C. dubia QA evaluation study

Stakeholder Committee Meeting

December 21, 2022

Agenda

1. Opening Remarks and Review of (5 min)

2. Baseline Intercalibration Study summary of results (30 min)

3. Expert Science Panel conclusions and recommendations for the scope of the next tasks (15 min)

4. Schedule and Next Steps (5 min)

Stakeholder Advisory Committee

- Katie Fong (SWRCB)
- Amelia Whitson (EPA Region IX)
  - Rochelle Cameron (today's alternate)
- Veronica Cuevas (RWQCB4)
- Mitch Mysliwiec (Larry Walker Assoc/CASA)
  - Jared Voskuhl (today's alternate)
- Jian Peng (Orange County Public Works/CASQA)
- Sarah Lopez (Central Coast Water Quality Preservation Inc)
- Peter Arth (Enthalpy Laboratories)
- Josh Westfall (Los Angeles County Sanitation Districts)
- Annelisa Moe (Heal the Bay)

Expert Science Panel

- Teresa Norberg-King (Formerly US EPA)
- Robert Brent (James Madison Univeristy)
- Howard Bailey (Nautilus Environmental)
- Leana Van der Vliet (Environment Canada)
- A. John Bailer (Miami University, Ohio)

Quick Review From Previous Meetings

Overall Project Tasks

- Task 1- Identify potential sources of variability within and among laboratories
  - Compile historical data
  - Conduct baseline intercalibration

COMPLETED

TODAY'S FOCUS

- Task 2- For potentially largest sources of variability, optimize test conditions and QA parameters to minimize variability

- Task 3- Evaluate efficacy of test conditions and QA refinements
  - Conduct second intercalibration

Findings of the Historical Data Analysis

- Compiled historical data from all accredited labs in California
  - Last 3 years/ 30 tests
- Quantified variability within and among laboratories
  - Substantial variability in biological responses between labs
- Several types of data analysis to identify lab techniques that might be responsible for variation within or among laboratories
  - In the end, no two labs conducted the test the exact same way

Baseline Intercalibration Study

Study Design Overview

- Conduct test using lab SOPs, but enforce similarity in data collection

Sample Type	Number of Samples	Dilution Series	Number of Rounds	Number of Labs
Sample 1- dilution water recipe #1 (EPA MH)	1	No	3	12
Sample 2A- dilution water recipe #2 (Perrier)	1	No	3	12
Sample 2B-F- NaCl in dilution water recipe #2 (Perrier)	5	No	3	12
Sample 3- NaCl solid sample (to be diluted by the labs using their own dilution water)	1	Yes	3	12

- Expected Outcome: 3 EPA MH Water, 3 Perrier, 3 dilution series in Perrier water, 3 dilution series in lab water, for each of the 12 labs

