0	0	7	NA
O	0	4	NA	1.2	4	0-3.3	0	4	NA	0	4	0-0	0	4	NA
P	0.2	5	0-3.3	2.9	5	0-13.3	0	5	NA	8.0	5	3.3-13.3	17.3	5	6.7-23.3
Q	0	3	0-0	2.8	3	0-5.0	0	3	NA	0.6	3	0-1.7	1.1	3	0-1.7

NA= not applicable

Figure B2. Box plot of number of neonates per adult female for the lab's own dilution water in the ILS. The plot shows the results of each control sample within a round by laboratory (N= 3 or 4). The black line within each box represents the median and the green is the mean. The lines extending horizontally for each lab is the mean neonates per adult female from the historical data previously submitted by the laboratories. Lab N reported culture issues prior to Round 1 and Lab L prior to Round 2.

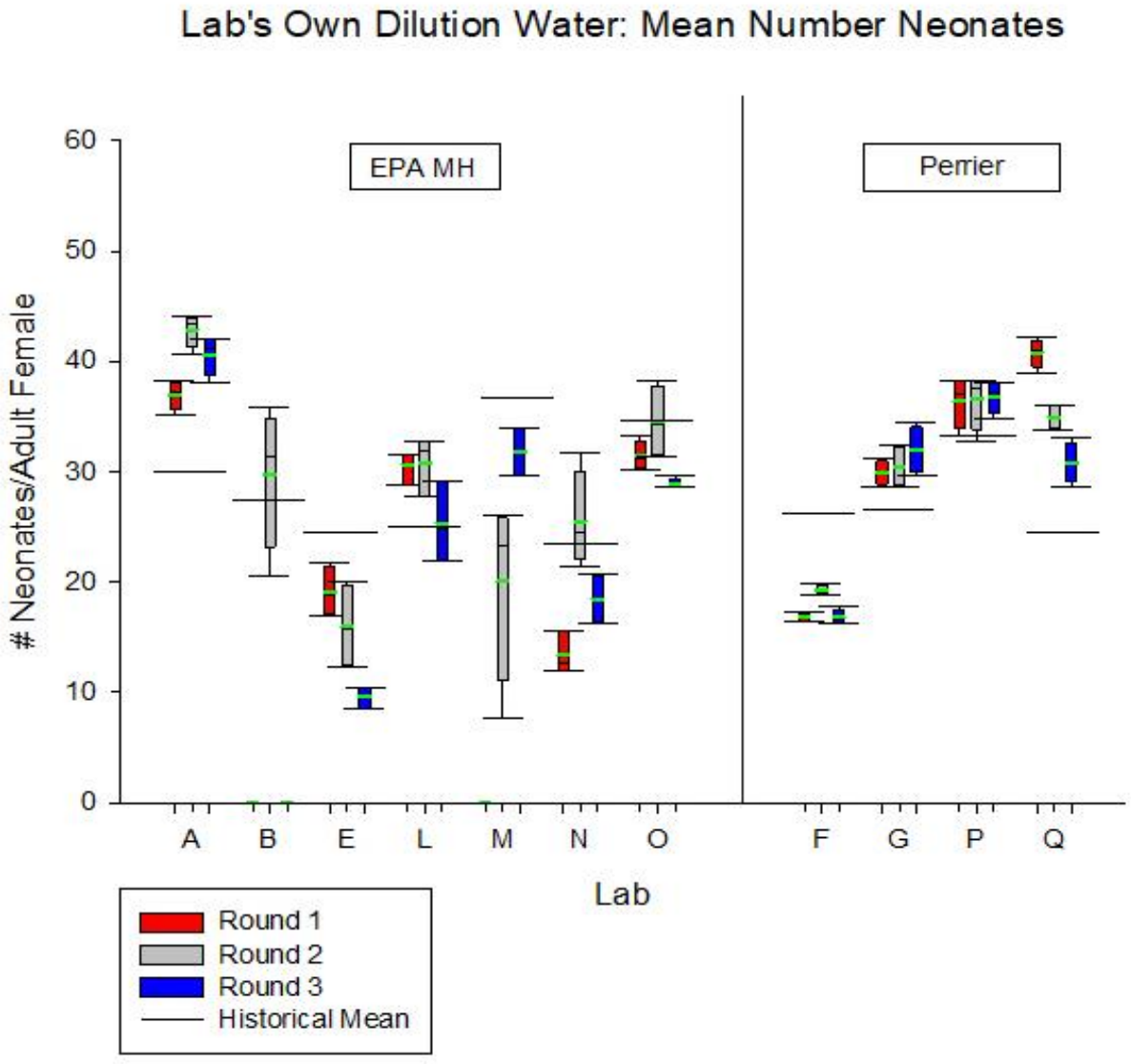


Figure B3. Box plot of coefficient of variation for the mean number of neonates per adult female in the lab's own dilution water in the ILS. The plot shows the results of each control sample within a round by laboratory (N= 3 or 4). The black line within each box represents the median and the green is the mean. The lines extending horizontally for each lab is the mean neonates per adult female from the historical data previously submitted by the laboratories. Lab N reported culture issues prior to Round 1 and Lab L prior to Round 2.

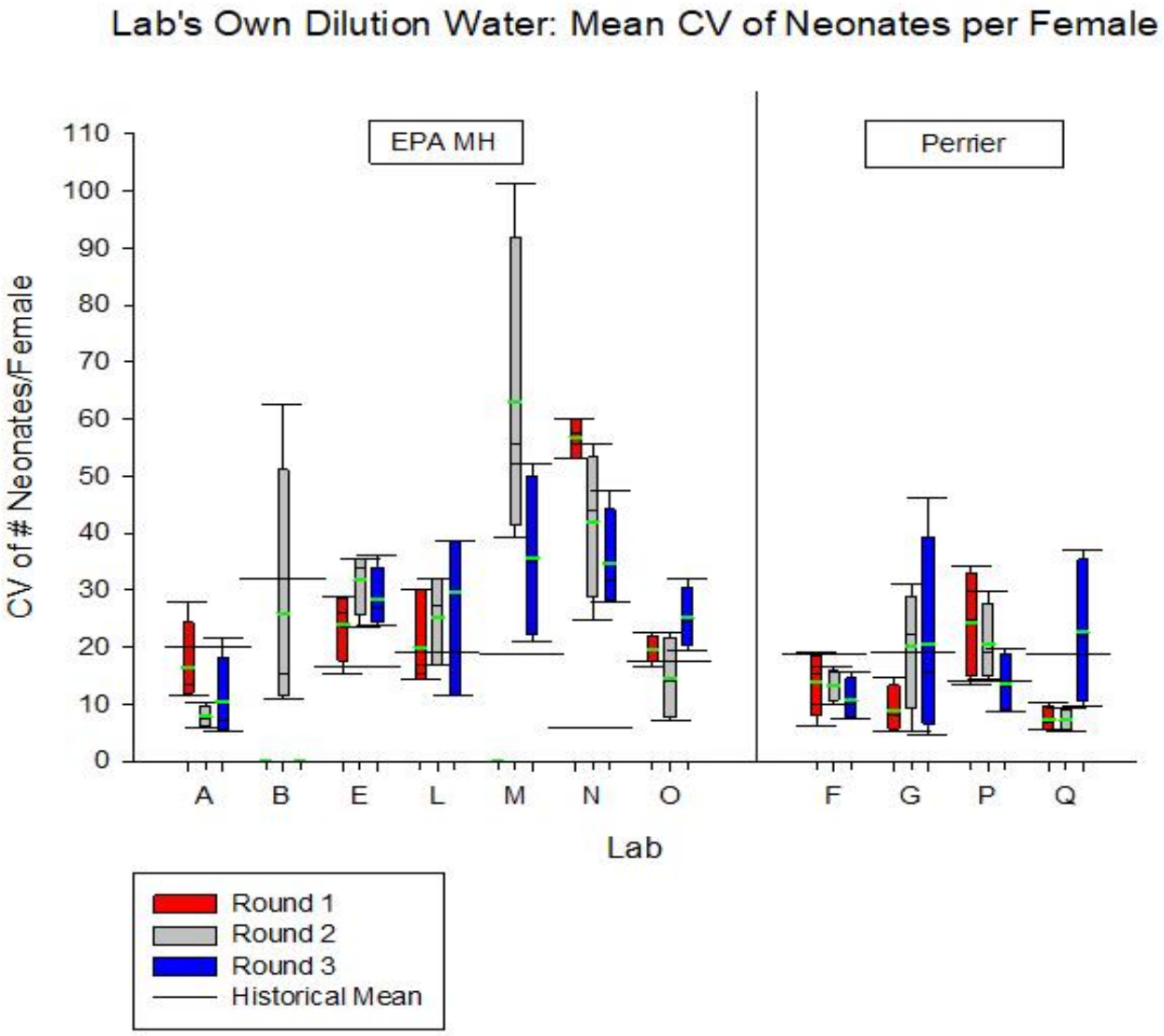


Figure B4. Box plot of number of neonates per adult female for Sample 1 (EPA MHW) in the ILS. Since there is only one sample per round, the plot is of individual replicates (N=10). The black line within each box represents the median and the green is the mean. The dots are the high and low outlier, if any. The lines extending horizontally for each lab is the mean neonates per adult female from the historical data previously submitted by the laboratories. Laboratories are ordered based on the water type typically used in their cultures. Lab N reported culture issues prior to Round 1 and Lab L prior to Round 2.

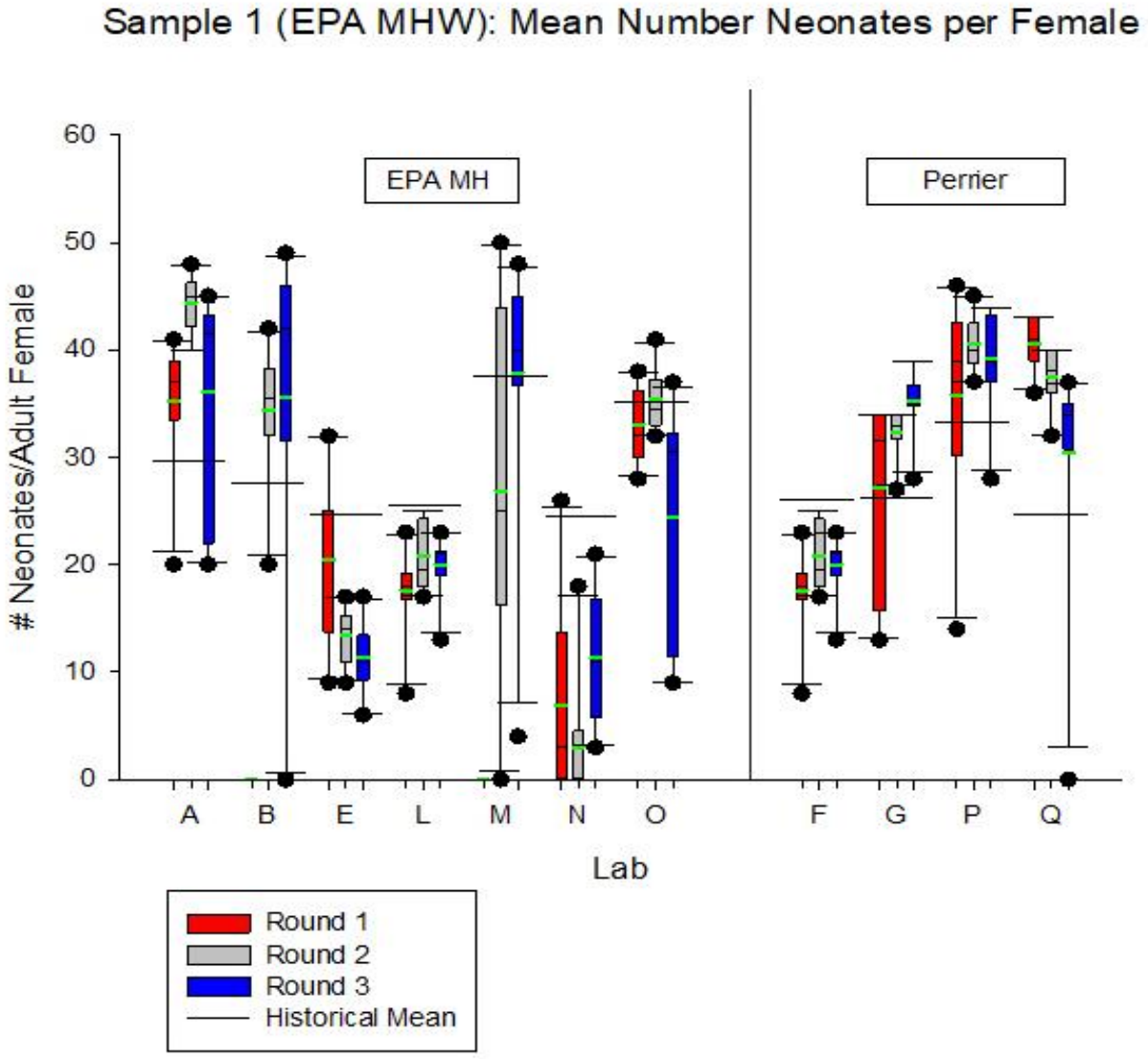


Figure B5. Box plot of number of neonates per adult female for Sample 2A (DMW Perrier®) in the ILS. Since there is only one sample per round, the plot is of individual replicates (N=10). The black line within each box represents the median and the green is the mean. The dots are the high and low outlier, if any. The lines extending horizontally for each lab is the mean neonates per adult female from the historical data previously submitted by the laboratories. Laboratories are ordered based on the water type typically used in their cultures. Lab N reported culture issues prior to Round 1 and Lab L prior to Round 2.

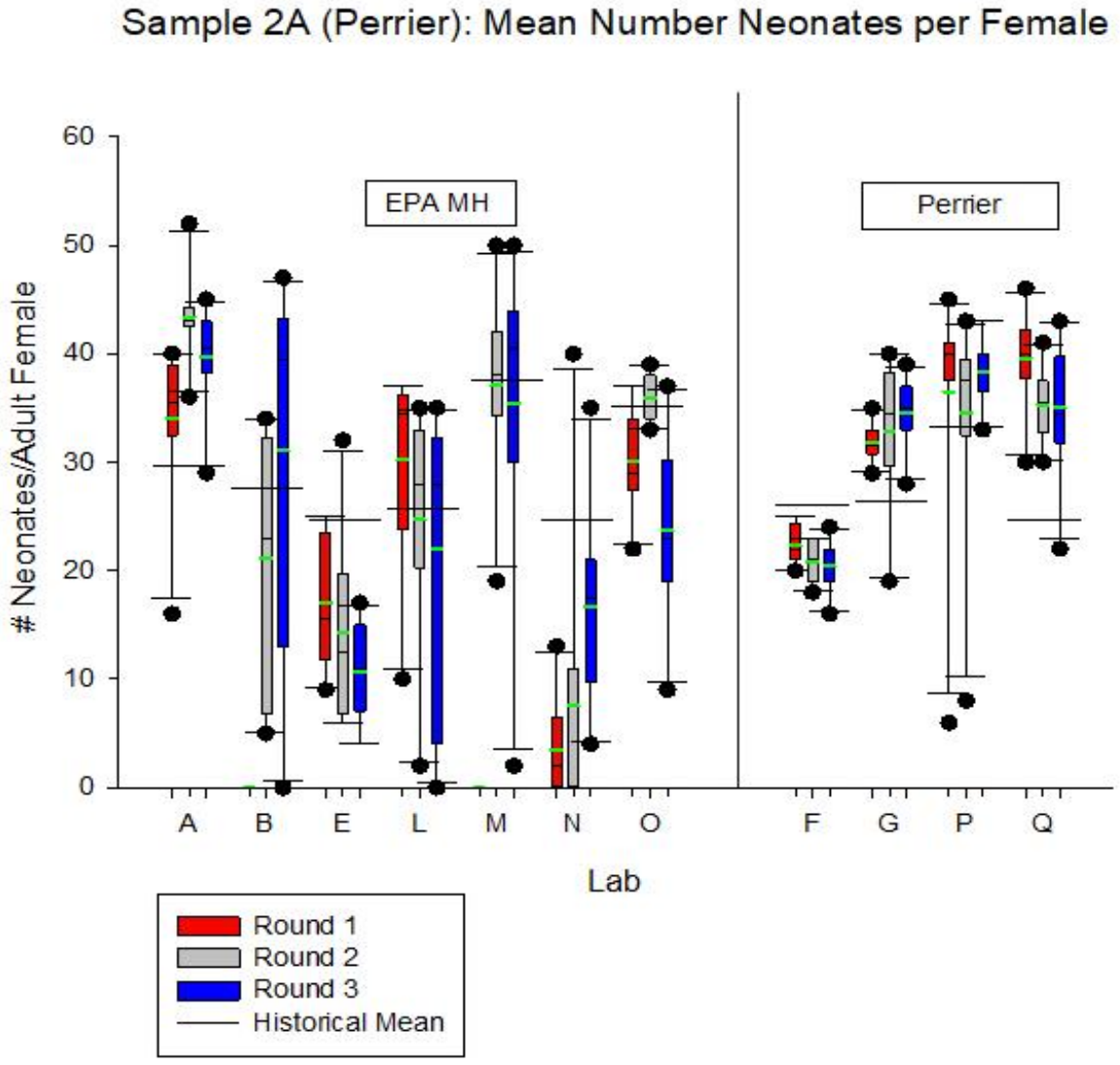


Figure B6. Plot of the age of the females whose neonates were used to initiate a test batch versus the number of neonates produced in dilution water controls.

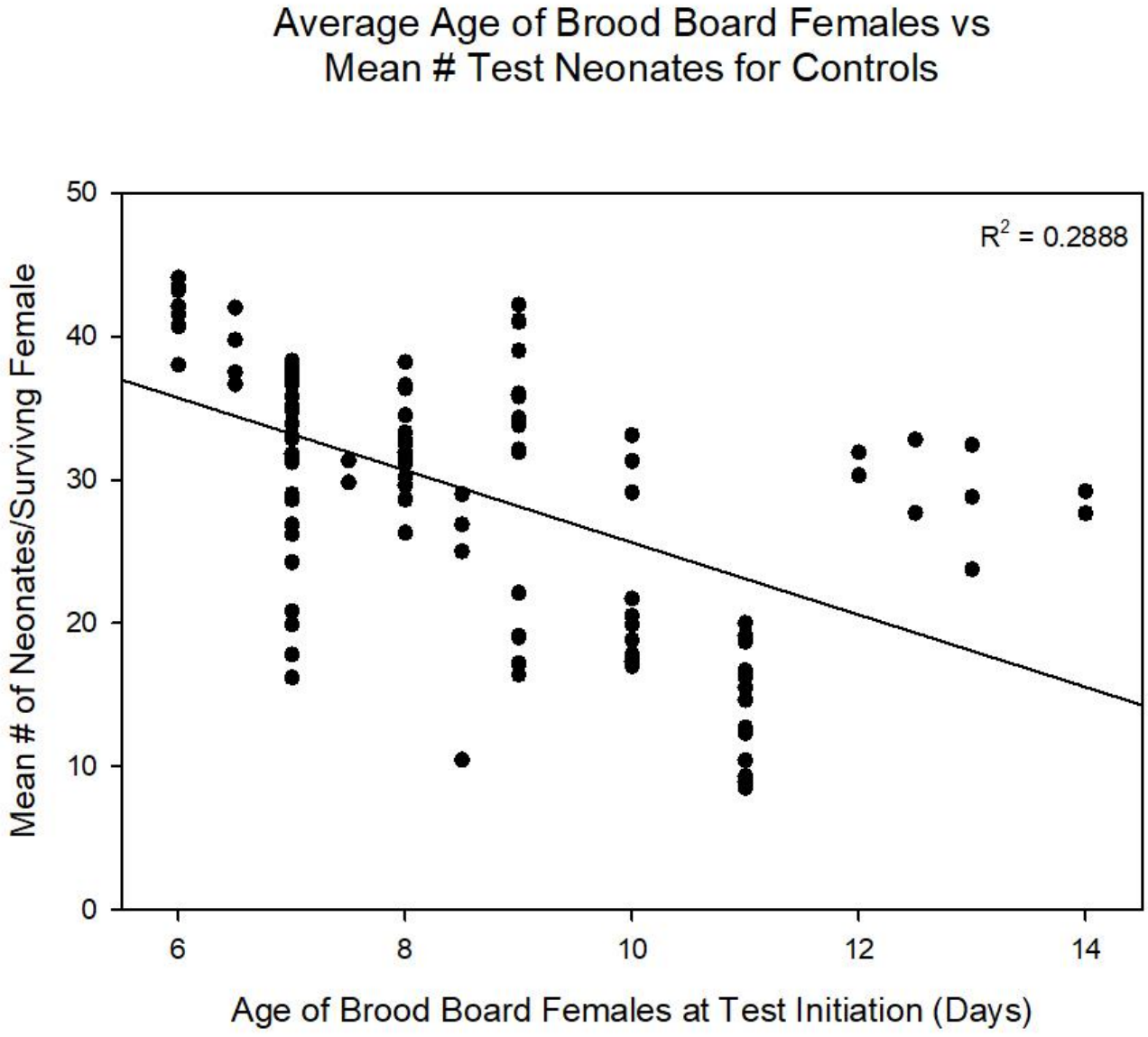


Figure B7. Plot of the age of the females whose neonates were used to initiate a test batch versus the number of neonates produced in Sample 1 (EPA MHW).

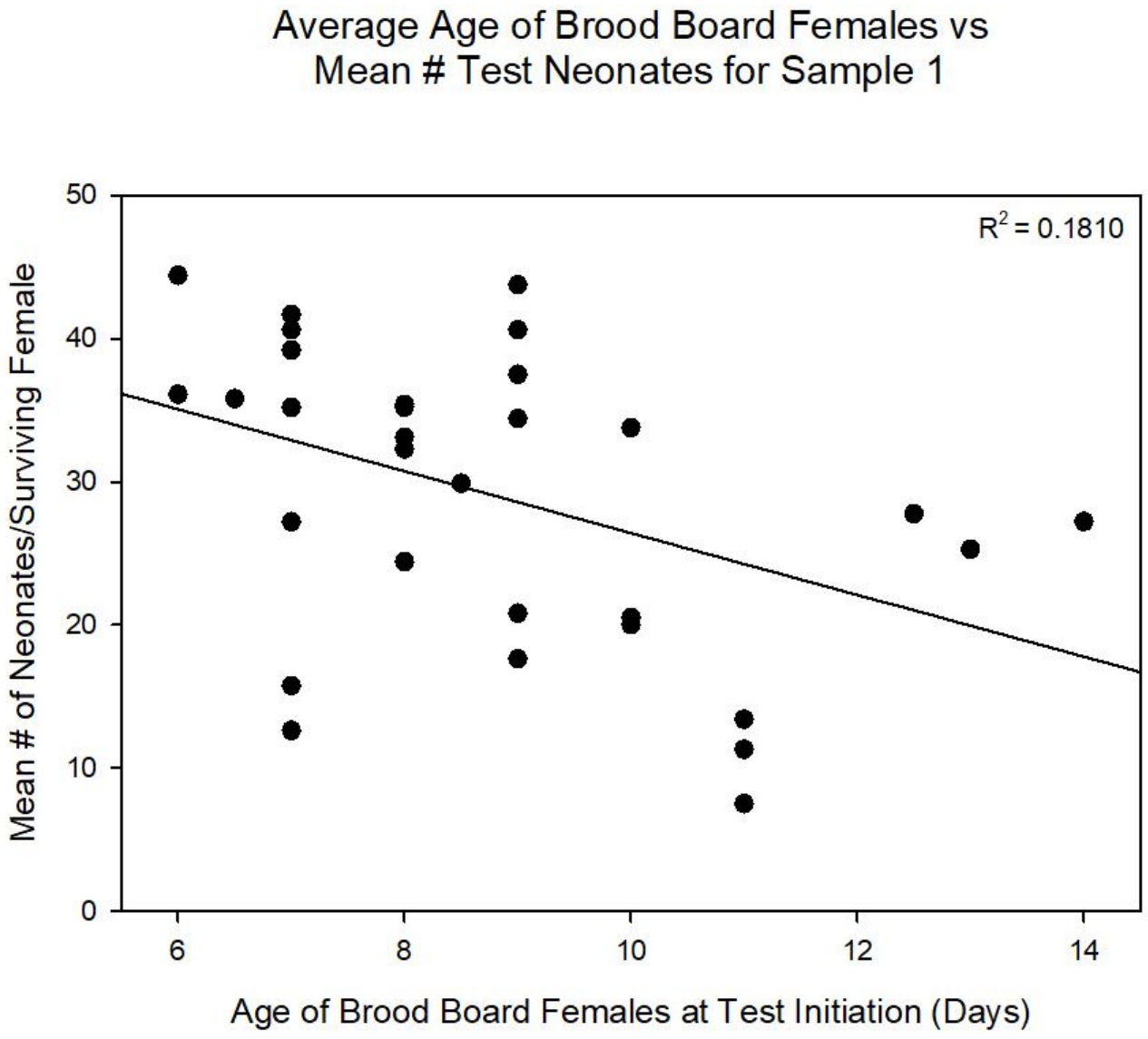


Figure B8. Plot of the age of the females whose neonates were used to initiate a test batch versus the number of neonates produced in Sample 2A (DMW with Perrier®).

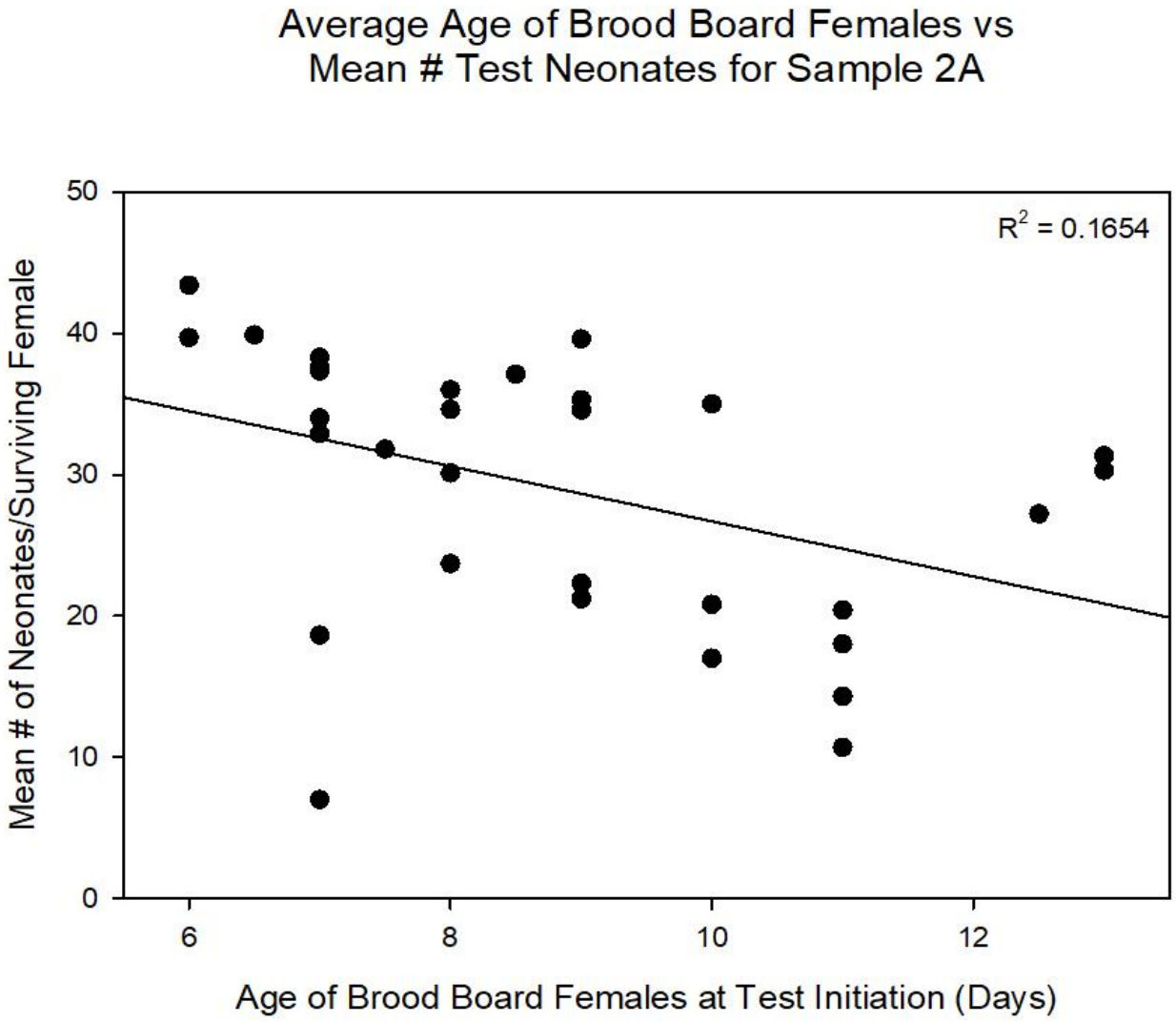


Figure B9. Plot of the age of the females whose neonates were used to initiate a test batch versus the IC25 for the Sample 2 series (DMW with Perrier®).

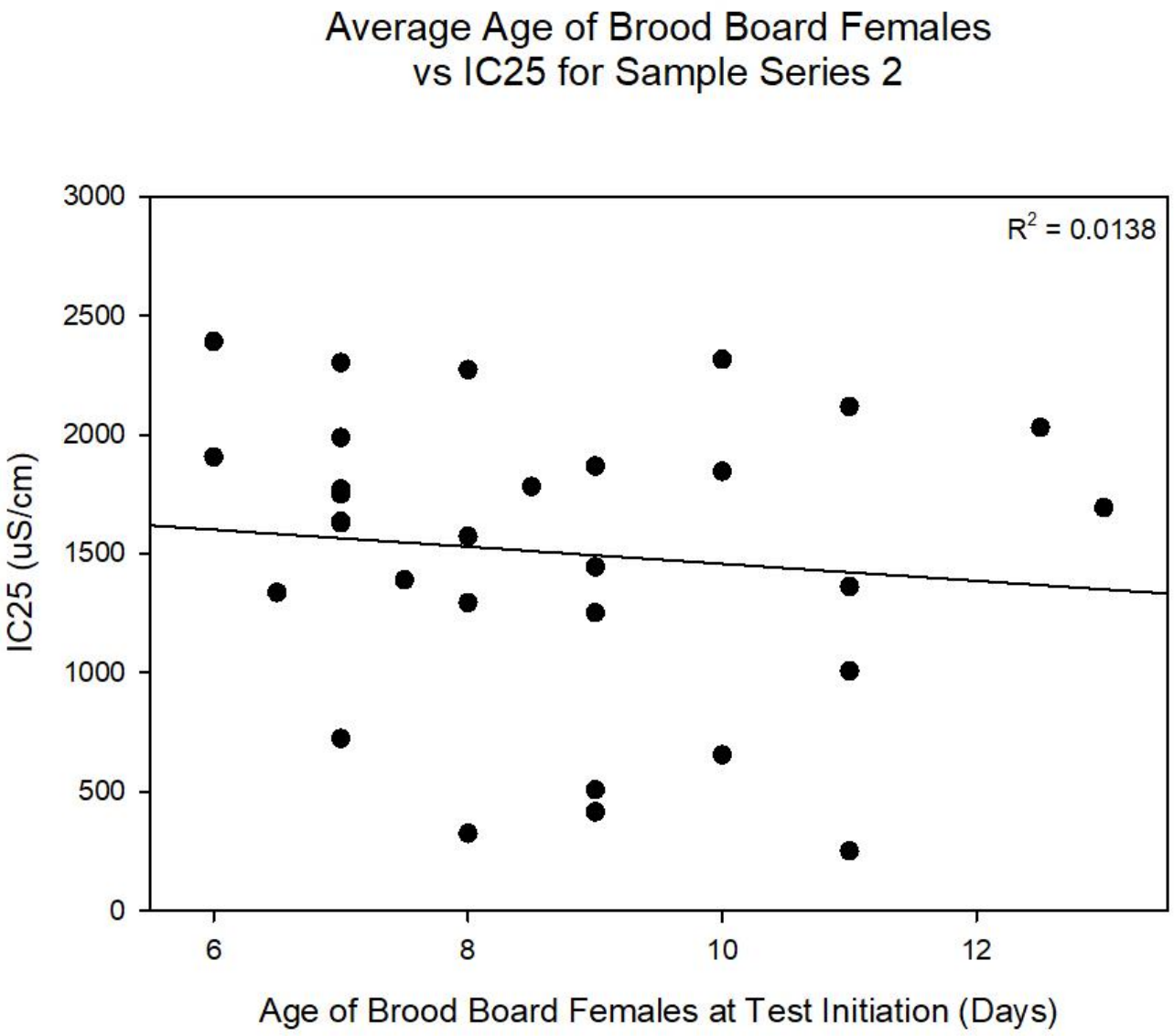
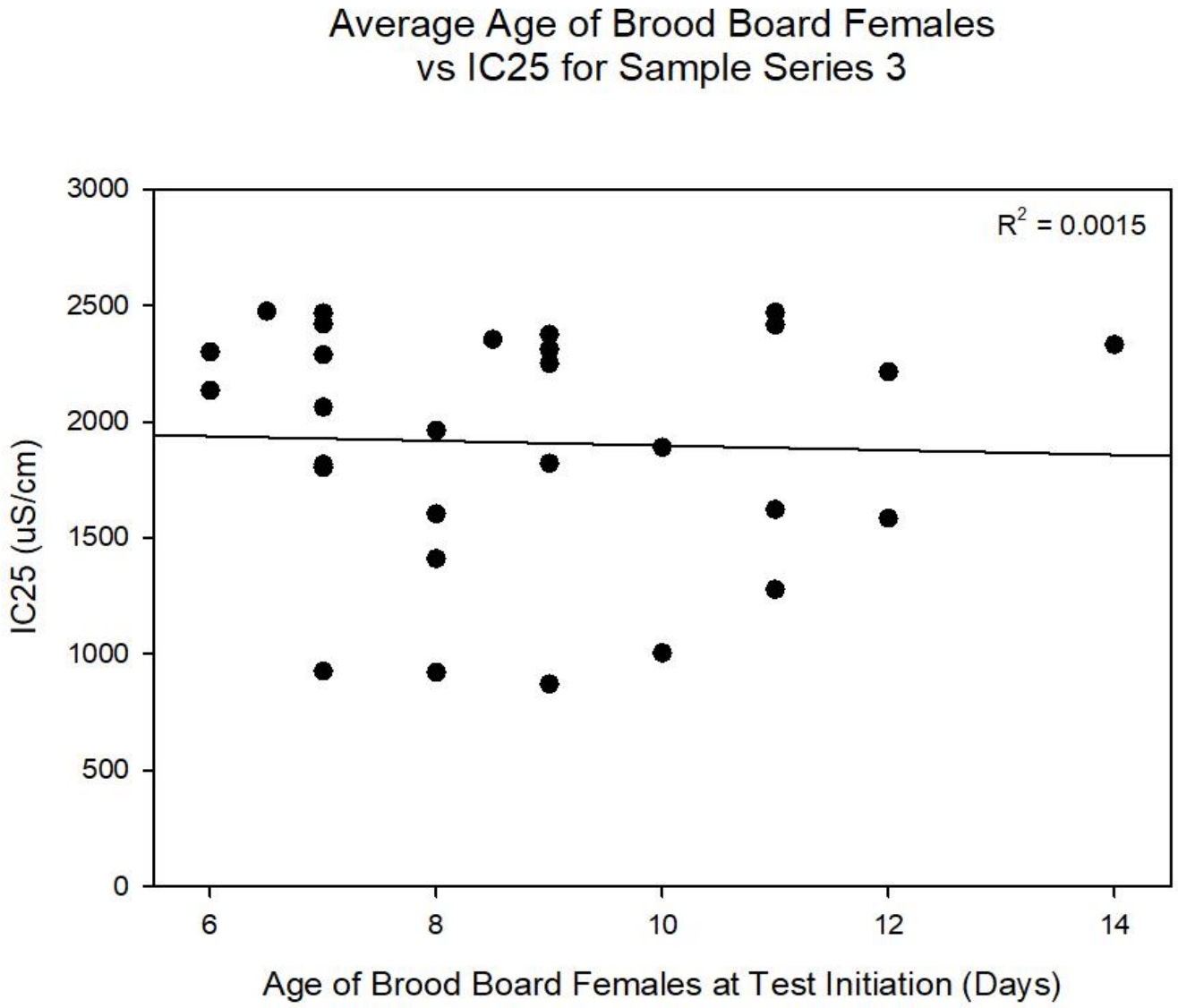


Figure B10. Plot of the age of the females whose neonates were used to initiate a test batch versus the number the IC25 for the Sample 3 series (lab's own dilution water).



Concentration-response data

This section summarizes the concentration-response data from samples consisting of dilution water spiked with different concentrations of sodium chloride. SCCWRP prepared Samples 2A-F in DMW (Perrier®) water, and split samples were shipped to the labs participating in the ILS. Sample 3 was shipped as a solid to the labs, and each lab prepared concentrations using their own lab water. Data are plotted as mean neonates per sample type per round against the measured conductivity and against the nominal sodium chloride concentration. Point estimates (IC and LC) were calculated in Python according to the EPA manual

Figure B11. Concentration-response plot of all three rounds of testing for Lab A series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend.

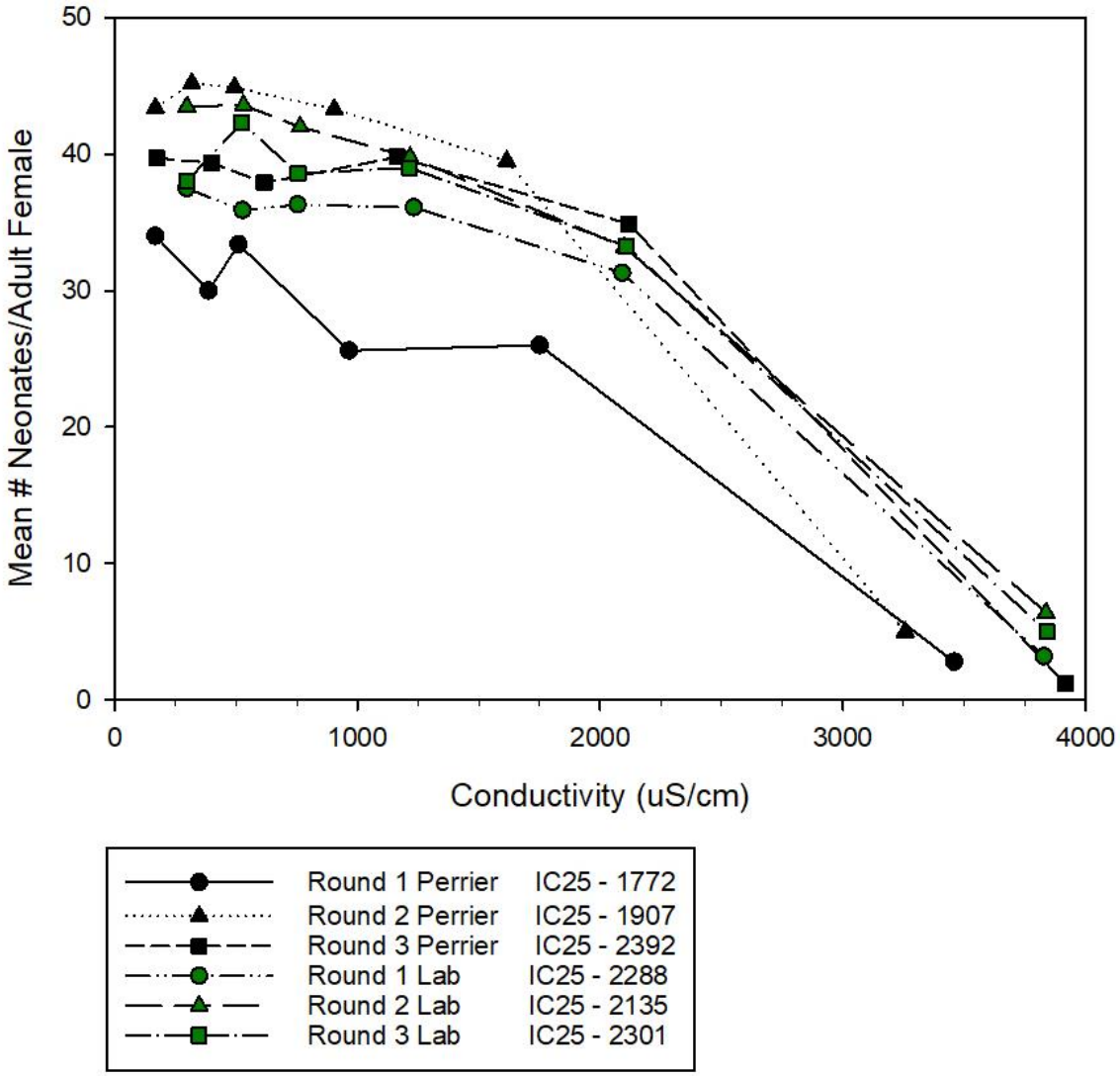


Figure B12. Concentration-response plot of all three rounds of testing for Lab A series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend.

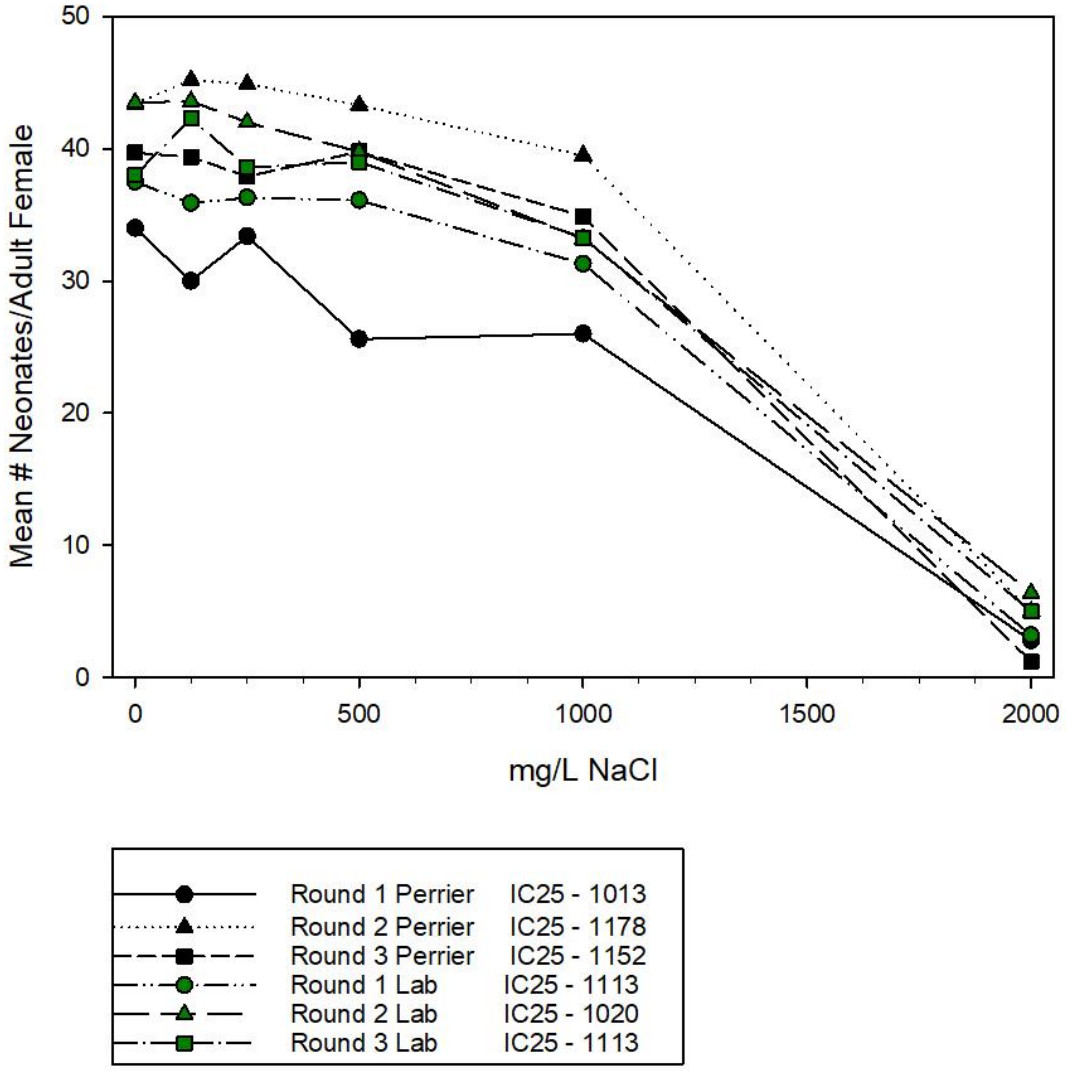


Figure B13. Concentration-response plot of all three rounds of testing for Lab B series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend. Note that for Lab B, Round 1 samples were not tested, and for Round 3, the Sample 3 series all had complete mortality on Day 1.

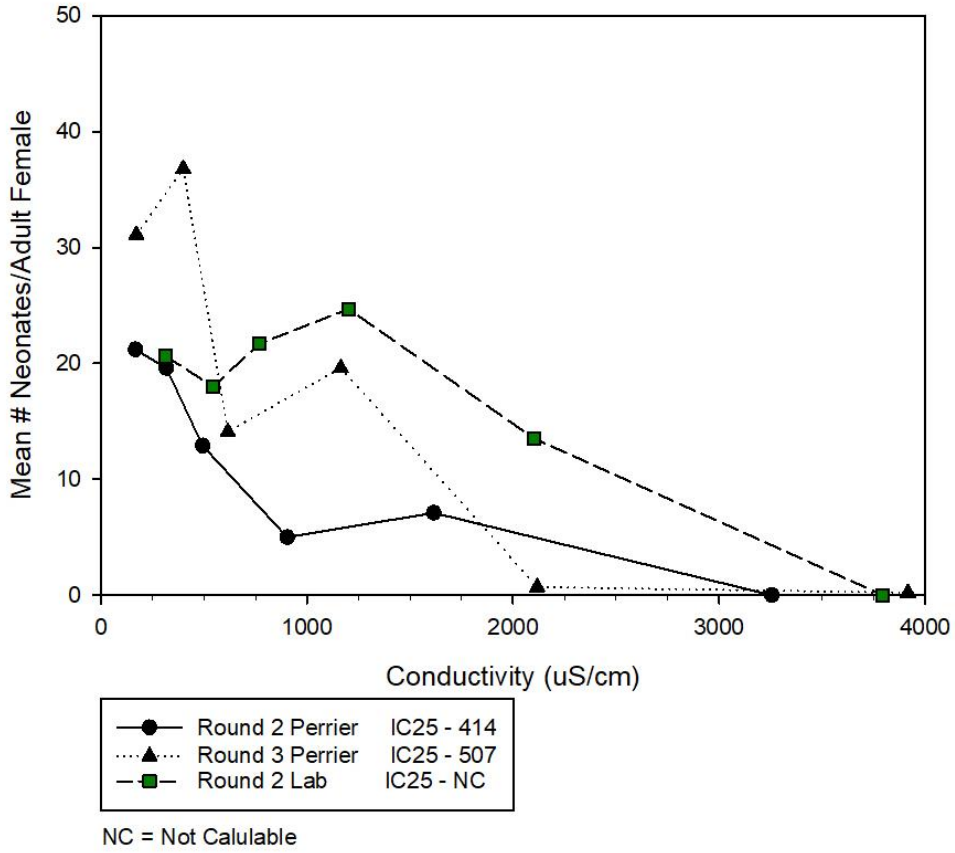


Figure B14. Concentration-response plot of all three rounds of testing for Lab B series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend. Note that for Lab B, Round 1 samples were not tested and for Round 3, the Sample 3 series all had complete mortality at Day 1.

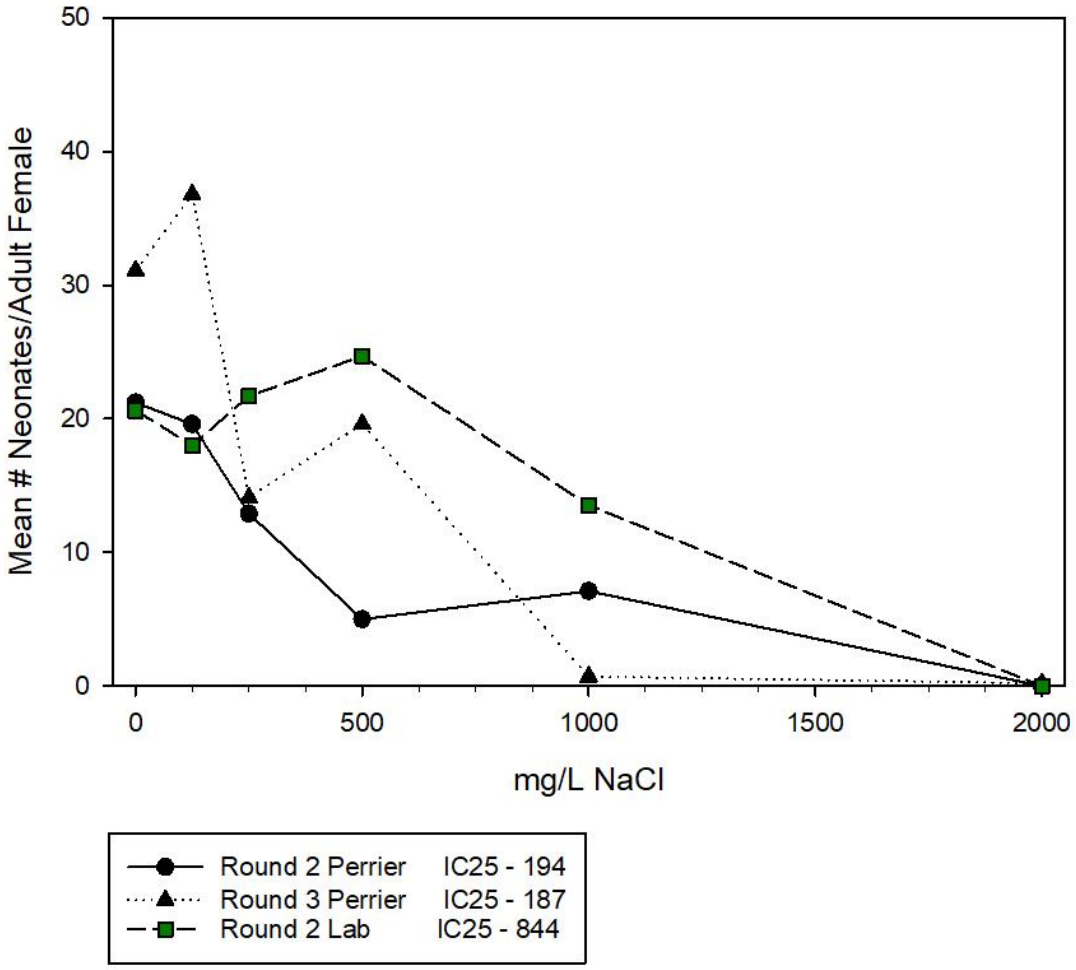


Figure B15. Concentration-response plot of all three rounds of testing for Lab E series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend.

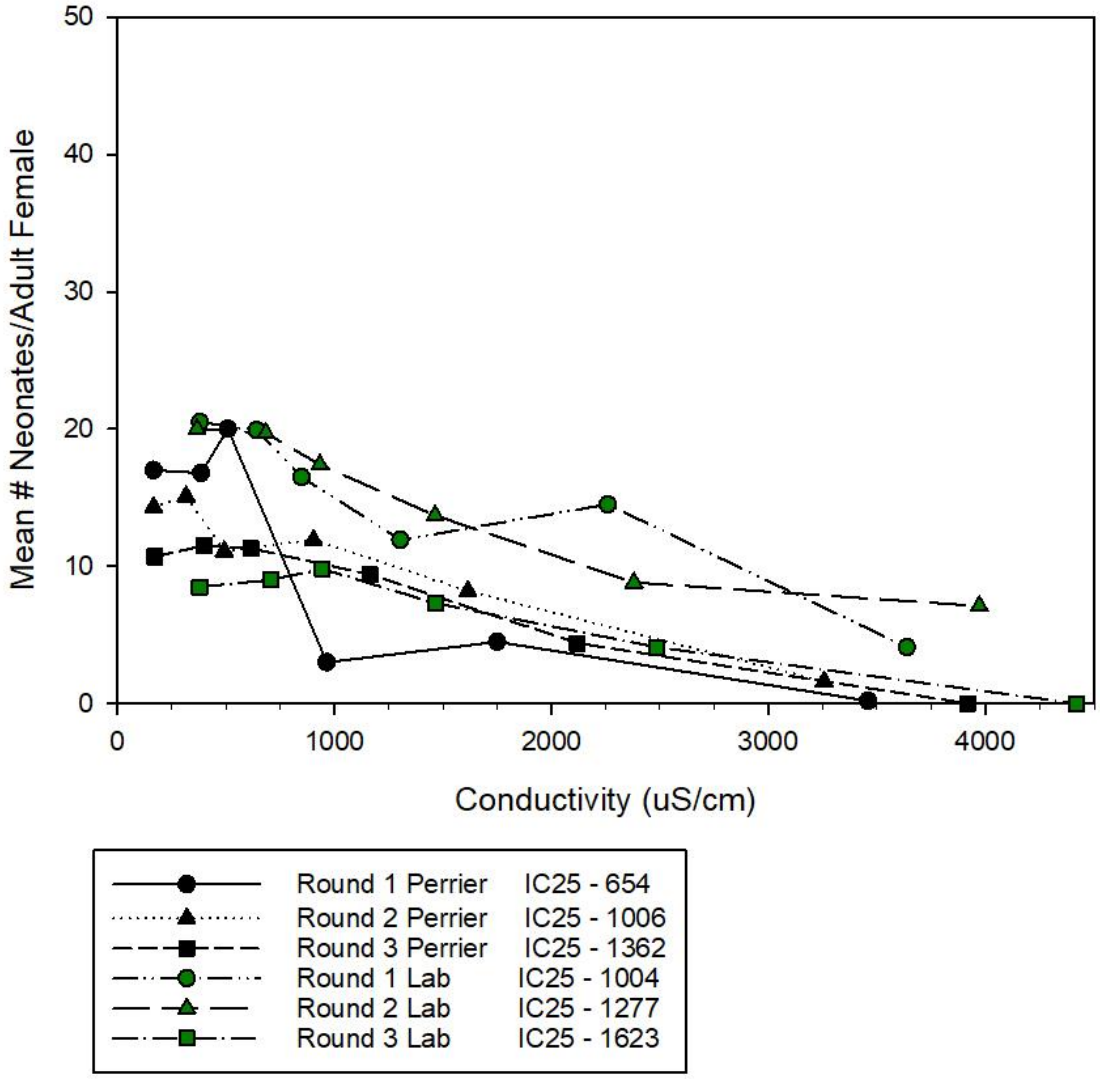


Figure B16. Concentration-response plot of all three rounds of testing for Lab E series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend.

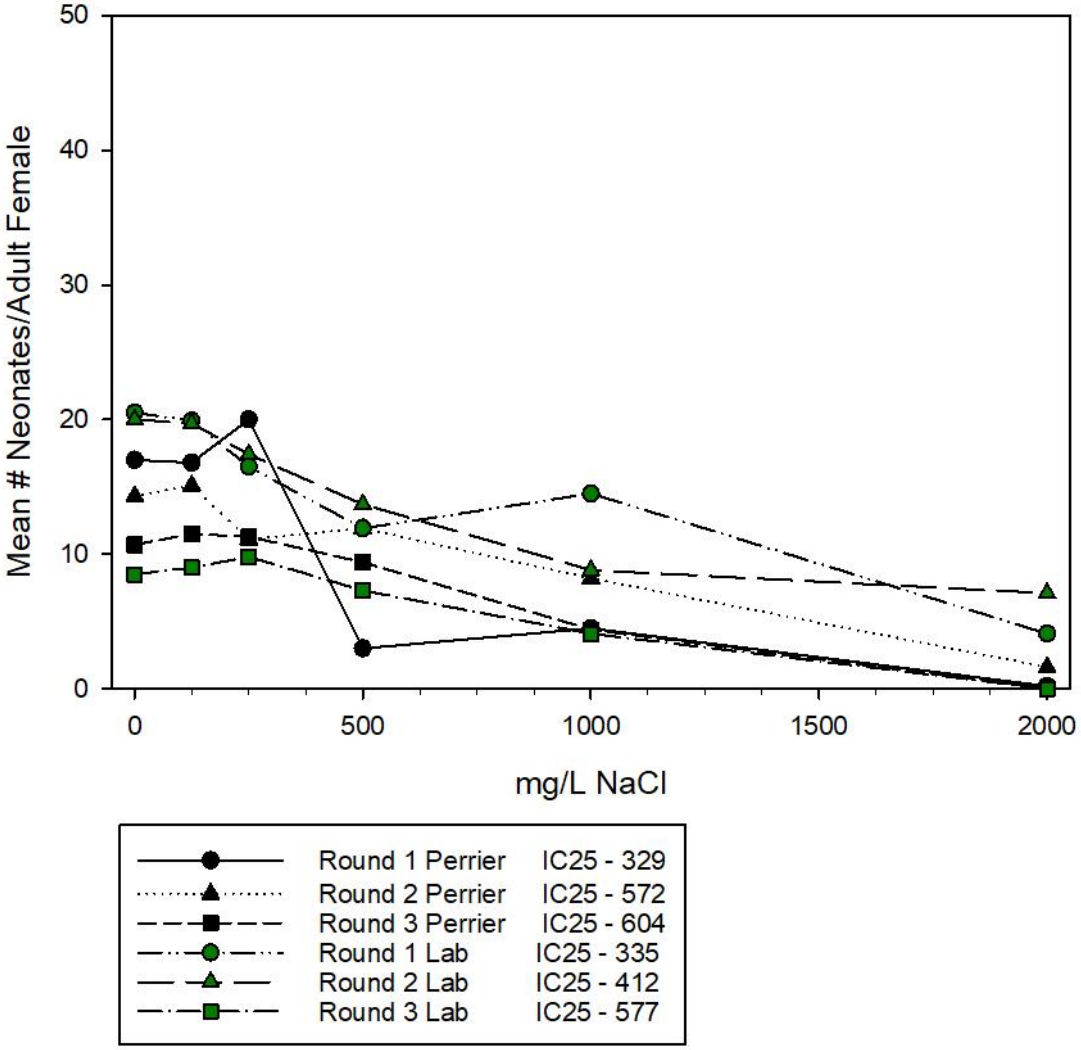


Figure B17. Concentration-response plot of all three rounds of testing for Lab F series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend.

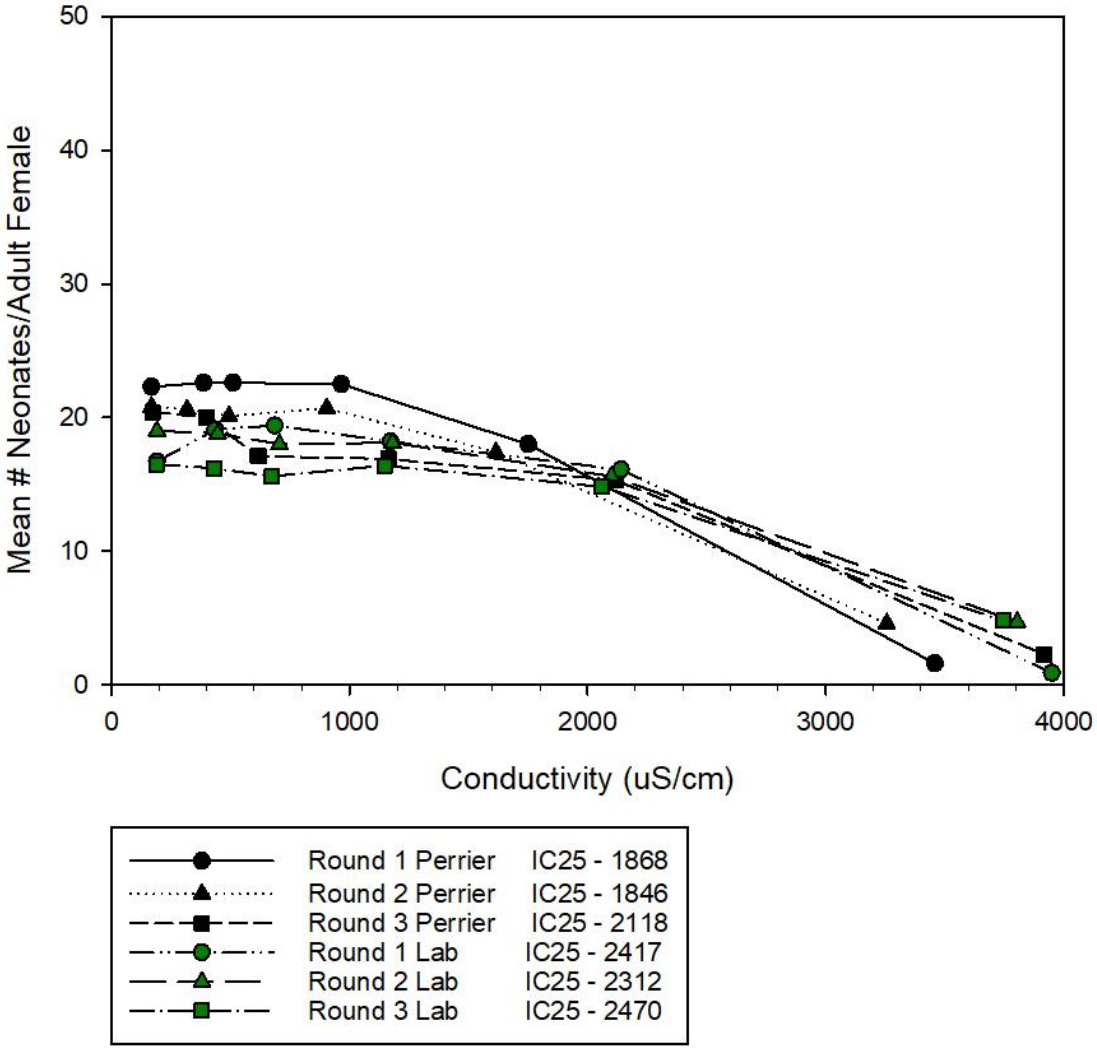


Figure B18. Concentration-response plot of all three rounds of testing for Lab F series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend.

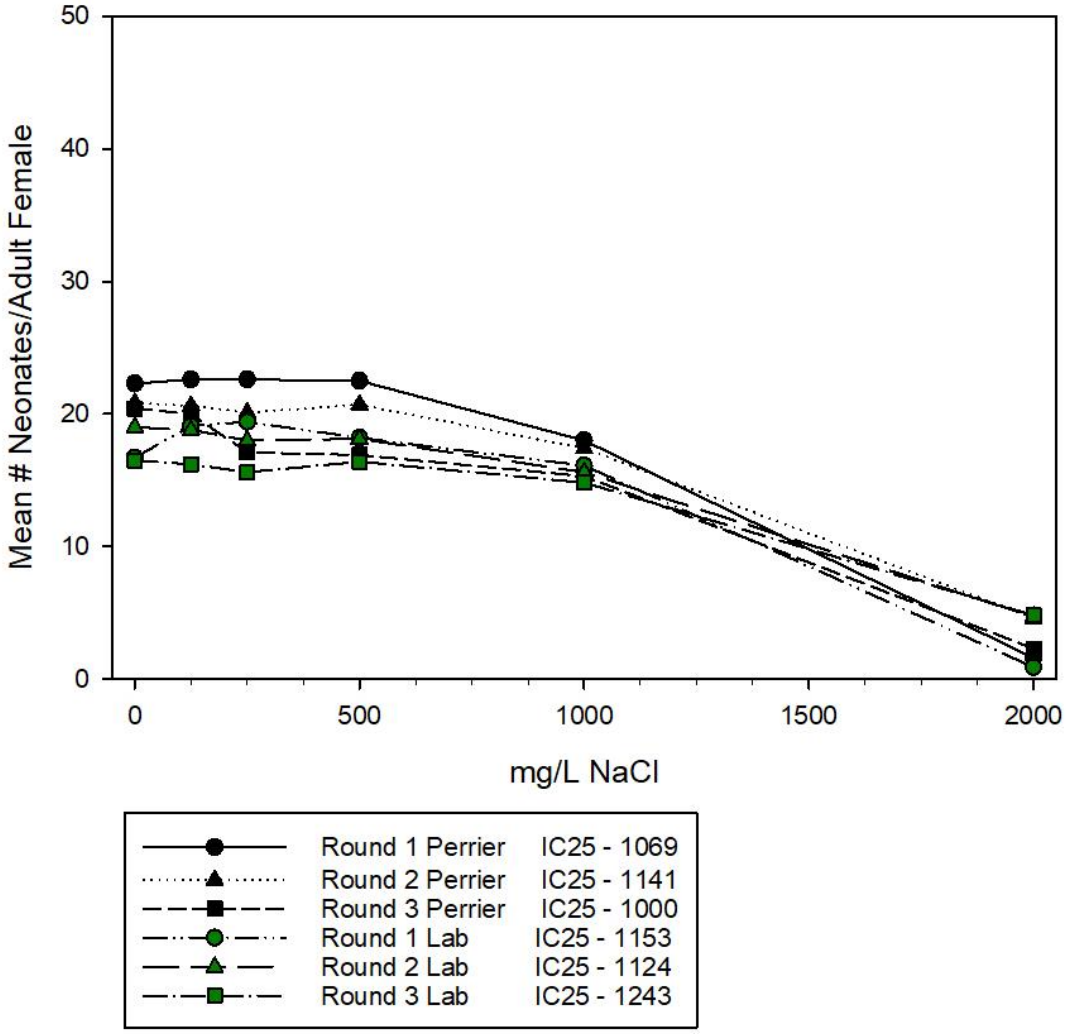


Figure B19. Concentration-response plot of all three rounds of testing for Lab G series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend.

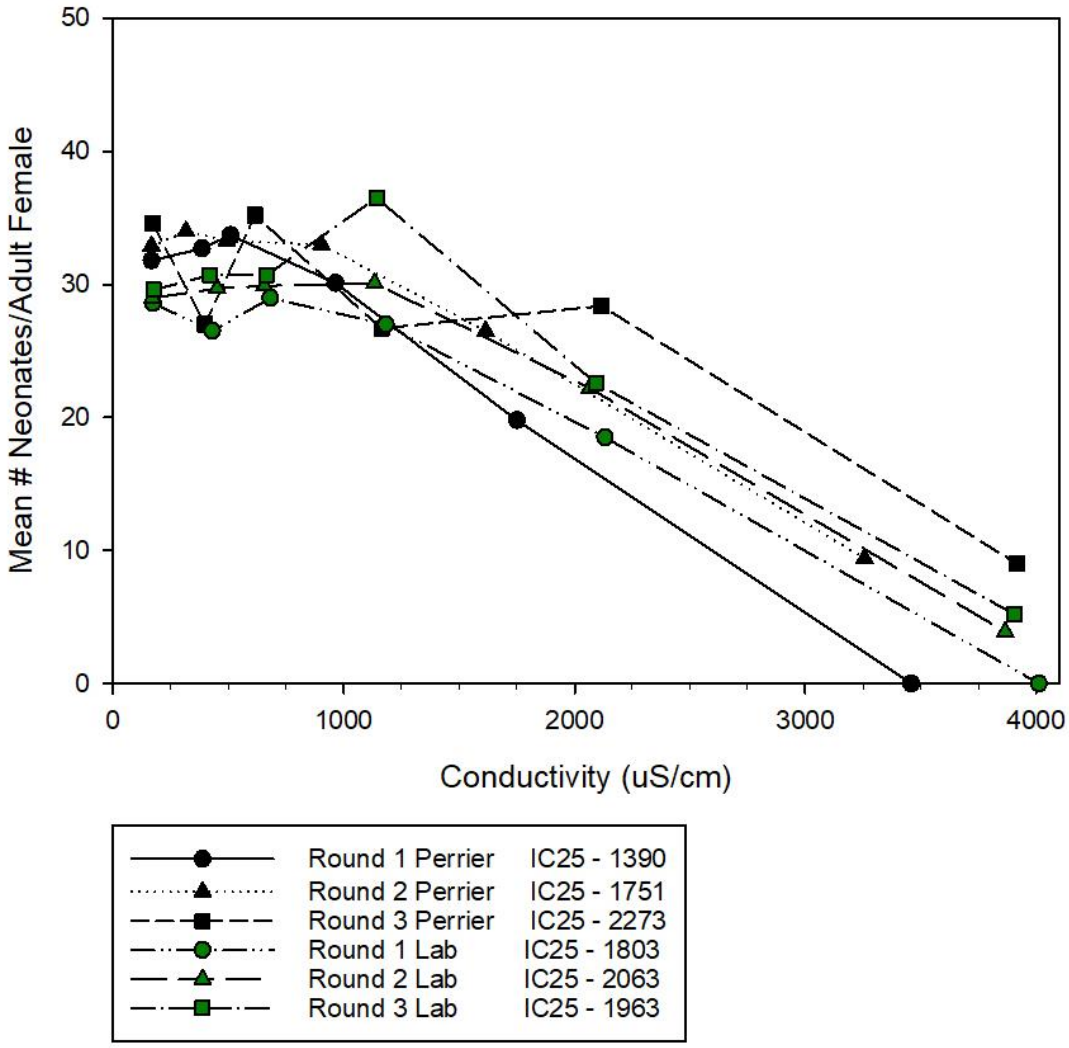


Figure B20. Concentration-response plot of all three rounds of testing for Lab G series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend.

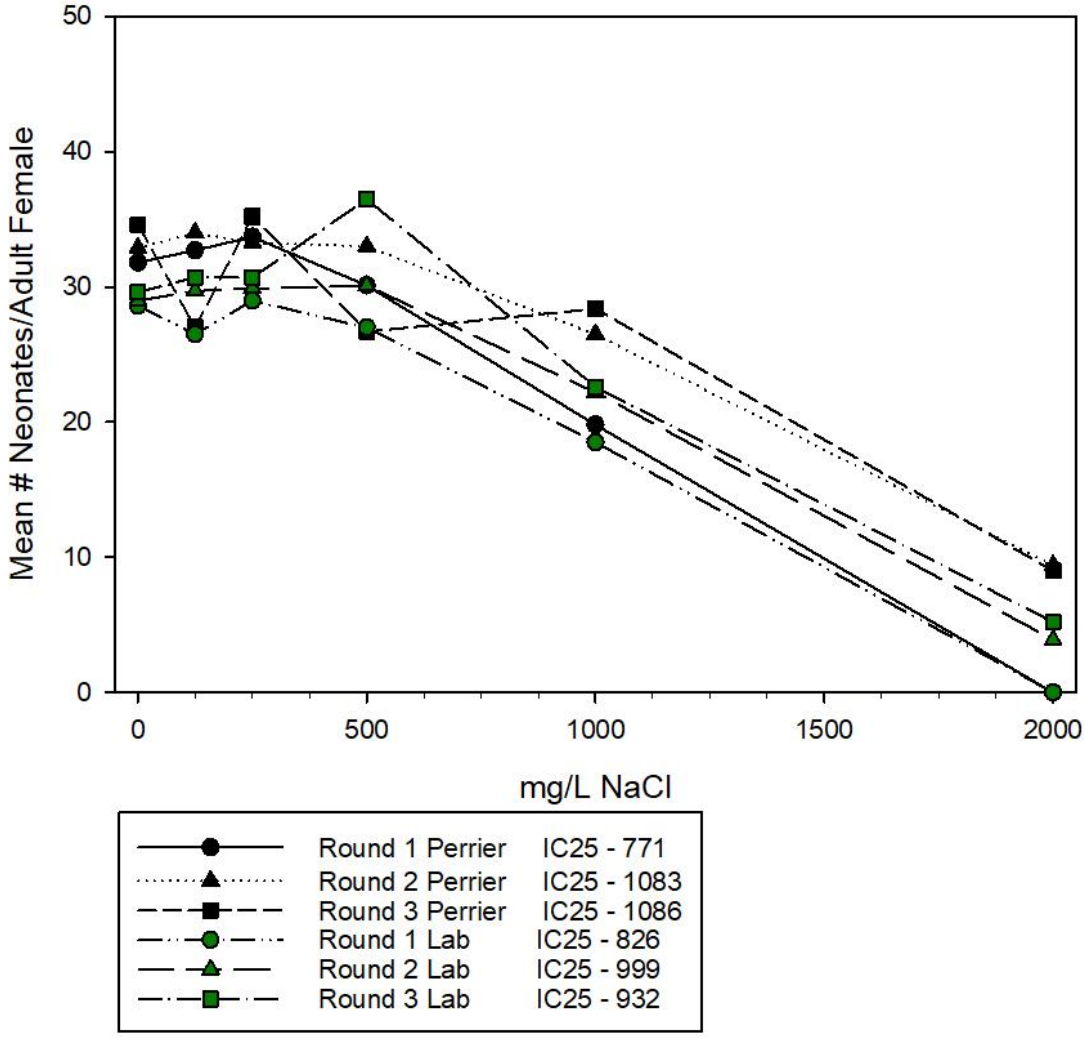


Figure B21. Concentration-response plot of all three rounds of testing for Lab L series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend. Lab L reported culture issues prior to Round 2.

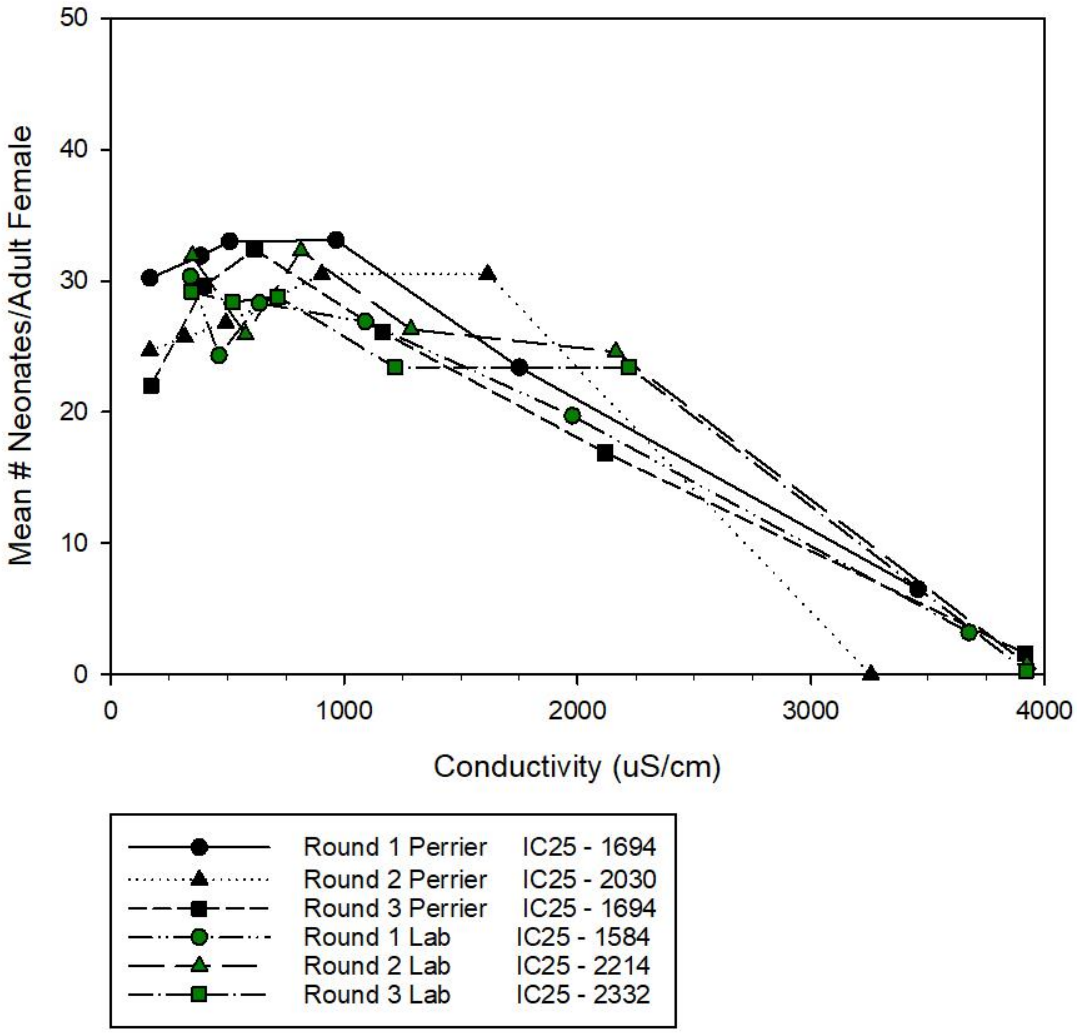


Figure B22. Concentration-response plot of all three rounds of testing for Lab L series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend. Lab L reported culture issues prior to Round 2.

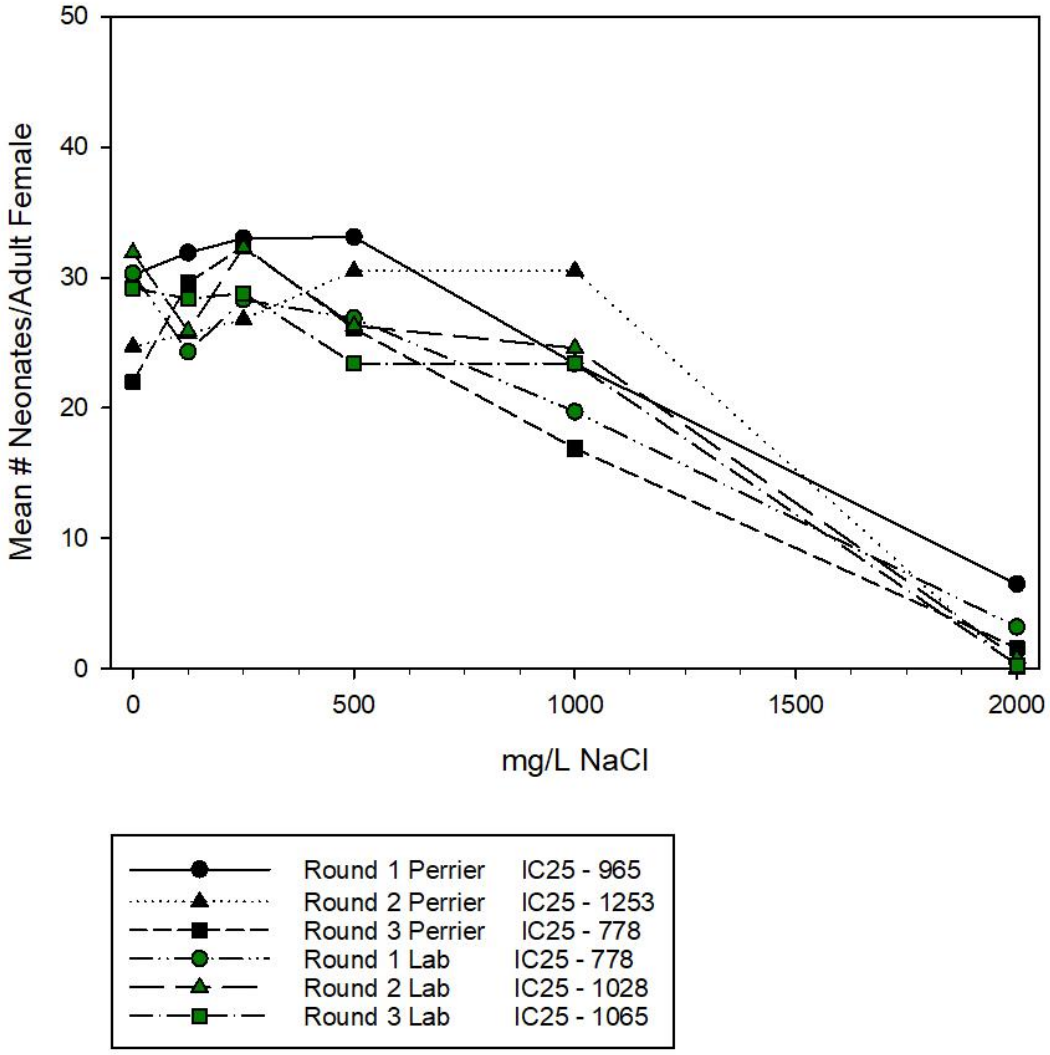


Figure B23. Concentration-response plot of all three rounds of testing for Lab M series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend. Note that Lab M did not participate in Round 1 due to sample delivery issues.

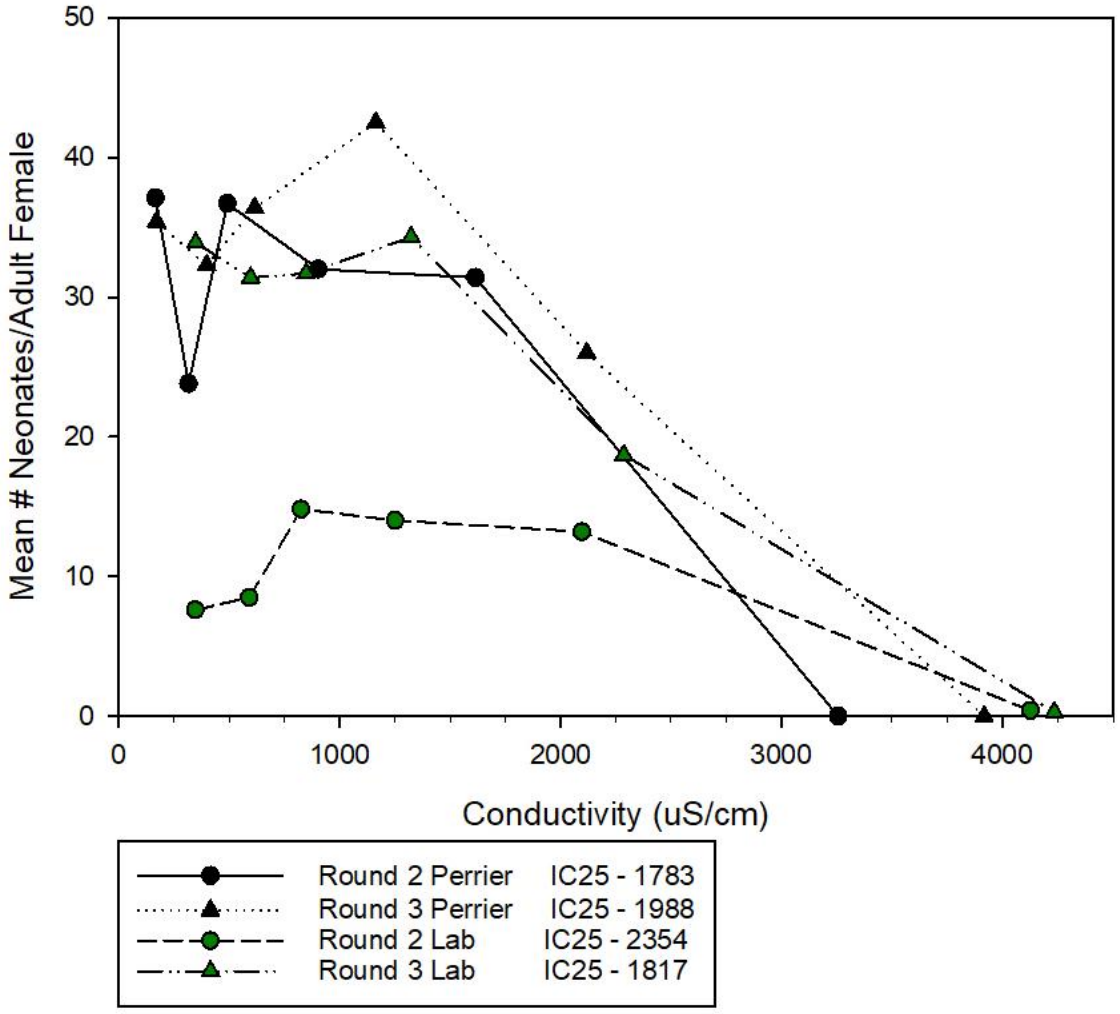


Figure B24. Concentration-response plot of all three rounds of testing for Lab M series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend. Note that Lab M did not participate in Round 1 due to sample delivery issues.

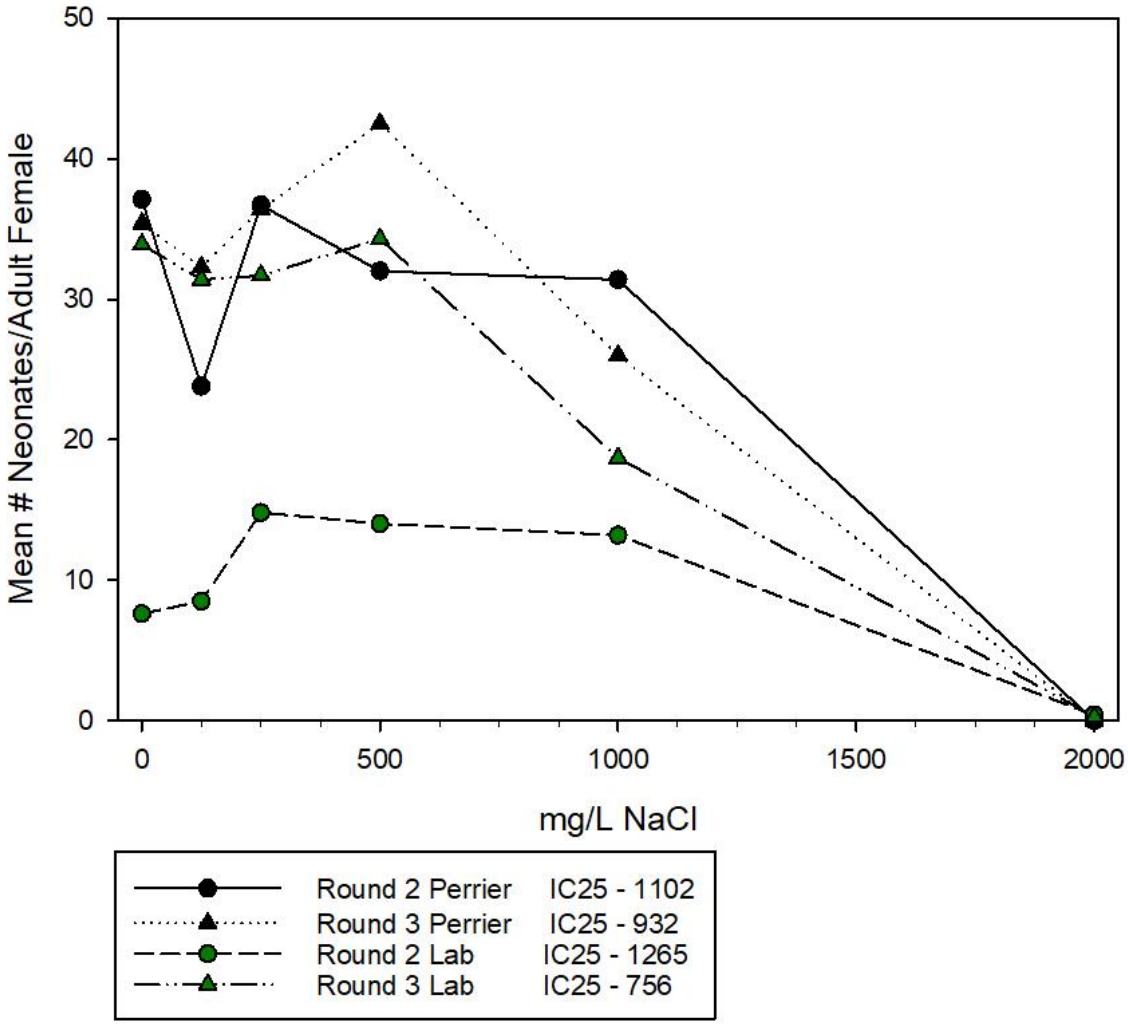


Figure B25. Concentration-response plot of all three rounds of testing for Lab N series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend. Note that Lab N did not analyze the Sample 3 series in Round 1 due to insufficient neonates. Lab N reported culture issues prior to Round 1.

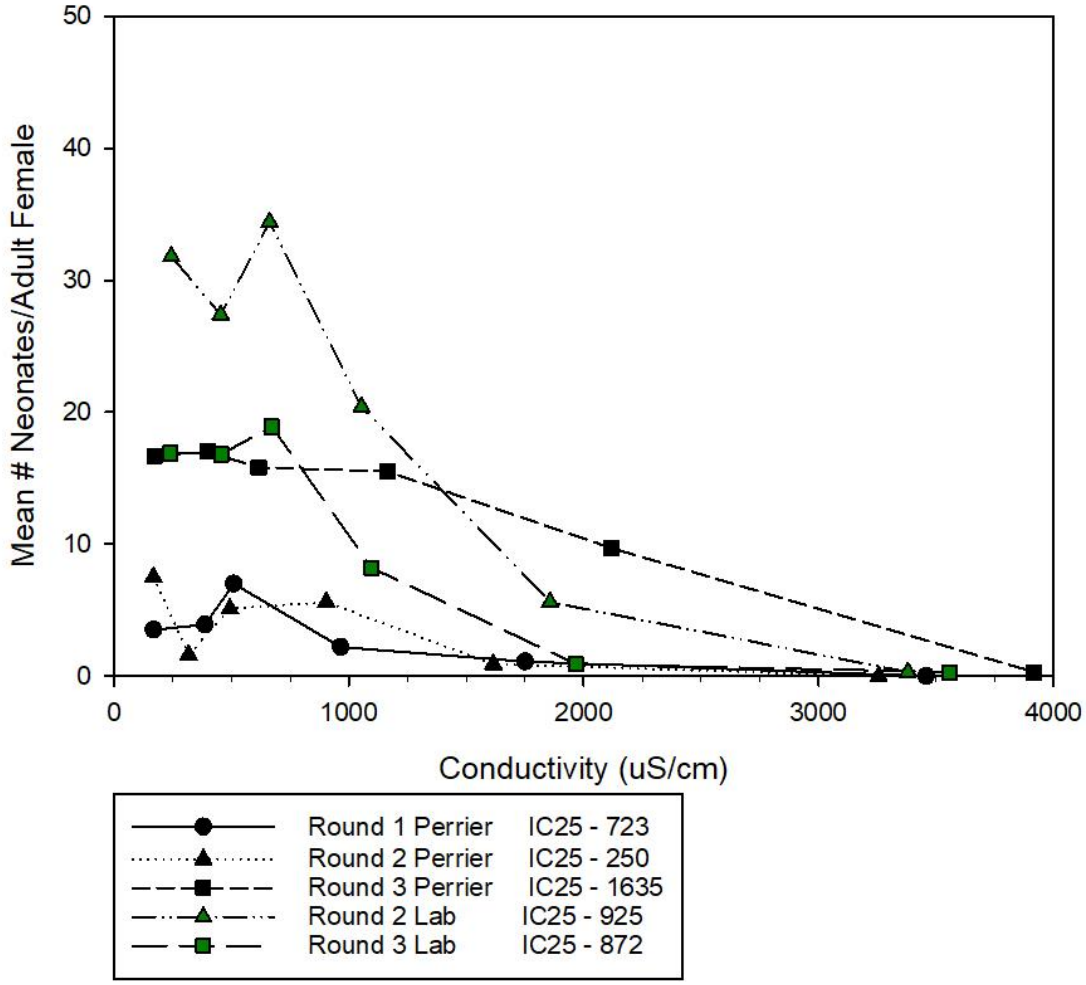


Figure B26. Concentration-response plot of all three rounds of testing for Lab N series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend. Note that Lab N did not analyze the Sample 3 series in Round 1 due to insufficient neonates. Lab N reported culture issues prior to Round 1.

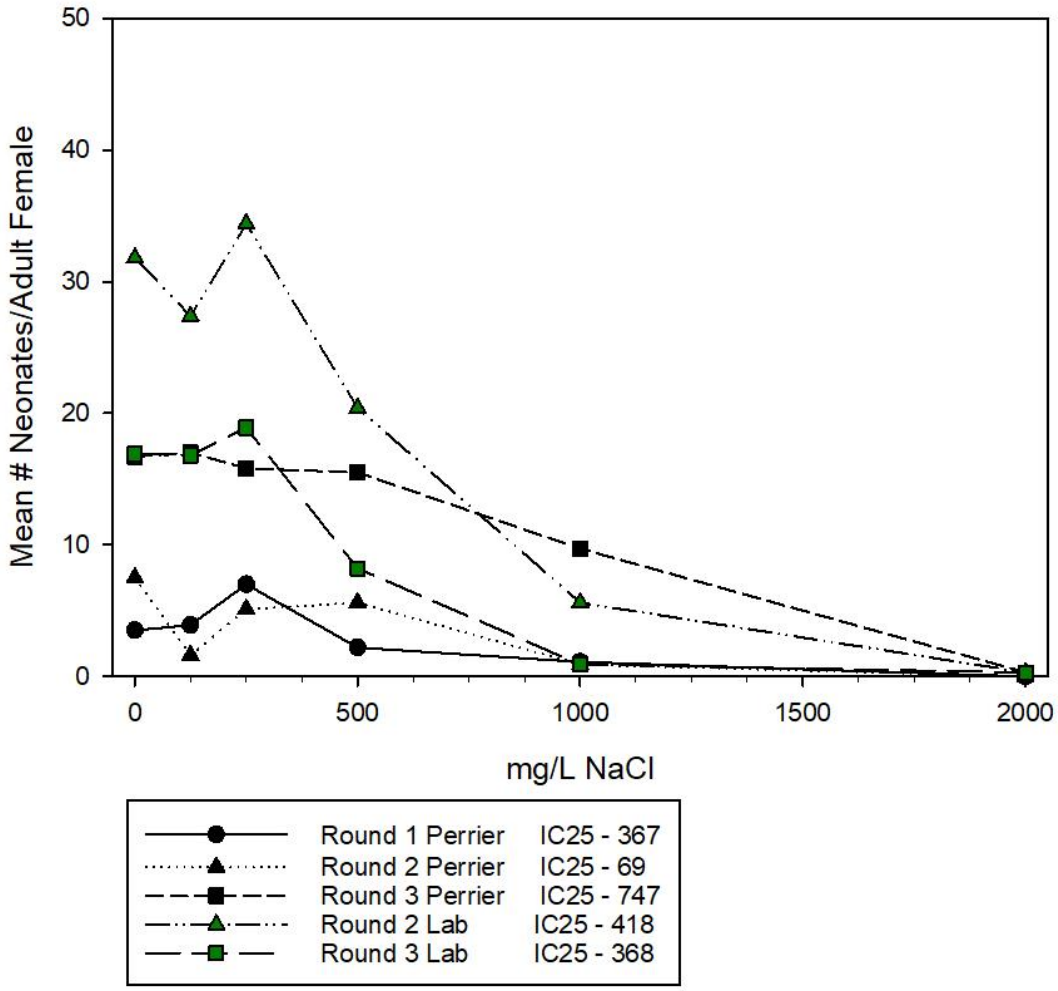


Figure B27. Concentration-response plot of all three rounds of testing for Lab O series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend.