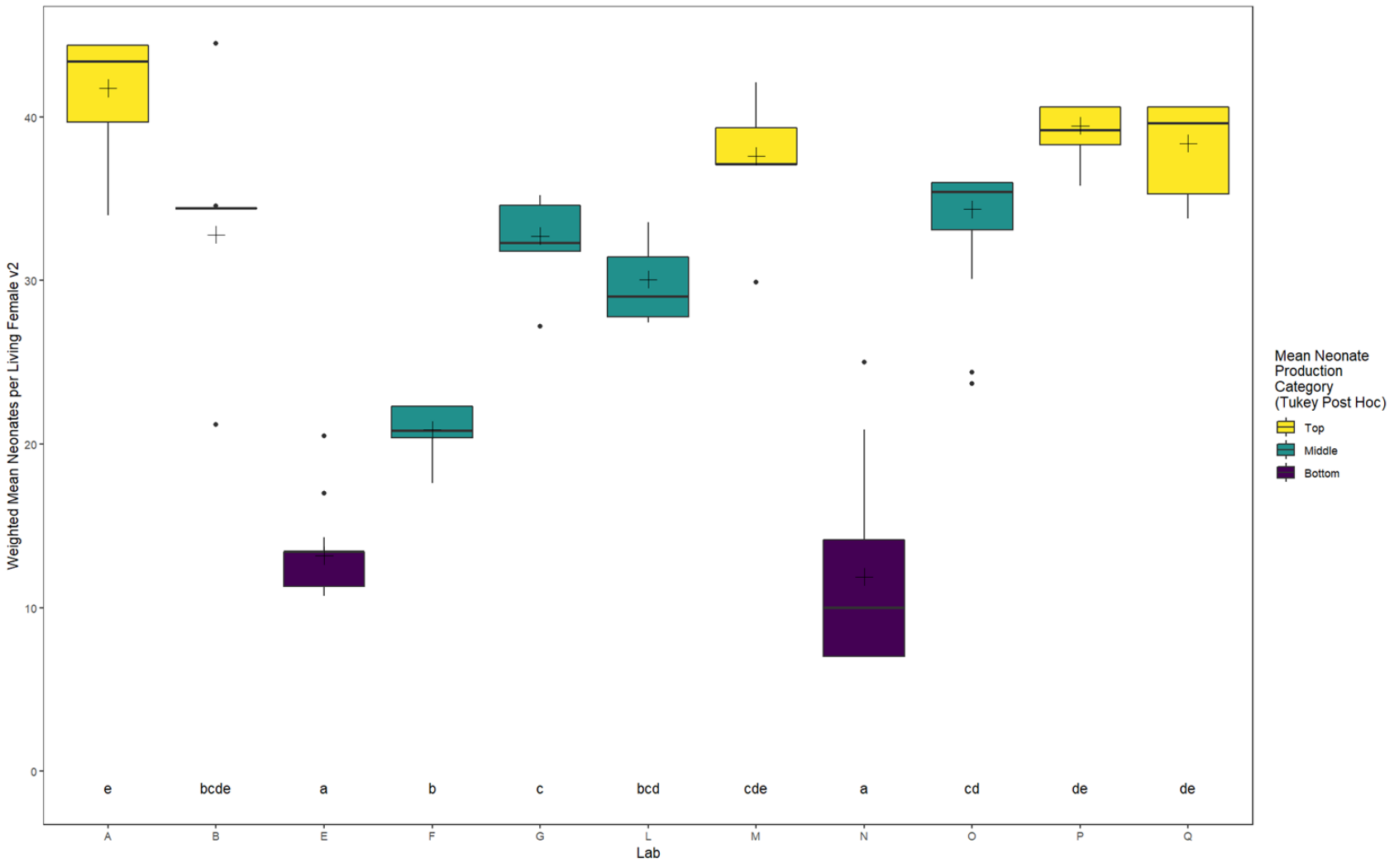
Testing Success

Lab ID	# Perrier	# EPA MHW	# Our Dose-Response	# Lab Dose-Response
A	3	3	3	3
B	2	2	2	1
E	3	3	3	3
F	3	3	3	3
G	3	3	3	3
L	3	3	3	3
M	2	2	2	2
N	3	3	3	2
O	3	3	3	3
P	3	3	3	3
Q	3	3	3	3
All Labs	31	31	31	29