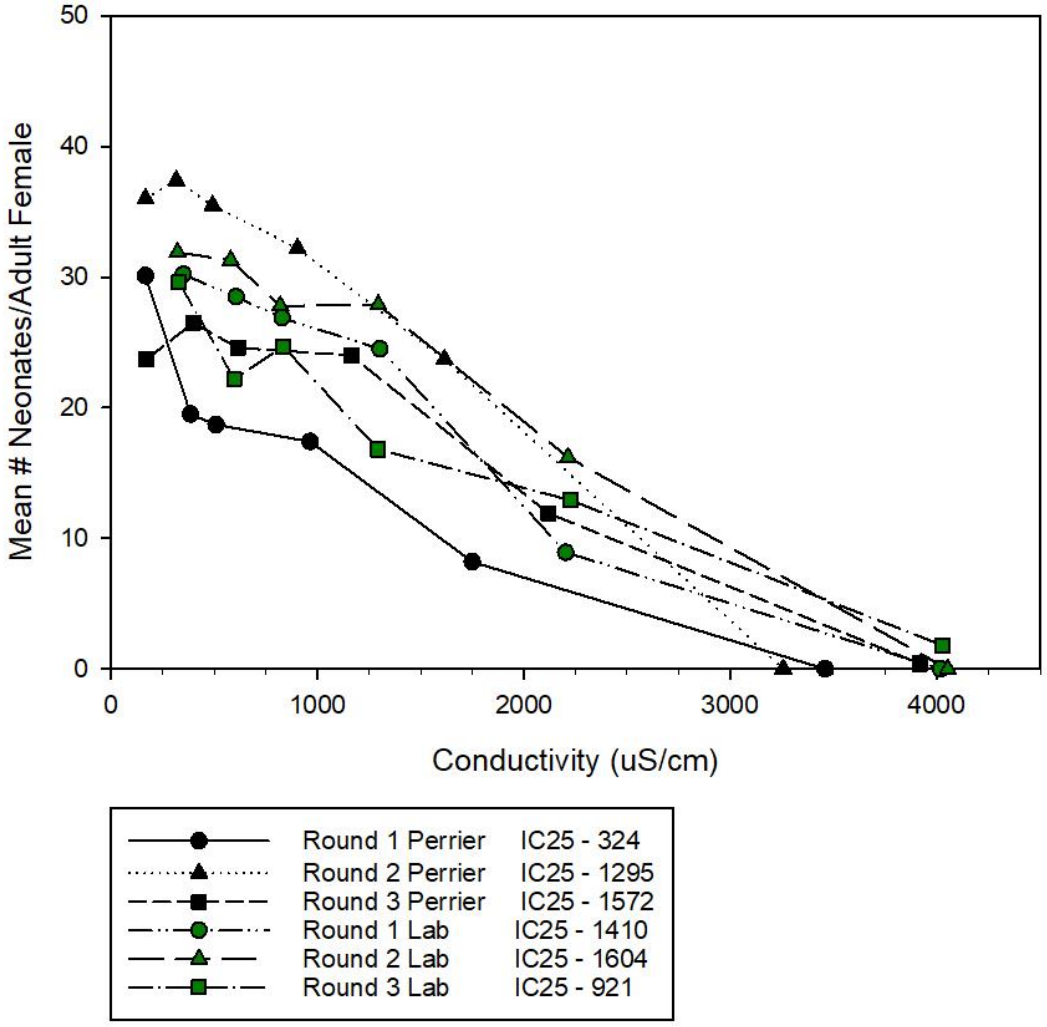


Figure B28. Concentration-response plot of all three rounds of testing for Lab O series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend.

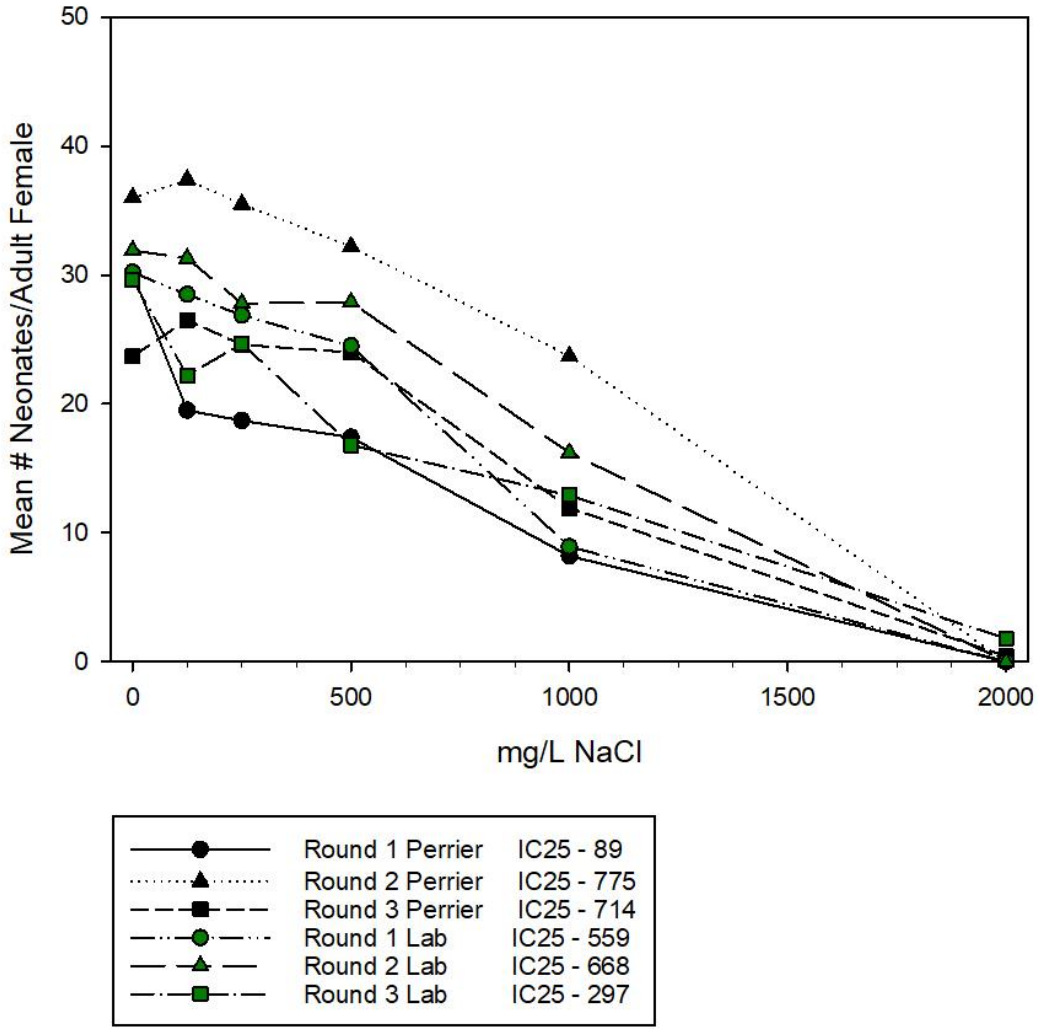


Figure B29. Concentration-response plot of all three rounds of testing for Lab P series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend.

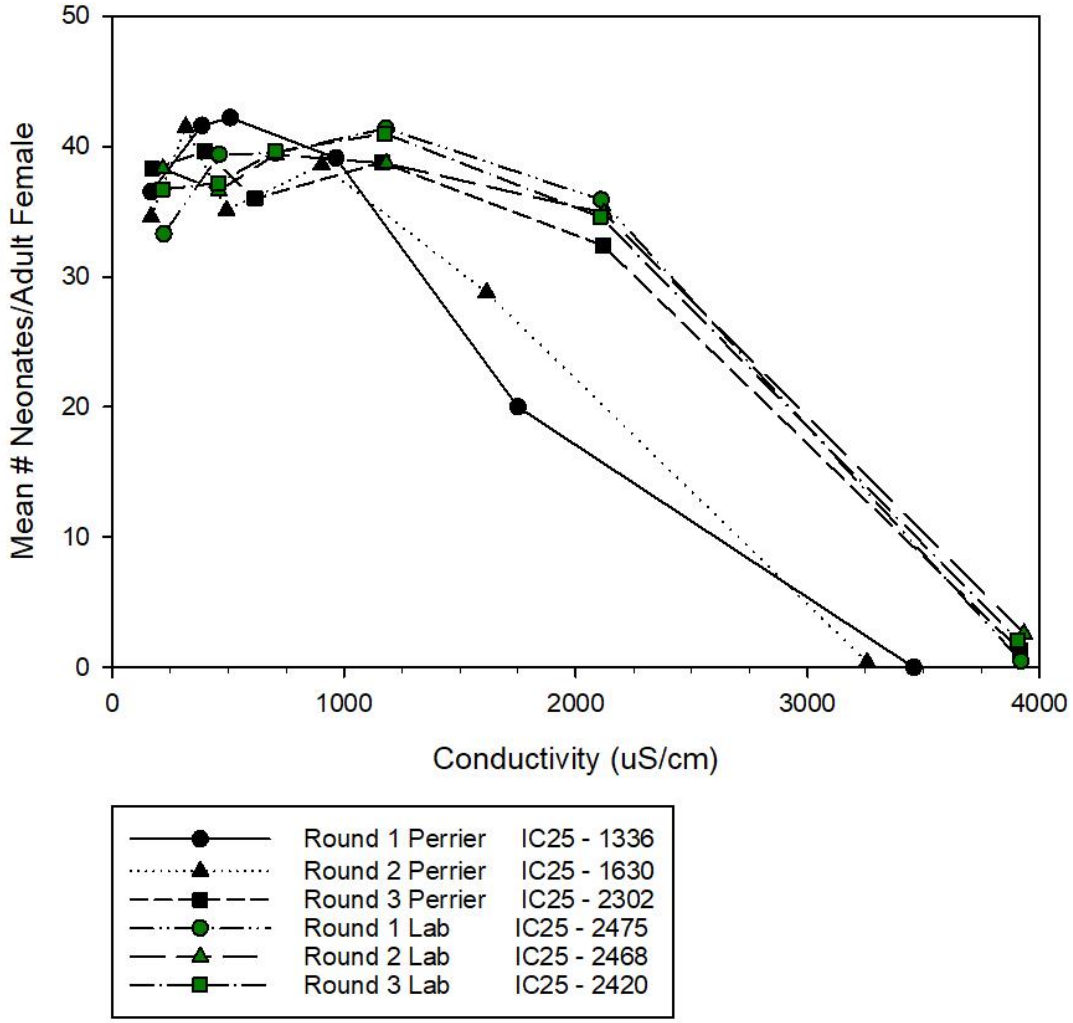


Figure B30. Concentration-response plot of all three rounds of testing for Lab P series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend.

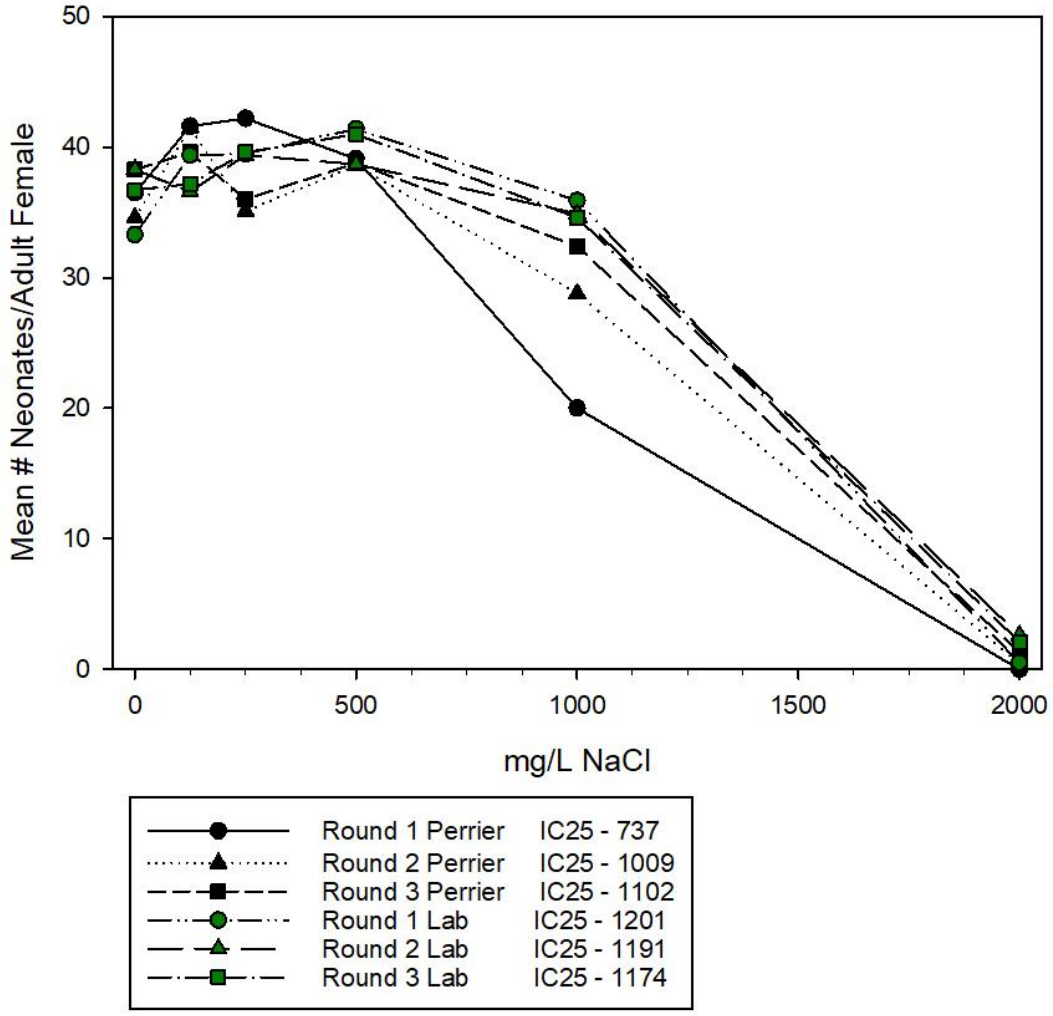


Figure B31. Concentration-response plot of all three rounds of testing for Lab Q series 2 and 3 dilutions based on the laboratories measured conductivity. The IC25 for each sample can be found in the legend.

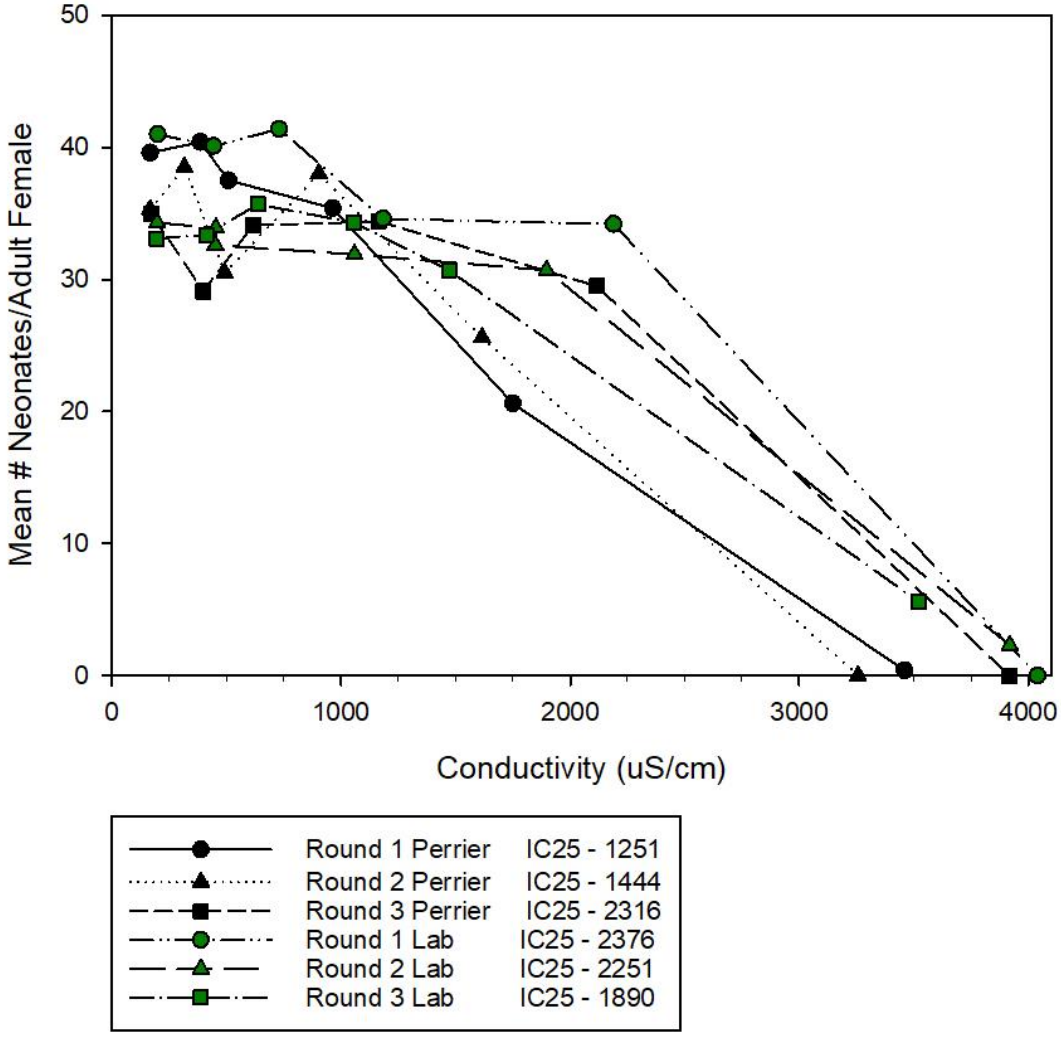


Figure B32. Concentration-response plot of all three rounds of testing for Lab Q series 2 and 3 dilutions based on the nominal sodium chloride concentration. The IC25 for each sample can be found in the legend.

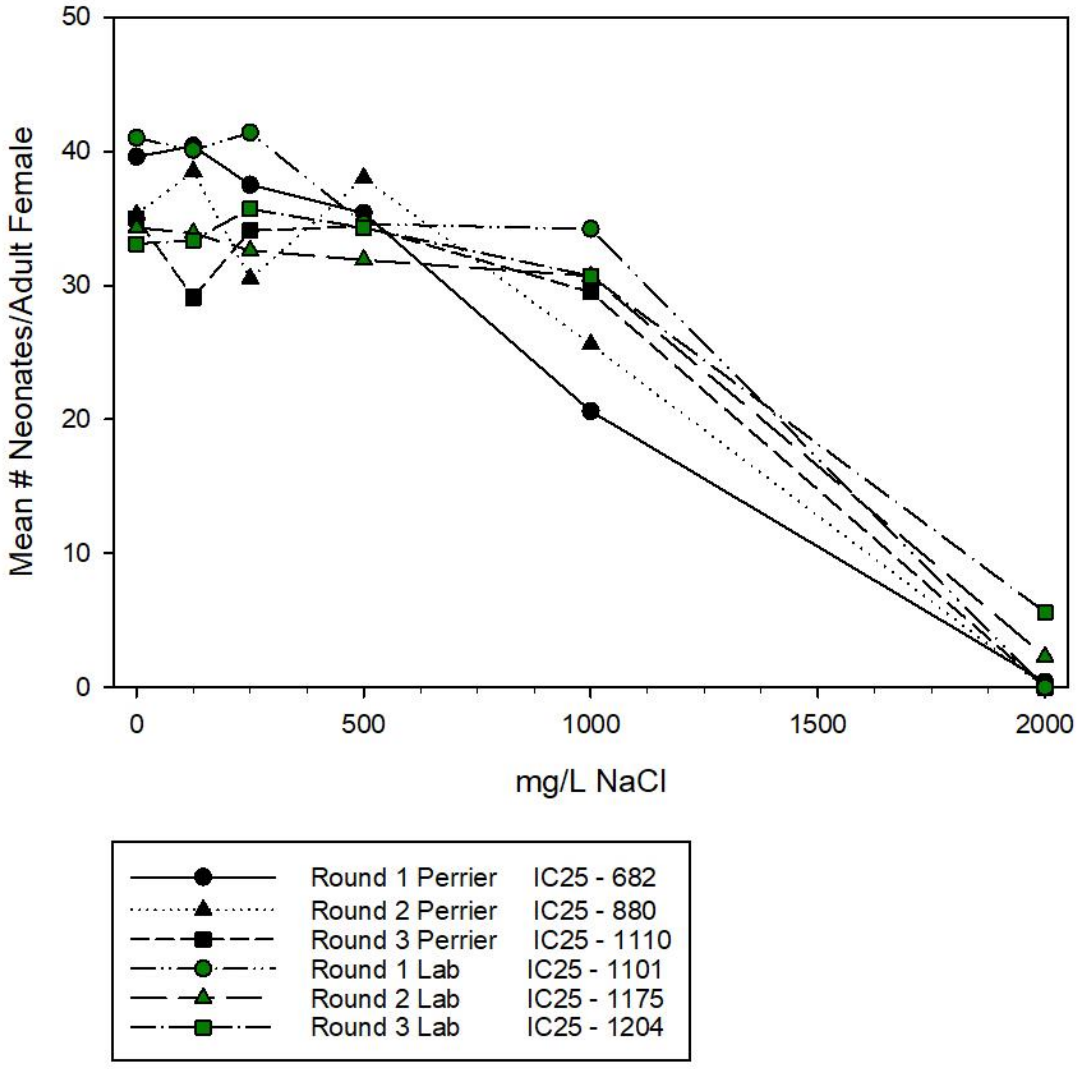


Table B15. Point estimate values based on measured conductivity (µS/cm) in samples 2 and 3 tested as dilution series during the baseline *C. dubia* ILS.

Lab Code	Round #	SCCWRP Perrier (Samples 2A-F)			Lab Dilution Water (Sample 3 series)		
		IC25	IC50	LC50	IC25	IC50	LC50
A	1	1772	2404	3459	2288	2867	3481
A	2	1907	2436	3257	2135	2841	NC
A	3	2392	2921	3242	2301	2917	3552
B	2	414	631	728	1822	2465	2272
B	3	507	614	1450	NC	NC	NC
E	1	654	798	889	1004	2707	NC
E	2	1006	1826	3257	1277	2157	NC
E	3	1362	1894	3117	1623	2347	4096
F	1	1868	2454	2971	2417	2964	3433
F	2	1846	2513	NC	2312	3054	NC
F	3	2118	2824	3617	2470	3165	3747
G	1	1390	2048	2510	1803	2293	2447
G	2	1751	2554	NC	2063	2792	NC
G	3	2273	3112	NC	1963	2790	NC
L	1	1694	2497	NC	1584	2446	NC
L	2	2030	2439	2201	2214	2804	3045
L	3	1694	2459	3197	2332	2869	3071
M	2	1783	2274	2231	2354	2605	2878
M	3	1988	2651	3467	1817	2473	3585
N	1	723	932	2320	NC	NC	NC
N	2	250	983	NC	925	1298	1455
N	3	1635	2362	3917	872	1071	2285
O	1	324	1166	2434	1410	1845	2807

Lab Code	Round #	SCCWRP Perrier (Samples 2A-F)			Lab Dilution Water (Sample 3 series)		
		IC25	IC50	LC50	IC25	IC50	LC50
O	2	1295	1986	2231	1604	2242	3134
O	3	1572	2067	2568	921	1772	3578
P	1	1336	1748	2711	2475	2965	3274
P	2	1630	2180	2671	2468	3005	3934
P	3	2302	2866	3917	2420	2955	3136
Q	1	1251	1801	2699	2376	2931	3012
Q	2	1444	2074	1889	2251	2862	3020
Q	3	2316	2850	2918	1890	2586	NC

*NC = Not Calculable

Table B16. Point estimate values based on nominal NaCl (mg/L) concentrations for samples 2 and 3 tested as dilution series during the baseline *C. dubia* ILS.

Lab Code	Round #	SCCWRP Perrier (Samples 2A-F)			Lab Dilution Water (Sample 3 series)		
		IC25	IC50	LC50	IC25	IC50	LC50
A	1	1010	1380	2000	1110	1450	1800
A	2	1180	1500	2000	1020	1430	NC
A	3	1150	1450	1630	1110	1470	1830
B	2	194	334	393	844	1210	1100
B	3	187	249	650	NC	NC	NC
E	1	329	408	458	335	1320	NC
E	2	572	1130	2000	412	878	NC
E	3	604	882	1560	577	931	1830
F	1	1070	1410	1710	1160	1460	1710
F	2	1140	1550	NC	1120	1560	NC
F	3	1000	1390	1830	1240	1660	2000
G	1	770	1170	1440	826	1230	1440
G	2	1080	1570	NC	999	1400	NC
G	3	1170	1760	NC	1040	1530	NC
L	1	965	1440	NC	778	1280	NC
L	2	1250	1500	1360	1030	1360	1500
L	3	778	1190	1600	1060	1380	1500
M	2	1100	1400	1380	1260	1520	1800
M	3	932	1300	1750	756	1100	1670
N	1	367	482	1330	NC	NC	NC
N	2	70	556	NC	418	652	750
N	3	747	1140	2000	368	485	1200
O	1	89	628	1400	559	801	1330

Lab Code	Round #	SCCWRP Perrier (Samples 2A-F)			Lab Dilution Water (Sample 3 series)		
		IC25	IC50	LC50	IC25	IC50	LC50
O	2	775	1230	1380	668	1020	1500
O	3	714	973	1250	297	756	1750
P	1	737	999	1560	1200	1470	1640
P	2	1010	1340	1640	1190	1490	2000
P	3	1100	1420	2000	1170	1470	1570
Q	1	682	1030	1560	1100	1400	1440
Q	2	880	1280	1170	1180	1480	1560
Q	3	1110	1410	1440	1200	1540	NC

*NC = Not Calculable

Table B16B. Point estimate values 95% confidence intervals based on nominal NaCl (mg/L) concentrations for samples 2 and 3 series. The confidence intervals were calculated by randomly resampling with replacement for the replicates within each control and concentration. The IC25/50 were calculated for each resampling The upper and lower bounds were then calculated based on the results from 100 resamples.

Lab Code	Round #	SCCWRP Perrier (Samples 2A-F)				Lab Dilution Water (Sample 3 series)			
		IC25 Lower Bound	IC25 Upper Bound	IC50 Lower Bound	IC50 Upper Bound	IC25 Lower Bound	IC25 Upper Bound	IC50 Lower Bound	IC50 Upper Bound
A	1	377	1180	1040	1500	1020	1200	1380	1530
A	2	1140	1220	1450	1560	884	1080	1350	1510
A	3	1080	1200	1390	1480	1040	1180	1380	1580
B	2	79	335	200	454	103	1150	860	1430
B	3	122	510	208	680	NC	NC	NC	NC
E	1	292	349	383	446	195	479	1140	1480
E	2	148	999	715	1410	190	653	541	1370
E	3	379	729	707	1090	424	804	753	1250
F	1	1010	1140	1360	1480	1040	1230	1380	1520
F	2	879	1260	1380	1650	810	1250	1400	1680
F	3	578	1080	1300	1480	1090	1350	1470	1890
G	1	535	992	967	1330	633	1060	961	1380
G	2	788	1310	1310	1710	826	1090	1270	1500
G	3	436	1200	1420	1690	655	1260	937	1590
L	1	815	1190	1260	1590	117	1140	893	1480
L	2	668	1170	946	1450	404	1140	1220	1440
L	3	460	1100	835	1420	451	802	917	1230
M	2	332	1200	1270	1470	997	1280	1360	1550
M	3	697	1090	1150	1400	559	833	943	1200

Lab Code	Round #	SCCWRP Perrier (Samples 2A-F)				Lab Dilution Water (Sample 3 series)			
		IC25 Lower Bound	IC25 Upper Bound	IC50 Lower Bound	IC50 Upper Bound	IC25 Lower Bound	IC25 Upper Bound	IC50 Lower Bound	IC50 Upper Bound
N	1	217	1050	825	1380	NC	NC	NC	NC
N	2	269	581	396	1020	276	539	493	741
N	3	36	724	71	1020	290	406	415	603
O	1	68	276	235	802	308	664	631	932
O	2	468	1060	997	1380	249	832	836	1270
O	3	506	871	825	1250	83	562	398	1170
P	1	718	1210	1050	1480	1150	1260	1430	1510
P	2	644	913	839	1290	1050	1260	1390	1540
P	3	975	1190	1330	1480	1050	1250	1390	1540

Table B16B continued. 95% confidence limits of IC25/50 data.

Lab Code	Round #	SCCWRP Perrier (Samples 2A-F)	Lab Dilution Water (Sample 3 series)	Lab Code	Round #	SCCWRP Perrier (Samples 2A-F)	Lab Dilution Water (Sample 3 series)	Lab Code	Round #
Q	1	566	778	854	1210	488	1130	1320	1420
Q	2	654	1120	995	1410	1120	1200	1420	1530
Q	3	1010	1190	1340	1460	1130	1270	1470	1610

*NC = Not Calculable

Table B17. Minimum significant difference (MSD) and percent minimum significant difference (PMSD) values for dilution series sample reproduction endpoint. Values exceeding EPA (2000) 90th percentile value of 37% are highlighted in italics.

		Round 1		Round 2		Round 3	
Lab ID	Sample Series	MSD	PMSD	MSD	PMSD	MSD	PMSD
A	2A-2F	9.88	29.1	3.83	8.82	3.35	8.43
A	3 Series	4.38	11.7	4.34	10.0	5.52	14.5
B	2A-2F	NA	NA	9.35	***44.1***	10.9	***52.8***
B	3 Series	NA	NA	15.4	***49.7***	NA	NA
E	2A-2F	5.55	32.7	5.54	***38.7***	3.91	36.6
E	3 Series	4.54	22.1	5.31	26.5	2.58	30.3
F	2A-2F	7.84	35.2	6.49	31.2	5.86	28.7
F	3 Series	2.82	16.9	3.81	20.0	2.75	16.6
G	2A-2F	4.26	13.4	5.00	15.2	7.09	20.5
G	3 Series	6.14	21.5	5.04	17.4	11.6	***39.2***
L	2A-2F	6.27	20.8	15.6	***63.1***	14.2	***64.4***
L	3 Series	7.78	25.7	7.63	23.9	5.44	18.6
M	2A-2F	NA	NA	15.7	***42.3***	16.8	***47.5***
M	3 Series	NA	NA	6.45	***84.8***	8.62	25.4
N	2A-2F	3.59	*102*	8.48	***113***	7.00	41.9
N	3 Series	NA	NA	8.11	25.5	5.06	29.9
O	2A-2F	6.73	22.4	6.16	17.1	7.21	30.4
O	3 Series	7.23	23.9	8.40	26.3	10.2	34.3
P	2A-2F	9.52	26.1	9.18	26.0	4.20	11.0
P	3 Series	6.33	19.0	6.05	15.8	5.15	14.0
Q	2A-2F	5.89	14.9	7.24	20.5	6.01	17.2
Q	3 Series	5.73	14.0	3.47	10.1	3.63	11.0

NA=Not available either due to sample not being tested or sample being extremely toxic.

Water quality data

Table B18. Summary of conductivity and pH data collected in control chambers from the eleven laboratories participating in the baseline C. dubia interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions). N values refer to the number of tests conducted and included in the means.

		Conductivity (µS/cm) before			Conductivity (µS/cm) after			pH- before			pH- after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
A	Lab Water	313	12	310-317	296	12	295-298	7.52	12	7.42-7.62	7.86	12	7.82-7.92
A	Perrier Water (2A)	196	3	190-199	181	3	176-183	7.64	3	7.60-7.70	7.76	3	7.74-7.79
A	EPA MHW (1)	386	3	385-390	369	3	365-371	7.53	3	7.44-7.59	7.72	3	7.71-7.73
B	Lab Water	338	8	292-387	335	8	312-360	7.57	8	7.38-7.80	7.43	8	7.16-7.70
B	Perrier Water (2A)	212	2	211-212	206	2	192-219	7.45	2	7.38-7.52	7.14	2	7.07-7.21
B	EPA MHW (1)	338	2	327-349	378	2	360-395	7.45	2	7.38-7.52	7.14	2	7.10-7.17
E	Lab Water	376	12	366-382	376	12	367-381	8.06	12	7.96-8.09	8.14	12	8.13-8.19
E	Perrier Water (2A)	207	3	203-210	209	3	206-212	8.07	3	8.02-8.14	8.15	3	8.09-8.21
E	EPA MHW (1)	413	3	406-421	417	3	412-423	8.04	3	8.01-8.10	8.12	3	8.03-8.24
F	Lab Water	207	12	200-218	189	12	186-191	8.17	12	8.15-8.19	8.15	12	8.10-8.17

		Conductivity (µS/cm) before			Conductivity (µS/cm) after			pH- before			pH- after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
F	Perrier Water (2A)	195	3	184-204	174	3	171-176	8.17	3	8.15-8.18	7.96	3	7.91-8.03
F	EPA MHW (1)	386	3	380-395	358	3	353-362	8.07	3	8.05-8.09	7.95	3	7.92-8.00
G	Lab Water	190	12	184-196	174	12	172-176	8.10	12	7.97-8.20	8.10	12	7.98-8.20
G	Perrier Water (2A)	192	3	188-194	177	3	173-181	8.10	3	7.98-8.19	7.87	3	7.80-7.98
G	EPA MHW (1)	372	3	362-386	364	3	357-371	8.05	3	7.92-8.13	7.89	3	7.81-8.02
L	Lab Water	351	9	340-360	345	9	341-350	8.13	9	7.99-8.21	8.12	9	8.10-8.17
L	Perrier Water (2A)	167	3	160-170	164	3	158-168	8.22	3	8.20-8.25	7.94	3	7.77-8.14
L	EPA MHW (1)	350	3	340-360	343	3	337-348	8.18	3	8.17-8.20	8.04	3	7.93-8.10
M	Lab Water	386	8	364-399	348	8	344-350	8.30	8	8.24-8.35	8.17	8	8.16-8.17
M	Perrier Water (2A)	210	2	202-218	188	2	187-188	8.44	2	8.43-8.45	8.00	2	7.84-8.16
M	EPA MHW (1)	421	2	418-425	378	2	376-380	8.25	2	8.25-8.25	7.99	2	7.93-8.05
N	Lab Water	254	11	248-262	244	11	239-251	7.82	11	7.02-8.08	8.11	11	8.08-8.15
N	Perrier Water (2A)	149	3	147-151	143	3	139-145	8.01	3	7.87-8.12	8.09	3	7.98-8.16

		Conductivity (µS/cm) before			Conductivity (µS/cm) after			pH- before			pH- after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
N	EPA MHW (1)	302	3	294-313	296	3	283-310	7.88	3	7.73-8.00	7.97	3	7.85-8.06
O	Lab Water	344	12	337-351	344	12	323-351	7.85	12	7.78-7.95	7.89	12	7.79-7.96
O	Perrier Water (2A)	183	3	179-186	176	3	174-180	7.98	3	7.91-8.04	7.80	3	7.70-7.87
O	EPA MHW (1)	380	3	377-384	360	3	359-362	7.86	3	7.78-7.94	7.77	3	7.70-7.88

Table B18 cont. Summary of conductivity and pH data collected in control chambers from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions).

		Conductivity (µS/cm) before			Conductivity (µS/cm) after			pH- before			pH- after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
P	Lab Water	219	12	216-223	218	12	215-224	8.02	12	7.88-8.27	7.89	12	7.74-7.99
P	Perrier Water (2A)	196	3	195-198	196	3	195-196	8.18	3	8.10-8.30	8.01	3	7.97-8.06
P	EPA MHW (1)	384	3	380-387	379	3	374-382	8.15	3	8.04-8.31	7.93	3	7.86-8.00
Q	Lab Water	180	12	177-186	202	12	193-250	8.08	12	7.98-8.17	8.19	12	8.16-8.24

		Conductivity (µS/cm) before			Conductivity (µS/cm) after			pH- before			pH- after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
Q	Perrier Water (2A)	173	3	168-175	185	3	181-187	7.76	3	7.72-7.79	8.20	3	8.18-8.23
Q	EPA MHW (1)	355	3	352-358	376	3	374-377	7.77	3	7.74-7.82	8.10	3	8.07-8.14

Table B19. Summary of dissolved oxygen (DO) and water temperature (water temp) data collected in control chambers from the eleven laboratories participating in the baseline C. dubia interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions). N values refer to the number of tests conducted and included in the means.

		DO (mg/L) before			DO (mg/L) after			Water temp (°C) before			Water temp (°C) after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
A	Lab Water	6.44	12	5.76-6.93	7.65	12	7.29-7.84	24.1	12	24.0-24.3	24.2	12	24.1-24.4
A	Perrier Water (2A)	6.48	3	5.88-6.79	8.01	3	7.96-8.04	24.1	3	24.1-24.2	24.2	3	24.1-24.3
A	EPA MHW (1)	6.32	3	5.79-6.68	7.88	3	7.77-7.97	24.2	3	24.0-24.3	24.5	3	24.4-24.6
B	Lab Water	8.67	8	8.20-8.40	9.12	8	8.51-9.80	24.8	8	24.7-24.9	24.7	8	24.5-24.9
B	Perrier Water (2A)	8.62	2	8.40-8.85	8.91	2	8.51-9.31	24.8	2	24.7-24.9	24.7	2	24.6-24.9
B	EPA MHW (1)	8.34	2	7.99-8.68	8.88	2	8.51-9.26	24.8	2	24.7-24.9	24.7	2	24.6-24.9
E	Lab Water	7.44	12	7.22-7.71	7.76	12	7.71-7.93	25.0	12	25.0-25.0	24.9	12	24.8-25.0
E	Perrier Water (2A)	7.51	3	7.36-7.66	7.89	3	7.79-7.96	25.0	3	24.9-25.0	25.0	3	25.0-25.0
E	EPA MHW (1)	7.40	3	7.29-7.56	7.92	3	7.73-8.15	25.0	3	24.9-25.0	25.0	3	25.0-25.0
F	Lab Water	8.10	12	7.90-8.29	8.38	12	7.81-8.79	25.0	12	24.7-25.4	25.2	12	24.9-25.5
F	Perrier Water (2A)	7.99	3	7.89-8.09	8.74	3	8.70-8.81	25.2	3	25.1-25.4	25.3	3	25.1-25.4

		DO (mg/L) before			DO (mg/L) after			Water temp (°C) before			Water temp (°C) after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
F	EPA MHW (1)	7.98	3	7.85-8.09	8.80	3	8.77-8.83	25.1	3	25.0-25.2	25.3	3	25.0-25.5
G	Lab Water	8.32	12	8.28-8.40	8.35	12	8.30-8.39	25.0	12	24.5-25.2	24.6	12	24.4-24.8
G	Perrier Water (2A)	8.30	3	8.26-8.35	9.07	3	8.91-9.20	25.1	3	25.0-25.2	24.8	3	24.7-24.9
G	EPA MHW (1)	8.30	3	8.26-8.35	8.96	2	8.76-9.10	25.0	3	25.0-25.2	24.9	3	24.9-25.0
L	Lab Water	8.71	9	8.61-8.83	9.10	9	8.96-9.31	24.2	9	24.0-24.6	24.1	9	24.0-24.3
L	Perrier Water (2A)	8.68	3	8.65-8.72	9.63	3	9.17-10.1	24.3	3	24.0-24.6	24.1	3	24.0-24.2
L	EPA MHW (1)	8.71	3	8.70-8.71	9.61	3	9.20-10.2	24.3	3	24.0-24.6	24.2	3	24.0-24.3
M	Lab Water	6.69	8	6.41-6.89	7.50	9	7.21-9.07	24.9	8	24.3-25.6	24.4	8	24.0-24.9
M	Perrier Water (2A)	6.70	2	6.62-6.79	8.04	2	7.84-8.25	24.9	2	24.8-25.1	24.7	2	24.5-24.9
M	EPA MHW (1)	6.62	2	6.50-6.74	8.10	2	7.80-8.39	24.9	2	24.8-25.0	24.6	2	24.4-24.7
N	Lab Water	7.46	11	6.56-8.01	8.32	11	8.31-8.36	25.1	11	24.7-25.4	24.9	11	24.7-25.1
N	Perrier Water (2A)	7.46	3	6.71-8.08	8.32	3	8.27-8.36	25.1	3	25.0-25.4	24.6	3	24.4-24.7
N	EPA MHW (1)	7.46	3	6.69-8.04	8.30	3	8.26-8.33	25.1	3	24.9-25.3	24.6	3	24.5-24.6

		DO (mg/L) before			DO (mg/L) after			Water temp (°C) before			Water temp (°C) after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
O	Lab Water	7.80	12	7.61-7.96	8.45	12	8.10-8.77	25.1	12	24.9-25.6	24.9	11	24.7-25.2
O	Perrier Water (2A)	7.87	3	7.78-7.94	10.2	3	9.79-10.5	25.2	3	25.0-25.4	24.6	3	24.4-24.7
O	EPA MHW (1)	7.89	3	7.75-8.02	10.1	3	9.93-10.3	25.3	3	25.2-25.4	24.6	3	24.5-24.6

Table B19 cont. Summary of dissolved oxygen (DO) and water temperature (water temp) data collected in control chambers from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions)

		DO (mg/L) before			DO (mg/L) after			Water temp (°C) before			Water temp (°C) after		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
P	Lab Water	8.30	12	8.22-8.38	8.27	12	8.14-8.34	25.0	12	24.7-25.4	25.2	12	25.0-25.4
P	Perrier Water (2A)	8.44	3	8.38-8.54	8.19	3	8.14-8.21	25.1	3	25.0-25.4	25.3	3	25.3-25.4
P	EPA MHW (1)	8.47	3	8.41-8.51	8.16	3	8.03-8.23	25.1	3	24.8-25.3	25.3	3	25.2-25.5
Q	Lab Water	7.93	12	7.84-8.01	8.33	12	8.25-8.43	25.1	12	24.9-25.6	23.5	12	22.8-24.0
Q	Perrier Water (2A)	8.63	3	8.59-8.69	8.27	3	8.22-8.35	25.2	3	25.0-25.4	22.8	3	22.0-23.3
Q	EPA MHW (1)	8.66	3	8.48-8.89	8.36	3	8.32-8.44	25.3	3	25.2-25.4	22.7	3	22.0-23.2

Table B20. Summary of air temperature data collected in control chambers from the eleven laboratories participating in the baseline C. dubia interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions). N values refer to the number of tests conducted and included in the means.

		(°C) Before			(°C) After		
Lab	Sample Type	Mean	N	Range	Mean	N	Range
A	Lab Water	24.7	12	24.6-24.7	24.5	12	24.4-24.7
A	Perrier Water (2A)	24.7	3	24.7-24.7	24.5	3	24.4-24.7
A	EPA MHW (1)	24.7	3	24.7-24.7	24.5	3	24.4-24.6
B	Lab Water	24.8	8	24.1-25.1	24.7	8	24.4-25.0
B	Perrier Water (2A)	24.8	2	24.6-25.0	24.8	2	24.5-25.0
B	EPA MHW (1)	24.8	2	24.6-25.0	24.8	2	24.5-25.0
E	Lab Water	25.0	12	25.0-25.0	24.9	12	24.9-25.0
E	Perrier Water (2A)	24.9	3	24.9-25	25.0	3	24.9-25.0
E	EPA MHW (1)	25.0	3	25.0-25.0	25.0	3	24.9-25.0
F	Lab Water	26.6	12	26.4-26.8	26.6	12	26.5-26.8
F	Perrier Water (2A)	26.6	3	26.4-26.8	26.6	3	26.5-26.7
F	EPA MHW (1)	26.6	3	26.4-26.8	26.6	3	26.5-26.7
G	Lab Water	25.5	12	25.2-25.7	25.4	6	25.2-25.7
G	Perrier Water (2A)	25.5	3	25.2-25.7	25.4	3	25.2-25.7
G	EPA MHW (1)	25.5	3	25.2-25.7	25.5	3	25.2-25.7
L	Lab Water	22.7	9	21.8-24.0	23.0	9	21.9-24.3
L	Perrier Water (2A)	22.7	3	21.8-24.1	23.0	3	22.0-24.3
L	EPA MHW (1)	22.7	3	21.8-24.0	23.1	3	22.0-24.3
M	Lab Water	25.8	8	25.6-26.3	26.0	8	25.3-26.5
M	Perrier Water (2A)	25.6	2	25.6-25.6	26.1	2	25.7-26.5
M	EPA MHW (1)	25.6	2	25.6-25.6	26.2	2	25.9-26.5
N	Lab Water	25.4	11	25.2-25.6	25.4	11	25.3-25.5
N	Perrier Water (2A)	25.5	3	25.4-25.6	25.3	3	25.2-25.4
N	EPA MHW (1)	25.4	3	25.3-25.6	25.4	3	25.2-25.5

		(°C) Before			(°C) After		
Lab	Sample Type	Mean	N	Range	Mean	N	Range
O	Lab Water	25.6	12	25.4-25.7	25.6	12	25.4-25.9
O	Perrier Water (2A)	NA	NA	NA	NA	NA	NA
O	EPA MHW (1)	NA	NA	NA	NA	NA	NA

NA: Not available

Table B20 cont. Summary of air temperature data collected in control chambers from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions)

		(°C) Before			(°C) After		
Lab	Sample Type	Mean	N	Range	Mean	N	Range
P	Lab Water	25.4	12	24.9-25.6	25.3	12	25.1-25.5
P	Perrier Water (2A)	25.4	3	25.2-25.5	25.4	3	25.2-25.5
P	EPA MHW (1)	25.4	3	25.1-25.6	25.3	3	25.1-25.5
Q	Lab Water	25.9	12	25.9-26.1	26.0	12	25.8-26.1
Q	Perrier Water (2A)	25.9	3	25.9-26.0	26.0	3	25.9-26.0
Q	EPA MHW (1)	25.9	3	25.9-26.0	26.0	3	25.9-26.0

Table B21. Summary of hardness and alkalinity data collected in control chambers from the eleven laboratories participating in the baseline C. dubia interlaboratory study. These water quality parameters were only measured after renewal of test solutions. N values refer to the number of tests conducted and included in the means.