SCCWRP Data Quality Steps

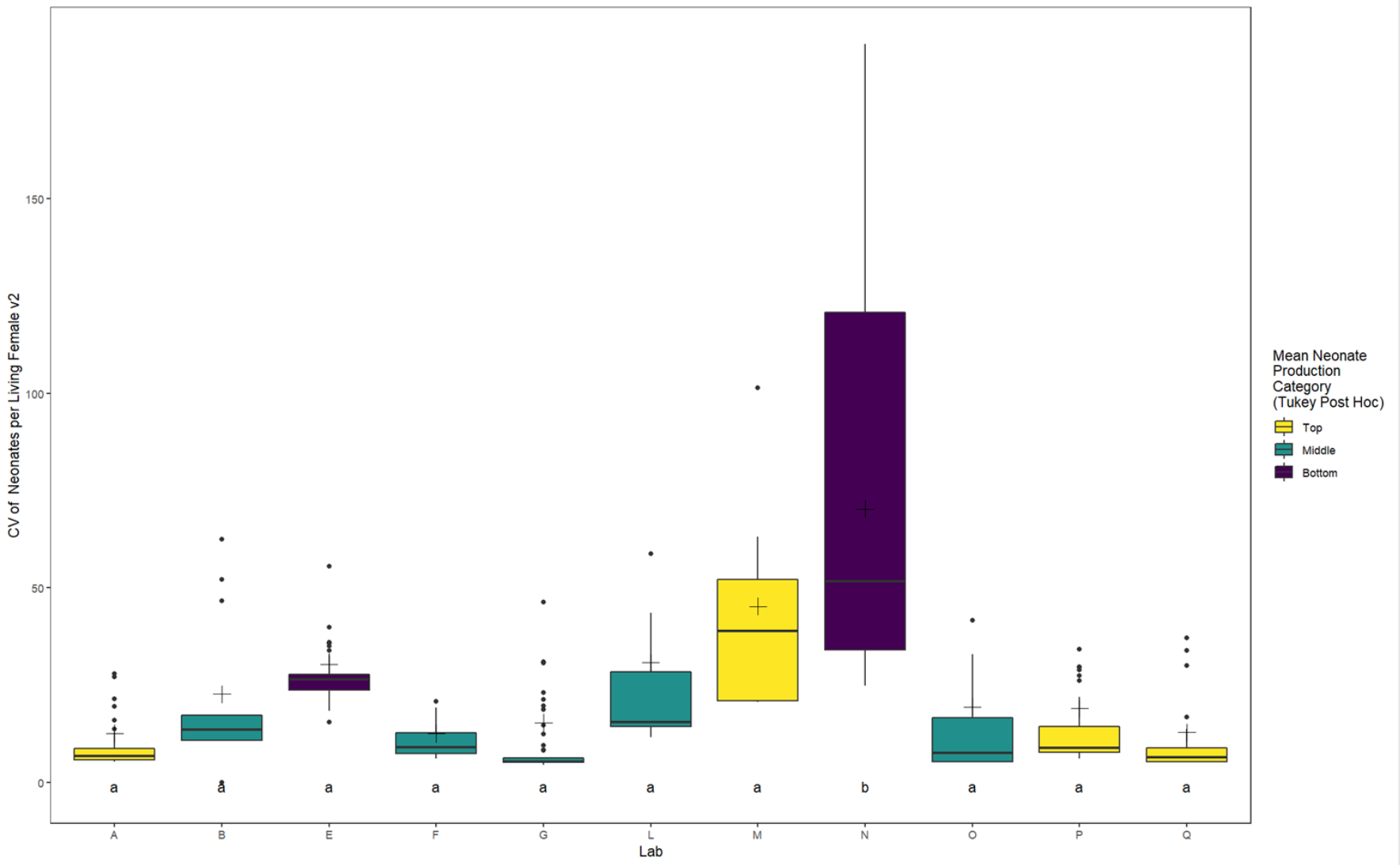
- 100% data check for completeness, DQO criteria
  - Flagged but did not exclude tests that failed Test Acceptability Criteria
- 5% random check for raw data entry errors (edata vs bench sheets)
  - Any failures results in 100% data audit for that lab
- Many hours of phone calls with each laboratory
  - Confirming corrections made by SCCWRP and approved by labs