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)		
Lab	Sample Type	Mean	N	Range	Mean	N	Range
A	Lab Water	82.7	3	82-84	64.0	3	64-64
A	Perrier Water (2A)	86.3	3	85-89	84.3	3	82-86
A	EPA MHW (1)	119	3	110-127	66.3	3	60-73
B	Lab Water	93.0	2	90-96	65.0	2	64-66
B	Perrier Water (2A)	82	2	74-90	76.0	2	74-78
B	EPA MHW (1)	121	2	118-124	57.0	2	52-62
E	Lab Water	95.5	2	95-96	61.0	2	60-62
E	Perrier Water (2A)	80.7	3	75-84	67.0	3	60-71
E	EPA MHW (1)	109	3	107-109	61.3	3	60-62
F	Lab Water	89.3	3	89-90	86.7	3	82-90
F	Perrier Water (2A)	85.7	3	84-87	76.7	3	74-79
F	EPA MHW (1)	120	3	116-126	58.7	3	49-70
G	Lab Water	81.7	3	80-84	78.8	4	76-81
G	Perrier Water (2A)	88.0	3	87-90	75.3	3	72-80
G	EPA MHW (1)	130	3	115-144	63.7	3	60-68
L	Lab Water	88.7	6	70-132	67.0	6	50-100
L	Perrier Water (2A)	100	3	80-144	80.3	3	62-108
L	EPA MHW (1)	104	3	82-138	65.3	3	54-84
M	Lab Water	100	2	100-100	67.0	3	66-68
M	Perrier Water (2A)	89	2	88-90	80.0	2	80-80
M	EPA MHW (1)	121	2	120-122	59.0	2	58-60
N	Lab Water	87.5	3	84-90	57.3	3	57-59
N	Perrier Water (2A)	86.3	3	81-90	74.3	3	73-75

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)		
Lab	Sample Type	Mean	N	Range	Mean	N	Range
N	EPA MHW (1)	122	3	118-125	56.0	3	55-57
O	Lab Water	88.8	12	54-98	66.5	3	61-83
O	Perrier Water (2A)	90.8	6	79-120	76.3	6	63-85
O	EPA MHW (1)	116	4	114-117	59.0	4	57-60

Table B21 cont. Summary of hardness and alkalinity data collected in control chambers from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study. These water quality parameters were only measured after renewal of test solutions.

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)		
Lab	Sample Type	Mean	N	Range	Mean	N	Range
P	Lab Water	94.3	12	92-97	91.7	12	90-93
P	Perrier Water (2A)	81.7	3	79-84	79.3	3	83-76
P	EPA MHW (1)	115	3	112-117	61.7	3	59-64
Q	Lab Water	89.8	12	86-97	86.2	12	83-89
Q	Perrier Water (2A)	85.3	3	81-89	80.7	3	77-83
Q	EPA MHW (1)	116	3	114-118	60.0	3	58-62

Table B22. Summary of water quality data collected from the brood boards used to initiate the tests during the baseline C. dubia interlaboratory study. The data is divided into two categories: 'before' defined as water in test chambers for 24 hours, and 'after' defined as water quality measurements recorded after renewal in the test chambers. N values refer to the number of water quality measurements of the brood boards reported by the laboratories.

		Hardness ($mg/L\ CaCO_3$)			Alkalinity ($mg/L\ CaCO_3$)			Conductivity ($\mu S/cm$)			pH		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
A	Before	NM	-	-	NM	-	-	312	5	310-313	7.6	5	7.5-7.7
A	After	85	5	82-87	61	5	59-64	304	5	302-306	7.7	5	7.6-7.8
B	Before	NM	-	-	NM	-	-	350	2	341-358	7.3	2	7.2-7.5
B	After	92	1	-	64	1	-	355	2	343-367	7.2	2	7.0-7.4
E	Before	NM	-	-	NM	-	-	370	3	363-376	8.0	3	7.8-8.1
E	After	95	3	95-95	60	3	60-60	369	3	363-374	8.0	3	7.8-8.1
F	Before	NM	-	-	NM	-	-	209	9	206-212	8.2	9	8.2-8.2
F	After	88	9	84-91	86	9	76-92	190	9	189-191	8.1	9	8.1-8.2
G	Before	NM	-	-	NM	-	-	188	6	183-191	8.2	6	8.1-8.2
G	After	80	6	79-81	76	6	75-78	172	6	171-173	8.0	6	8.0-8.1
L	Before	NM	-	-	NM	-	-	348	10	340-358	8.1	10	8.1-8.2

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)			Conductivity (µS/cm)			pH		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range	Mean	N	Range
L	After	82.8	5	71-90	60.0	5	58-64	341	7	330-346	8.1	7	8.1-8.2
M	Before	NM	-	-	NM	-	-	391	3	386-397	8.3	3	8.1-8.4
M	After	99.7	2	99-100	66.3	2	65-67	344	3	342-346	8.2	3	8.1-8.2
N	Before	NM	-	-	NM	-	-	260	7	251-306	8.0	7	8.0-8.2
N	After	85.4	7	83-88	57.3	7	57-59	254	7	243-273	8.1	7	8.0-8.1
O	Before	NM	-	-	NM	-	-	339	4	335-347	7.7	4	7.7-7.9
O	After	91.0	4	88-92	61.8	4	61-64	327	4	322-336	7.8	4	7.8-7.9
P	Before	NM	-	-	NM	-	-	218	3	216-219	8.0	3	7.9-8.0
P	After	92.8	5	91-94	91.4	5	90-94	210	5	202-216	7.7	5	7.6-7.8
Q	Before	NM	-	-	NM	-	-	174	3	168-181	8.0	3	8.0-8.0
Q	After	89.3	3	86-93	85.9	3	84-89	201	3	197-207	8.1	3	8.1-8.1

Table B22 cont. Summary of water quality data collected from the brood boards used to initiate the tests during the baseline *C. dubia* interlaboratory study. The data is divided into two categories: 'before' defined as water in test chambers for 24 hours, and 'after' defined as water quality measurements recorded after renewal in the test chambers. N values refer to the number of water quality measurements of the brood boards reported by the laboratories.

		Dissolved Oxygen (mg/L)			Water Temperature (°C)			Air Temperature (°C)		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range
A	Before	6.33	5	6.16-6.70	24.6	5	24.2-25.0	25.4	5	24.8-26.2
A	After	7.21	5	7.04-7.66	24.8	5	24.2-24.5	25.7	5	25.2-26.4
B	Before	8.76	2	8.71-8.80	24.8	2	24.8-24.9	24.8	2	24.7-24.9
B	After	9.49	2	8.68-10.3	24.9	2	24.8-25.0	24.8	2	24.7-24.9
E	Before	7.70	3	7.42-8.08	24.9	3	24.9-25.0	25.1	3	25.0-25.1
E	After	7.76	3	7.47-8.14	25.0	3	24.9-25.0	25.1	3	25.0-25.1
F	Before	7.99	9	7.82-8.18	25.0	9	24.8-25.1	26.4	9	26.2-26.6
F	After	8.00	9	7.82-8.62	24.9	9	24.8-25.2	26.5	9	26.4-26.6
G	Before	8.30	6	8.20-8.50	25.2	6	25.1-25.6	25.4	6	25.0-25.7
G	After	8.49	6	8.44-8.54	25.0	6	24.7-25.3	25.8	4	25.5-26.0

		Dissolved Oxygen (mg/L)			Water Temperature (°C)			Air Temperature (°C)		
Lab	Sample Type	Mean	N	Range	Mean	N	Range	Mean	N	Range
L	Before	8.70	10	8.58-8.81	24.4	10	24.1-24.8	23.4	10	22.0-25.2
L	After	8.92	7	8.81-9.17	24.1	7	23.6-24.6	23.7	7	22.1-25.7
M	Before	6.52	3	6.41-6.73	24.2	3	23.8-24.8	24.2	3	23.7-24.8
M	After	6.95	3	6.89-7.06	24.4	3	23.7-24.9	24.1	3	23.6-24.6
N	Before	7.93	7	7.78-8.22	24.9	6	24.7-25.2	25.0	7	24.8-25.3
N	After	8.08	7	8.02-8.22	24.9	6	24.8-24.9	25.0	7	24.8-25.3
O	Before	7.77	4	7.45-8.07	24.9	4	24.8-24.9	25.6	3	25.3-26.1
O	After	7.74	4	7.53-8.04	24.7	2	24.5-24.9	25.7	2	25.5-25.9
P	Before	8.26	3	8.26-8.26	25.2	3	25.2-25.3	25.7	3	25.5-26.1
P	After	8.49	5	8.27-8.84	25.2	5	25.1-25.2	25.5	3	25.4-25.8
Q	Before	7.98	3	7.97-7.99	24.6	3	23.6-25.3	25.8	3	25.8-25.9
Q	After	8.40	3	8.36-8.43	22.9	3	22.1-23.2	25.8	3	25.7-25.9

Table B23. Water chemistry data for lab dilution waters and split samples (samples 1 and 2A) used in Round 1.

Lab	Sample	Hardness (mg/L CaCO3)	Total Alk. (mg/L CaCO3)	Chloride (mg/L)	Nitrate (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)	Se (ug/L)	Ca (mg/L)	Mg (mg/L)	K (mg/L)	Na (mg/L)
A	EPA MHW (1)	87	55	1.96	0.498	111	55	0.0243	33.8	0.71	6.6	68.6
E	EPA MHW (1)	93	56	1.92	0.498	109	56	ND	36.1	0.75	6.34	66.5
F	EPA MHW (1)	89	55	1.96	0.501	120	55	0.031	34.3	0.71	6.12	67.8
G	EPA MHW (1)	91	55	1.91	0.501	111	55	0.023	35.4	0.73	ND	26.1
L	EPA MHW (1)	87	55	1.95	0.494	108	55	0.026	33.6	0.71	ND	33
N	EPA MHW (1)	69	55	1.9	0.499	113	55	0.0302	26.5	0.80	ND	25.3
O	EPA MHW (1)	72	55	1.97	0.498	113	55	0.0345	27.4	0.83	ND	25.6
P	EPA MHW (1)	72	56	1.97	0.497	112	56	0.0355	27.5	0.76	ND	25.6
Q	EPA MHW (1)	90	54	1.97	0.497	108	54	ND	34.7	0.70	6.48	67.2
A	Lab Water	103	55	1.95	0.493	74.7	55	0.027	16.3	15.1	ND	29.6
E	Lab Water	112	60	4.21	ND	83.4	60	ND	17.9	16.4	5.59	33
F	Lab Water	98	83	4.27	0.763	4.52	83	0.13	37.8	0.81	ND	ND
G	Lab Water	86	74	3.88	0.728	4.12	74	0.092	33.4	0.72	ND	ND
L	Lab Water	148	55	1.88	ND	111	55	1.67	33.2	15.7	ND	30.5
N	Lab Water	103	56	1.95	0.495	74.2	56	3.06	16	15.2	ND	30.6
O	Lab Water	116	56	4.52	ND	84.1	56	1.46	18.4	17.0	ND	33.3

Lab	Sample	Hardness (mg/L CaCO3)	Total Alk. (mg/L CaCO3)	Chloride (mg/L)	Nitrate (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)	Se (ug/L)	Ca (mg/L)	Mg (mg/L)	K (mg/L)	Na (mg/L)
P	Lab Water	91	85	4.66	0.715	4.68	85	0.14	35.3	0.80	ND	ND
Q	Lab Water	91	80	4.18	0.756	4.48	80	0.133	35.2	0.73	ND	6.31
A	Perrier (2A)	90	73	3.94	0.747	4.14	73	0.116	34.8	0.72	ND	5.39
E	Perrier (2A)	121	74	3.92	0.746	4.22	74	0.112	28.4	12.2	ND	5.43
F	Perrier (2A)	90	73	3.9	0.75	4.14	73	0.114	34.9	0.70	ND	5.66
G	Perrier (2A)	82	73	3.89	0.75	4.1	73	0.109	31.5	0.80	ND	ND
L	Perrier (2A)	87	74	3.91	0.75	4.11	74	0.149	33.6	0.72	ND	ND
N	Perrier (2A)	82	71	3.91	0.75	4.09	71	0.126	31.5	0.81	ND	ND
O	Perrier (2A)	82	72	3.93	0.75	4.11	72	0.165	31.4	0.78	ND	ND
P	Perrier (2A)	83	72	3.93	0.752	4.12	72	0.164	31.9	0.79	ND	ND
Q	Perrier (2A)	75	73	3.93	0.745	4.14	73	0.116	28.9	0.69	ND	5.58

ND= Not detected.

Table B24. Water chemistry data for lab dilution waters and split samples (samples 1 and 2A) used in Round 2.

Lab	Sample	Hardness (mg/L CaCO3)	Total Alk. (mg/L CaCO3)	Chloride (mg/L)	Nitrate (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)	Se (ug/L)	Ca (mg/L)	Mg (mg/L)	K (mg/L)	Na (mg/L)
A	EPA MHW (1)	135	56	1.88	ND	111	56	0.0421	31.4	13.8	ND	28.7
B	EPA MHW (1)	128	57	1.89	ND	111	57	ND	30.0	13	ND	27.8
E	EPA MHW (1)	145	55	1.88	ND	112	55	0.025	32.5	15.4	ND	30.1
F	EPA MHW (1)	136	58	1.88	ND	112	58	ND	31.6	13.8	ND	28.6
G	EPA MHW (1)	126	56	1.85	ND	111	56	0.0867	29.1	12.9	ND	27.5
L	EPA MHW (1)	129	57	1.89	ND	112	57	ND	30.0	13.2	ND	27.8
M	EPA MHW (1)	134	56	1.88	ND	113	56	0.022	31.1	13.7	ND	28.7
N	EPA MHW (1)	148	56	1.83	ND	112	56	ND	33.4	15.6	ND	30.5
O	EPA MHW (1)	128	57	1.89	ND	112	57	ND	30.0	12.8	ND	27.7
P	EPA MHW (1)	130	56	1.90	ND	113	56	0.030	30.3	13.1	ND	27.9
Q	EPA MHW (1)	134	57	1.87	ND	113	57	0.037	31.4	13.6	ND	28.8
A	Lab Water	88.2	59	3.11	ND	76.2	59	0.0373	14.7	12.5	ND	27.3
B	Lab Water	15.4	10	0.595	ND	12.8	10	ND	2.78	2.05	ND	ND
E	Lab Water	97.0	62	4.19	ND	84.7	62	ND	16.0	13.9	ND	30.5
F	Lab Water	97.8	83	4.29	0.76	4.59	83	0.117	37.8	0.834	ND	ND
G	Lab Water	84.6	74	3.84	0.729	4.08	74	0.118	32.7	0.700	ND	ND
L	Lab Water	127	57	1.87	ND	111	57	1.92	29.7	12.8	ND	27
M	Lab Water	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
N	Lab Water	89.9	58	1.94	ND	78.4	58	ND	15.1	12.7	ND	27.0
O	Lab Water	97.0	60	3.02	ND	83.2	60	1.46	16.4	13.6	ND	29.5
P	Lab Water	104	87	4.72	0.733	4.93	87	0.057	40.1	0.851	ND	ND
Q	Lab Water	92	79	4.06	0.743	4.33	79	0.150	35.6	0.768	ND	ND
A	Perrier (2A)	97.7	77	3.99	0.737	4.32	77	0.0968	37.8	0.835	ND	ND

Lab	Sample	Hardness (mg/L CaCO3)	Total Alk. (mg/L CaCO3)	Chloride (mg/L)	Nitrate (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)	Se (ug/L)	Ca (mg/L)	Mg (mg/L)	K (mg/L)	Na (mg/L)
B	Perrier (2A)	90.5	77	4.02	0.736	4.34	77	0.103	35.0	0.769	ND	ND
E	Perrier (2A)	103	77	3.98	0.737	4.31	77	0.122	39.8	0.988	ND	ND

Table B24. Water chemistry from Round 2 continued.

Lab	Sample	Hardness (mg/L CaCO3)	Total Alk. (mg/L CaCO3)	Chloride (mg/L)	Nitrate (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)	Se (ug/L)	Ca (mg/L)	Mg (mg/L)	K (mg/L)	Na (mg/L)
F	Perrier (2A)	96.9	76	4.05	0.738	4.42	76	0.141	37.4	0.856	ND	ND
G	Perrier (2A)	89.1	77	4.01	0.735	4.35	77	0.165	34.5	0.749	ND	ND
L	Perrier (2A)	92.8	78	4.01	0.736	4.36	78	0.122	35.9	0.773	ND	ND
M	Perrier (2A)	95.3	76	3.97	0.735	4.30	76	0.137	36.8	0.812	ND	ND
N	Perrier (2A)	102	78	3.98	0.737	4.31	78	0.106	39.2	0.944	ND	ND
P	Perrier (2A)	90.5	77	4.01	0.736	4.33	77	0.123	35.0	0.761	ND	ND
Q	Perrier (2A)	95.9	77	3.97	0.734	4.32	77	0.168	37.1	0.820	ND	ND
O	Perrier (2A)	92.8	78	4.03	0.736	4.35	78	0.145	35.8	0.794	ND	ND

NA= Not analyzed due to labeling error.

ND= Not detected.

Table B25. Water chemistry data for lab dilution waters and split samples (samples 1 and 2A) used in Round 3.

Lab	Sample	Hardness (mg/L CaCO3)	Total Alk. (mg/L CaCO3)	Chloride (mg/L)	Nitrate (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)	Se (ug/L)	Ca (mg/L)	Mg (mg/L)	K (mg/L)	Na (mg/L)
A	Lab Water	84.0	56	3.03	ND	75.7	56	ND	14.7	11.5	ND	26.1
A	Perrier (2A)	93.1	77	3.92	0.740	4.16	77	0.0985	35.9	0.861	ND	ND
A	EPA MHW (1)	122	56	2.12	0.499	104	56	ND	29.2	12.0	ND	26.3
B	Lab Water	111	65	2.17	ND	98.3	65	ND	0.175	26.8	ND	29.8
B	Perrier (2A)	92.5	77	3.96	0.738	4.17	77	0.154	35.5	0.923	ND	ND
B	EPA MHW (1)	122	55	2.14	ND	104	55	0.0643	29.4	11.8	ND	26.1
E	Lab Water	95.3	61	4.00	ND	84.6	61	ND	16.8	13.0	ND	29.2
E	Perrier (2A)	91.2	75	3.88	0.740	4.16	75	0.128	35.1	0.829	ND	ND
E	EPA MHW (1)	121	54	2.07	ND	109	54	0.030	29.0	11.9	ND	25.7
F	Lab Water	97.8	84	4.11	0.756	4.42	84	0.0975	37.7	0.869	ND	ND
F	Perrier (2A)	89.5	77	3.86	0.737	4.12	77	0.130	34.6	0.764	ND	ND
F	EPA MHW (1)	122	55	2.09	ND	106	55	ND	29.0	12.0	ND	26.7
G	Lab Water	90.1	74	3.82	0.728	4.07	74	0.133	34.7	0.859	ND	ND
G	Perrier (2A)	92.8	76	3.97	0.741	4.18	76	0.124	35.8	0.847	ND	ND
G	EPA MHW (1)	124	55	2.15	ND	105	55	0.0638	29.9	11.9	ND	25.7
L	Lab Water	124	57	2.06	ND	105	57	1.92	30.0	12.0	ND	26.3
L	Perrier (2A)	92.6	76	3.79	0.736	4.09	76	0.128	35.7	0.855	ND	ND
L	EPA MHW (1)	102	55	2.08	0.500	103	55	0.0699	24.2	10.1	ND	22.5

Lab	Sample	Hardness (mg/L CaCO3)	Total Alk. (mg/L CaCO3)	Chloride (mg/L)	Nitrate (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)	Se (ug/L)	Ca (mg/L)	Mg (mg/L)	K (mg/L)	Na (mg/L)
M	Lab Water	102	66	1.94	ND	91.3	66	1.95	18.4	13.7	ND	29.6
M	Perrier (2A)	93.2	77	3.82	0.737	4.08	77	0.164	35.9	0.885	ND	ND
M	EPA MHW (1)	120	56	2.08	ND	105	56	0.0511	28.3	12.0	ND	26.6
N	Lab Water	92.0	58	2.28	ND	79.9	58	3.15	16.5	12.3	ND	26.5
N	Perrier (2A)	93.8	76	3.92	0.741	4.16	76	0.0993	36.1	0.923	ND	ND
N	EPA MHW (1)	124	55	2.14	ND	105	55	ND	29.6	12.1	ND	26.3
O	Lab Water	97.1	57	5.76	ND	87.9	57	1.46	17.1	13.2	ND	29.7
O	Perrier (2A)	90.8	75	3.82	0.735	4.09	75	0.112	35.1	0.781	ND	ND
O	EPA MHW (1)	120	55	2.05	ND	104	55	0.0224	28.6	11.8	ND	25.8

Table B25. Water chemistry from Round 3 continued

Lab	Sample	Hardness (mg/L CaCO3)	Total Alk. (mg/L CaCO3)	Chloride (mg/L)	Nitrate (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)	Se (ug/L)	Ca (mg/L)	Mg (mg/L)	K (mg/L)	Na (mg/L)
P	Lab Water	106	86	4.74	0.739	4.85	86	0.168	40.7	0.922	ND	ND
P	Perrier (2A)	75.4	76	3.84	0.736	4.11	76	0.158	29.2	0.628	ND	ND
P	EPA MHW (1)	102	55	2.09	ND	105	55	0.062	24.2	10.1	ND	22.0
Q	Lab Water	96.9	81	4.14	0.773	4.09	81	0.121	37.4	0.838	ND	ND
Q	Perrier (2A)	91	76	3.82	0.736	4.08	76	0.161	35.1	0.801	ND	ND
Q	EPA MHW (1)	122	54	2.08	ND	105	54	0.089	29.0	12.0	ND	25.6

ND= Not detected.

Table B26. Neonate production and adult female survival data for Sample 1 (EPA Moderately Hard Water) collected from the eleven laboratories participating in the baseline C. dubia interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
A	1	35.2	0.160	10
A	2	44.4	0.058	10
A	3	36.1	0.271	9
B	2	34.4	0.171	10
B	3	35.6	0.466	8
E	1	20.5	0.339	10
E	2	13.4	0.183	10
E	3	11.3	0.277	10
F	1	17.6	0.208	10
F	2	20.8	0.142	10
F	3	20	0.130	10
G	1	27.2	0.306	10
G	2	32.3	0.062	10
G	3	35.2	0.082	10
L	1	21.8	0.410	7
L	2	25	0.386	9
L	3	26.1	0.267	9
M	2	26.9	0.574	9
M	3	37.9	0.314	9
N	1	6.9	1.26	4
N	2	3.0	1.90	3
N	3	11.3	0.515	8
O	1	33.1	0.100	10
O	2	35.4	0.075	10
O	3	24.4	0.417	10

Table B26 continued. Neonate production and adult female survival data for Sample 1 (EPA Moderately Hard Water) collected from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
P	1	35.8	0.261	10
P	2	40.6	0.061	10
P	3	39.2	0.118	10
Q	1	40.6	0.052	10
Q	2	37.5	0.075	10
Q	3	30.4	0.339	9

Table B27. Neonate production and survival data for Sample 2A (Perrier based Moderately Hard Water) collected from the eleven laboratories participating in the baseline C. dubia interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
A	1	34.0	0.195	10
A	2	43.4	0.086	10
A	3	39.7	0.108	10
B	2	21.2	0.521	10
B	3	31.1	0.522	9
E	1	17.0	0.331	10
E	2	14.3	0.556	10
E	3	10.7	0.399	10
F	1	22.3	0.075	10
F	2	20.8	0.091	10
F	3	20.4	0.101	10

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
G	1	31.8	0.052	10
G	2	32.9	0.196	10
G	3	34.6	0.084	10
L	1	30.2	0.285	9
L	2	24.7	0.436	9
L	3	22.0	0.588	7

Table B27 continued. Neonate production and survival data for Sample 2A (Perrier based Moderately Hard Water) collected from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
M	2	37.1	0.207	10
M	3	35.4	0.390	9
N	1	3.5	1.21	5
N	2	7.5	1.77	3
N	3	16.7	0.517	8
O	1	30.1	0.147	10
O	2	36.0	0.053	10
O	3	23.7	0.330	10
P	1	36.5	0.289	9
P	2	34.6	0.274	9
P	3	38.3	0.078	10
Q	1	39.6	0.102	10
Q	2	35.3	0.088	10
Q	3	35.0	0.169	10

Table B28. Neonate production and survival data for laboratory controls collected from the eleven laboratories participating in the baseline C. dubia interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
A	1	37.1	0.117	10
A	1	35.1	0.279	9
A	1	38.3	0.129	10
A	1	37.5	0.137	10
A	2	44.1	0.067	10
A	2	43.2	0.060	10
A	2	40.7	0.103	10
A	2	43.5	0.084	10
A	3	40.8	0.054	10
A	3	41.5	0.086	10
A	3	42.1	0.059	10
A	3	38.0	0.215	10
B	2	35.9	0.108	10
B	2	31.9	0.135	9
B	2	30.8	0.173	9
B	2	20.6	0.625	9
B	3	0	0	0
B	3	0	0	0
B	3	0	0	0
B	3	0	0	0

Table B28 continued. Neonate production and survival data for laboratory controls collected from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
E	1	17.3	0.289	10
E	1	21.7	0.154	10
E	1	17.0	0.277	10
E	1	20.5	0.240	10
E	2	12.3	0.325	10
E	2	18.7	0.356	10
E	2	12.7	0.236	10
E	2	20.0	0.351	10
E	3	8.9	0.238	10
E	3	10.4	0.276	10
E	3	9.3	0.360	10
E	3	8.5	0.264	10
F	1	16.4	0.191	10
F	1	17.1	0.061	10
F	1	17.2	0.135	10
F	1	16.7	0.170	10
F	2	19.1	0.166	10
F	2	18.8	0.100	9
F	2	19.9	0.128	10
F	2	19.0	0.135	10
F	3	17.8	0.075	10
F	3	16.2	0.110	10
F	3	16.6	0.158	10
F	3	16.5	0.087	10

Table B28 continued. Neonate production and survival data for laboratory controls collected from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
G	1	31.3	0.054	10
G	1	30.0	0.146	9
G	1	29.8	0.063	10
G	1	28.6	0.095	10
G	2	28.7	0.309	9
G	2	31.8	0.213	9
G	2	32.5	0.054	10
G	2	29.0	0.230	10
G	3	32.8	0.124	10
G	3	34.5	0.045	10
G	3	31.1	0.187	9
G	3	29.6	0.463	7
L	1	28.8	0.143	10
L	1	31.6	0.155	9
L	1	30.3	0.301	10
L	2	32.8	0.272	10
L	2	27.7	0.319	9
L	2	31.9	0.169	10
L	3	24.9	0.387	9
L	3	21.9	0.387	8
L	3	29.2	0.115	10
M	2	25.0	0.393	10
M	2	26.1	0.483	9
M	2	21.5	0.632	8
M	2	7.6	1.01	7
M	3	33.9	0.258	9
M	3	29.6	0.521	8

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
M	3	30.0	0.438	9
M	3	33.9	0.209	10

Table B28 continued. Neonate production and survival data for laboratory controls collected from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
N	1	15.5	0.601	8
N	1	12.7	0.575	8
N	1	11.9	0.531	8
N	2	21.4	0.462	8
N	2	23.9	0.557	9
N	2	25.0	0.415	8
N	2	31.8	0.248	9
N	3	20.8	0.280	9
N	3	16.2	0.476	8
N	3	19.9	0.292	10
N	3	16.9	0.340	8
O	1	31.1	0.225	10
O	1	31.3	0.199	10
O	1	33.3	0.167	10
O	1	30.2	0.194	10
O	2	31.4	0.226	10
O	2	38.2	0.073	10
O	2	36.6	0.087	10
O	2	31.9	0.194	9
O	3	26.3	0.32	9
O	3	28.7	0.257	10
O	3	28.6	0.233	10

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
O	3	29.6	0.195	10

Table B28 continued. Neonate production and survival data for laboratory controls collected from the eleven laboratories participating in the baseline *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
P	1	37.5	0.133	10
P	1	37.2	0.203	8
P	1	38.3	0.296	9
P	1	33.3	0.341	9
P	2	37.1	0.143	10
P	2	36.9	0.219	10
P	2	32.8	0.297	9
P	2	38.3	0.165	10
P	3	38.1	0.163	10
P	3	37.8	0.100	10
P	3	34.8	0.196	10
P	3	36.7	0.088	10
Q	1	41.1	0.058	10
Q	1	39.0	0.103	10
Q	1	42.2	0.055	10
Q	1	41.0	0.076	10
Q	2	33.8	0.094	10
Q	2	36.0	0.053	10
Q	2	34.3	0.064	10
Q	2	35.8	0.077	10
Q	3	30.2	0.300	9
Q	3	31.3	0.136	10
Q	3	28.7	0.372	9
Q	3	33.1	0.096	10

Figure B33. Plot of the age of the females whose neonates were used to initiate a test batch versus the IC50 value expressed as measured conductivity for the Sample 2 series (Perrier based Moderately Hard Water). Symbols represent the laboratory code.

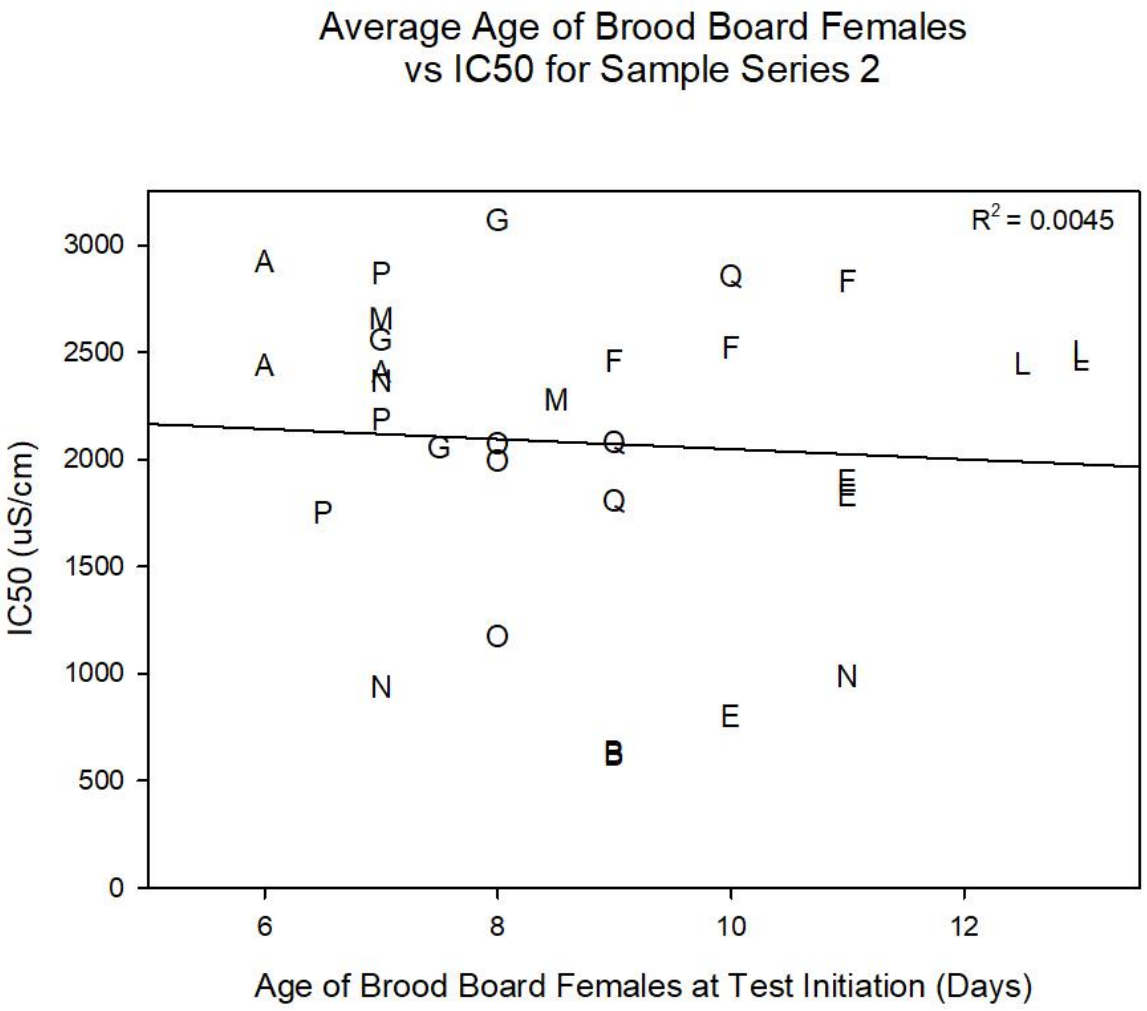


Figure B34. Plot of the age of the females whose neonates were used to initiate a test batch versus the IC50 value expressed as measured conductivity for the Sample 3 series (lab's own dilution water). Symbols represent the laboratory code.

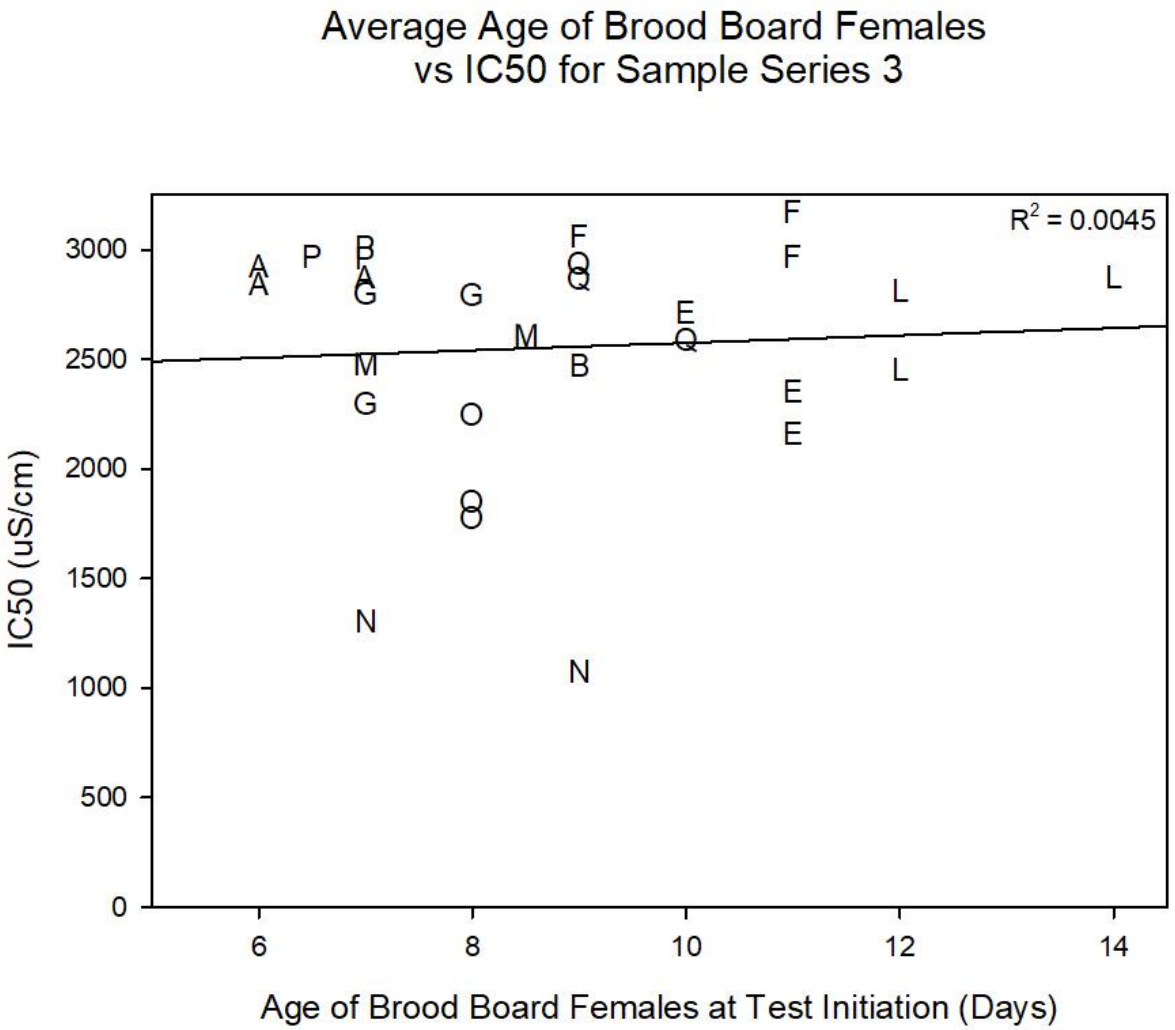


Figure B35. Plot of the age of the females whose neonates were used to start the test versus mean neonate production for labs that culture in EPA Moderately Hard Water for Sample 1 (EPA Moderately Hard Water. Symbols represent the laboratory code.

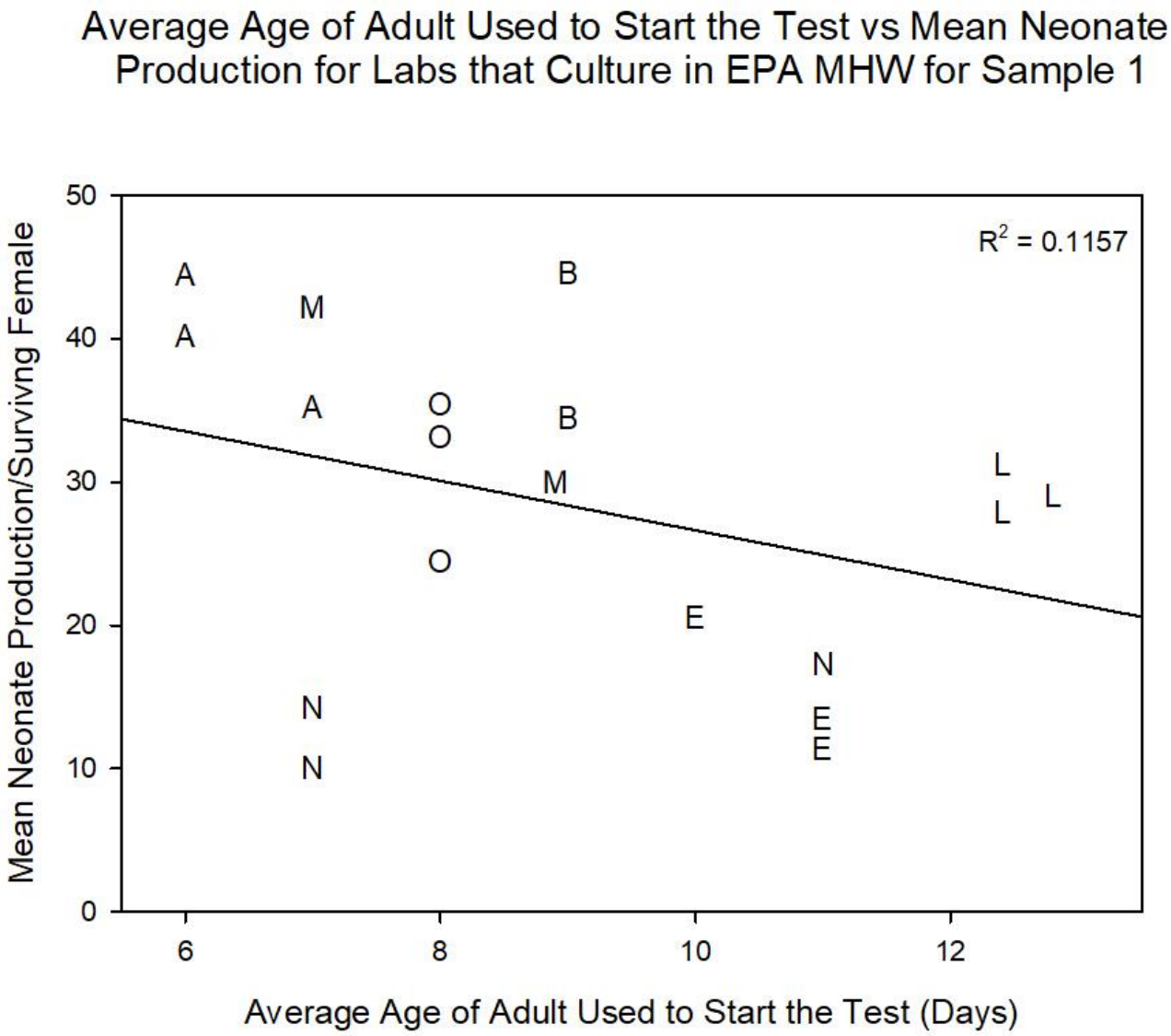


Figure B36. Plot of the age of the females whose neonates were used to start the test versus mean neonate production for labs that culture in Perrier based Moderately Hard Water vs. neonate production for Sample 2A (Perrier based Moderately Hard Water). Symbols represent the laboratory code.

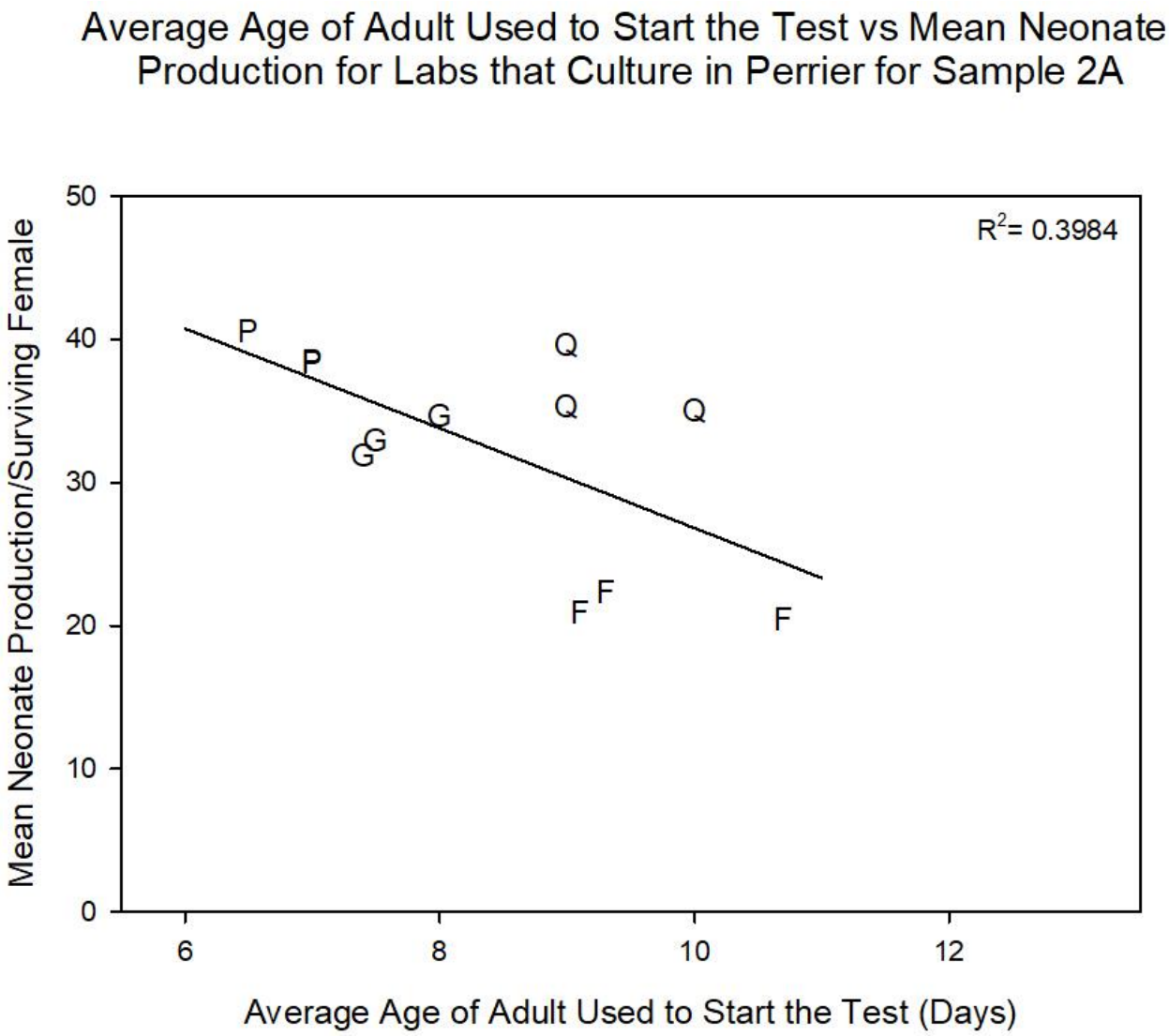
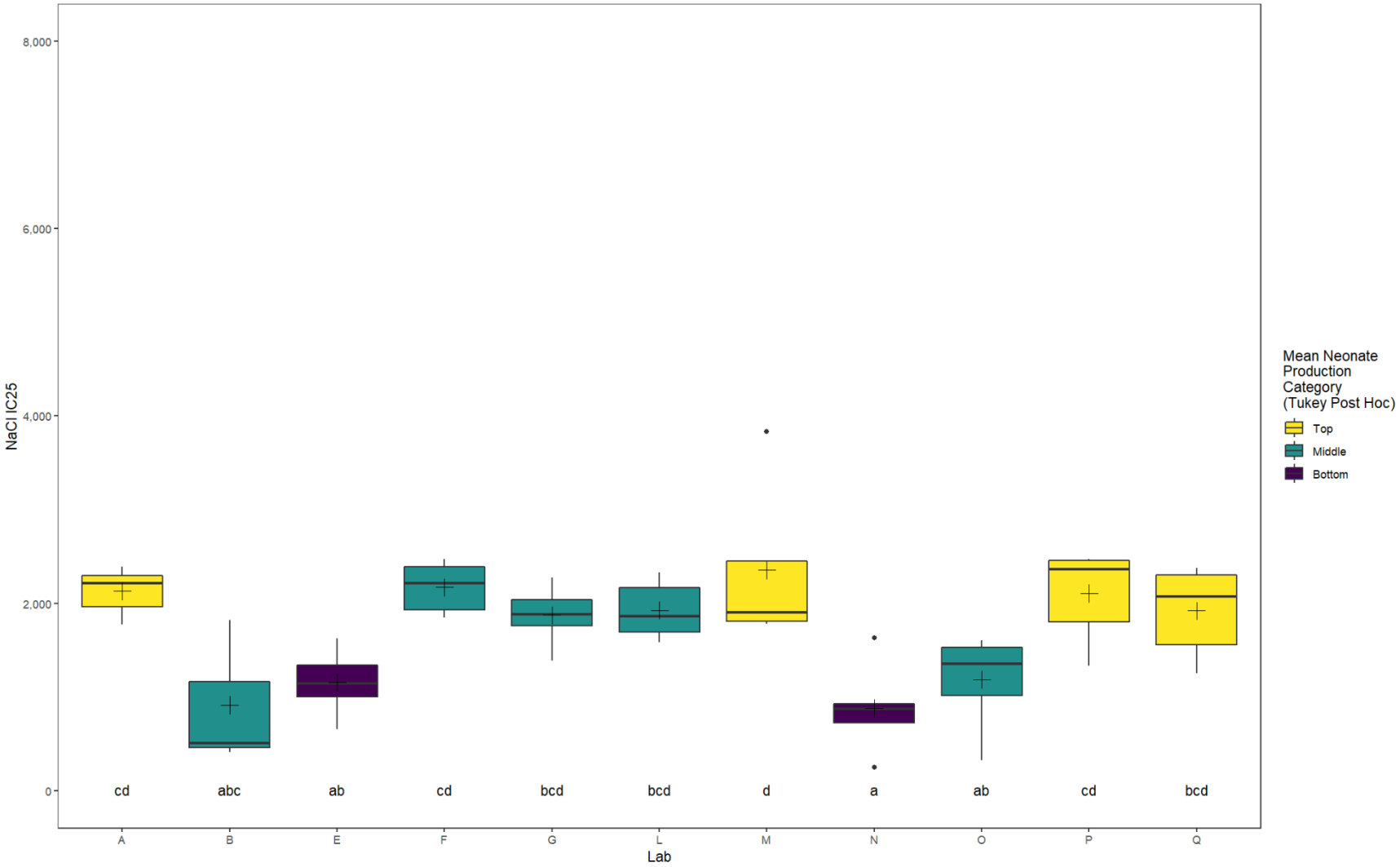
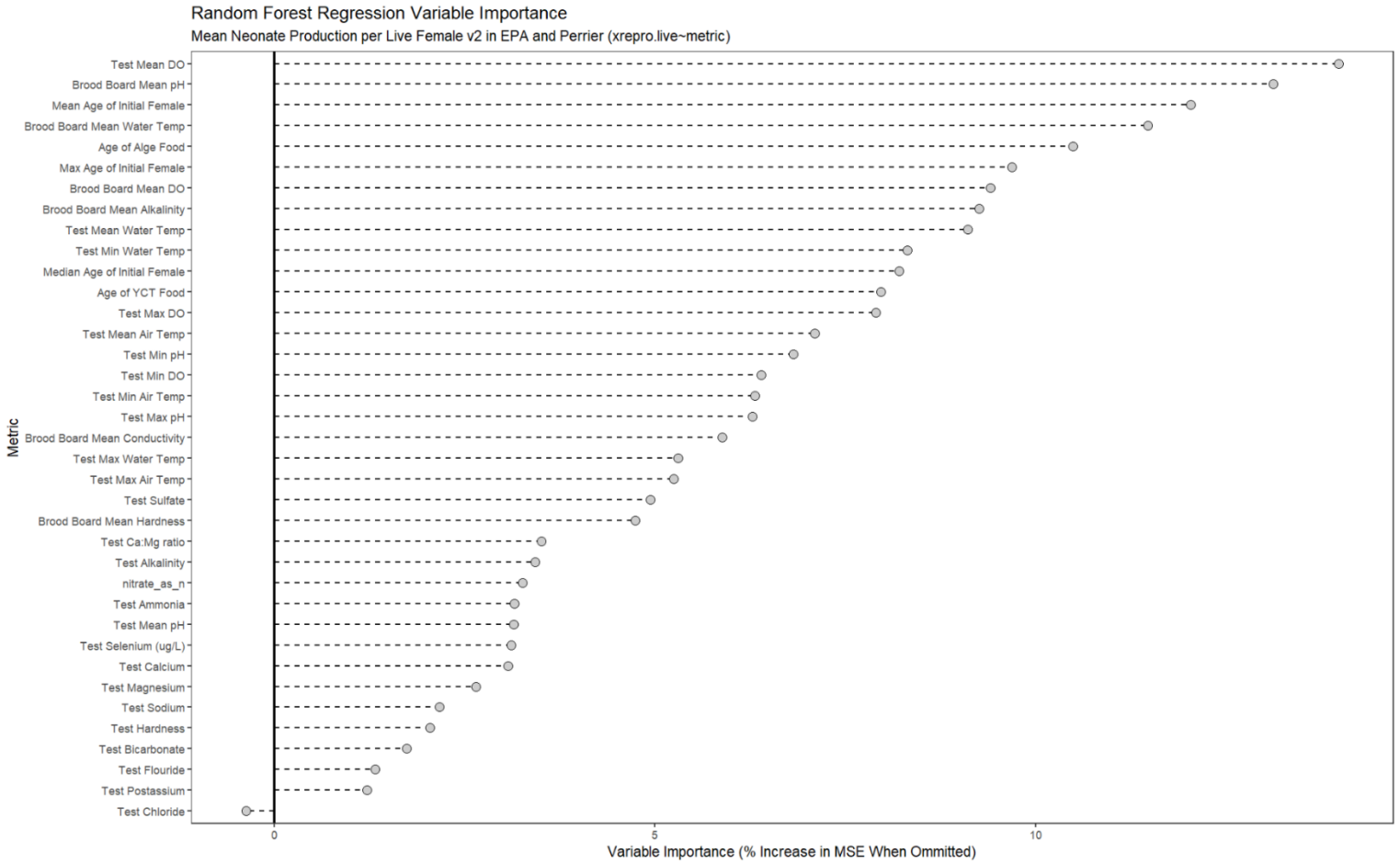


Figure B37. Box plots of IC25 by laboratory for nominal NaCl concentrations during the baseline ILS using combined SCCWRP-supplied and laboratory-supplied dose response samples in rounds 1 through 3. Letters at 0 on the y-axis denotes post-hoc Tukey tests for significant differences between laboratories. Lab N reported culture issues prior to Round 1 and Lab L prior to Round 2.