Comparison of mean neonates per female among labs using SCCWRP supplied water



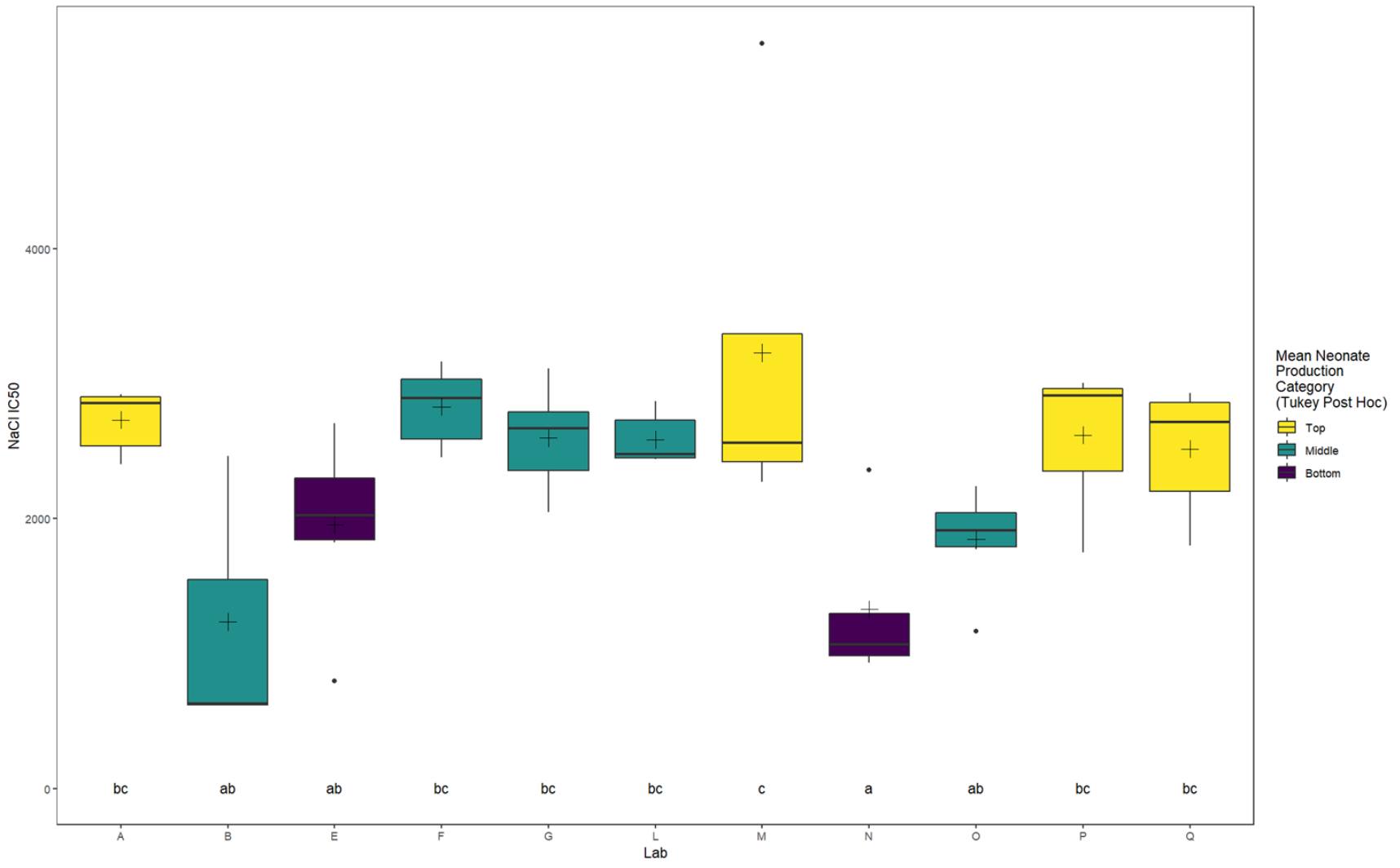
- Look at effect of lab and water type on mean or CV of neonate production
  - Includes EPA and Perrier data, N=5 or 6 per lab
- Significant effect of lab on mean reproduction
  - No effect of water or lab*water interaction
- Post hoc tests classifies labs in 3 general categories
  - With a little BPJ

Comparison of CV of mean neonates per female among labs using SCCWRP supplied water



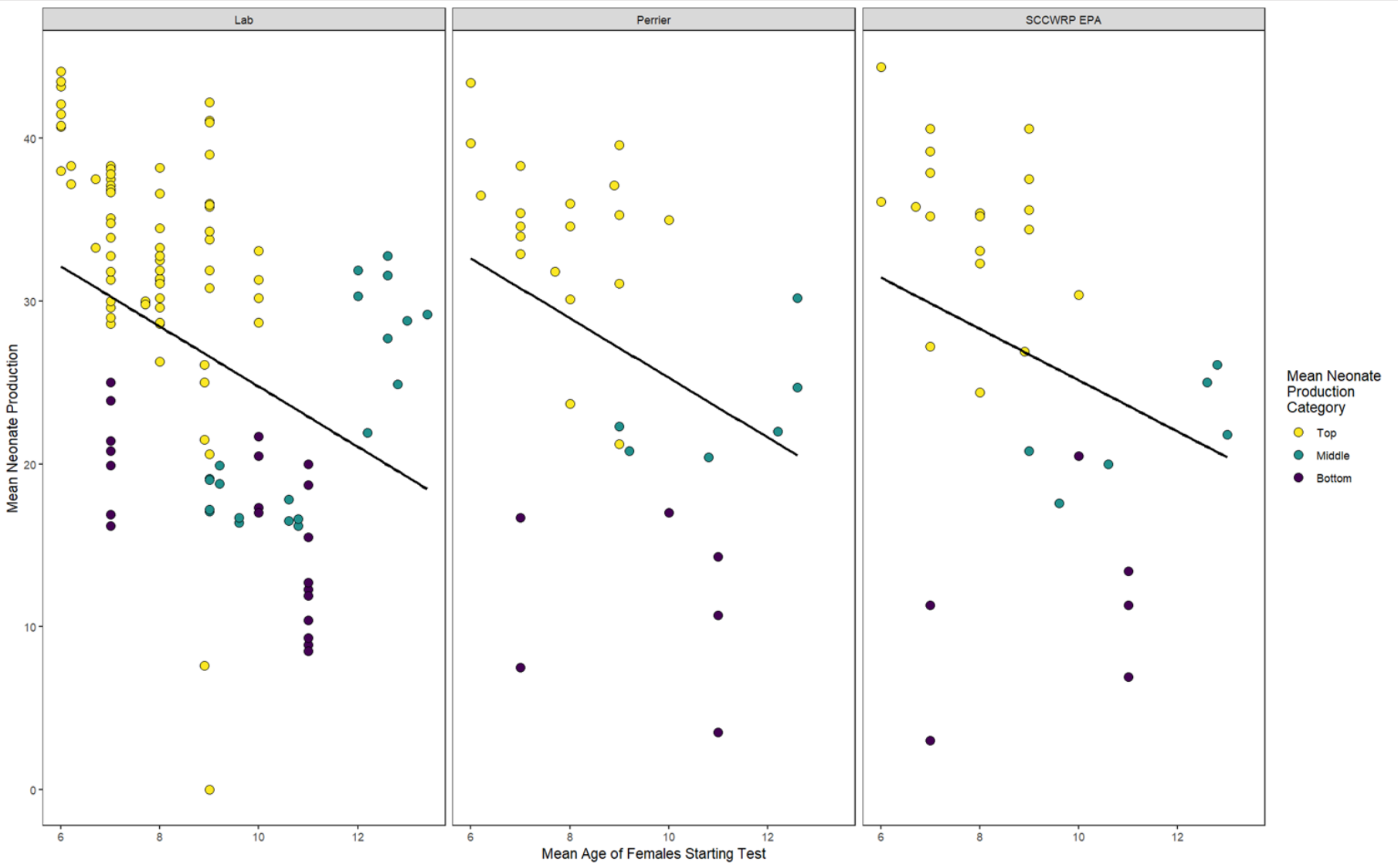
- Look at effect of lab and water type on mean or CV of neonate production
  - Only included EPA and Perrier data
- Significant effect of lab on CV of mean reproduction
  - No effect of water or lab*water interaction
- Post hoc tests classifies labs in 2 general categories
  - Color coding based on the mean neonate production