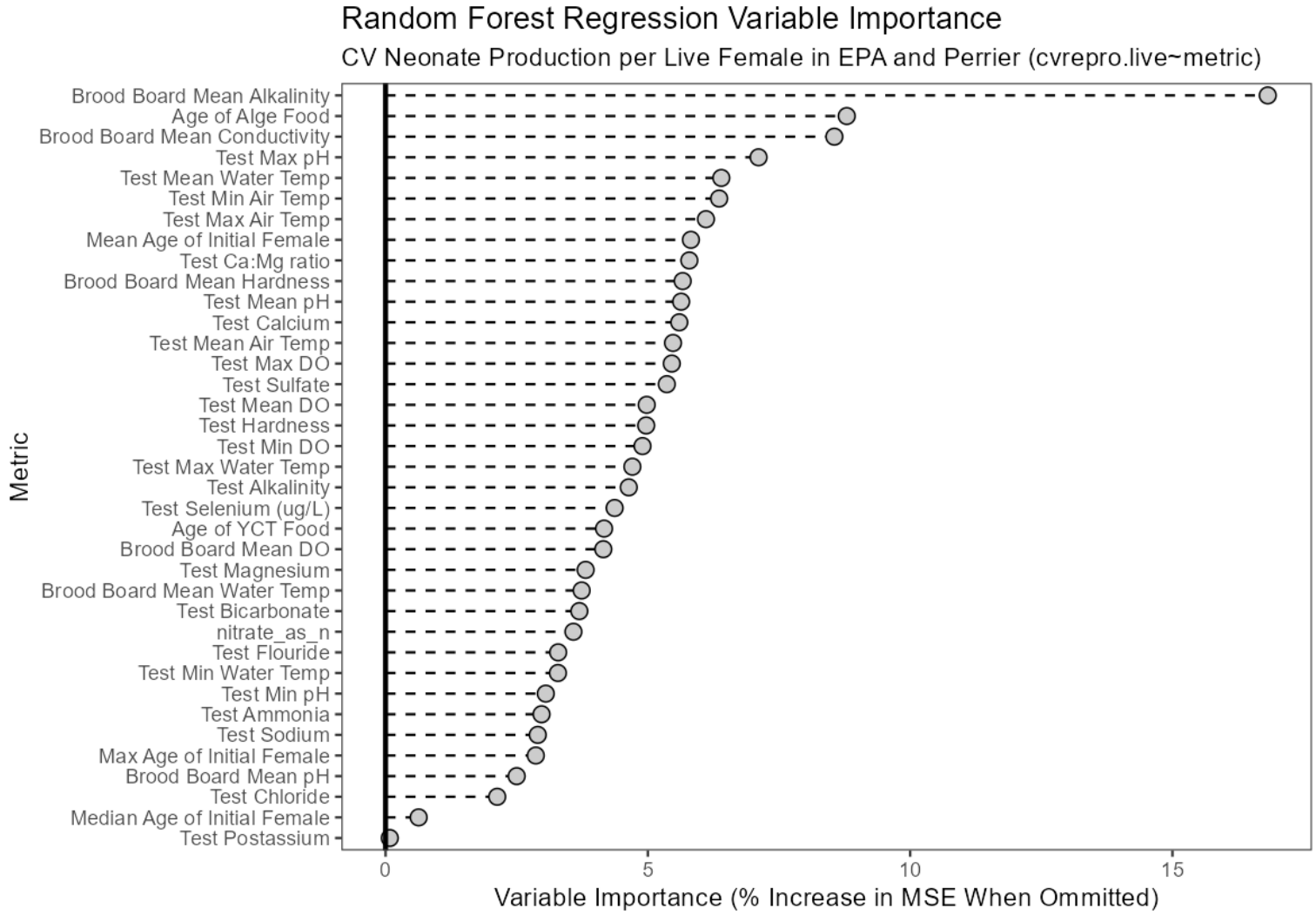
Potential sources of variance in test performance

Figure B38. A plot showing the importance of each test, brood board, and feeding variable to explain the patterns in mean neonate production per living female across all labs in Samples 1 and 2A from the Baseline Intercalibration tests (rounds 1, 2, and 3). Mean neonate production was modeled using random forest regression with each of the metric on the y-axis as a predictor variable. Importance is expressed as the % increase in model Mean Square Error when the variable was removed from the model



.

Figure B39. A plot showing the importance of each test, brood board, and feeding variable to explain the patterns in mean neonate production per living female across all labs in Samples 1 and 2A from the Baseline Intercalibration tests (rounds 1, 2, and 3). CV of neonate production was modeled using random forest regression with each of the metric on the y-axis as a predictor variable. Importance is expressed as the % increase in model Mean Square Error when the variable was removed from the model.

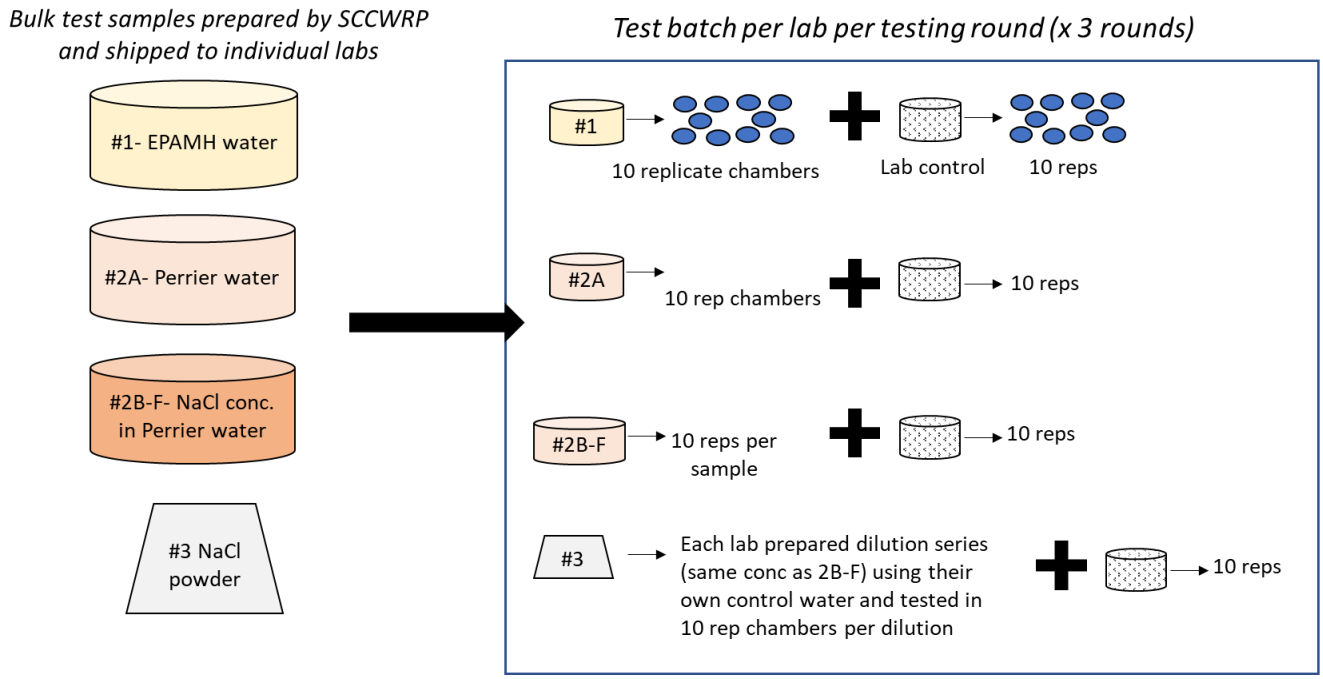


Appendix C - Study plan and summary data for the second intercalibration study.

Overview of baseline testing proedure

The specific objective of the second ILS was to collect additional C. dubia chronic toxicity data and a more complete/consistent lab technique dataset across California-accredited laboratories. Ten (10) laboratories participated in an intercomparison exercise consisting of several split samples tested in three separate testing batches within a ~ 8-week window. Testing design (Figure C1), sample type, preparation and distribution was done as described in the baseline ILS study plan (**Appendix B**). However, laboratories followed a more consistent set of laboratory techniques (see below).

Figure C1: Overview of the *C. dubia* baseline study design.



Summary of standardized parameters

- All laboratories were required to meet the following specifications:
- Limit the age of adults used to start the test to 6-10 days old.
- Use < 24 hr old neonates produced within an 8-hr window.
    - Record which brood number (e.g., 3rd, 4th, or 5th) is used to start the test.
    - Record the specific beginning and end time window for age of neonates at test initiation (hours and minutes).
- Evaluate brood board health for 2 weeks prior to testing using a common set of health criteria provided by SCCWRP.
- Use randomization by blocking of known parentage AND randomization of cups on the test board. Test set-up must use the randomized blocking by known parentage, using 10 randomly selected brood board chambers with a minimum of 8 neonates from the adult on test initiation day. Each test will be treated as independent for blocking by known parentage except for samples 2A and 2B-F. These two tests will be treated as one for the purposes of blocking by known parentage.
- Renew or terminate test boards daily at 24 hours within a 2-hr window from test initiation.
- Conduct the tests for 8 days (i.e., 192 hours). Record of when the tests would have ended if the lab was using their own SOP.
- Independently quantify initial food density (both algae and YCT, do not use numbers provided by the supplier) and describe feeding procedure (incl. quantification method and volume dispensed) to estimate food density in test cups and ensure that it meets EPA protocols requirements. Record all information and measurements. Use repeating pipettor or volumetric pipette for accurate volumes.
- Follow holding times of YCT ≤ 8 days after thawing and algae ≤ 21 days (from time of receipt if purchased or from production if cultured in-house).
- Document split broods on bench sheets at the time of observation including number and size of neonates, and observations of females' movement and presence of eggs in pouch.

Inventory of data collected

- Ten labs participated in testing.
- For Round 1 nine laboratories tested samples
    - Lab M could not participate in the first round due to staffing issues.
- For Round 2 ten laboratories tested samples
    - Half of the samples for Lab B were returned to SCCWRP by the shipping company, but were turned around the next day and all samples were tested together and within holding time
    - Lab G had a mix up of their concentrations for Series 3 samples leading to data for that test being unusable.
- For Round 3 ten laboratories tested samples, all on time
- Lab I consistently had high mortality and low reproduction in the samples sent them by SCCWRP, but generally had good success with their laboratory controls.
- Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.
- Out of 120 expected laboratory controls from 10 labs, 115 were tested to completion.
- Of the 115 controls completed, 4 did not pass test acceptability criteria.

Table C1. Inventory of data collected in the second ILS. Here, a dataset is defined as one sample tested with 10 replicates.

Lab ID	Lab controls	#1- MHW	#2A- DMW Perrier®	#2B-F- SCCWRP concentration-response	#3- Lab concentration-response
A	12	3	3	3	3
B	12	3	3	3	3
E	12	3	3	3	3
F	12	3	3	3	3
G	11	3	3	3	2
I	12	3	3	3	3
L	-	-	-	-	-
M	8	2	2	2	2
N	12	3	3	3	3
O	12	3	3	3	3
P	-	-	-	-	-
Q	12	3	3	3	3
All Labs	**115**	**29**	**29**	**29**	**28**

Table C2. Individual test batches not meeting test acceptability criteria for reproduction (≥ 15 neonates/surviving female) or survival (>90% in controls) in the second ILS. Cells with dashes indicate that the endpoint met test acceptability criteria for that sample.

Lab	Test Round	Mean Neonates/Surviving Female	Survival
N	Round 2	1.2	-
N	Round 2	5.9	-
N	Round 2	0	-
N	Round 2	2.4	-
I	Round 2	-	70%

Table C3. Individual test batches not meeting acceptable brood board mortality (< 20%). Note that all labs had at least 8 neonates per female used to initiate tests and no females were older than 14 days.

Lab	Test Round	Brood Board Percent Mortality	Mean Neonates per Female in Lab Control
I	Round 1	23	33.9[a]
I	Round 1	22	42.0[b]
I	Round 2	23	33.4[c]
G	Round 1	25	22.3[c]

[a]Mean of three controls that were initiated from that brood board.
[b]Brood board used to initiate one test.
[c]Mean of two controls that were initiated from that brood board.

Biological response data

Figure C2. Schematic box plot of the mean number of neonates per female in laboratory control water for each test in both ILS. Dots represent the individual test values and the + symbol represents the mean of all the tests. Data represents protocol specified test termination of the day when 60% of control females achieved three broods. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

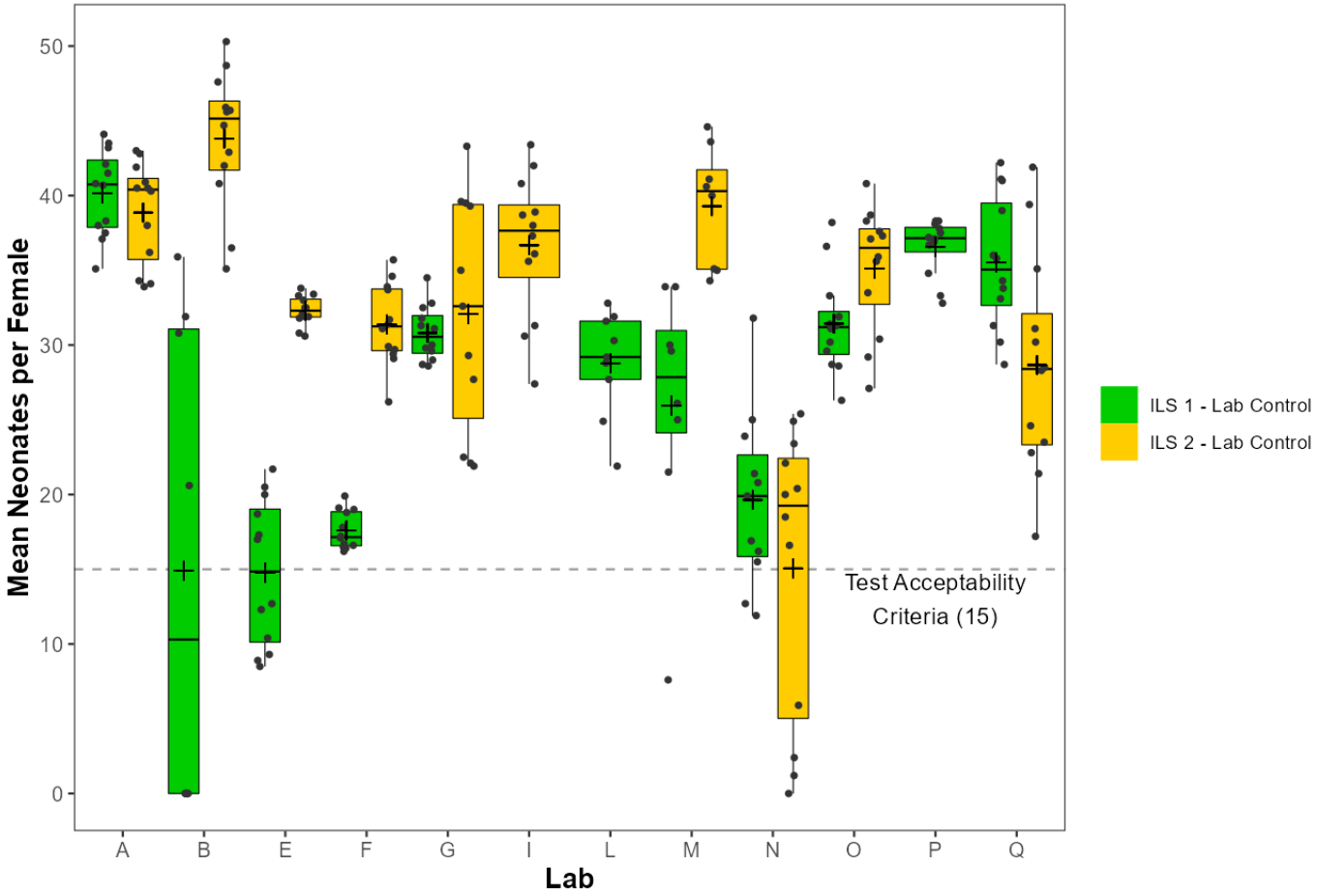


Figure C3. Schematic box plot of the mean number of neonates per female in laboratory in the sample waters provided by SCCWRP for each test in both ILS. The dots represent the individual test values and the + symbol represents the mean of all the tests. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

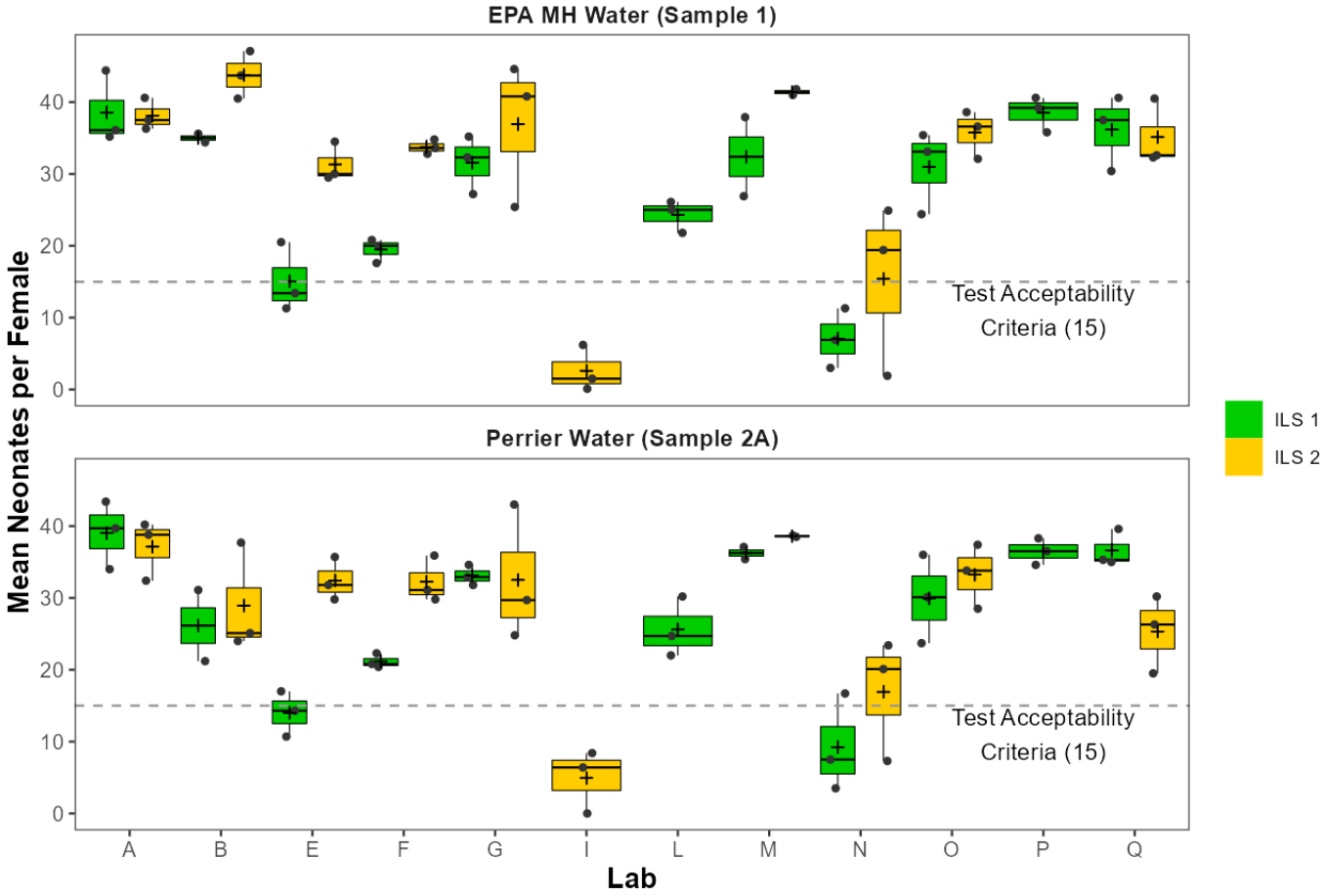


Figure C4. Schematic plot of the coefficient of variation for the number of neonates per female in laboratory control water for each test in both ILS. The dots represent the individual test values and the + symbol represents the mean of all the tests. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

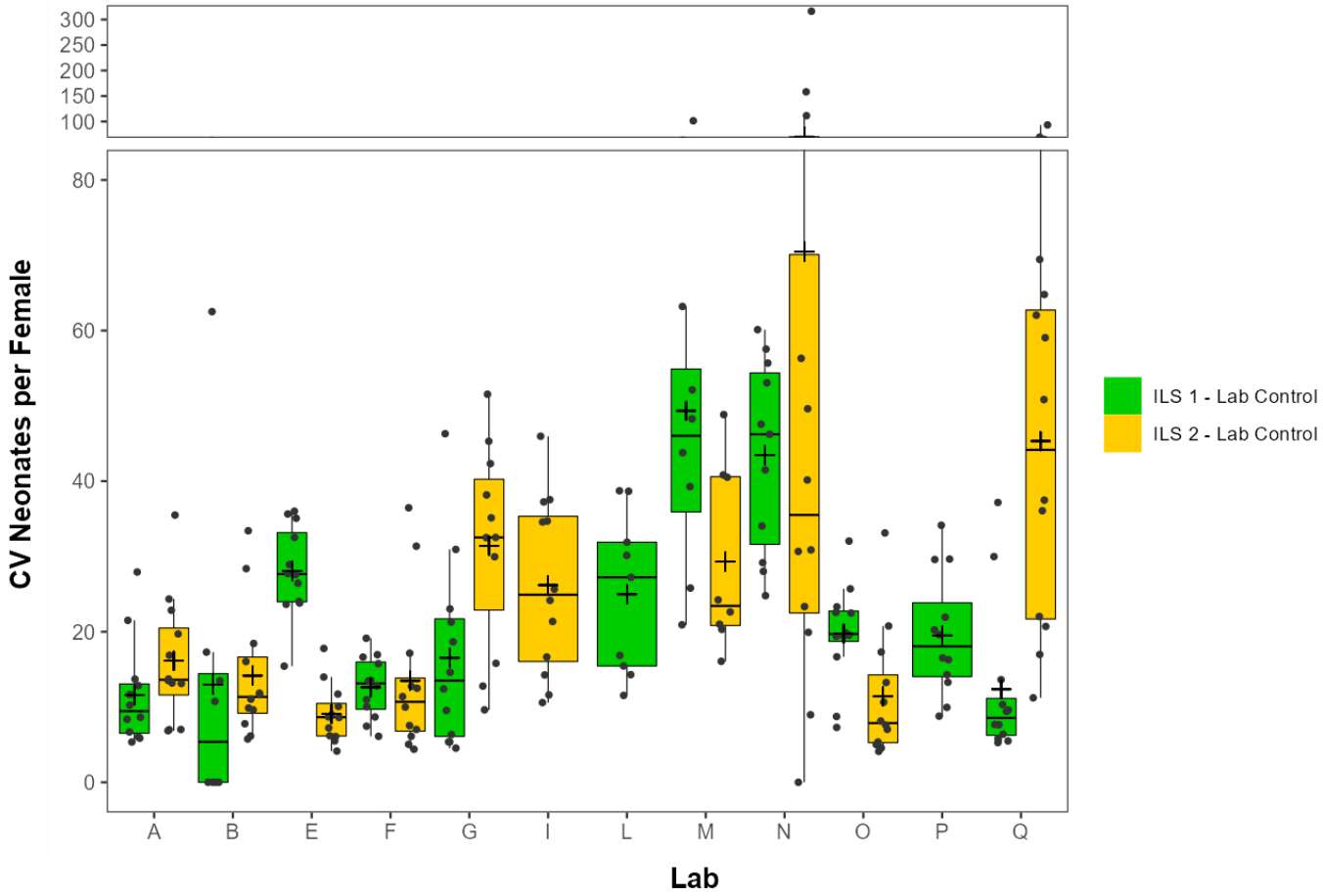


Figure C5. Schematic box plot of the coefficient of variation for the number of neonates per female in laboratory in the sample waters provided by SCCWRP for each test in both ILS. The dots represent the individual test values and the + symbol represents the mean of all the tests. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

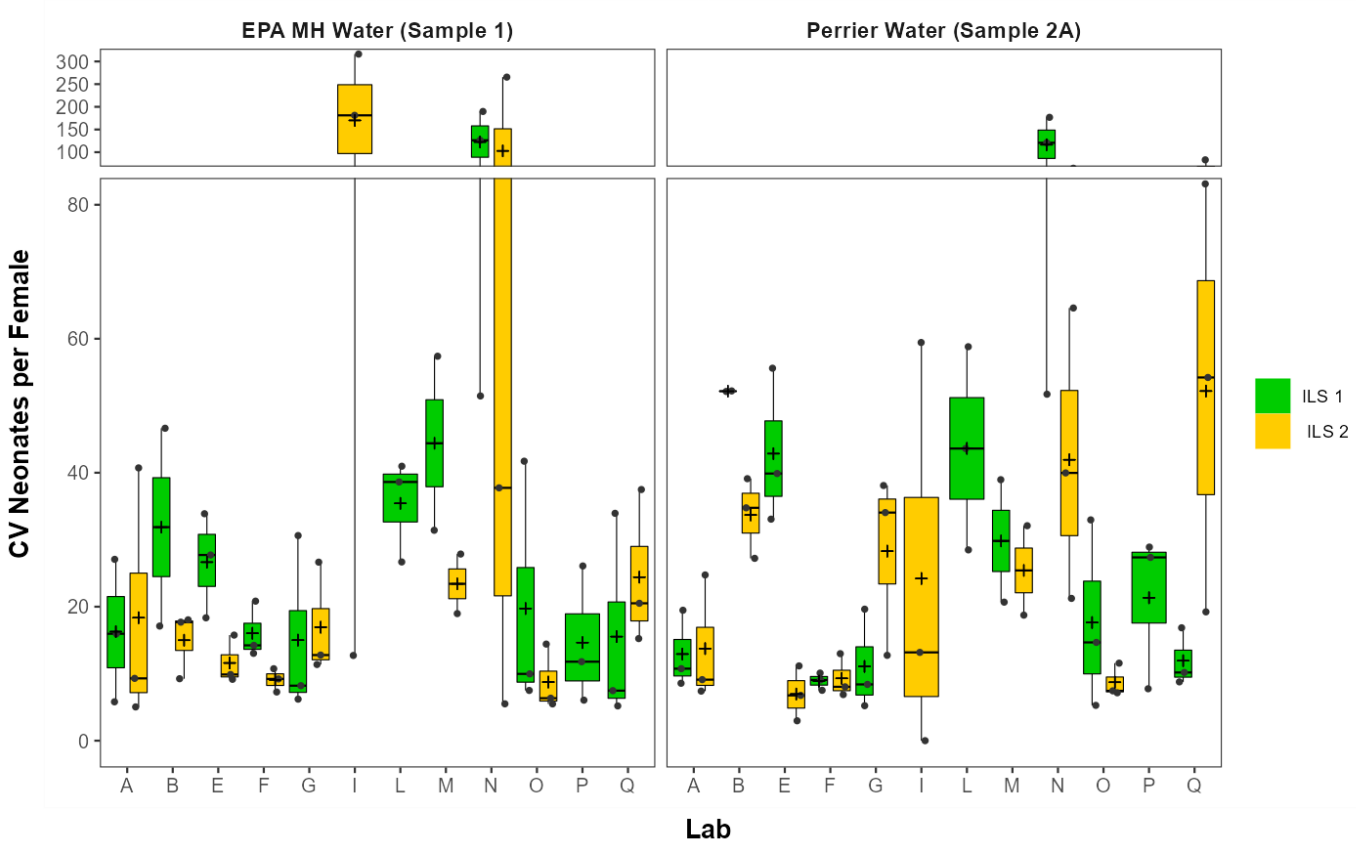


Table C4. Summary of reproduction for Sample 1 from the laboratories participating in the baseline and second C. dubia interlaboratory studies, based on the number of neonates in three broods. For each lab, the data is presented as mean of 3 rounds (except for lab M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Neonates/Adult Female				Neonates/Adult Female			
	Baseline				Second			
Lab	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	38.6	3	35-44	0.13	38.1	3	36-41	0.058
B	35.0	2	34-36	0.024	43.8	3	40-47	0.075
E	16.0	3	13-21	0.25	31.3	3	30-34	0.088
F	20.1	3	18-21	0.072	33.7	3	33-35	0.030
G	31.6	3	27-35	0.13	36.9	3	25-45	0.285
I					2.6	3	0-6	1.23
L	24.3	3	22-26	0.092				
M	32.4	2	27-38	0.24	41.4	2	41-42	0.014
N	7.1	3	3-11	0.59	15.4	3	2-25	0.780
O	31.0	3	24-35	0.19	35.8	3	32-39	0.093
P	38.6	3	36-41	0.065				
Q	36.2	3	30-41	0.14	35.1	3	32-40	0.132

Test acceptability criterion is a mean of 15 neonates/surviving female.

Table C5. Summary of biological data for Sample 2A (diluted mineral water with Perrier®) collected from the laboratories participating in the baseline and second C. dubia interlaboratory studies, based on the number of neonates in three broods. For each lab, the data is presented as mean of 3 rounds (except for lab M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Neonates/Adult Female				Neonates/Adult Female			
	Baseline				Second			
Lab	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	39.0	3	34-43	0.12	37.1	3	32-40	0.112
B	26.2	2	21-31	0.27	28.9	3	24-38	0.263
E	15.4	3	11-21	0.34	32.4	3	30-36	0.093
F	21.2	3	20-22	0.047	32.3	3	30-36	0.100
G	33.1	3	32-35	0.043	32.5	3	25-43	0.290
I					4.9	3	0-8	0.889
L	25.6	3	22-30	0.16				
M	35.2	2	35-35	0.008	38.6	2	39-39	0.004
N	9.2	3	4-17	0.73	16.9	3	7-23	0.502
O	29.9	3	24-36	0.21	33.2	3	28-37	0.125
P	37.0	3	35-38	0.056				
Q	36.6	3	35-40	0.07	25.3	3	20-30	0.214

Test acceptability criterion is a mean of 15 neonates/surviving female.

Table C6. Summary of biological data for laboratory dilution water from the laboratories participating in the baseline and second C. dubia interlaboratory studies, based on the number of neonates in three broods. For each lab, the data is presented as mean of 3 rounds (except for lab M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

		Neonates/Adult Female				Neonates/Adult Female			
		Baseline				Second			
Lab	Sample Type	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	Lab Water	40.2	12	35-44	0.072	38.9	12	34-43	0.088
B	Lab Water	29.8	4	21-36	0.22	43.8	12	35-50	0.105
E	Lab Water	14.9	12	8-22	0.32	32.3	12	31-34	0.031
F	Lab Water	17.6	12	16-20	0.072	31.4	12	26-36	0.087
G	Lab Water	30.8	12	29-35	0.06	32.1	11[a]	22-43	0.246
I	Lab Water					36.7	12	27-43	0.131
L	Lab Water	28.9	9	22-33	0.13				
M	Lab Water	26.0	8	8-34	0.33	39.3	8	34-45	0.102
N	Lab Water	19.6	11	12-32	0.3	15.1	12	0-25	0.649
O	Lab Water	31.4	12	26-38	0.106	35.1	12	27-41	0.120
P	Lab Water	36.6	12	33-38	0.054				
Q	Lab Water	35.5	12	29-42	0.13	28.7	12	17-42	0.258

Test acceptability criterion is a mean of 15 neonates/surviving female.

[a]One test lost due to technical error.

Figure C6. Dot plot of the mean number of neonates per female expressed as a percentage of control response for Sample 1 (EPA Moderately Hard Water) for each test in both ILS. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

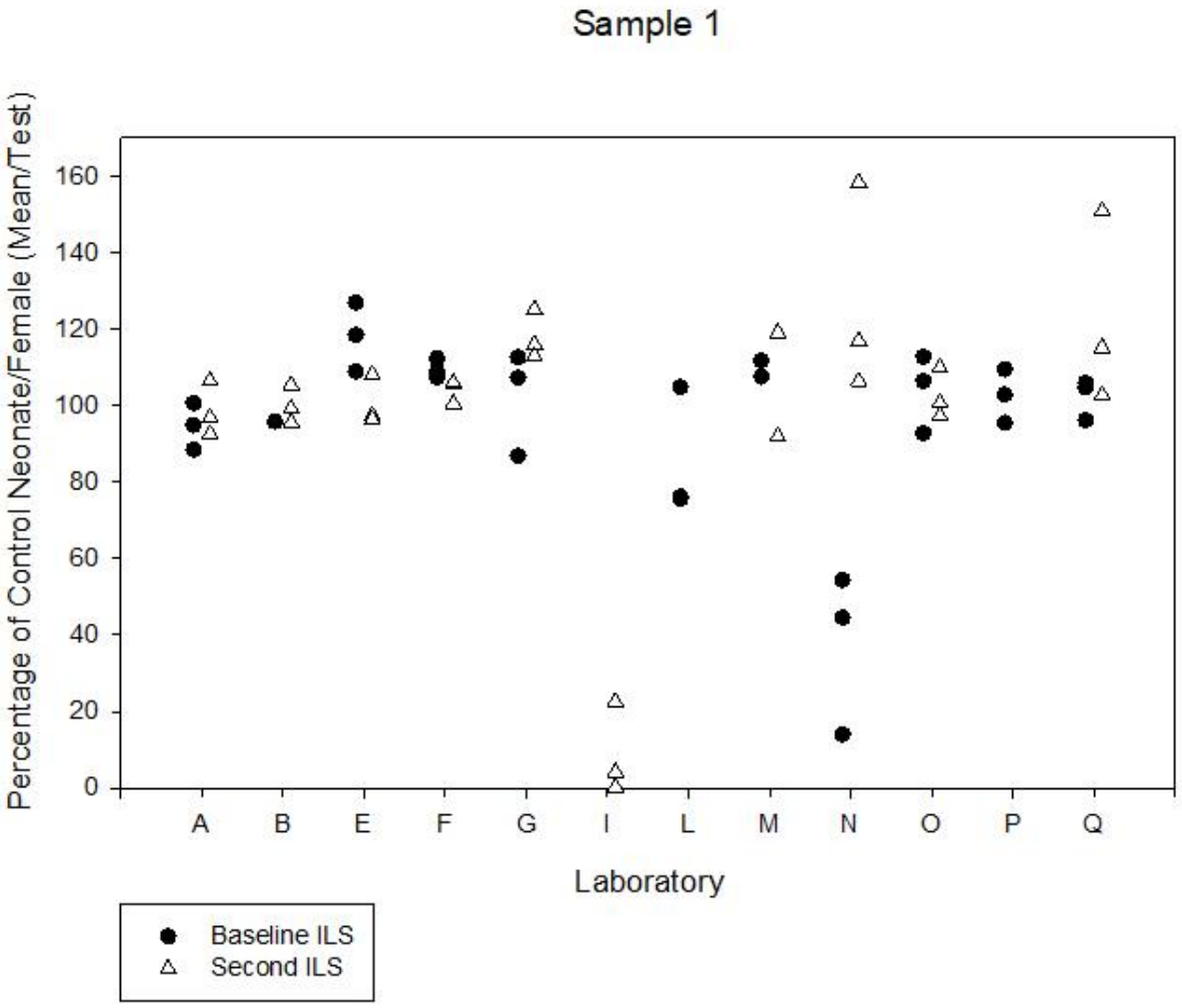


Figure C7. Dot plot of the mean number of neonates per female expressed as a percentage of control response for Sample 2A (DMW Perrier®) for each test in both ILS. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

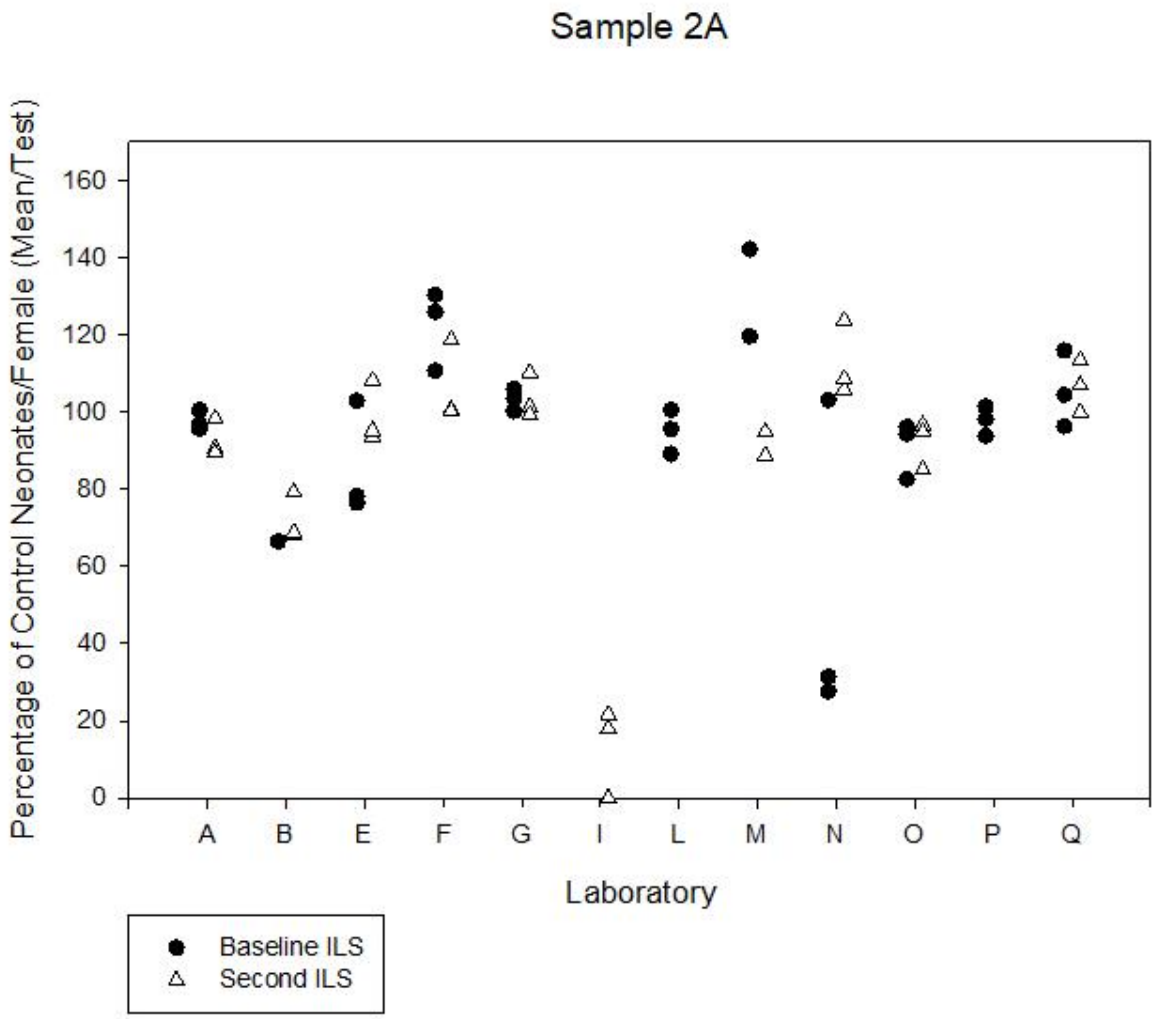


Table C7. Summary of reproduction endpoint for Sample 1 (EPA MHW) from both ILS, expressed as a percentage of control reproduction. For each lab, the data is presented as mean of 3 rounds (except for labs B and M for the baseline and lab M for the second who could not participate in round 1 of testing). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Baseline				Second			
Lab	Control Adjusted Mean (%)	N	Mean Range	CV of Mean	Control Adjusted Mean (%)	N	Mean Range	CV of Mean
A	94.6	3	88-101	0.064	98.6	3	93-106	0.072
B	95.8	1[a]	-	-	100	3	96-105	0.049
E	118	3	109-127	0.076	101	3	96-108	0.065
F	110	3	107-112	0.024	104	3	101-106	0.029
G	102	3	87-113	0.133	118	3	113-125	0.054
I					9.0[b]	3	0.3-23	1.32
L	85.6	3	77-105	0.195				
M	110	3	108-112	0.027	106	2	92-119	0.182
N	37.6	3	14-54	0.559	127	3	106-158	0.216
O	104	3	93-113	0.098	103	3	97-110	0.063
P	103	3	95-109	0.068				
Q	102	3	96-106	0.052	123	3	103-151	0.203

[a]The laboratory control had zero reproduction for round 3, therefore a control adjusted mean could not be calculated.

[b]The laboratory had low reproduction and survival in Sample 1, but normal values in their laboratory controls.

Table C8. Summary of reproduction endpoint for Sample 2A (diluted mineral water with Perrier®) second ILS, expressed as a percentage of control reproduction. For each lab, the data is presented as mean of 3 rounds (except for lab M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

		Baseline				Second			
Lab	**Sample Type**	**Control Adjusted Mean (%)**	**N**	**Mean Range**	**CV of Mean**	**Control Adjusted Mean (%)**	**N**	**Mean Range**	**CV of Mean**
A	DMW Perrier® (2A)	97.7	3	96-100	0.026	92.8	3	90-98	0.051
B	DMW Perrier® (2A)	66.5	1[a]	-	-	72.1	3	68-79	0.085
E	DMW Perrier® (2A)	85.9	3	76-103	0.172	98.9	3	93-108	0.081
F	DMW Perrier® (2A)	122	3	111-130	0.085	107	3	100-119	0.099
G	DMW Perrier® (2A)	103	3	100-106	0.028	104	3	99-110	0.056
I	DMW Perrier® (2A)					13.2[b]	3	0-22	0.877
L	DMW Perrier® (2A)	95.1	3	89-100	0.060				
M	DMW Perrier® (2A)	131	2	120-142	0.122	91.8	2	89-95	0.047
N	DMW Perrier® (2A)	54.0	3	28-103	0.788	113	3	106-124	0.085
O	DMW Perrier® (2A)	91.0	3	83-96	0.081	92.2	3	85-97	0.068
P	DMW Perrier® (2A)	97.7	3	94-101	0.038				

		Baseline				Second			
Lab	**Sample Type**	**Control Adjusted Mean (%)**	**N**	**Mean Range**	**CV of Mean**	**Control Adjusted Mean (%)**	**N**	**Mean Range**	**CV of Mean**
Q	DMW Perrier® (2A)	106	3	96-116	0.093	107	3	100-113	0.063

[a]The laboratory control had zero reproduction for round 3, therefore a control adjusted mean could not be calculated.

[b]The laboratory had low reproduction and survival in Sample 1, but normal values in their laboratory controls.

Table C9. Neonate production and adult female survival data for Sample 1 (EPA Moderately Hard Water) collected from the ten laboratories participating in the second C. dubia interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
A	1	37.5	0.407	9
A	2	36.3	0.050	10
A	3	40.6	0.093	10
B	1	40.5	0.180	10
B	2	43.7	0.177	9
B	3	47.1	0.092	10
E	1	30.0	0.092	10
E	2	29.5	0.158	10
E	3	34.5	0.099	10
F	1	34.8	0.073	10
F	2	32.8	0.107	10
F	3	33.6	0.092	10
G	1	25.4	0.266	10
G	2	40.8	0.114	10
G	3	44.6	0.128	10
I	1	6.2	0.127	0
I	2	1.5	1.81	3
I	3	0.1	3.16	0
M	2	41.0	0.279	8
M	3	41.8	0.190	10
N	1	24.9	0.055	9
N	2	1.9	2.65	10
N	3	19.4	0.377	10
O	1	36.6	0.063	10
O	2	32.1	0.144	10

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
O	3	38.6	0.055	10
Q	1	32.3	0.152	9
Q	2	32.6	0.205	9
Q	3	40.5	0.375	10

Table C10. Neonate production and survival data for Sample 2A (diluted mineral water with Perrier®) collected from the ten laboratories participating in the second *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
A	1	38.8	0.248	9
A	2	32.4	0.091	10
A	3	40.2	0.074	10
B	1	24.0	0.272	10
B	2	37.7	0.391	9
B	3	25.1	0.348	9
E	1	29.8	0.112	10
E	2	31.8	0.068	10
E	3	35.7	0.030	10
F	1	35.9	0.069	10
F	2	31.1	0.130	9
F	3	29.8	0.080	10
G	1	24.8	0.340	9
G	2	29.7	0.381	8
G	3	43.0	0.127	10
I	1	6.4	0.132	1
I	2	8.4	0.594	8
I	3	0	-	2
M	2	38.7	0.321	9
M	3	38.5	0.187	10
N	1	23.4	0.212	10
N	2	7.3	0.646	10
N	3	20.1	0.400	9
O	1	37.4	0.072	10
O	2	28.5	0.074	10

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
O	3	33.8	0.116	10
Q	1	19.5	0.831	8
Q	2	30.2	0.192	9
Q	3	26.3	0.542	9

Table C11. Neonate production and survival data for laboratory controls collected from the ten laboratories participating in the second *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
A	1	40.5	0.228	9
A	1	42.8	0.131	10
A	1	33.9	0.355	9
A	1	38.0	0.169	9
A	2	34.1	0.137	10
A	2	36.2	0.070	10
A	2	34.3	0.132	10
A	2	40.3	0.070	10
A	3	41.9	0.135	10
A	3	40.9	0.197	10
A	3	40.5	0.243	10
A	3	43.0	0.068	10
B	1	35.1	0.284	10
B	1	40.8	0.184	10
B	1	42.9	0.099	10
B	1	42.0	0.078	10
B	2	45.7	0.058	10
B	2	47.6	0.160	9
B	2	50.3	0.062	10
B	2	48.7	0.096	10
B	3	44.7	0.118	10
B	3	36.5	0.334	10
B	3	45.9	0.116	10
B	3	45.6	0.110	10

Table C11 continued. Neonate production and survival data for laboratory controls collected from the ten laboratories participating in the second *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
E	1	30.8	0.061	10
E	1	31.9	0.101	10
E	1	32.5	0.086	10
E	1	31.8	0.072	10
E	2	30.6	0.178	10
E	2	33.4	0.055	10
E	2	33.8	0.041	10
E	2	32.1	0.090	10
E	3	31.9	0.140	10
E	3	33.0	0.087	10
E	3	33.3	0.062	10
E	3	32.5	0.117	10
F	1	34.6	0.061	10
F	1	35.7	0.044	10
F	1	33.7	0.050	10
F	1	33.9	0.070	10
F	2	31.1	0.100	10
F	2	26.2	0.314	10
F	2	29.9	0.128	10
F	2	29.4	0.365	9
F	3	31.7	0.172	9
F	3	29.7	0.125	10
F	3	29.1	0.114	10
F	3	31.4	0.075	10

Table C11 continued. Neonate production and survival data for laboratory controls collected from the ten laboratories participating in the second *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
G	1	21.9	0.382	8
G	1	22.5	0.423	10
G	1	22.1	0.325	10
G	1	27.7	0.351	8
G	2	32.6	0.300	8
G	2	29.3	0.515	8
G	2	39.6	0.128	10
G	2	NT	NT	NT
G	3	39.5	0.158	10
G	3	43.3	0.096	10
G	3	39.3	0.325	9
G	3	35.0	0.453	8
I	1	27.4	0.375	8
I	1	35.6	0.242	9
I	1	38.7	0.346	9
I	1	42.0	0.106	9
I	2	36.1	0.166	10
I	2	38.9	0.143	10
I	2	31.3	0.214	10
I	2	30.6	0.460	7
I	3	37.3	0.256	9
I	3	43.4	0.116	9
I	3	38.0	0.372	8
I	3	40.8	0.347	9

NT= Not tested. Technical error led to loss of sample.

Table C11 continued. Neonate production and survival data for laboratory controls collected from the ten laboratories participating in the second *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
M	2	44.6	0.203	10
M	2	43.6	0.226	10
M	2	41.1	0.242	10
M	2	40.0	0.161	8
M	3	35.1	0.405	9
M	3	40.6	0.210	9
M	3	34.3	0.488	8
M	3	35.0	0.408	9
N	1	23.4	0.233	9
N	1	22.1	0.309	10
N	1	24.9	0.090	10
N	1	25.4	0.199	9
N	2	1.2	3.16	10
N	2	5.9	1.58	10
N	2	0	-	10
N	2	2.4	1.11	10
N	3	16.6	0.563	10
N	3	18.5	0.402	10
N	3	20.4	0.307	10
N	3	20	0.496	9

Table C11 continued. Neonate production and survival data for laboratory controls collected from the ten laboratories participating in the second *C. dubia* interlaboratory study presented by test. Each test had 10 replicates.

Lab	Round	Mean Neonates/Adult Female	CV of Neonates/Female	# Surviving Females
O	1	37.6	0.054	10
O	1	38.7	0.046	10
O	1	35.9	0.081	10
O	1	37.1	0.041	10
O	2	29.2	0.173	10
O	2	33.5	0.106	10
O	2	30.4	0.208	10
O	2	27.1	0.331	10
O	3	38.3	0.076	10
O	3	35.6	0.070	10
O	3	37.3	0.133	10
O	3	40.8	0.050	10
Q	1	21.4	0.620	9
Q	1	17.2	0.933	8
Q	1	22.8	0.590	9
Q	1	28.5	0.220	9
Q	2	28.3	0.508	9
Q	2	30.2	0.207	10
Q	2	31.1	0.170	10
Q	2	23.5	0.648	9
Q	3	35.1	0.361	10
Q	3	24.6	0.694	9
Q	3	41.9	0.112	10
Q	3	39.4	0.375	9

Table C12. Age of female *Ceriodaphnia dubia* when their neonates were used to initiate testing in baseline and second ILS. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

Lab	Mean Age of Female at Test Initiation (Days)	N	Mean Range	Mean Age of Female at Test Initiation (Days)	N	Mean Range
	Baseline			Second		
A	6.3	3	6 - 7	6.2	4	6 - 7
B	9.0	2	9 - 9	7.0	3	7 - 7
E	10.7	3	10 - 11	9.0	3	7 - 10
F	9.7	9	9 - 11	9.0	6	7 - 10
G	7.7	6	7 - 8	7.0	6	7 - 7
L	12.4	14	12 - 14			
I				8.7	6	8 - 9
M	8.0	3	7 - 9	8.6	5	8 - 9
N	7.6	7	7 - 11	8.5	8	8 - 9
O	8.0	4	8 - 8	7.8	6	7 - 9
P	6.8	5	6 - 7			
Q	9.3	3	9 - 10	9.0	5	9 - 9

Table C13. Summary of the number of broods produced during eight days of testing in laboratory dilution water collected from the 12 laboratories participating in the baseline and second *C. dubia* interlaboratory study. For each lab, the data is presented as mean of 3 rounds (except for lab M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

		Number of Broods (8 days)							
		Baseline				Second			
Lab	Sample Type	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	Lab Water	4.0	12	3.7-4.1	0.025	4.0	12	3.7-4.0	0.025
B	Lab Water	3.2	4	2.4-3.8	0.19	4.0	12	3.9-4.0	0.011
E	Lab Water	4.5	12	4.1-5.0	0.067	4.3	12	3.8-5.0	0.10
F	Lab Water	4.4	12	4.0-4.9	0.071	4.5	12	4.0-5.0	0.093
G	Lab Water	4.0	12	3.2-4.2	0.070	3.8	11[a]	3.4-4.0	0.052
I	Lab Water					3.8	11[b]	3.6-4.0	0.040
L	Lab Water	3.1	9	2.6-3.8	0.12				
M	Lab Water	2.7	8	1.1-3.4	0.26	4.3	8[c]	4.0-4.6	0.046
N	Lab Water	3.4	11	2.9-4.2	0.12	2.9	8[b]	2.6-3.2	0.066
O	Lab Water	4.3	12	4.0-4.7	0.045	4.2	12	4.0-5.0	0.069
P	Lab Water	3.9	12	3.7-4.0	0.036				
Q	Lab Water	3.7	12	3.1-4.0	0.094	3.3	12	2.6-4.0	0.15

[a]Test lost due to technical error.

[b]Test(s) that did not meet test acceptability criteria not included.

[c]Unable to participate in Round 1 due to staffing issues.

Table C14. Summary of the number of days to the first brood produced in laboratory dilution water collected from the 12 laboratories participating in the baseline and second *C. dubia* interlaboratory study. For each lab, the data is presented as mean of 3 rounds (except for lab M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

		Time to First Brood							
		Baseline				Second			
Lab	Sample Type	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	Lab Water	4.0	12	4.0-4.0	0	3.6	12	3.0-4.0	0.13
B	Lab Water	4.3	4	4.0-5.1	0.13	3.9	12	3.6-4.0	0.044
E	Lab Water	4.3	12	3.8-4.8	0.075	4.2	12	3.6-4.7	0.083
F	Lab Water	3.4	12	3.0-3.9	0.096	3.3	12	3.0-3.9	0.10
G	Lab Water	3.5	12	3.0-4.0	0.13	3.8	11[a]	3.0-4.0	0.10
I	Lab Water					3.7	11[b]	3.0-4.0	0.085
L	Lab Water	4.3	9	4.0-5.0	0.088				
M	Lab Water	3.7	8	3.0-4.2	0.13	3.2	8[c]	3.0-3.6	0.068
N	Lab Water	4.0	11	3.3-4.4	0.094	4.1	8[b]	3.3-5.0	0.18
O	Lab Water	3.0	12	3.0-3.2	0.019	3.4	12	3.0-4.0	0.095
P	Lab Water	4.0	12	4.0-4.1	0.007				
Q	Lab Water	4.0	12	4.0-4.0	0	4.1	12	4.0-4.4	0.030

[a]Test lost due to technical error.

[b]Test(s) that did not meet test acceptability criteria not included.

[c]Unable to participate in Round 1 due to staffing issues.

Table C15. Summary of minimum age of neonates used to initiate testing from the baseline and second ILS. The N value represents the total number of samples tested. Note that Lab M has a lower number because they were unable to participate in Round 1 of the second ILS. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Minimum Neonate Age at Test Start (h)							
	Baseline				Second			
Lab	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	6.7	12	6-8	0.098	8.8	12	5-15	0.48
B	0	4	0-0	0	7.7	12	8-8	0.031
E	4.3	12	4-5	0.11	5.0	12	3-6	0.28
F	10.8	12	8-16	0.34	10.0	12	7-15	0.34
G	12.2	12	6-20	0.53	8.1	12	4-18	0.74
I					10.2	12	4-17	0.51
L	4.0	9	1-8	0.53				
M	12.7	8	6-18	0.45	17.5	8	14-20	0.14
N	6.4	11	4-13	0.67	3.7	12	1-5	0.51
O	14.3	12	7-22	0.36	12.8	12	8-21	0.35
P	1.8	12	1-3	0.42				
Q	2.1	12	0.5-5	0.59	2.0	12	0.4-3	0.40

Table C16. Summary of maximum age of neonates used to initiate testing from the baseline and second ILS. The N value represents the total number of test performed. Note that Lab B has a lower number because they were unable to participate in Round 1. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Maximum Neonate Age at Test Start (h)							
	Baseline				Second			
Lab	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	14.6	12	14-15	0.035	16.5	12	13-22	0.23
B	6.8	4	6.8-6.8	0	22.3	12	21-24	0.046
E	8.0	12	8-8	0	21.0	12	19-22	0.067
F	18.5	12	15-24	0.20	17.6	12	15-22	0.19
G	18.7	12	14-25	0.26	15.3	12	12-23	0.31
I	15.5	9	5-24	0.49	17.4	12	12-23	0.25
L	19.5	8	14-24	0.23				
M	14.4	11	12-21	0.30	23.3	8	22-24	0.024
N	20.8	12	15-24	0.18	9.7	12	9-10	0.049
O	6.3	12	5-8	0.13	19.6	12	16-24	0.17
P	6.6	12	5-10	0.20				
Q	14.6	12	14-15	0.035	5.1	12	3-7	0.27

Table C17. Calculated day testing would have ended using the protocol trigger of 60% of control females having produced three broods for the baseline and second ILS. The N value refers to the number of tests performed. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Calculated Test Duration (Days to 60% of females having 3 broods)					
	Baseline			Second		
Lab	Mean	N	Mean Range	Mean	N	Mean Range
A	7.0	12	7-7	6.2	12	6-7
B	7.2	4	7-8	6.8	12	6-7
E	6.5	12	6-7	6.7	12	6-7
F	6.1	12	6-7	6.0	12	6-6
G	6.2	12	6-7	6.6	11[a]	6-7
I				6.3	11[b]	6-7
L	7.2	9	7-8			
M	6.6	8	6-8	6.0	8[c]	6-6
N	6.7	11	6-8	7.5	8[b]	7-8
O	6.0	12	6-6	6.0	12	6-6
P	6.7	12	6-7			
Q	7.0	12	7-7	7.1	12	7-8

[a]Test lost due to technical error.

[b]Test(s) that did not meet test acceptability criteria not included.

[c]Unable to participate in Round 1 due to staffing issues.

Table C18. Average number of males observed per test in the baseline and second ILS. The N value refers to the number of tests successfully performed. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Number of Males per Test			
	Baseline		Second	
Lab	Mean	N	Mean	N
A	0	12	0	12
B	0	4	0	12
E	0	12	0	12
F	0	12	0	12
G	0	12	0.36	11[a]
I			0	11[b]
L	0	9		
M	0	8	0	8[c]
N	0	11	0	8[b]
O	0	12	0	12
P	0	12		
Q	0	12	0.42	12

[a]Test lost due to technical error.

[b]Test(s) that did not meet test acceptability criteria not included.

[c]Unable to participate in Round 1 due to staffing issues.

Table C19. Summary of number of neonates per surviving female in EPA Moderately Hard Water (Sample 1) from the 12 laboratories participating in the baseline *C. dubia* interlaboratory study. The mean was calculated by taking the total number of neonates produced for all females and dividing by the number of surviving females. For each lab, the data is presented as mean of 3 rounds (except for labs M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Neonates/ Surviving Female							
	Baseline				Second			
Lab	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	39.9	3	35-44	0.16	39.5	3	36-42	0.072
B	39.5	2	34-44	0.12	45.4	3	40-49	0.095
E	15.1	3	11-20	0.27	31.3	3	30-34	0.088
F	19.5	3	18-21	0.16	33.7	3	33-35	0.030
G	31.6	3	27-35	0.15	36.9	3	25-45	0.28
I					*	*	*	*
L	29.3	3	28-31	0.18				
M	36.0	2	30-42	0.27	46.5	2	42-51	0.14
N	13.8	3	10-17	0.62	16.3	3	2-28	0.81
O	31.0	3	24-35	0.20	35.8	3	32-39	0.093
P	38.5	3	36-41	0.15				
Q	37.3	3	34-41	0.062	37.5	3	36-40	0.069

*Very low or no surviving females.

Table C20. Summary of number of neonates per surviving female in diluted mineral water with Perrier® (Sample 2A) from the 12 laboratories participating in the baseline and second *C. dubia* interlaboratory studies. The mean was calculated by taking the total number of neonates produced for all females and dividing by the number of surviving females. For each lab, the data is presented as mean of 3 rounds (except for labs M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Neonates/ Surviving Female							
	Baseline				Second			
Lab	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	39.0	3	34-43	0.13	38.6	3	32-43	0.14
B	27.9	2	21-35	0.45	31.3	3	24-42	0.30
E	14.0	3	11-17	0.43	32.4	3	30-36	0.093
F	21.2	3	20-22	0.089	33.4	3	30-36	0.096
G	33.1	3	32-35	0.11	35.9	3	28-43	0.22
I					*	*	*	*
L	30.8	3	27-34	0.22				
M	38.2	2	37-39	0.28	40.8	2	38-43	0.078
N	17.6	3	7-25	0.61	17.7	3	7-23	0.51
O	29.9	3	24-36	0.18	33.2	3	28-37	0.13
P	39.1	3	38-41	0.082				
Q	36.6	3	35-40	0.12	29.1	3	24-34	0.16

*Very low or no surviving females.

Table C21. Summary of number of neonates per surviving female in laboratory dilution water from the 12 laboratories participating in the baseline and second *C. dubia* interlaboratory studies. The mean was calculated by taking the total number of neonates produced for all females and dividing by the number of surviving females. For each lab, the data is presented as mean of 3 rounds (except for labs M who could not participate in round 1). N values refer to the number of tests included in the mean and CV. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

		Neonates/ Surviving Female							
		Baseline				Second			
Lab	Sample Type	Mean	N	Mean Range	CV of Mean	Mean	N	Mean Range	CV of Mean
A	Lab Water	40.5	12	37-44	0.099	39.9	12	34-45	0.089
B	Lab Water	32.1	4	23-36	0.23	44.3	12	35-53	0.12
E	Lab Water	14.8	12	8-22	0.28	32.3	12	31-34	0.031
F	Lab Water	17.8	12	16-21	0.13	31.9	12	26-36	0.090
G	Lab Water	33.0	12	29-42	0.11	35.8	11	22-44	0.23
I	Lab Water					41.1	11	31-48	0.14
L	Lab Water	30.4	9	27-35	0.21				
M	Lab Water	29.2	8	11-38	0.37	43.1	8	39-50	0.084
N	Lab Water	23.1	11	15-35	0.34	15.7	12	0-28	0.66
O	Lab Water	32.0	12	29-38	0.20	35.1	12	27-41	0.12
P	Lab Water	38.3	12	35-47	0.14				
Q	Lab Water	36.1	12	31-42	0.11	30.8	12	22-44	0.22

Figure C8. Schematic box plot of the IC25 of NaCl spiked waters based on the nominal concentration in the second ILS. The dots represent the values of the individual tests and the + represents the mean of all the tests. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

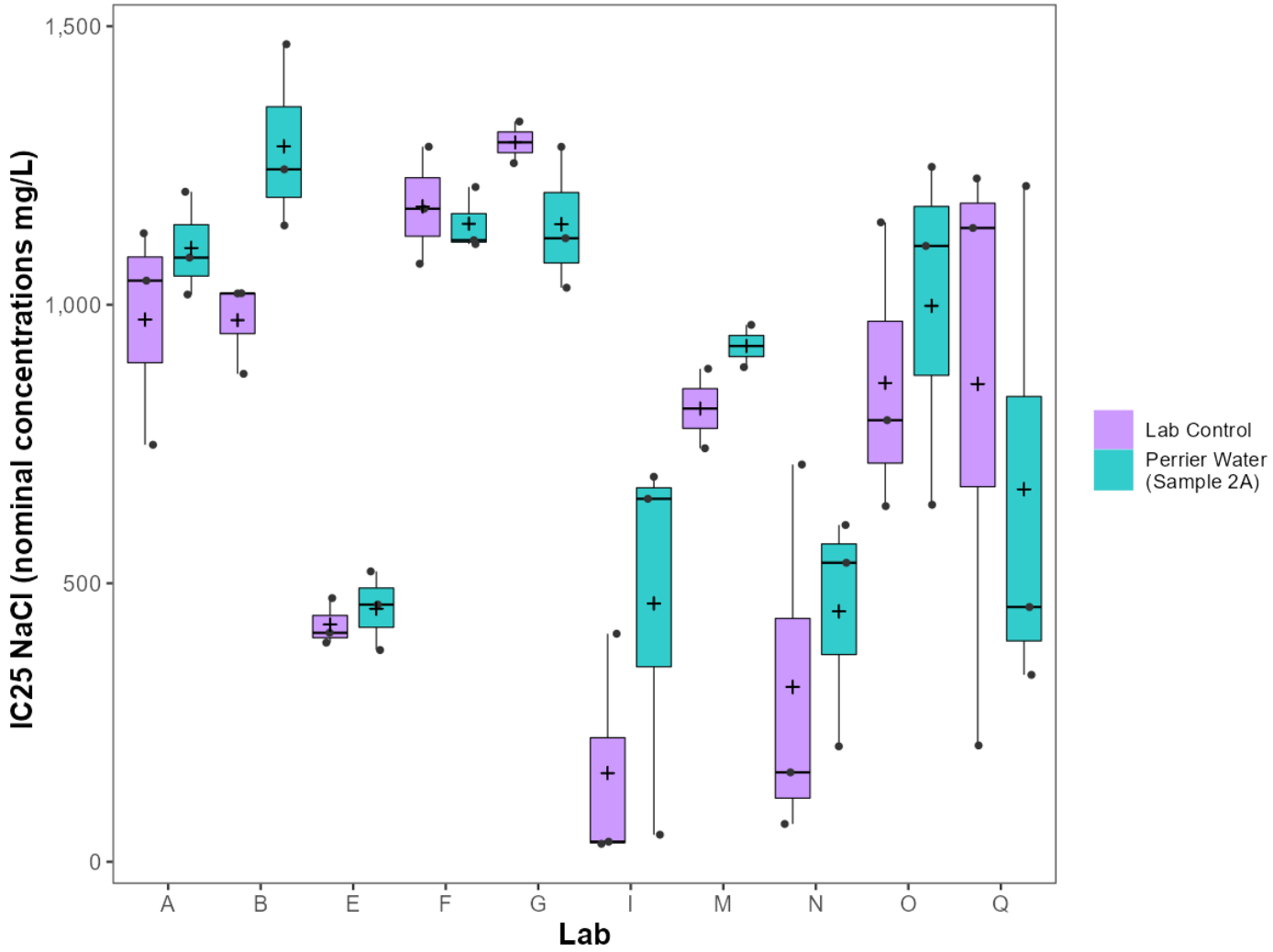


Figure C9. Schematic box plot of the IC50 of NaCl spiked waters based on the nominal concentration in the second ILS. The dots represent the values of the individual tests and the + represents the mean of all the tests. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

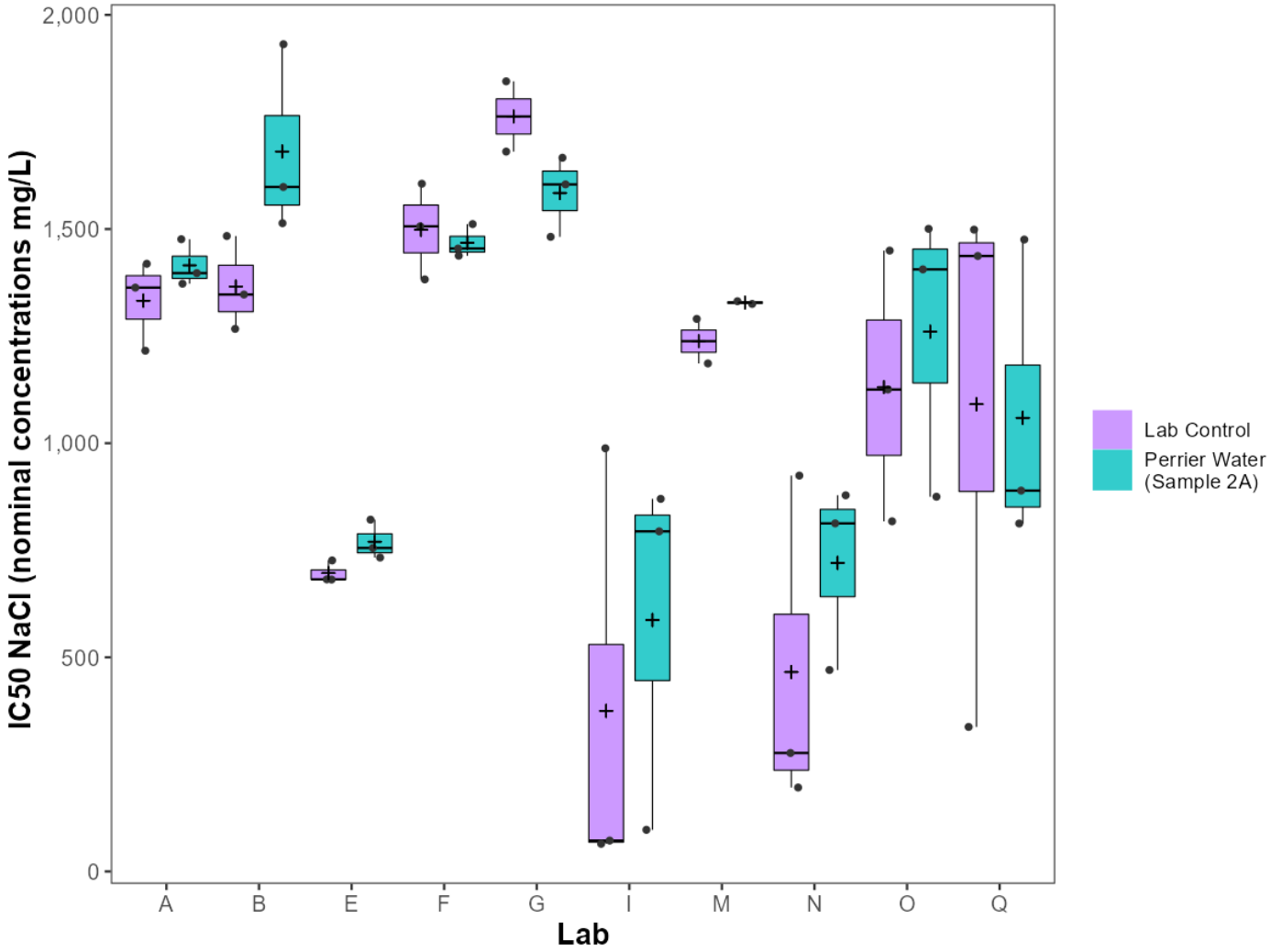


Figure C10. Schematic box plot of the IC50:IC25 of NaCl spiked waters based on the nominal concentration in the second ILS. The dots represent the values of the individual tests and the + represents the mean of all the tests. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

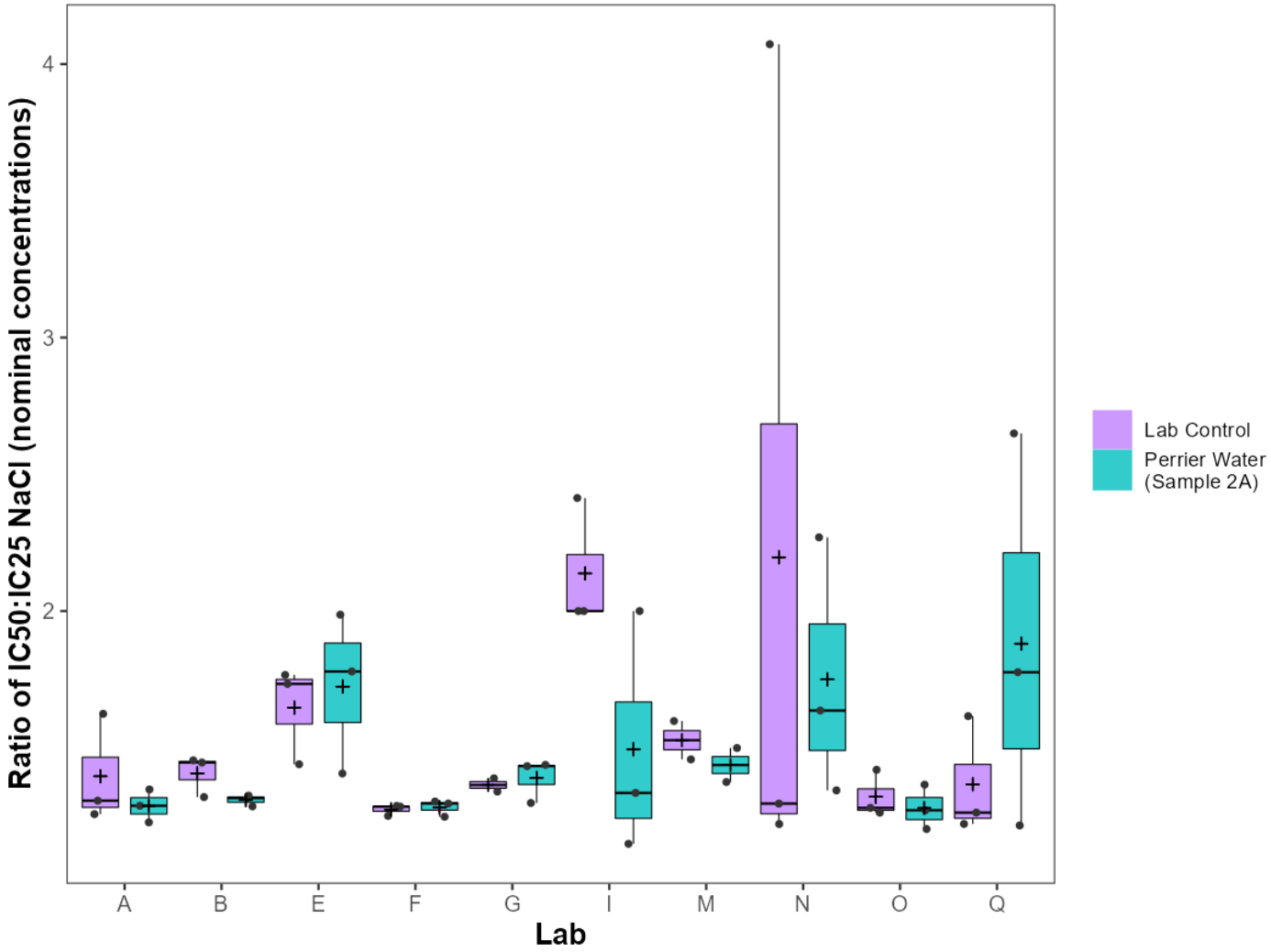


Table C22. IC25/50 and LC50 from NaCl spiked samples from the second ILS. Data are expressed as measured conductivity in µS/cm.

Lab Code	Round #	SCCWRP DMW Perrier® (Samples 2A-F)			Lab Dilution Water (Sample 3 series)		
		IC25	IC50	LC50	IC25	IC50	LC50
A	1	2151	2788	NC	2312	2905	3274
A	2	2270	2832	NC	1725	2612	NC
A	3	2483	2975	3117	2282	2124	2080
B	1	2555	3194	NC	2021	2762	3342
B	2	2959	3794	NC	2445	3057	NC
B	3	2374	3042	NC	2282	3101	4014
E	1	1205	1608	NC	1307	1753	NC
E	2	1080	1777	NC	1209	1690	NC
E	3	901.3	1651	NC	1113	1795	NC
F	1	2314	2905	3242	2230	2782	2992
F	2	2498	3038	3274	2509	3091	NC
F	3	2327	2936	3557	2383	2994	NC
G	1	2628	3317	NC	2645	3427	NC
G	2	2173	2985	NC	NT	NT	NC
G	3	2333	3205	NC	2754	3728	NC
I	1	NC	NC	NC	NC	NC	NC
I	2	NC	NC	NC	NC	NC	NC
I	3	NC	NC	NC	NC	NC	NC
M	2	2049	2703	3377	1960	2690	NC
M	3	1904	2714	2838	1738	2564	NC
N	1	1364	1761	NC	1298	1600	NC
N	2	NC	NC	NC	NC	NC	NC
N	3	542	1099	NC	377	720	NC
O	1	2308	2848	3917	1552	1886	2071
O	2	2563	3018	2918	1852	2112	1747

Lab Code	Round #	SCCWRP DMW Perrier® (Samples 2A-F)			Lab Dilution Water (Sample 3 series)		
		IC25	IC50	LC50	IC25	IC50	LC50
O	3	1433	1880	3617	2465	3027	4050
Q	1	1071	1760	1760	2573	3080	3549
Q	2	2323	2576	3130	615	871	1484
Q	3	804	1907	2868	2424	2987	3106

*NC = Not Calculable
*NT = Not tested

Table C23. IC25, IC50 and LC50 results from NaCl spiked samples from the second ILS. Data are expressed as nominal NaCl concentrations in mg/L.

Lab Code	Round #	SCCWRP DMW Perrier® (Samples 2A-F)			Lab Dilution Water (Sample 3 series)		
		IC25	IC50	LC50	IC25	IC50	LC50
A	1	1020	1370	NC	1040	1360	1560
A	2	1080	1400	NC	749	1220	NC
A	3	1200	1480	1560	1130	1420	1500
B	1	1240	1600	NC	876	1270	1570
B	2	1470	1930	NC	1020	1350	NC
B	3	1140	1510	NC	1020	1480	2000
E	1	521	733	NC	473	682	NC
E	2	462	821	NC	393	682	NC
E	3	380	755	NC	411	726	NC
F	1	1110	1440	1620	1070	1380	1500
F	2	1210	1510	1640	1280	1610	NC
F	3	1120	1460	1800	1170	1510	NC
G	1	1280	1670	NC	1250	1680	NC
G	2	1030	1480	NC	NT	NT	NT
G	3	1120	1600	NC	1330	1840	NC
I*	1	652	870	62	36	72	62
I*	2	49	97	83	409	988	62
I*	3	691	794	62	32	65	62
M	2	964	1320	1700	885	1290	NC
M	3	888	1330	1400	742	1190	NC
N	1	1360	1760	NC	1300	1600	NC
N	2	1230	1890	NC	509	572	NC
N	3	542	1100	NC	377	720	NC
O	1	2310	2850	3920	1550	1890	2070

Lab Code	Round #	SCCWRP DMW Perrier® (Samples 2A-F)			Lab Dilution Water (Sample 3 series)		
		IC25	IC50	LC50	IC25	IC50	LC50
O	2	2560	3020	2920	1850	2110	1750
O	3	1430	1880	3620	2460	3030	4050
Q	1	1070	1760	1760	2570	3080	3550
Q	2	2500	2970	3130	615	871	1480
Q	3	804	1910	2870	2420	2990	3110

*Values are presented for this laboratory, however the number of neonates was very low and the response plots are flat. NC= Not calculable. NT= Not tested.

Table C23B **Point estimate values 95% confidence intervals based on nominal NaCl (mg/L) concentrations for samples 2 and 3 series. The confidence intervals were calculated by randomly resampling with replacement for the replicates within each control and concentration. The IC25/50 were calculated for each resampling The upper and lower bounds were then calculated based on the results from 100 resamples.**

Lab Code	Round #	SCCWRP DMW Perrier® (Samples 2A-F)				Lab Dilution Water (Sample 3 series)			
		IC25 Lower Bound	IC25 Upper Bound	IC50 Lower Bound	IC50 Upper Bound	IC25 Lower Bound	IC25 Upper Bound	IC50 Lower Bound	IC50 Upper Bound
A	1	454	1192	972	1482	838	1114	1309	1412
A	2	911	1173	1307	1455	608	1042	910	1362
A	3	1127	1252	1418	1511	961	1210	1315	1473
B	1	934	1361	1450	1753	714	1048	968	1369
B	2	1317	1680	1665	1996	904	1081	1294	1387
B	3	800	1317	1251	1666	858	1099	1371	1586
E	1	453	564	685	761	393	531	643	716
E	2	357	562	745	875	365	424	493	820
E	3	345	412	677	942	375	437	658	791
F	1	1033	1167	1376	1500	820	1160	1291	1440
F	2	1107	1269	1430	1578	1217	1355	1524	1736
F	3	1051	1160	1397	1520	1091	1224	1440	1580
G	1	897	1411	1434	1847	850	1390	1380	1830
G	2	823	1245	1253	1680	NT	NT	NT	NT
G	3	345	1231	1439	1779	1120	1520	1570	1980
I	1	562	724	795	994	36	37	71	73
I	2	31	133	63	172	120	545	496	1210
I	3	659	691	773	794	32	33	63	67
M	2	772	1081	1206	1396	670	1099	1032	1427
M	3	231	1160	941	1480	114	1089	890	1357
N	1	469	672	743	879	659	759	857	1041
N	2	76	880	453	1442	32	177	65	229

Lab Code	Round #	SCCWRP DMW Perrier® (Samples 2A-F)				Lab Dilution Water (Sample 3 series)			
		IC25 Lower Bound	IC25 Upper Bound	IC50 Lower Bound	IC50 Upper Bound	IC25 Lower Bound	IC25 Upper Bound	IC50 Lower Bound	IC50 Upper Bound
N	3	87	406	215	1069	52	269	104	379
O	1	1061	1135	1374	1426	599	701	746	990
O	2	1108	1256	1405	1508	678	864	953	1249
O	3	460	709	790	968	1046	1210	1382	1497
Q	1	204	719	372	1221	57	320	114	402
Q	2	1153	1252	1435	1501	1159	1267	1453	1536
Q	3	186	734	413	1171	1031	1224	1363	1498

NT= Not tested.

Figure C11. Schematic box plot of the PMSD values of NaCl spiked waters. The dots represent individual test values and the + symbol represents the mean value of all the tests. Note: Data have not been censored based on the shape of the concentration-response plot. Lab Q reported issues with their culture before Round 1, and Lab N before Round 2.

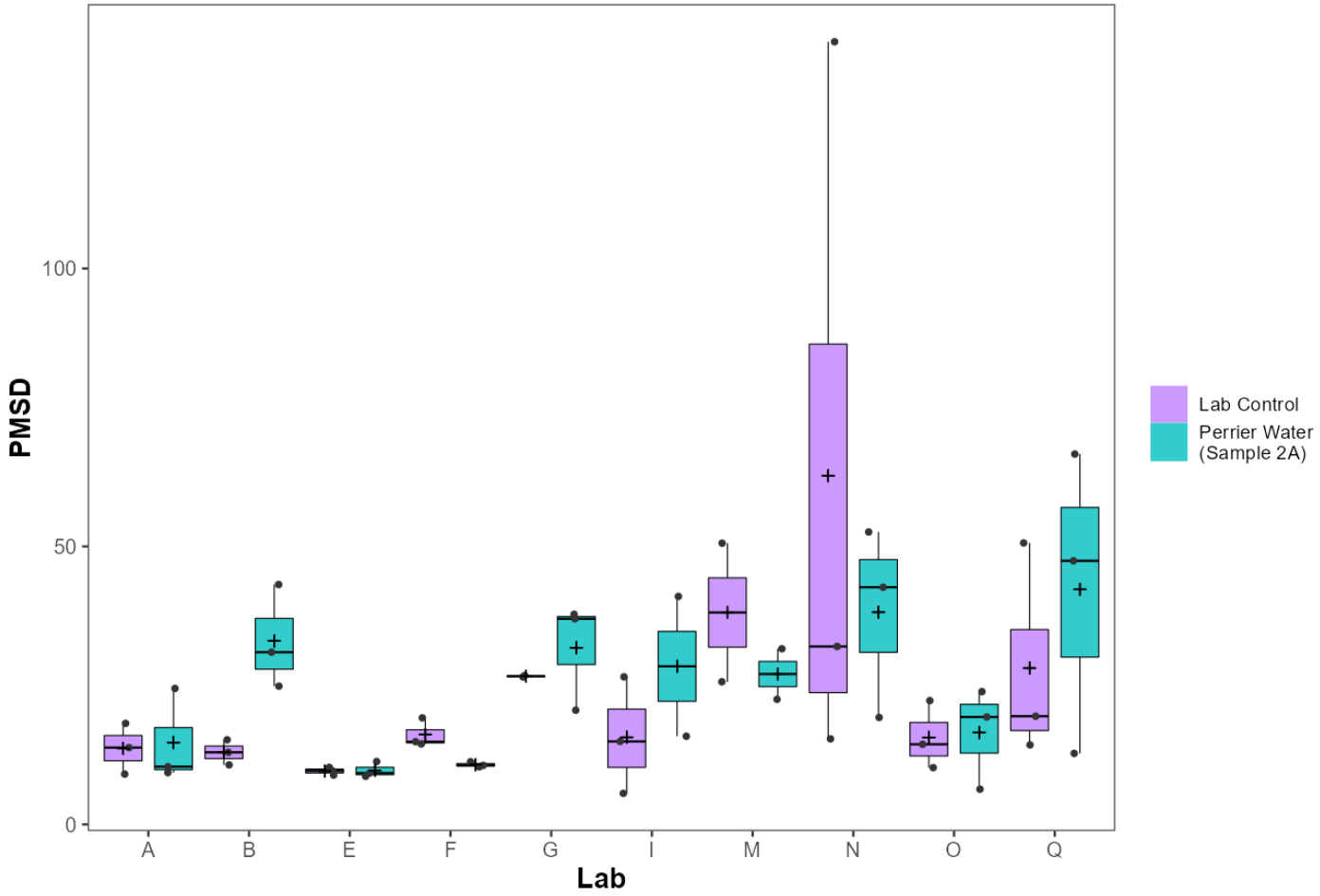


Table C24. Minimum significant difference (MSD)and percent minimum significant difference (PMSD) values for dilution series sample reproduction endpoint. Values exceeding USEPA (2000) 90th percentile value of 37 are highlighted in italics. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

		Round 1		Round 2		Round 3		Round 1		Round 2		Round 3	
		Baseline						Second					
Lab ID	Sample Series	MSD	PMSD	MSD	PMSD	MSD	PMSD	MSD	PMSD	MSD	PMSD	MSD	PMSD
A	2A-2F	9.88	29.1	3.83	8.82	3.35	8.43	9.49	24.5	3.37	10.3	3.74	9.31
A	3 Series	4.38	11.7	4.34	10.0	5.52	14.5	6.91	18.2	5.57	13.8	3.89	9.05
B	2A-2F	NA	NA	9.35	44.1	10.9	***52.8***	7.43	31.0	9.37	24.9	10.8	43.1
B	3 Series	NA	NA	15.4	***49.7***	NA	NA	5.45	13.0	7.41	15.2	4.88	10.7
E	2A-2F	5.55	32.7	5.54	***38.7***	3.91	36.6	2.75	9.24	3.60	11.3	3.09	8.66
E	3 Series	4.54	22.1	5.31	26.5	2.58	30.3	2.82	8.86	3.30	10.3	3.13	9.64
F	2A-2F	7.84	35.2	6.49	31.2	5.86	28.7	3.72	10.4	3.30	10.6	3.36	11.3
F	3 Series	2.82	16.9	3.81	20.0	2.75	16.6	5.04	14.9	5.63	19.2	4.54	14.5
G	2A-2F	4.26	13.4	5.00	15.2	7.09	20.5	9.37	***37.8***	11.0	37.0	8.83	20.5
G	3 Series	6.14	21.5	5.04	17.4	11.6	***39.2***	7.42	26.8	NT	NT	9.29	26.5
I	2A-2F							NA	NA	NA	NA	NA	NA
I	3 Series							NA	NA	8.12	26.6	NA	NA
L	2A-2F	6.27	20.8	15.6	***63.1***	14.2	***64.4***						
L	3 Series	7.78	25.7	7.63	23.9	5.44	18.6						
M	2A-2F	NA	NA	15.7	***42.3***	16.8	***47.5***	NT	NT	8.71	22.5	12.2	31.6
M	3 Series	NA	NA	6.45	***84.8***	8.62	25.4	NT	NT	10.3	25.7	17.7	***50.6***
N	2A-2F	3.59	***102***	8.48	***113***	7.00	***41.9***	4.50	19.2	3.84	***52.6***	8.58	***42.7***
N	3 Series	NA	NA	8.11	25.5	5.06	29.9	3.91	15.4	3.38	***140***	6.40	32.0
O	2A-2F	6.73	22.4	6.16	17.1	7.21	30.4	2.37	6.33	6.81	23.9	6.53	19.3
O	3 Series	7.23	23.9	8.40	26.3	10.2	34.3	5.35	14.4	6.04	22.3	4.16	10.2

		Round 1		Round 2		Round 3		Round 1		Round 2		Round 3	
		Baseline						Second					
Lab ID	Sample Series	MSD	PMSD	MSD	PMSD	MSD	PMSD	MSD	PMSD	MSD	PMSD	MSD	PMSD
P	2A-2F	9.52	26.1	9.18	26.0	4.20	11.0						
P	3 Series	6.33	19.0	6.05	15.8	5.15	14.0						
Q	2A-2F	5.89	14.9	7.24	20.5	6.01	17.2	13.0	***66.6***	3.9	12.8	12.5	***47.4***
Q	3 Series	5.73	14.0	3.47	10.1	3.63	11.0	4.1	14.3	11.9	***50.6***	8.2	19.5

NA=Not available either due to sample not being tested or sample being extremely toxic. NT=Not tested

Table C25. Brood board health parameters recorded in boards that were used to initiate the tests during the baseline and second C. dubia interlaboratory studies. Data is expressed as percentage of brood board cups exhibiting a health issue category per brood board. N values refers to the number of brood boards used by each lab to initiate all their tests. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Unhealthy Adult						Dead Adult					
	Baseline			Second			Baseline			Second		
Lab	Mean	N	MeanRange	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
A	0	5	0-0	0	4	0-0	0	5	0-0	0	4	0-0
B	0	2	0-0	10.8	3	5-20	11.2	2	7.5-15.0	2.5	3	0-7.5
E	0	3	0-0	0	3	0-0	0	3	0-0	0	0	0-0
F	1.1	9	0-10	0.24	7	0-1.7	1.5	9	0-6.7	6.7	7	0-13.3
G	3.3	6	0-20	2.1	6	0-5	0.6	6	0-3.3	14.3	6	3.3-25.0
I				4.2	6	1.7-8.3				17.5	6	6.7-23.3
L	3.3	14	0-10				17.6	14	0-46.7			
M	0	3	0-0	0	5	0-0	5.6	3	0-11.7	7.7	5	1.7-15.0
N	0	7	0-0	1.9	8	0-15.0	11.0	7	0-41.7	1.2	8	0-6.7
O	0	4	0-0	0	5	0-0	1.2	4	0-3.3	2.0	5	0-6.7
P	0.2	5	0-3.3				2.9	5	0-13.3			
Q	0	3	0-0	2.7	5	0-11.7	2.8	3	0-5.0	7.0	5	3.3-15.0

Table C26. Additional brood board health parameters recorded in boards that were used to initiate the tests during the baseline and second C. dubia interlaboratory studies. Data is expressed as percentage of brood board cups exhibiting a health issue category per brood board. N values refers to the number of brood boards used by each lab to initiate all their tests. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Male						Unhealthy Neonates					
	Baseline			Second			Baseline			Second		
Lab	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
A	0	5	0-0	0	4	0-0	0	5	0-0	0	4	0-0
B	0	2	0-0	0	3	0-0	0	2	0-0	0	3	0-0
E	0	3	0-0	0	3	0-0	0	3	0-0	0	3	0-0
F	0	9	0-0	0	7	0-0	0.2	9	0-1.7	0	7	0-0
G	1.1	6	0-6.7	0	6	0-0	0	6	0-0	0	6	0-0
I				0	6	0-0				33.0	6	3.3-73.3
L	1.9	14	0-10.0				0	14	0-0			
M	0	3	0-0	0	5	0-0	9.5	3	1.7-21.7	20.3	5	6.7-31.7
N	0	7	0-0	0	8	0-0	0	7	0-0	0	8	0-0
O	0	4	0-0	0	5	0-0	0	4	0-0	0	5	0-0
P	0	5	0-0				8.0	5	3.3-13.3			
Q	0	3	0-0	0	5	0-0	0.6	3	0-1.7	0	5	0-0

Table C27. Additional brood board health parameters recorded in boards that were used to initiate the tests during the baseline and second *C. dubia* interlaboratory studies. Data is expressed as percentage of brood board cups exhibiting a health issue category per brood board. N values refers to the number of brood boards used by each lab to initiate all their tests. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

	Dead Neonates						Other Abnormalities*		
	Baseline			Second			Second		
Lab	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
A	0	5	0-0	0	4	0-0	0	4	0-0
B	0	2	0-0	0	3	0-0	0	3	0-0
E	0	3	0-0	0	3	0-0	0	3	0-0
F	0.2	9	0-1.7	0	7	0-0	0.2	7	0-1.7
G	0	6	0-0	0	6	0-0	2.1	6	0-10
I				13.6	6	0-36.7	1.1	6	0-3.3
L	0	14	0-0						
M	20.0	3	11.7-33.3	34.3	5	15-51.7	0	5	0-0
N	0	7	0-0	0	8	0-0	0	8	0-0
O	0	4	0-0	0	5	0-0	0.3	5	0-1.7
P	17.3	5	6.7-23.3						
Q	1.1	3	0-1.7	0	5	0-0	0	5	0-0

*This was a new code for the Second ILS.

Table C28. Summary of water treatment methods from the ten laboratories participating in the second *C. dubia* interlaboratory study.

Lab	Source Water	Treatment Step 1	Treatment Step 2	Treatment Step 3	Treatment Step 4	Treatment Step 5	Treatment Step 6	Final DI Water Type
A	Municipal	Carbon cartridge	Ion exchange	Ion exchange	Filtration	Disinfection: UV	N/A	1
B	Municipal	Carbon cartridge	Ion exchange	Ion exchange	Other organic clean-up	Other organic clean-up	N/A	1
E	Municipal	Filtration	Ion exchange	Carbon cartridge	N/A	N/A	N/A	1
F	Municipal	Filtration	Carbon cartridge	Reverse osmosis	Ion exchange	N/A	N/A	2
G	Municipal	Filtration	Carbon cartridge	Reverse osmosis	Ion exchange	N/A	N/A	2
I	Municipal	Carbon cartridge	Reverse osmosis	Ion exchange	Filtration	Disinfection: UV	N/A	1
M	Municipal	Ion exchange	Carbon cartridge	Filtration	Disinfection: UV	Disinfection: 0.2 µm filter	N/A	1
N	Municipal	Filtration	Carbon cartridge x3	Filtration	Reverse osmosis	Ion exchange	Disinfection: UV	2+
O	Municipal	Carbon cartridge	Water softener	Filtration	Reverse osmosis	Ion exchange	N/A	1
Q	Arrowhead Distilled Water	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table C29. Information on the algae used for feeding Ceriodaphnia during ILS testing. Values for the baseline represent a single value reported by the laboratory that may have been measured by the laboratory or the vendor. Values for the second ILS represent the mean of values for the three rounds and were measured by the laboratories in all cases. Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

Lab	Algae Source	Algae Concentration (cells/mL)		
		Baseline	Second ILS (Mean)*	Second ILS (Range)
A	In-house	250,000	277,308	249,709 - 332,388
B	ABS	695,600	429,600	187,200 – 550,800
E	ABS	233,333	211,000	200,000 – 233,000
F	ARO	233,333	233,333	233,333 – 233,333
G	ABS	200,000	219,667	216,000 – 226,333
I			235,800	233,400 – 247,000
L	ABS	220,000		
M	ABS	210,000	220,000	209,000 – 231,000
N	In-house	213,000	214,000	213,000 – 216,000
O	In-house	245,000	252,000	245,000 – 255,000
P	ABS	300,000		
Q	ABS	215,000	222,667	220,000 – 228,000

*200,000 to 230,000 cells/ml are recommended in the EPA manual.

Table C30. Information on the YCT used for feeding Ceriodaphnia during ILS testing. Note that for the baseline, feeding data was provided only on a volume-to-volume basis (volume of YCT stock to volume in test chamber). Lab I did not participate in the Baseline and labs L and P did not participate in the Second ILS.

Lab	YCT Source	YCT Recipe	Feeding Method	YCT Volume In Test Chamber (ml/ml)	YCT Volume In Test Chamber (ml/ml)	YCT Mass In Test Chambers (mg/l)
				Baseline	Second	Second
A	ARO	Fleishman's Yeast+Blue Seal Alfalfa+Zeigler #1 Finfish Crumble Trout Chow	In test solution	0.0075	0.0088	0.012
B	ABS	NA	In test solution	0.0168	0.012	0.014
E	ABS	NA	Direct addition	0.0067	0.0067	0.012
F	In-house	Fleischmann's Yeast + Pines Wheatgrass + Thomas Fish Co Trout Chow	Direct addition	0.0067	0.0067	0.010
G	ABS	NA	Direct addition	0.0067	0.0089	0.015
I	In-house	Fleischmann's Yeast + Frontier Co-0p Wheeatgrass + Skretting Trout Chow	In test solution		0.007	0.013
L	ABS	NA	Direct addition	0.0067		
M	ARO	NA	Direct addition	0.0067	0.0067	0.012
N	In-house	Trout Chow (Purina Aquamax Fry Starter 100) / Carolina Daphnia Food (4 oz) + Fleischmann's baker's yeast (one pouch 7 grams) + Cerophyl (Wards Cereal Grass Media)	Direct addition	0.005	0.0067	0.012

Lab	YCT Source	YCT Recipe	Feeding Method	YCT Volume In Test Chamber (ml/ml)	YCT Volume In Test Chamber (ml/ml)	YCT Mass In Test Chambers (mg/l)
				Baseline	Second	Second
O	In-house	Fleishman's active dry yeast + Pines Wheatgrass + Purina Trout Chow (supplied by ABS)	In test solution	0.007	0.007	0.013
P	ARO	NA	In test solution	0.0067		
Q	ABS	NA	Direct addition	0.0067	0.0067	0.012

NA= Not applicable, lab purchases YCT.

Water quality data

Table C31. Summary of conductivity and pH data collected in control chambers from the ten laboratories participating in the second *C. dubia* interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions). N values refer to the number of tests conducted and included in the means.

		Conductivity (µS/cm) Before			Conductivity (µS/cm) After			pH: Before			pH: After		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
A	Lab Water	313	12	309-317	289	12	288-291	7.66	12	7.51-7.74	7.97	12	7.90-8.06
A	DMW Perrier® (2A)	205	3	203-208	184	3	182-186	7.83	3	7.78-7.88	7.79	3	7.62-7.92
A	MHW (1)	398	3	373-417	374	3	347-395	7.55	3	7.47-7.64	7.77	3	7.70-7.82
B	Lab Water	380	12	358-433	376	12	358-388	7.51	12	7.21-7.64	6.90	12	6.73-7.10
B	DMW Perrier® (2A)	235	3	234-237	218	3	220-230	7.58	3	7.51-7.69	6.75	3	6.53-6.96
B	MHW (1)	414	3	399-442	407	3	377-450	7.41	3	7.25-7.49	6.98	3	6.87-7.11
E	Lab Water	370	12	368-371	369	12	368-370	7.98	12	7.95-8.00	8.00	12	7.99-8.01
E	DMW Perrier® (2A)	236	3	224-250	227	3	213-243	7.95	3	7.89-8.00	7.95	3	7.91-7.99
E	MHW (1)	405	3	367-471	399	3	378-441	7.93	3	7.89-7.95	7.99	3	7.91-8.06
F	Lab Water	209	12	203-223	192	12	189-194	8.22	12	8.18-8.30	8.16	12	0.04-0.10

		Conductivity (µS/cm) Before			Conductivity (µS/cm) After			pH: Before			pH: After		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
F	DMW Perrier® (2A)	198	3	194-203	177	3	174-181	8.23	3	8.18-8.30	7.87	3	7.67-8.11
F	MHW (1)	380	3	368-392	362	3	337-378	8.07	3	8.03-8.14	7.95	3	7.94-7.96
G	Lab Water	198	11	194-218	184	11	182-186	8.16	11	8.10-8.26	8.17	11	8.09-8.24
G	DMW Perrier® (2A)	199	3	198-199	187	3	184-190	8.17	3	8.10-8.28	7.70	3	7.56-7.94
G	MHW (1)	380	3	340-406	381	3	351-403	8.07	3	8.05-8.12	7.81	3	7.73-7.87
I	Lab Water	403	12	361-583	338	12	324-358	7.91	12	7.83-7.99	7.89	12	7.78-7.94
I	DMW Perrier® (2A)	204	3	198-217	179	3	177-181	8.01	3	7.96-8.04	7.67	3	7.46-7.90
I	MHW (1)	406	2	391-422	381	3	337-433	7.88	3	7.81-7.99	7.82	3	7.74-7.92
M	Lab Water	357	8	350-367	329	8	326-334	8.25	8	8.21-8.31	8.03	8	7.98-8.12
M	DMW Perrier® (2A)	200	2	197-203	182	2	179-186	8.35	2	8.34-8.36	7.74	2	7.74-7.74
M	MHW (1)	418	2	410-425	379	2	371-387	8.17	2	8.17-8.18	7.95	2	7.90-8.00
N	Lab Water	291	12	274-309	271	12	260-292	8.15	12	8.05-8.22	8.15	12	8.05-8.30
N	DMW Perrier® (2A)	179	3	170-185	161	3	155-169	8.21	3	8.18-8.27	8.11	3	8.01-8.24
N	MHW (1)	333	3	284-361	328	3	288-350	8.13	3	8.12-8.14	7.96	3	7.85-8.12

		Conductivity (µS/cm) Before			Conductivity (µS/cm) After			pH: Before			pH: After		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
O	Lab Water	353	12	347-358	341	12	334-348	7.71	12	7.63-7.76	7.87	12	7.81-7.92
O	DMW Perrier® (2A)	198	3	195-203	188	3	186-190	7.91	3	7.89-7.92	7.65	3	7.46-7.93
O	MHW (1)	389	3	363-393	374	3	350-396	7.74	3	7.72-7.76	7.81	3	7.80-7.82
Q	Lab Water	213	12	209-221	191	12	191-191	8.31	12	8.14-8.44	8.22	12	8.07-8.34
Q	DMW Perrier® (2A)	203	3	202-204	187	3	185-189	8.34	3	8.21-8.42	7.73	3	7.59-7.94
Q	MHW (1)	410	3	366-441	381	3	353-403	8.17	3	8.04-8.23	7.87	3	7.82-7.90

Table C32. Summary of dissolved oxygen (DO) and water temperature (water temp) data collected in control chambers from the ten laboratories participating in the second *C. dubia* interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions). N values refer to the number of tests conducted and included in the means.

		DO (mg/L) Before			DO (mg/L) After			Water Temp (°C) Before			Water Temp (°C) After		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
A	Lab Water	6.95	12	6.62-7.15	7.63	12	7.37-7.90	24.3	12	24.1-25.0	24.3	12	24.0-24.5
A	DMW Perrier®(2A)	6.92	3	6.82-6.98	7.89	3	7.77-8.01	24.3	3	24.2-24.3	24.3	3	24.0-24.4
A	MHW (1)	6.44	3	6.21-6.56	7.85	3	7.64-8.01	24.7	3	24.1-25.1	24.5	3	24.2-24.8
B	Lab Water	8.59	12	8.38-8.91	8.57	12	8.29-8.79	25.5	12	25.3-25.6	25.5	12	25.2-25.6
B	DMW Perrier®(2A)	8.68	3	8.62-8.73	9.61	3	8.98-10.10	25.4	3	25.3-25.5	25.4	3	25.2-25.6
B	MHW (1)	8.63	3	8.46-8.79	9.41	3	9.10-9.94	25.5	3	25.3-25.6	25.5	3	25.2-25.5
E	Lab Water	7.37	12	7.34-7.41	7.93	12	7.89-8.00	24.7	12	24.5-24.7	24.8	12	24.8-24.9
E	DMW Perrier®(2A)	7.40	3	7.29-7.51	7.94	3	7.87-8.01	24.7	3	24.6-24.8	24.8	3	24.8-24.8
E	MHW (1)	7.41	3	7.31-7.51	7.98	3	7.91-8.04	24.7	3	24.6-24.7	24.8	3	24.8-24.8
F	Lab Water	8.25	12	8.05-8.43	8.05	12	7.83-8.14	24.8	12	24.3-25.6	24.9	12	24.7-25.0
F	DMW Perrier®(2A)	8.09	3	8.06-8.12	8.83	3	8.77-8.87	25.1	3	24.5-25.6	25.0	3	24.8-25.3

		DO (mg/L) Before			DO (mg/L) After			Water Temp (°C) Before			Water Temp (°C) After		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
F	MHW (1)	8.15	3	8.09-8.21	8.86	3	8.77-8.94	24.9	3	24.6-25.1	25.1	3	24.9-25.3
G	Lab Water	7.96	11	7.66-8.31	8.18	11	8.11-8.26	24.8	11	24.6-25.0	24.4	11	24.2-24.6
G	DMW Perrier®(2A)	8.05	3	7.84-8.28	8.89	3	8.74-8.98	24.9	3	24.7-25.0	24.8	3	24.6-25.0
G	MHW (1)	7.91	3	7.61-8.16	8.95	3	8.85-9.05	24.9	3	24.8-25.0	24.6	3	24.6-24.8
I	Lab Water	6.74	12	6.07-7.60	8.32	12	7.63-9.10	24.8	12	24.5-25.4	24.9	12	24.4-25.2
I	DMW Perrier®(2A)	6.76	3	6.62-6.99	9.15	3	8.87-9.66	24.7	3	24.6-24.8	24.8	3	24.8-24.9
I	MHW (1)	6.51	2	6.22-7.05	9.33	3	8.81-9.60	25.0	3	24.5-25.5	24.9	3	24.6-25.2
M	Lab Water	6.64	8	6.50-6.74	7.27	8	7.11-7.40	22.9	8	22.1-24.1	24.8	8	24.3-25.5
M	DMW Perrier®(2A)	6.61	2	6.55-6.66	7.73	2	7.48-7.99	22.1	2	21.6-22.6	24.7	2	24.6-24.7
M	MHW (1)	6.58	2	6.54-6.62	7.69	2	7.36-8.02	22.8	2	22.8-22.9	24.6	2	24.4-24.7
N	Lab Water	8.42	12	8.12-8.59	8.44	12	8.33-8.61	24.9	12	24.5-25.3	25.2	12	24.6-25.6
N	DMW Perrier®(2A)	8.38	3	8.20-8.54	8.44	3	8.31-8.57	24.9	3	24.7-25.1	24.6	3	24.5-24.8
N	MHW (1)	8.39	3	8.22-8.49	8.41	3	8.34-8.51	24.9	3	24.7-25.1	24.8	3	24.5-25.0
O	Lab Water	7.74	12	7.54-7.86	8.92	12	8.46-9.33	24.7	12	24.4-25.0	24.7	12	24.3-25.0

		DO (mg/L) Before			DO (mg/L) After			Water Temp (°C) Before			Water Temp (°C) After		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
O	DMW Perrier®(2A)	7.77	3	7.69-7.90	10.45	3	9.77-10.93	24.8	3	24.6-25.0	24.9	3	24.7-25.1
O	MHW (1)	7.77	3	7.56-7.90	10.05	3	9.86-10.26	24.8	3	24.6-24.9	24.7	3	24.4-24.9
Q	Lab Water	8.27	12	8.22-8.35	7.87	12	7.83-7.90	24.2	12	23.9-24.3	25.6	12	25.4-25.8
Q	DMW Perrier®(2A)	8.24	3	8.24-8.25	8.52	3	8.47-8.57	24.1	3	24.0-24.3	25.3	3	25.2-25.4
Q	MHW (1)	8.25	3	8.24-8.27	8.51	3	8.46-8.59	24.1	3	24.1-24.1	25.4	3	25.3-25.4

Table C33. Summary of air temperature data collected in control chambers from the ten laboratories participating in the second *C. dubia* interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions). N values refer to the number of tests conducted and included in the means.

		(°C) Before			(°C) After		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range
A	Lab Water	25.2	12	25.0-25.4	25.1	12	25.0-25.4
A	DMW Perrier®(2A)	25.2	3	25.1-25.3	25.2	3	25.1-25.2
A	MHW (1)	25.3	3	25.2-25.5	25.2	3	25.1-25.2
B	Lab Water	24.2	12	23.8-24.4	24.2	12	23.9-24.4
B	DMW Perrier®(2A)	24.2	3	23.9-24.4	24.2	3	23.9-24.4
B	MHW (1)	24.2	3	23.8-24.4	24.2	3	23.9-24.4
E	Lab Water	25.0	12	25.0-25.0	25.0	12	25.0-25.0
E	DMW Perrier®(2A)	25.0	3	25.0-25.0	25.0	3	25.0-25.0
E	MHW (1)	25.0	3	25.0-25.0	25.0	3	25.0-25.0
F	Lab Water	25.4	12	25.1-25.9	25.5	12	25.0-26.0
F	DMW Perrier®(2A)	25.4	3	25.1-25.8	25.5	3	25.0-25.9
F	MHW (1)	25.4	3	25.1-25.8	25.5	3	25.0-25.9
G	Lab Water	25.6	11	25.3-26.0	25.3	11	24.8-25.9
G	DMW Perrier®(2A)	25.5	3	25.3-25.9	25.3	3	24.8-25.9

		(°C) Before			(°C) After		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range
G	MHW (1)	25.7	3	25.3-25.9	25.3	3	24.8-25.9
I	Lab Water	26.7	12	26.1-27.0	26.7	12	26.4-26.9
I	DMW Perrier®(2A)	26.7	3	26.1-27.0	26.7	3	26.5-26.9
I	MHW (1)	26.5	3	25.9-27.0	26.7	3	26.6-26.9
M	Lab Water	24.9	8	24.2-26.1	25.5	8	24.9-26.0
M	DMW Perrier®(2A)	24.6	2	24.2-25.0	25.6	2	25.4-25.8
M	MHW (1)	24.7	2	24.4-25.0	25.6	2	25.4-25.8
N	Lab Water	25.2	12	25.0-25.7	25.4	12	24.7-25.8
N	DMW Perrier®(2A)	25.2	3	25.0-25.6	25.4	3	25.1-25.7
N	MHW (1)	25.2	3	25.0-25.6	25.5	3	25.1-25.7
O	Lab Water	25.2	12	25.0-25.3	25.2	12	25.0-25.3
O	DMW Perrier®(2A)	25.2	3	25.2-25.3	25.2	3	25.2-25.3
O	MHW (1)	25.2	3	25.4-25.5	25.2	3	25.0-25.3
Q	Lab Water	25.3	12	25.1-25.5	25.3	12	25.1-25.5
Q	DMW Perrier®(2A)	25.4	3	25.1-25.5	25.3	3	25.1-25.5
Q	MHW (1)	25.3	3	25.1-25.5	25.3	3	25.4-25.6

Table C34. Summary of hardness and alkalinity data collected in control chambers from the ten laboratories participating in the second *C. dubia* interlaboratory study. Water quality parameters are presented in two categories (before and after renewal of test solutions). N values refer to the number of tests conducted and included in the means.

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range
A	Lab Water	74	12	60-82	69.6	12	58-92
A	DMW Perrier®(2A)	82.3	3	69-91	90.3	3	88-93
A	MHW (1)	110	3	67-132	79.6	3	59-110
B	Lab Water	85.3	12	78-96	62	12	60-64
B	DMW Perrier®(2A)	85.3	3	82-88	79.3	3	78-80
B	MHW (1)	116.7	3	108-126	56	3	52-62
E	Lab Water	99.3	12	98-100	63.6	12	60-70
E	DMW Perrier®(2A)	103	3	100-107	87	3	85-90
E	MHW (1)	136.7	3	130-145	61.3	3	54-70
F	Lab Water	91.2	12	86-95	91.5	12	88-99
F	DMW Perrier®(2A)	87	3	84-89	83	3	80-89
F	MHW (1)	119	3	113-129	61	3	57-64
G	Lab Water	79.8	6	73-85	77.8	7	73-82
G	DMW Perrier®(2A)	91.3	6	81-104	77.7	4	72-82
G	MHW (1)	114.3	6	85-144	66.8	6	54-90
I	Lab Water	89.3	12	85-93	71.6	12	68-74
I	DMW Perrier®(2A)	93	3	84-98	66	3	56-79
I	MHW (1)	129.3	3	118-139	45	3	34-58
M	Lab Water	102	4	98-108	70.5	4	66-78
M	DMW Perrier®(2A)	90.5	4	86-94	82	3	80-84
M	MHW (1)	131	2	130-132	60	3	58-62
N	Lab Water	85.0	12	81-88.5	59.5	12	57-63
N	DMW Perrier®(2A)	86.3	3	82-91	75.3	3	73-78

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range
N	MHW (1)	121	3	109-124	53.6	3	52-55
O	Lab Water	91.6	12	84-97	63.3	12	61-67
O	DMW Perrier®(2A)	91.6	3	88-96	85.3	3	80-94
O	MHW (1)	127.6	3	112-144	64.3	3	56-69
Q	Lab Water	92	10	91-93	89.1	10	86-91
Q	DMW Perrier®(2A)	87.7	3	87-88	92	1	92
Q	MHW (1)	124	3	111-133	63	3	61-65

Table C35. Summary of water quality data collected from the brood boards used to initiate the tests during the second *C. dubia* interlaboratory study. The data is divided into two categories: 'before' defined as water in test chambers for 24 hours, and 'after' defined as water quality measurements recorded after renewal in the test chambers. N values refer to the number of water quality measurements of the brood boards reported by the laboratories.

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)			Conductivity (µS/cm)			pH		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
A	Before	NM	-	-	NM	-	-	307	4	304-309	7.6	4	7.6-7.7
A	After	75	3	64-81	69	3	60-84	298	3	295-300	7.7	3	7.6-7.9
B	Before	NM	-	-	NM	-	-	366	3	360-378	7.2	3	7.2-7.4
B	After	86	6	78-96	63	6	56-70	361	3	353-368	7.0	3	6.9-7.1
E	Before	100	2	100-100	68	2	66-70	366	3	364-368	8.0	3	8.0-8.1
E	After	99	3	98-100	62	3	60-66	364	3	362-365	8.0	3	7.9-8.1
F	Before	101	7	96-112	96	7	83-111	210	7	207-214	8.2	7	8.1-8.3
F	After	90	9	86-94	91	9	87-93	195	7	194-197	8.2	7	8.1-8.2
G	Before	81	1	81	73	1	73	199	6	196-210	8.2	6	8.1-8.2
G	After	81	3	81-82	77	4	68-82	186	4	183-194	8.2	4	8.1-8.2
I	Before	NM	-	-	NM	-	-	367	6	361-375	7.9	6	7.8-8.0

		Hardness (mg/L $CaCO_3$)			Alkalinity (mg/L $CaCO_3$)			Conductivity (µS/cm)			pH		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
I	After	93	3	90-97	72	3	55-87	342	6	340-343	7.9	6	7.8-8.0
M	Before	NM	-	-	NM	-	-	363	5	311-411	8.3	5	8.2-8.4
M	After	101	4	98-105	68	4	65-70	322	5	310-326	8.2	5	8.1-8.2
N	Before	NM	-	-	NM	-	-	277	8	264-297	8.0	8	8.0-8.4
N	After	85	7	83-87	59	7	58-60	269	5	255-286	8.1	5	7.9-8.3
O	Before	101	4	99-103	66	4	62-68	356	5	353-362	7.7	5	7.6-7.8
O	After	96	5	93-99	63	5	61-63	340	5	334-348	7.8	5	7.8-7.9
Q	Before	97	4	82-100	93	3	86-97	211	4	203-225	8.3	4	8.1-8.4
Q	After	91	7	79-98	87	7	83-92	187	8	182-193	8.3	8	8.1-8.4

Table C35. continued.

		Dissolved Oxygen (mg/L)			Water Temperature (°C)			Air Temperature (°C)		
Lab	Sample Type	Mean	N	Mean Range	Mean	N	Mean Range	Mean	N	Mean Range
A	Before	7.0	4	6.8-7.2	24.3	4	24.0-25.0	26.6	4	26.5-26.6
A	After	7.4	4	7.0-7.7	24.2	4	24.0-24.4	26.6	4	26.5-26.7
B	Before	8.8	3	8.8-8.9	25.4	3	25.2-25.5	25.5	3	25.4-25.5
B	After	8.8	3	8.6-9.1	25.4	3	25.2-25.5	25.5	3	25.4-25.5
E	Before	7.9	3	7.8-7.9	24.8	3	24.7-24.9	24.9	3	24.7-25.0
E	After	7.9	3	7.8-8.0	24.8	3	24.6-24.8	24.9	3	24.9-25.0
F	Before	8.2	7	8.2-8.3	24.4	7	24.2-24.6	25.3	7	25.2-25.4
F	After	8.0	7	7.9-8.0	25.0	7	24.8-25.3	25.3	7	25.2-25.3
G	Before	7.8	6	7.5-8.1	24.8	6	24.6-25.1	25.8	6	25.1-26.3
G	After	8.3	6	8.2-8.5	24.6	6	24.2-24.9	25.7	6	25.1-26.1
I	Before	7.0	6	6.4-7.4	24.9	6	24.8-25.1	25.5	6	25.2-25.7
I	After	7.8	6	7.6-7.8	24.9	6	24.8-25.0	25.3	6	25.0-25.7
M	Before	6.8	5	6.4-6.9	25.6	5	24.4-27.2	25.3	5	24.6-25.9
M	After	6.8	5	6.7-7.1	24.8	5	24.6-25.0	24.4	5	23.7-25.0
N	Before	8.2	8	8.2-8.3	24.5	8	24.3-24.9	25.0	8	24.6-25.4
N	After	8.2	8	8.1-8.4	24.7	8	24.4-25.0	25.0	8	24.7-25.6
O	Before	7.7	5	7.5-8.0	24.9	5	24.7-25.2	25.3	5	25.2-25.4
O	After	8.5	5	8.1-8.8	25.0	5	24.6-25.4	25.2	5	25.1-25.3
Q	Before	8.4	4	8.3-8.4	24.0	4	23.9-24.3	25.0	4	24.9-25.3
Q	After	7.9	4	7.9-8.0	25.3	4	25.0-25.7	25.0	4	24.9-25.2

Concentration-response data

Figure C12. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab A.

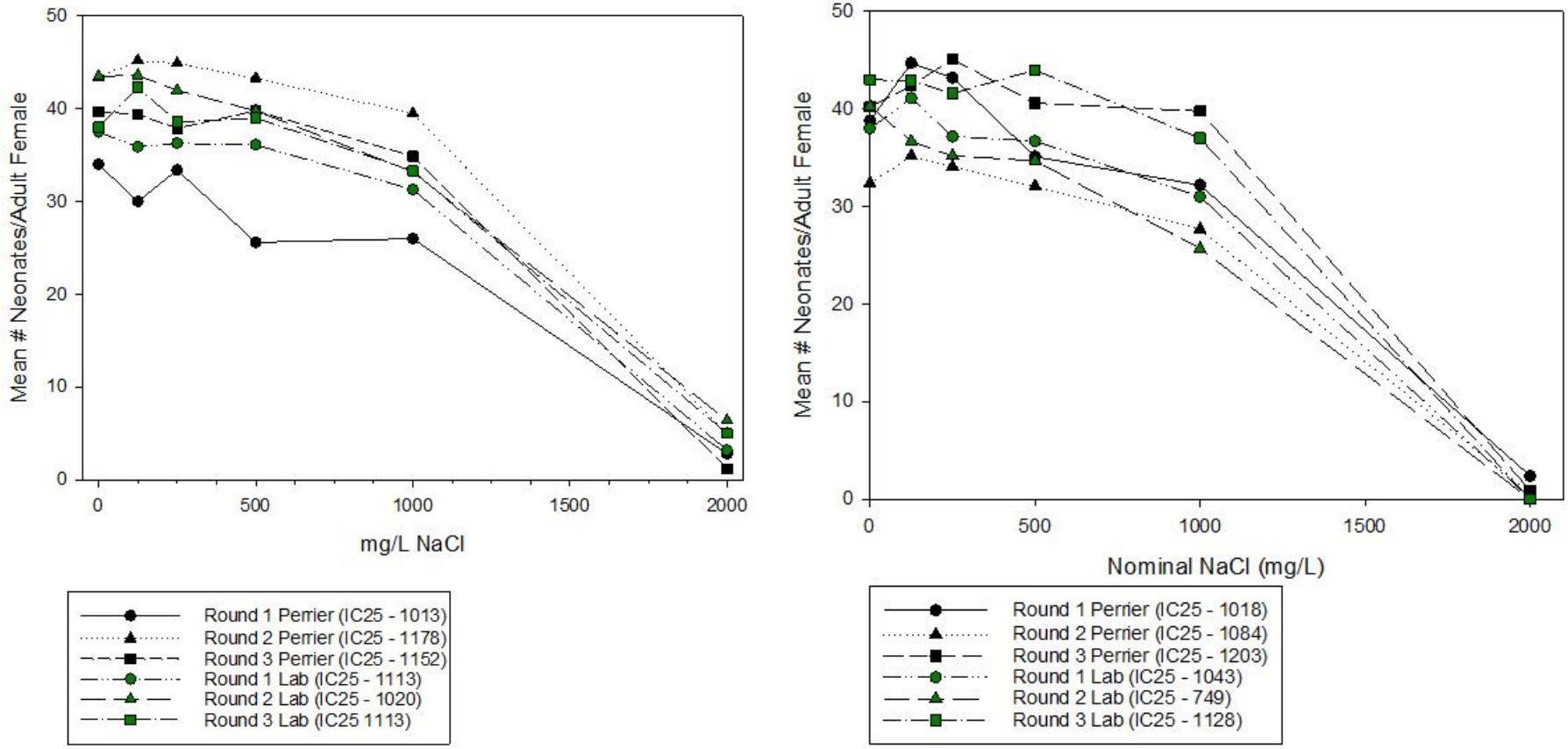


Figure C13. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab B.

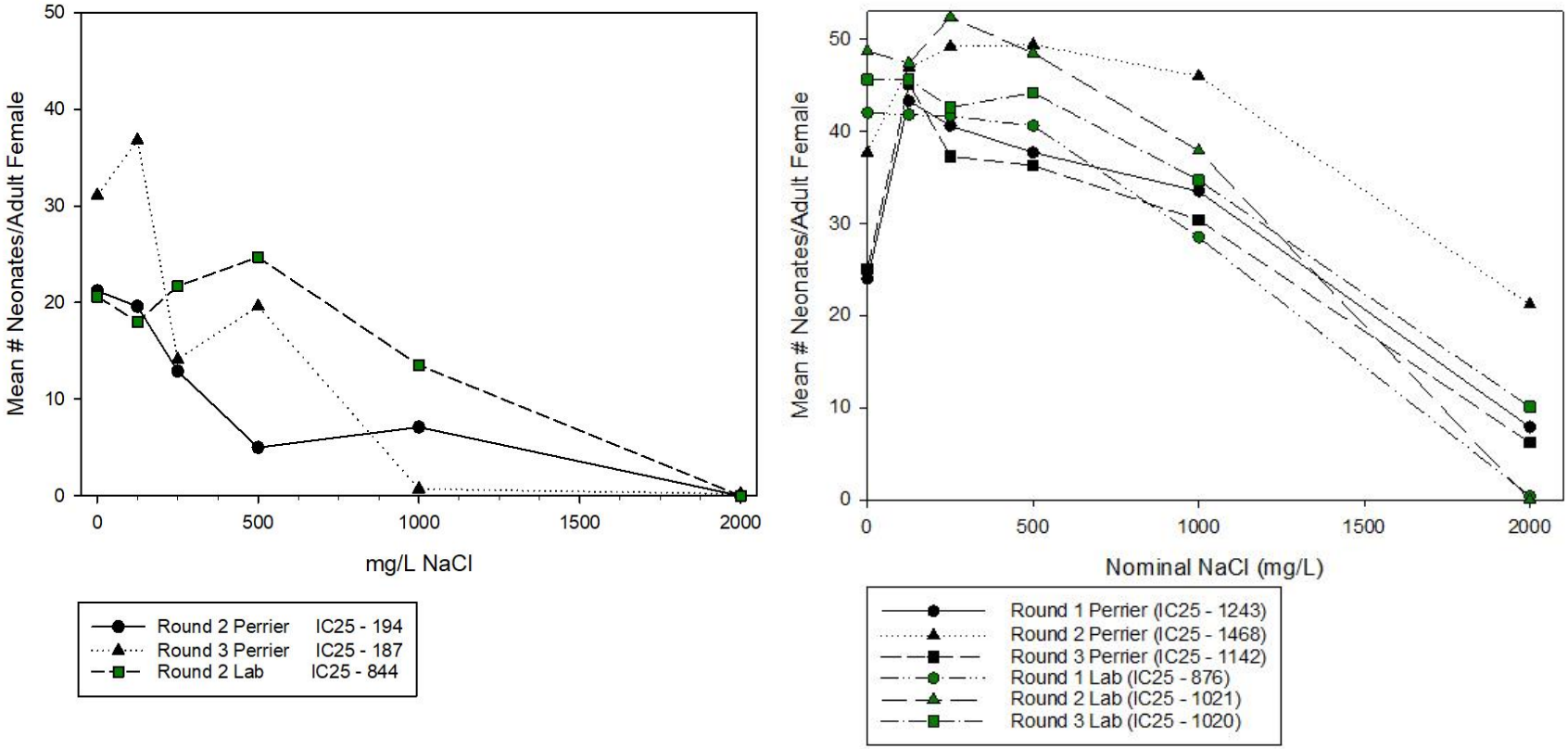


Figure C14. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab E.

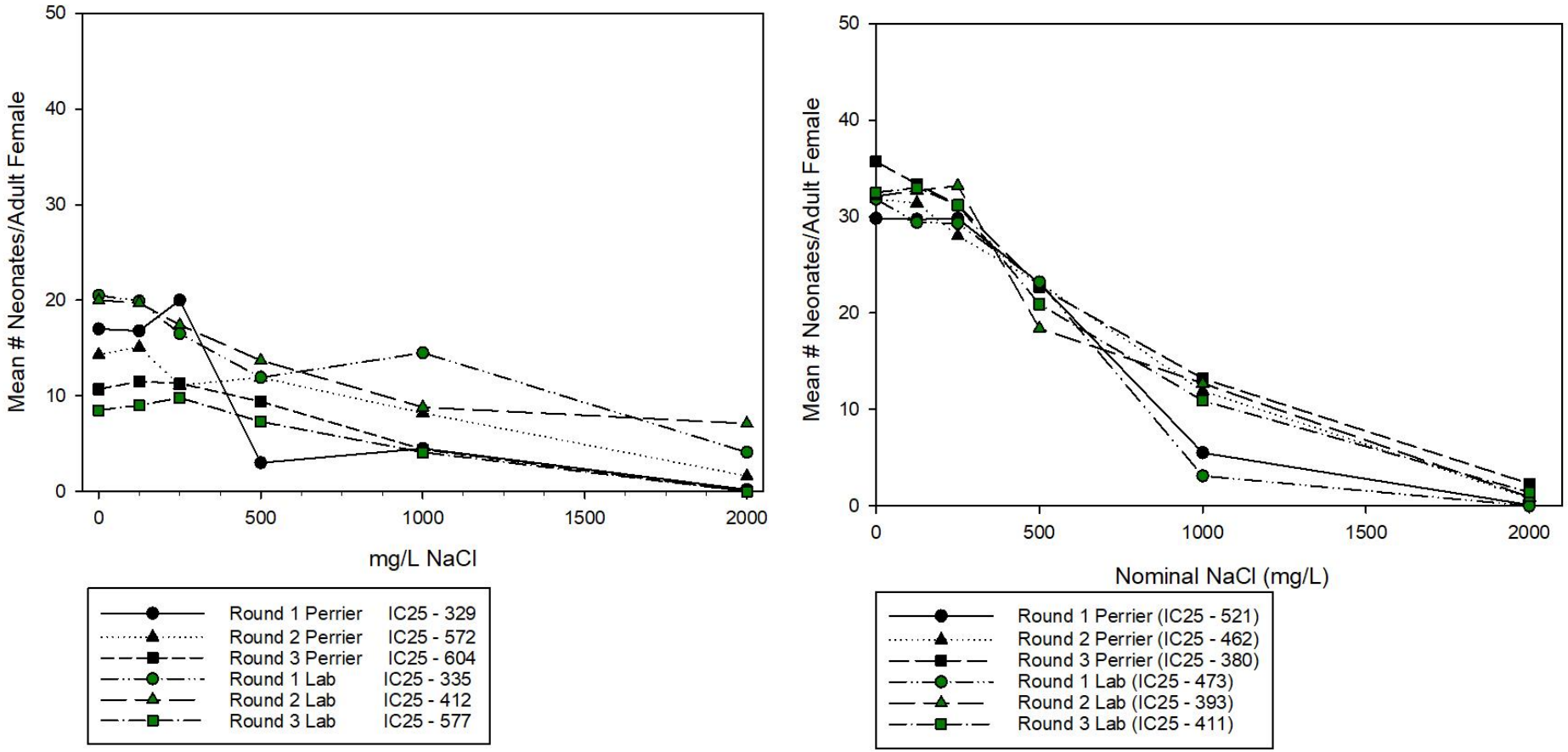
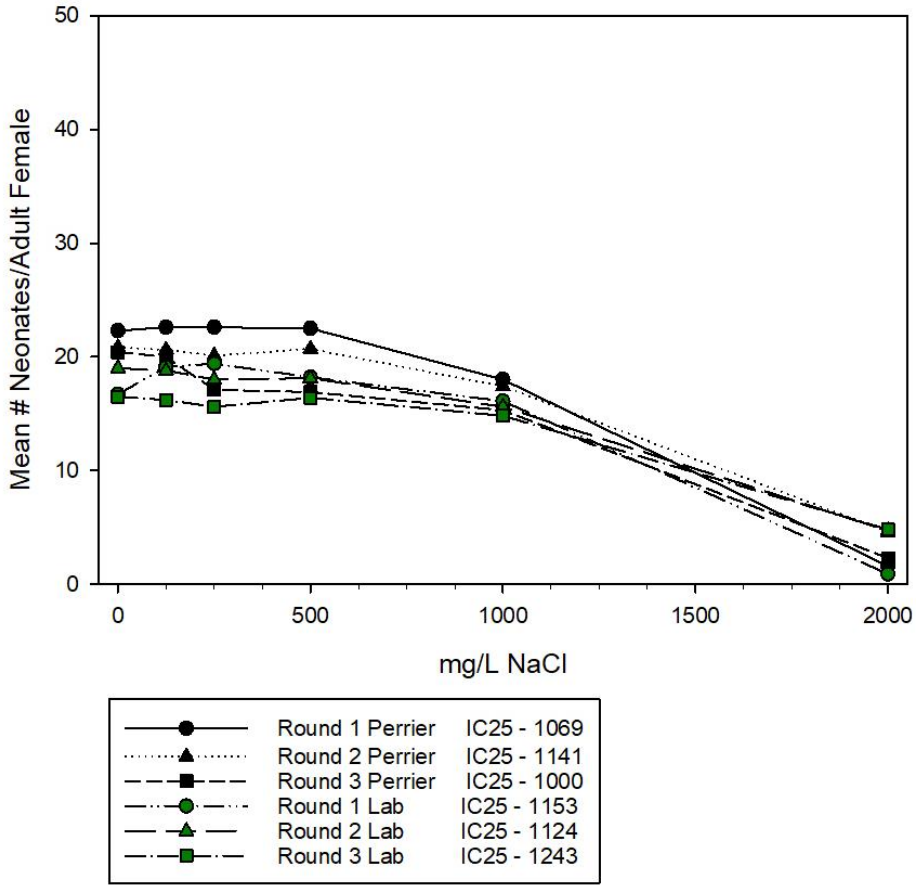


Figure C15. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab F.



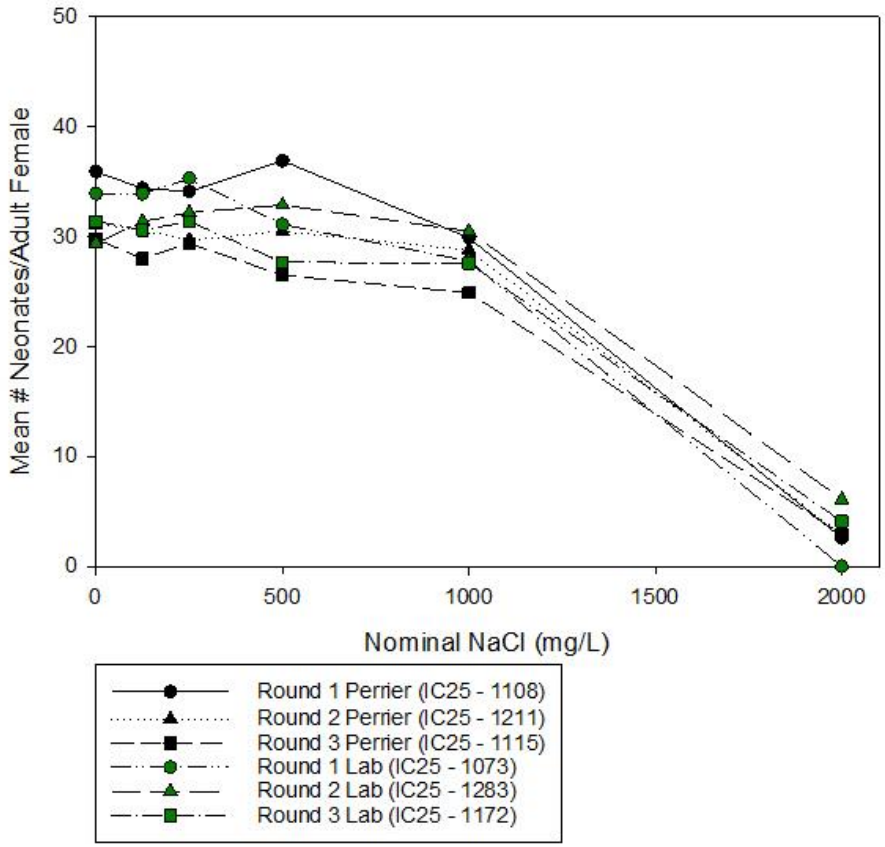


Figure C16. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab G.

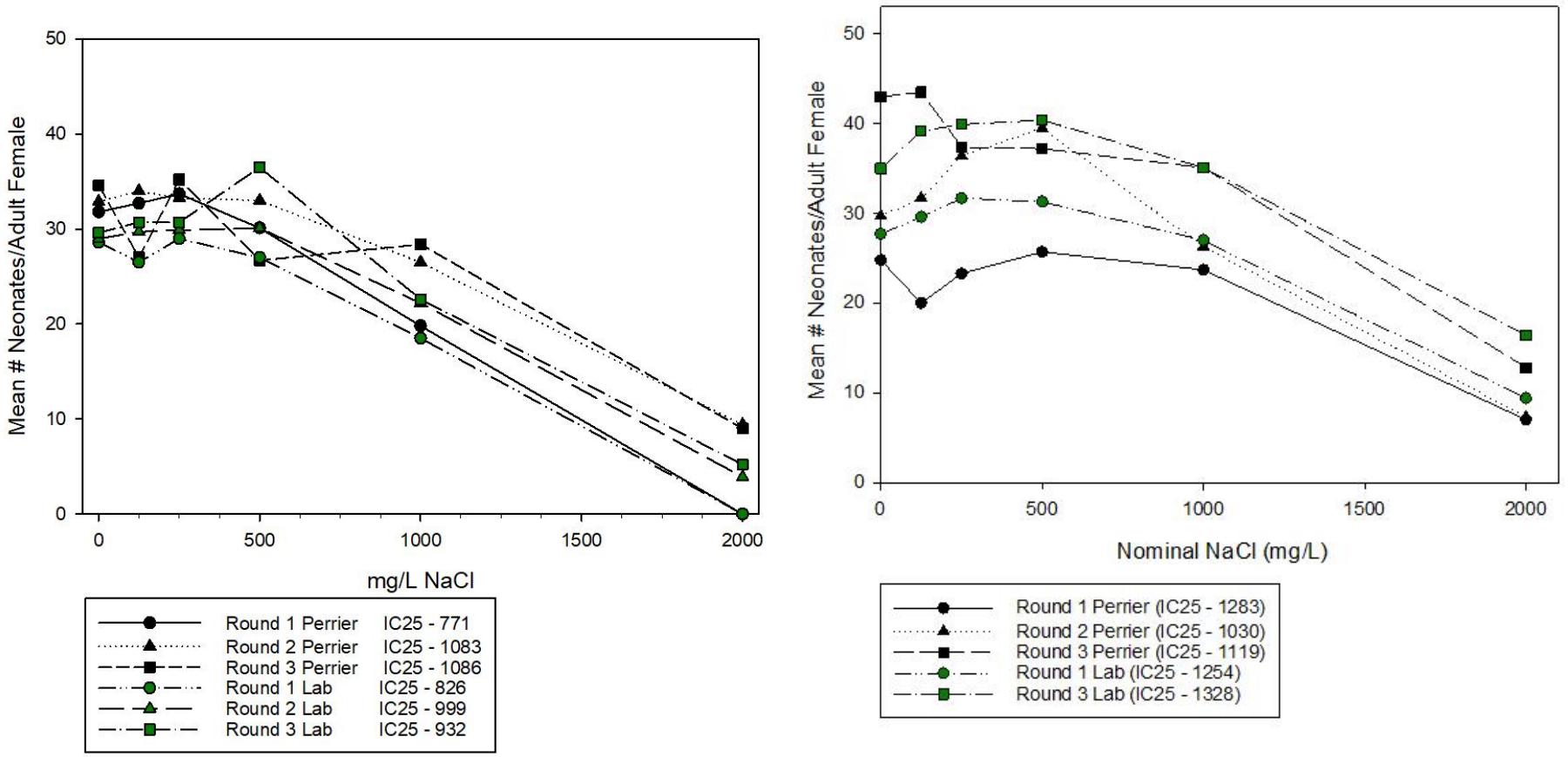


Figure C17: Second ILS NaCl Concentration Responses for Lab I. Reproduction in all the Perrier based concentrations was very low, including the unspiked sample. This lab did not participate in the Baseline ILS.

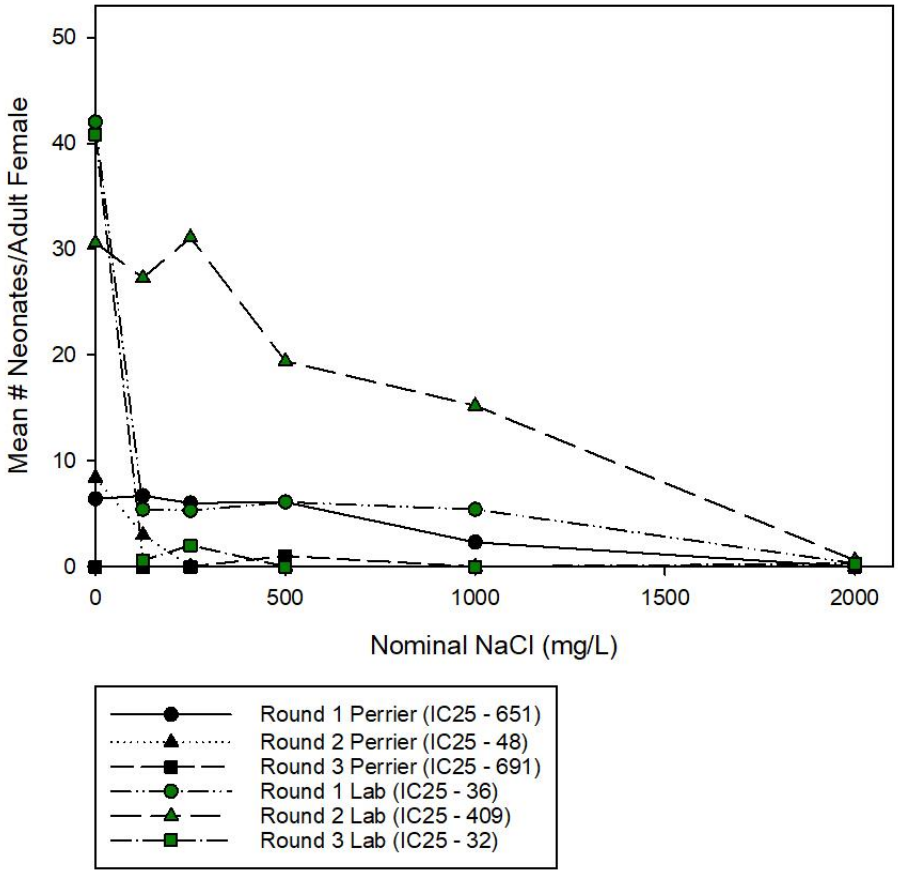
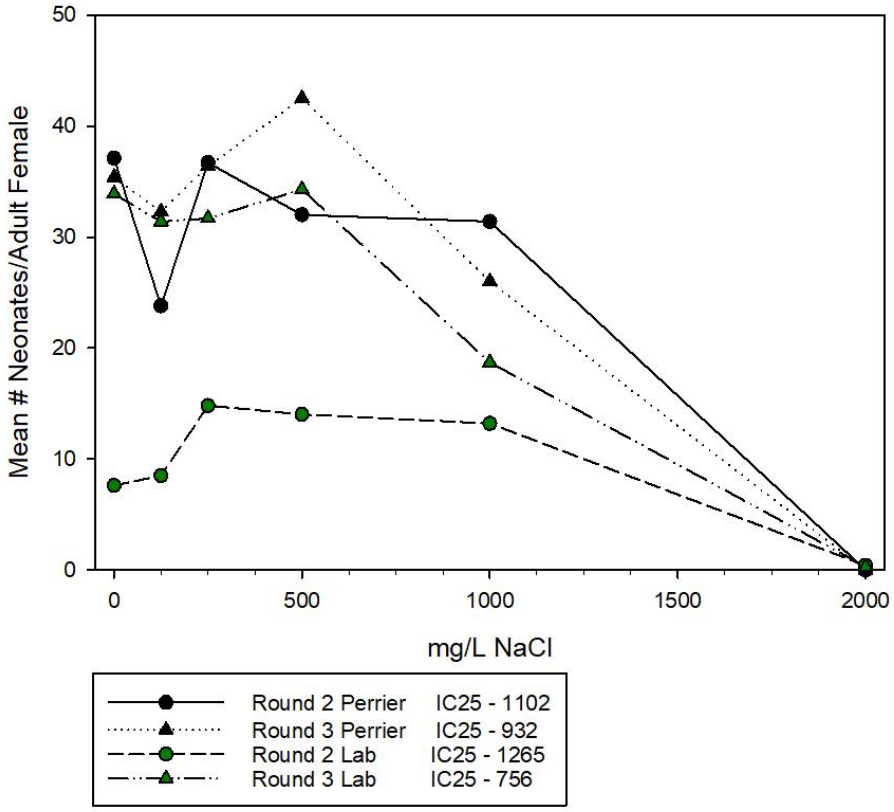


Figure C18. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab M.



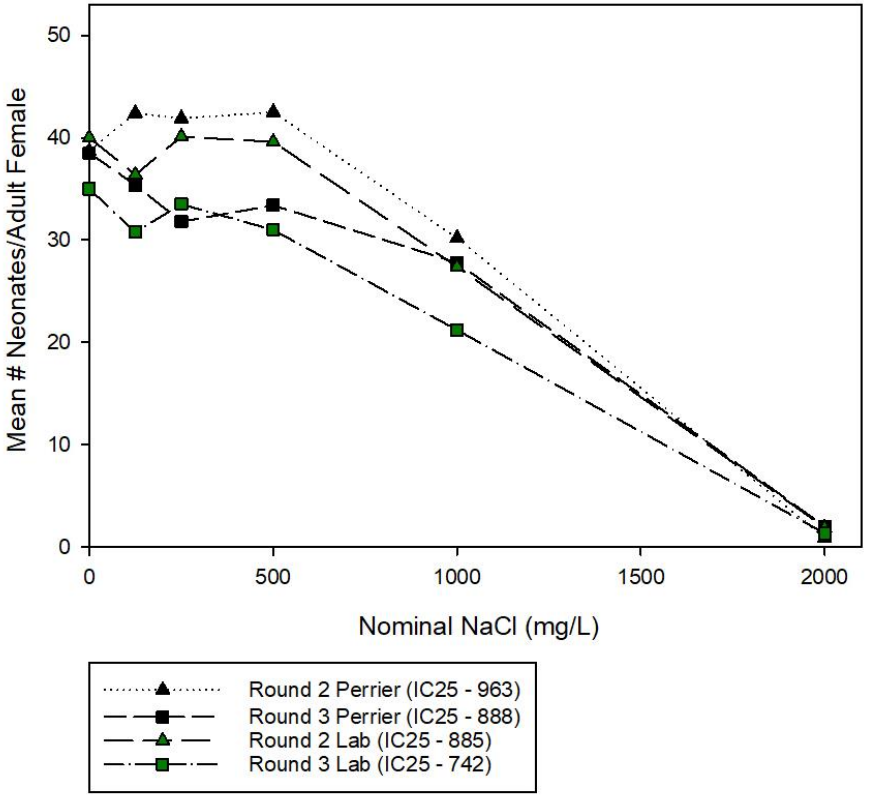


Figure C19. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab N. This lab reported issues with their culture before Round 2.

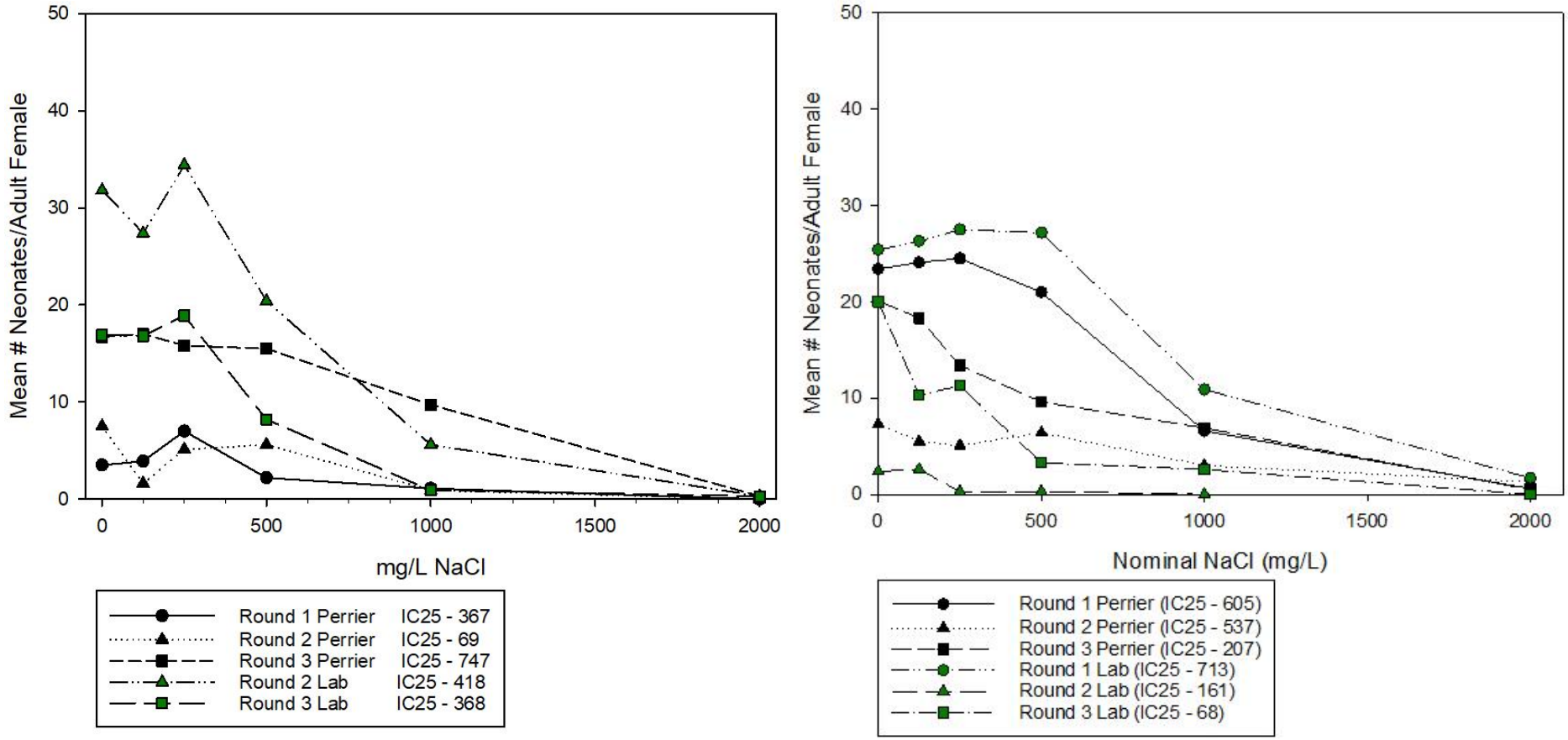


Figure C20. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab O.

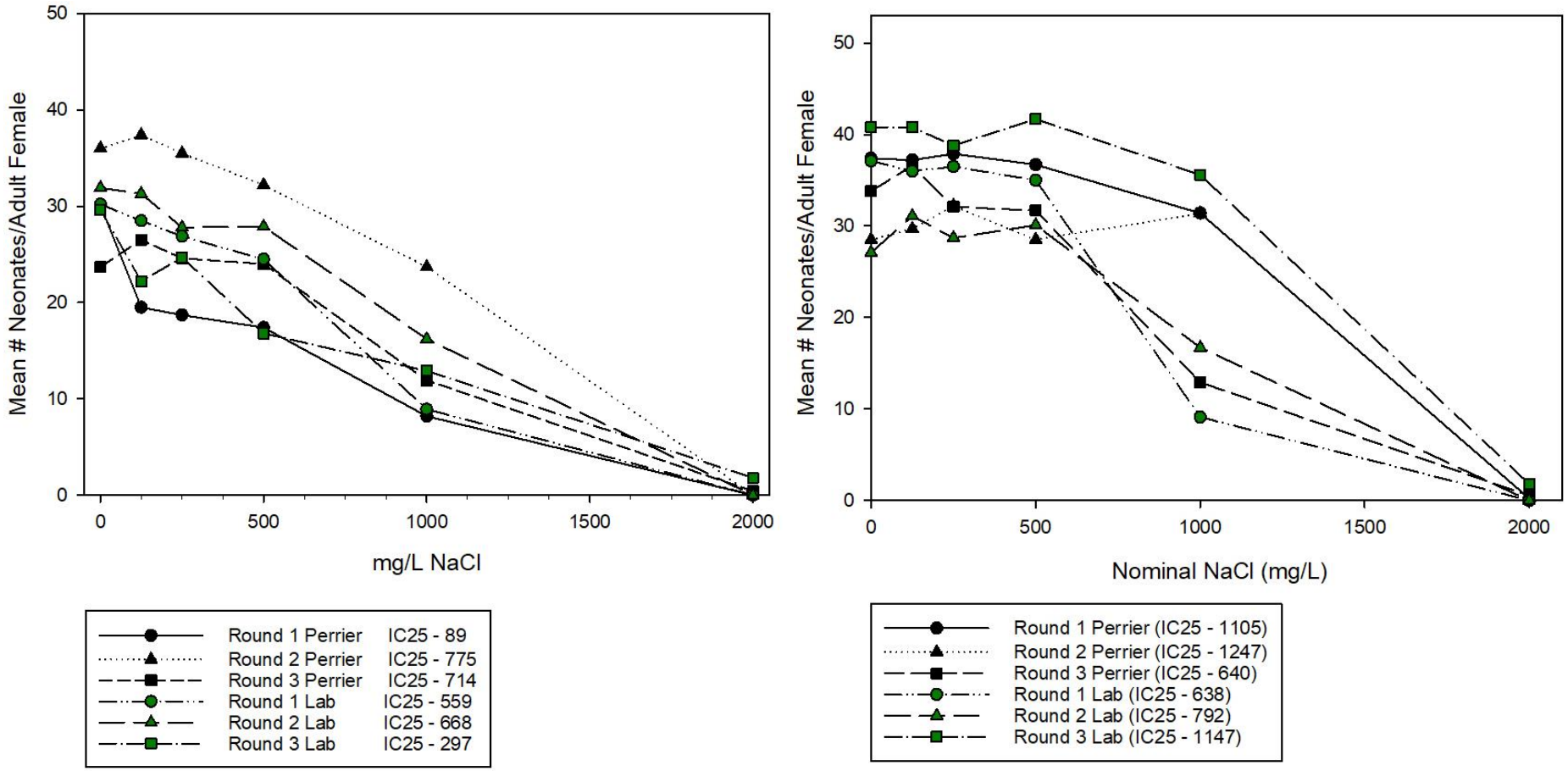
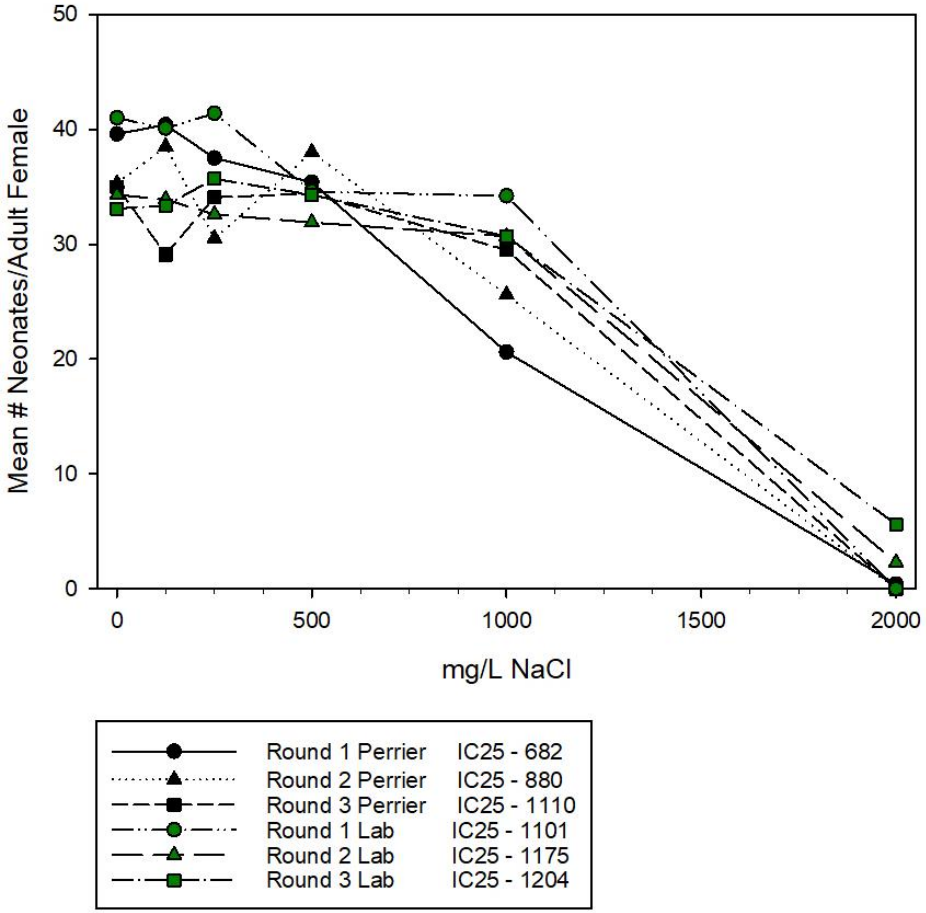
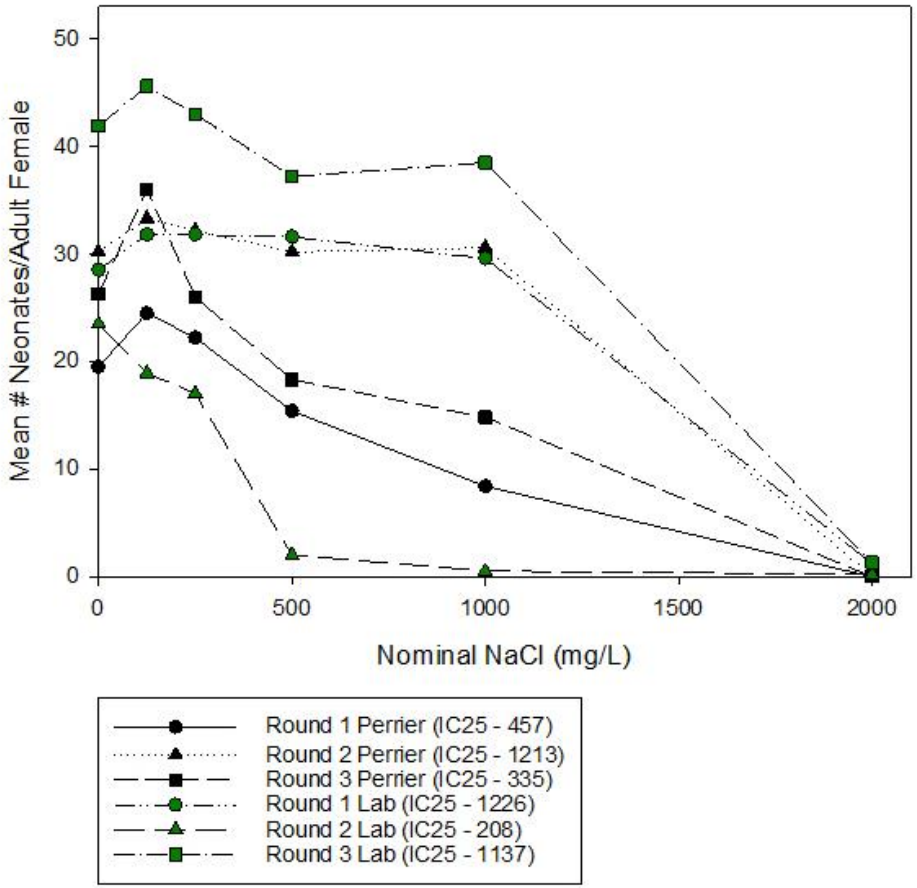


Figure C21. Baseline (left) and second (right) ILS NaCl Concentration Responses for Lab Q. This lab reported issues with their culture before Round 1.





Appendix D – Guidance materials developed during this project to improve documentation of laboratory practices for individual tests.

D1. Guidance for Documenting Brood Board Health

This is intended to help with the determination of abnormal occurrences in the brood board for the second interlaboratory study (ILS). If any of the notable observations illustrated below occur in the brood boards used in the second ILS, they should be noted with appropriate brood board health code (**Table D1**) on the daily observation sheets and included in the electronic data submission. Notes are given about how some of these appear to the naked eye. If any of these are suspected greater magnification should be used for verification.

Table D1. Health codes and descriptions that must be used for documenting brood board health for the second ILS.

Health Code	Parameter	Description of Parameter
A	Unhealthy adult	Lack of normal movement, not normal shape, smaller than normal, or atypical coloration. Presence of ephippia in brood pouch. Empty brood pouch, lack of clear gutline, and undersized brood based on age (i.e. only 3-6 eggs present 5 or more day post initiation).
D	Dead adult	Self-explanatory.
K	All OK	Nothing in any of the other categories to report.
M	Male	An animal that produces no young and is microscopically examined to determine sex. Triangular abdomen and fast/irregular swimming.
N	Dead neonates	Self-explanatory. Can be some or all of brood.
U	Unhealthy neonates	Lack of normal movement, not normal shape, ephippia or atypical coloration.
O	Other occurrences	Growth on adults or neonates, biofilm in brood cups, foreign species, aborted broods, flocculent material, etc. Describe in comments section.
Y	Neonates used to initiate a test	Neonates from this brood board chamber were used to initiate a test.

Abnormal, non-reproductive female.These will sometimes appear as having an empty and translucent brood pouch or contain a small number of eggs (A) that are aborted or do not become viable neonates. This would fall under health code A, as an unhealthy adult.

Photo courtesy of Alison Briden Pacific EcoRisk

Adult female with ephippia (A) (resting or diapause stage embryo). Ephippia are not normally observed in healthy cultures. This would receive health code A.

Photo courtesy of Alison Briden Pacific EcoRisk

Normal adult females will have visible eggs in the brood pouch (A) that are clearly differentiated, especially early in the brood cycle. As the organism gets closer to releasing the brood, unborn neonates should have visible eye spots. The adult should have a visible gut line and usually has a visible green coloration from feeding on the algal suspension.This would receive health code K.

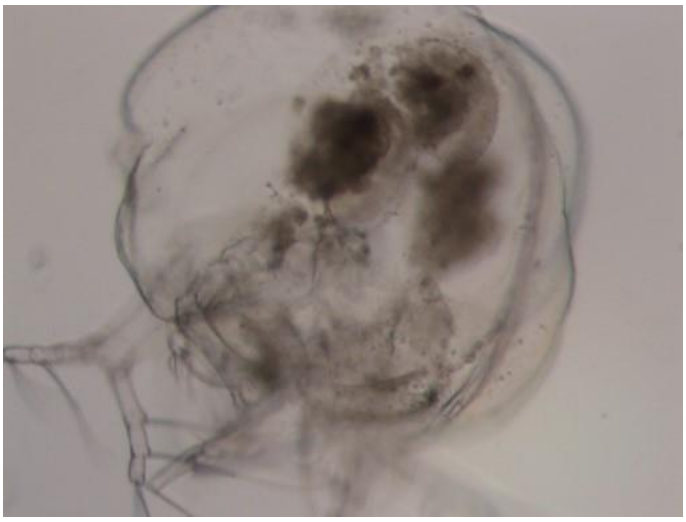
Photo courtesy of Alison Briden Pacific EcoRisk

Male. Note the flatter back with absence of brood pouch (A) and longer antennules (B) than the female. Males are noticeably smaller than reproductive age females. Males also tend to have more erratic swimming behavior.This would be health code M.

Photo courtesy of Alison Briden Pacific EcoRisk

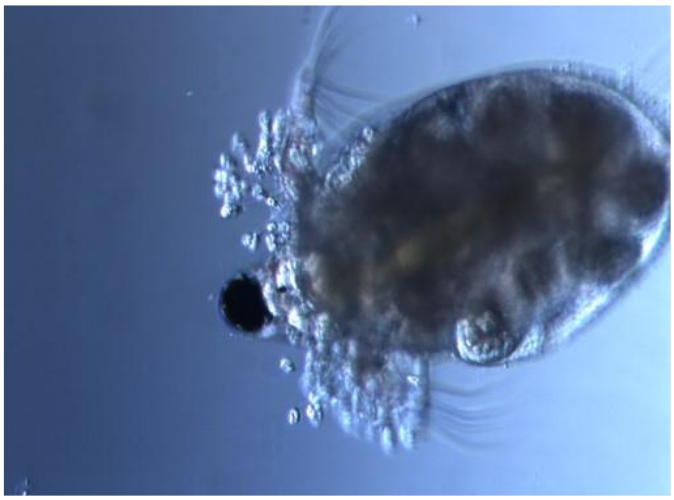
Male left and female right. Note the difference in size and shape (A) and difference in antennules (A) from the female Health code K for the female, M for the male.

Photo courtesy of Alison Briden Pacific EcoRisk



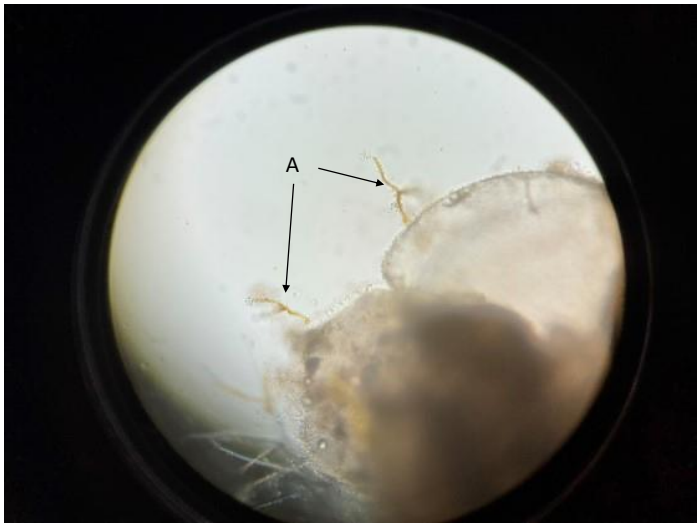
Partially aborted brood with eggs in molted carapace. The aborted brood was visible to the naked eye as a small black dot inside the molted carapace. Unhealthy females will sometimes partially or completely abort underdeveloped eggs without molting the carapace; if a female is observed with an empty brood pouch, the replicate should be examined for aborted eggs which appear as small black dots at the bottom of the chamber under a microscope. This would receive health code O.

Photo courtesy of Alison Briden Pacific EcoRisk



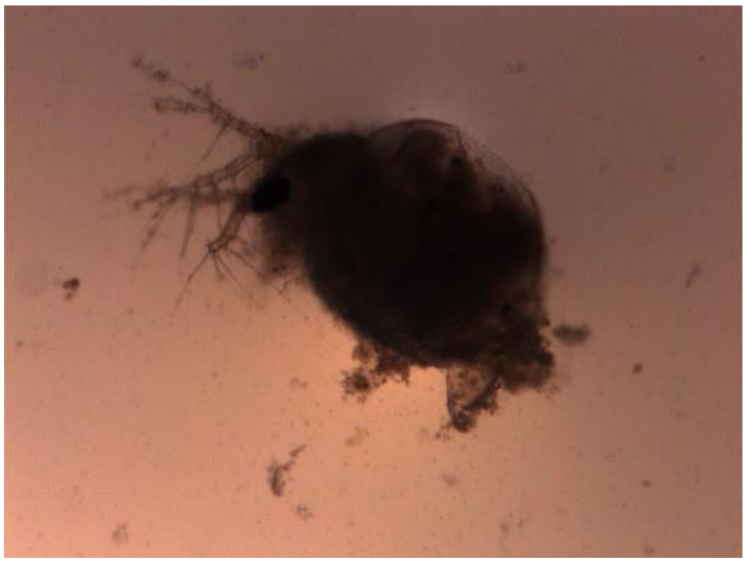
Female with stalked ciliate infection This infection affected reproduction but was not fatal. The ciliates were shed with the carapace but quickly returned This would receive health code Q

Photo courtesy of Alison Briden Pacific EcoRisk



Adult female withPeritrichs (stalked ciliates) on carapace. These appeared as reddish spots on the animal and bottom of cup to the naked eye. This infection was usually fatal. This would receive health code O.

Photo courtesy of Alison Briden Pacific EcoRisk

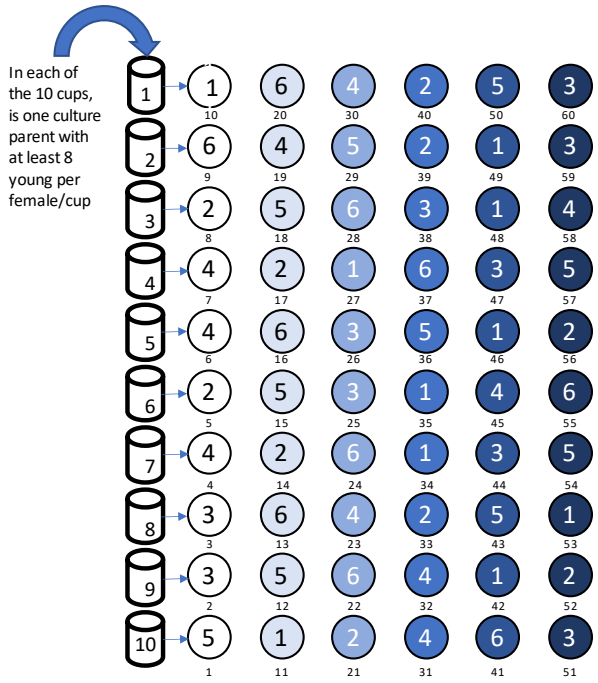


Dead adult covered with biofilm. This receive health code D,O.

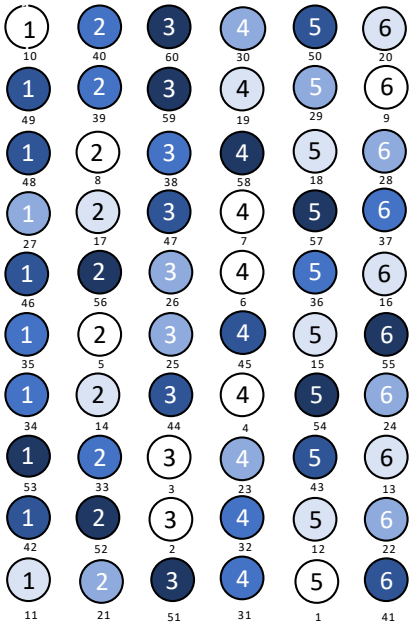
Photo courtesy of Alison Briden Pacific EcoRisk

D2. Guidance for Randomization of Test Chambers

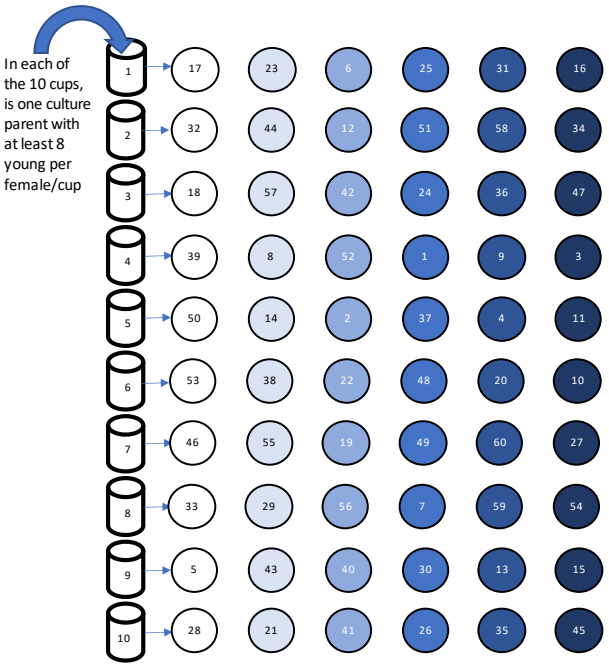
Randomization of the test chambers is a requirement in the test method. The manner of randomization is left to the individual laboratories. If a laboratory is currently randomizing their tests, they are encouraged to continue with their method. For laboratories not currently randomizing, below are two possible methods that can be used. Other acceptable methods may also be used. Either multiple sets of random numbers should be assigned for use in testing or new numbers can be chosen for each test performed. A single set of random numbers must not be used for each test.



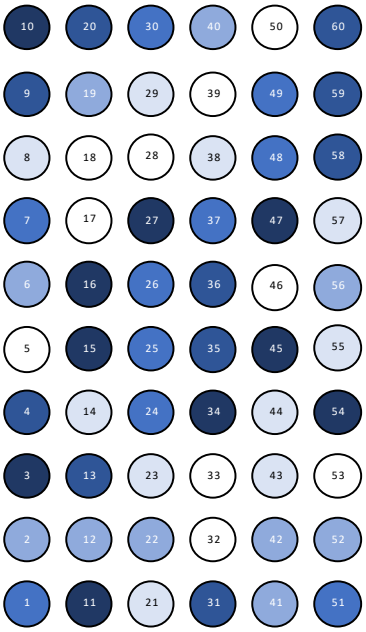
Test board with chambers in each row assigned random numbers. Each row represents a random block from a single female. Colors and columns represent separate samples or concentrations. The small numbers identify each chamber.



Test board adjusted using the random numbers assigned in the left graphic to randomize within each row. This method represents the minimum level of randomization required.



In this form of randomization, each chamber is assigned a random number. The chambers are arranged in treatment blocks for addition of the neonates.



After addition of the neonates, the chambers are arranged in numerical order which randomizes the entire test. The use of colored labels will help to keep the treatments straight during subsequent water changes.

D3. Ceriodaphnia Food and Feeding Data Sheet for Second ILS

Test Start Date_____

Test Batch Numbers_____

Sheet ____ of _____ for this set of tests

Feeding Method: Direct or In Test Solution (Circle)

Each food individually Or Foods are Mixed and Fed (Circle)

YCT

Source of Prepared YCT: inhouse or purchased_____ (Circle; give supplier if purchased)

Is YCT filtered before use_____ _ or does the YCT have solids strained out_____

Production Date_____ YCT Batch ID_____

Laboratory Measured Solids_____ mg/ml

Concentration fed to tests _____ (x ml/x ml of _____ mg/L TSS YCT)

Concentration in each test cup is calculated to be: _____

Algae

Source of Prepared Algae: inhouse or purchased_____ (Circle; give supplier if purchased)

Production Date_____ Algae Batch ID_____

Supplier Cells/ml_____

Laboratory Measured Cells/ml_____ Measurement Method_____

Procedure used to determine cells/ml_____

Concentration fed to tests _____ (x ml/x ml) of _____ algae concentration, cells/ml)

Concentration in each test cup is calculated to be: _____

D4. Guidance for Identifying and Documenting Split Broods

Initial identification of possible split broods must be conducted at the time of each daily check. Notations of possible split broods must be made on the datasheets daily. Notes for indicators of split broods should be made on bench sheets or an observation recording sheet. Final determination of split broods may be made by a more senior employee, but again must be done on a daily basis and the data sheet initialed. Identification of split broods by bench analysts can be reviewed as part of the laboratory's data QA/QC process and any changes to a determination for split brood made by the reviewer should be done in conjunction with the analyst who performed the test and the notes that they took during testing. Determination of split broods cannot be conducted after the test has ended unless detailed bench notes are available from the daily observations.

The process for identifying split broods should involve multiple steps.

- When a female produces a small brood relative to other brood sizes on that test day (within the same test concentration), she should be examined to determine if there are remaining eyed neonates in her brood pouch which, in combination with observing the neonates present in the replicate look newly released, is an indication that the organism is actively releasing the brood during the changeover time. If so, this should be noted on the bench sheet. The presence of additional neonates can be seen with the naked eye and is easy to observe under a dissecting scope.

- On the day following a small brood, the size of the individual neonates should be compared to those released by other females. If the neonates are larger than those from other females, this would indicate that they are from the part of the previous day's brood.

- If a split brood is detected, it should be indicated on the bench sheet and the neonate counts from the two days must be circled to denote that they should be treated as a single brood.

Appendix E – Performance metrics for *Ceriodaphnia dubia* survival and reproduction toxicity test.

E1. Goal of this Appendix

This Appendix aims to clarify the rationale and approach used to assess laboratory performance during the study (section 5 of the report). This approach represents a framework for developing and evaluating lab performance metrics but more discussion is needed before implementing such metrics beyond the scope of this study. Additional vetting of the performance metrics, criterion values, desired frequency distributions, scoring, and weighting may be needed.

E2. Approach to Developing the Draft Metrics

Performance metrics for laboratory tests are an important component of any monitoring program to ensure that desired levels of sensitivity, consistency and comparability are attained by all providers of similar services that support such programs. In the context of this study, such metrics benefit multiple entities, including:

- Testing laboratories: performance metrics provide immediately available feedback in terms of attaining desired performance goals, and early indications of organism condition or operations that may not conform to QA/QC standards.
- Regulated community: metrics identify laboratories with demonstrated ability to consistently perform tests that reflect the desired level of competence and organism condition, thereby increasing confidence in lab results and associated management decisions.
- Regulators: tests performed by laboratories that consistently achieve the desired level of competence facilitate data interpretation and improve compliance with water quality monitoring objectives.
- Lab accreditation programs: standard metrics applied across laboratories support the evaluation and attainment of objectives for consistency and comparability.

Proposed metrics were grouped into three categories:

- Biological metrics, representing general status of test organisms and procedures, and also basic Test Acceptability Criteria (TAC)
    - Survival
    - Mean number of young per *surviving* female in control group

- Variability and uncertainty metrics, representing consistency of output in a laboratory
    - CV associated with mean number of young per female in control group
    - PMSD associated with individual reference toxicant tests
- Toxicity / Potency Endpoint metrics
    - IC50, estimated concentration of a reference toxicant associated with a 50% decrease in young produced estimated using the ICp procedure
    - IC25, analogous to IC50 but with a 25% decrease (not applied in the current study)
    - Ratio IC25/IC50, informing shape of the concentration-response curves and expected to be similar among laboratories using the same toxicant (not applied in the current study)

Biological metrics were evaluated within a laboratory, while variability metrics and potency estimates (focus was on IC50 for this project) were assessed across laboratories. The Panel emphasizes these metrics are not intended to be final. While the approach used to derive the performance-thresholds has precedent for some of the values (see EPA 2000a and 2001a), more work is recommended to refine these values, the desired frequency of attainment and to investigate the other metrics (e.g., IC25, ratio IC25:IC50) to ensure that the benchmarks are representative, unbiased, and reliable.

Table E1 provides further details on the performance metrics described above and applied to the historical and ILS data. Test acceptability criteria were used as the "thresholds" (referred to as meeting expectations) for the biological metrics. Variability metrics were derived based on analysis of the data generated in this project and compared to those in the EPA (2000 and 2001) and Fox et al. (2019) studies. Toxicity potency metrics were assessed using a percentile approach although the traditional approach of comparing data to grand mean +standard deviation (SD) was also explored. In addition to the threshold-values, the Panel conducted preliminary investigations of acceptable frequency of attainment that would not penalize laboratories occasionally producing marginal data due to documented sub-optimal conditions (e.g., culture or brood board health issues, sample condition, etc.). There is no approved precedent or standard for frequency of occurrence and the values utilized in this example are largely based on best professional judgement.

E3. Application of Metrics

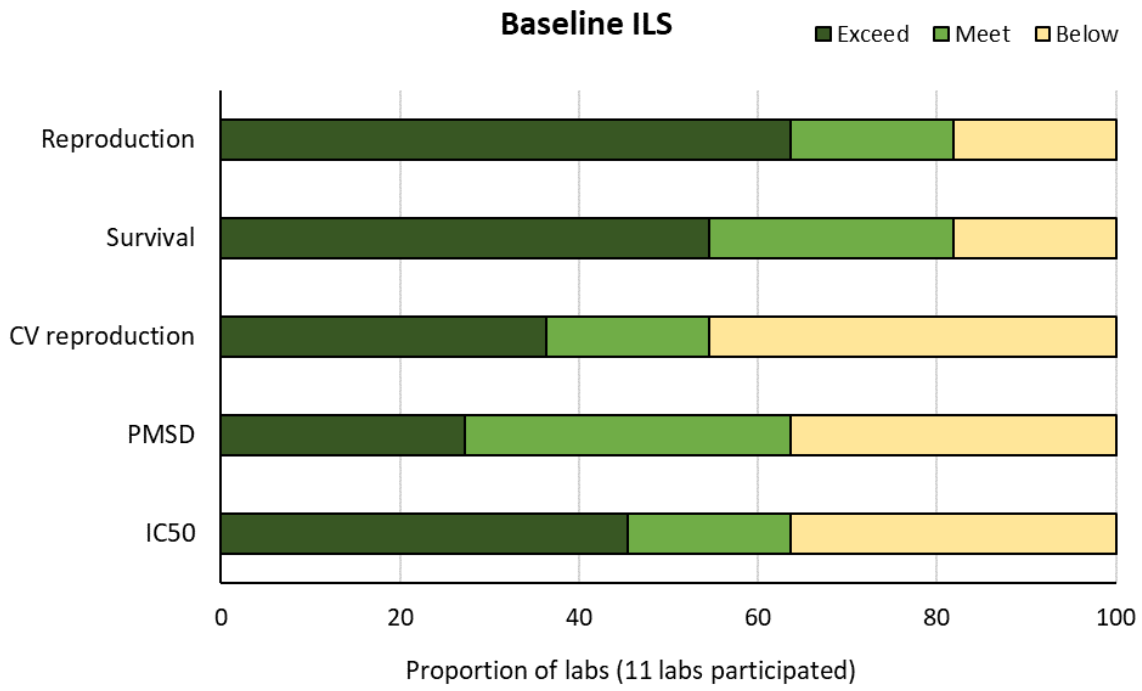
The Panel applied the performance metrics as an example in the report (section 5) to assess the performance of California-accredited laboratories. Using these metrics, the Science Panel observed that a third of the laboratories that generated historical and ILS data were consistently exceeding expectations, while the other half were either at meeting expectation levels or inconsistent.

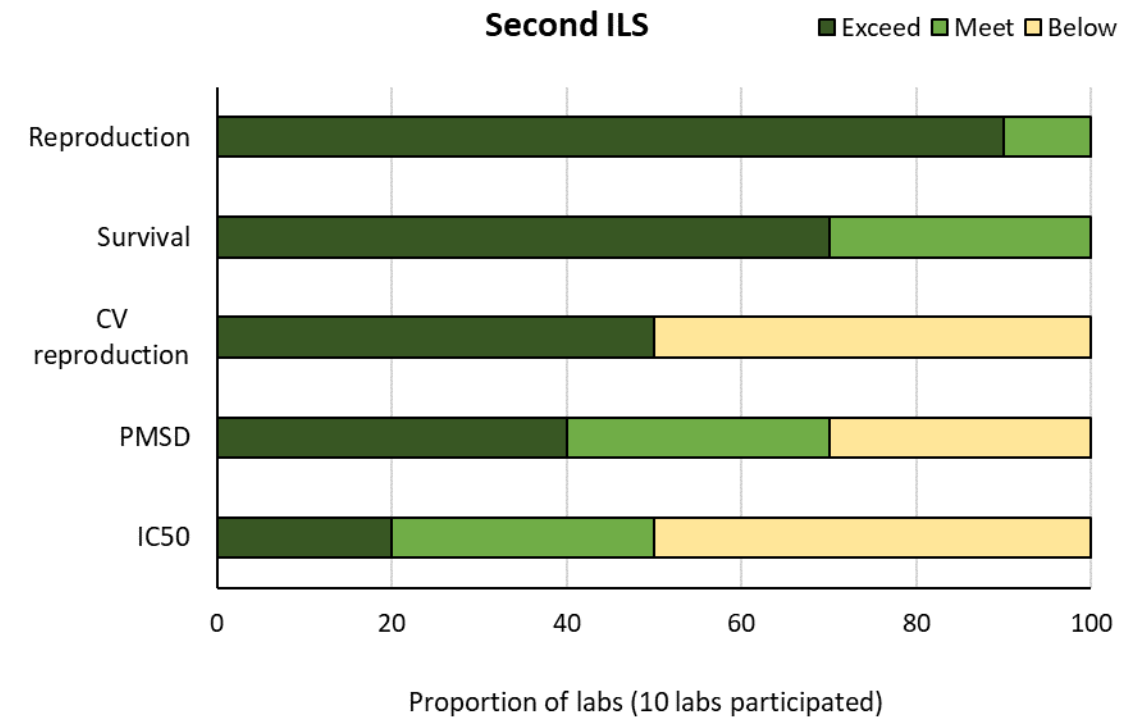
It should be noted that performance metrics are valuable for large datasets. Nevertheless, the Panel did use them for the two ILS that had limited data with the goal of determining whether comparability had improved within and among laboratories. **Figure E1** illustrates the findings of the analysis. Overall, laboratory performance improved with more laboratories exceeding expectations in the second ILS for all but one category compared to the baseline ILS. For the IC50 performance metric, the proportion of laboratories exceeding expectations decreased marginally from 64% to 50%.

Table E1. Draft performance metrics to assess laboratory consistency and comparability.

Lab performance	Survival	Reproduction	CV neonates per female	PMSD	IC50
Exceed expectations	≥ 90% in ≥ 90% of tests	≥ 20 in ≥ 90% of tests	≤ 0.2 in > 75% of tests	≤ 25% in > 75% of tests	within 25-75th >75% of tests
Meet expectations	≥ 80% in ≥ 90% of tests	≥15 in ≥ 90% of tests	≤ 0.2 in 50-74% of tests	≤ 25% in>50% of tests	within 25-75th in > 50% of tests and within 10-90th in > 75% of tests
Below expectations	≥ 80% in < 90% of tests	≥ 15 in < 90% of tests	≤ 0.2 in < 50% of tests	≤ 25% in < 50 % of tests	Outside of 25-75th in > 50% of tests

Figure E1. Example application of the draft performance metrics to Baseline ILS and Second ILS. Scoring and laboratories identity are purposely not reported.





While IC25 is an endpoint used in regulatory programs and proficiency testing, the Panel primarily used the more robust IC50 metric for evaluating interlaboratory variability because it better reflects the central tendency of the toxic response. The Panel did perform exploratory investigations for the IC25 and IC25:IC50 using the baseline ILS data only and a standard "grand mean and standard deviation approach" (for Sample 2 and Sample 3). The analysis assessed how many data and laboratories fell within the following categories:

- The average IC25 value for a given lab fell within +/- 1 SD of the grand mean for the 7 laboratories that met expectations.
- The majority of (at least 75%) datapoints for a given lab fell within +/- 1 SD of the grand mean; and
- No data points fell outside of 2SD of the grand mean.

The average IC25 values generally fell within 1SD of the grand mean except for Lab L. In terms of consistency, only 50 percent of the datapoints for Lab L, M and P fell within 1SD of the grand mean, and one datapoint from Lab L fell outside of 2SD from the grand mean. Water effects were judged on the basis of lack of overlapping values between the two waters; point estimates for Lab P did not overlap between waters, but the absolute differences were generally small (e.g., ≈15%).

An additional metric associated with the point estimates was also evaluated, i.e., the ratio of the estimated IC50 to the estimated IC25. From a toxicology perspective, this ratio reflects the slope of the concentration-response curve and is generally considered to be consistent relative to a specific toxicant tested under a specific set of conditions. Factors that might affect the viability or condition of the test organisms (e.g., a disease challenge, an additional toxicant, changes in water chemistry or test parameters) and otherwise impair their response to a chemical stressor would be expected to increase the degree of response in the most sensitive sublethal indicator, with a concomitant alteration of the concentration-response curve. Conversely, the absence of such a mechanism would be characterized by consistent IC25/IC50 ratios across all laboratories and waters.

This hypothesis was further investigated by comparing IC25/IC50 ratios for laboratories that met expectations against laboratories that were below expectations for the IC50 performance metric. Interestingly, laboratories that met expectations generally exhibited an average ratio of 0.73 ± 0.06, indicating a relatively steep concentration-response curve characterized by a high degree of consistency (CV=0.06; n= 40). Conversely, laboratories that "tested out" exhibited a lower average ratio, as well as greater variability (0.58 ± 0.20; CV=0.35; n= 20). Thus, laboratories below expectation for the IC50 metric exhibited a wider range of ratios.

From a distribution perspective, separating the laboratories on the basis of the 10th percentile cut-off (i.e., a ratio of 0.66) showed over 70% agreement with other metrics used to assess test performance at the IC50 and IC25 level (data not shown). Thus, these data suggest that the ratio may be a good predictor of lab performance as well as potentially informing overall data quality. Applying this criterion as a diagnostic is attractive as it is readily available and provides a more quantitative assessment of the concentration-response in addition to visual assessments of the curves. Importantly, this assessment was based on a sample size of three testing events and 11 laboratories and could still benefit from further validation and possible refinement.