Comparison of IC50s Among Laboratories Using Sodium Chloride



- IC50 for Ceriodaphnia reproduction
  - Included spiked samples using both SCCWRP Perrier water and Lab dilution water at the same nominal concentrations
- Significant effect of lab on IC50
  - No effect of water or lab*water interaction
- Post hoc tests classifies labs in 3 general categories
- Color coding based on mean neonate production

Mean neonates per female vs mean initial age of female in brood board – by water type



- Labs are color coded by mean neonate production in previous analysis
- Statistically significant relationship infers reduced mean neonate production during testing with older mean age of brood board females
  - Could be confounding or spurious
  - Current guidance is to use females ≤ 14 days old

Summary of Results

- There are significant differences among labs for mean neonate production using the same sample
- There is no clear “smoking gun” of the primary cause for variability in mean neonate production
- A contender is the brood board, which has some method guidance, but is rarely reported
  - Age of female used to start the brood board
  - Water quality in the brood board

Some Lab Techniques That Don’t Appear To Be Primary Drivers of Neonate Production Variability Between Labs

- Water type (Perrier vs Reconstituted Moderately Hard Water)

- Test Water Quality parameters (Temperature, DO, pH)

Agenda

1. Opening Remarks and Review of the Agenda (5 min)

2. Baseline Intercalibration Study summary of results (30 min)

3. Expert Science Panel conclusions and recommendations for the scope of the next tasks (15 min)

4. Schedule and Next Steps (5 min)

Expert Science Panel Findings and Preliminary Recommendations for the Ceriodaphnia dubia Toxicity Test

December 21, 2022

Topic: Baseline Intercalibration Study

***Ceriodaphnia dubia* Three Brood Toxicity Test**

Test variability is intrinsic to performing studies. Reliability and reproducibility are issues faced by many labs, not only for the *C. dubia test.*

The EPA Short-Term Test Method procedural document, after decades of use, is widely used by many laboratories. The test method provides required procedures and recommended procedures.

Designed for laboratories with an existing quality management framework.

Preliminary Findings

- Some labs did not pass Test Acceptability Criteria
- Variability in mean neonate production between labs is very large
  - Variability between labs was roughly similar to the variability between labs from historical data
- For labs with consistent quality, the IC25s are reasonably consistent

Additional Findings

- When labs are “in control” (demonstrating consistent culture health and consistently meeting Test Acceptability Criteria (TAC) variability in neonate production does not appear to greatly influence toxic endpoint variability
  - “It’s OK if one lab averages 25 neonates and another lab averages 40 neonates”
  - “It’s not OK when neonate averages bounce around 15”

- Dilution water type did not seem to be a major source of variability
  - Given the information generated using NaCl as tested in this study.

Source of the Variability is Still Uncertain

- Historical data analysis did not find one or more factors driving the source(s) of observed variability in responses
  - Meaning it did not find “one test” condition that explained variability
- Baseline intercalibration study found
  - When labs demonstrated consistent performance metrics and culture and organism health, consistent results can be achieved within and across labs

Source of the Variability is Still Uncertain

- Performing the tests using test concentrations to obtain a dose-response, determined the Standard Deviation in the controls is not the only factor that impacts the outcome of the test.
  - As suspected, the biological sensitivity of the test organisms to the toxicant affects it as well.

- However, from the study to date, the cause of poor culture health was not identified.

Preliminary Findings and Recommendations

Additional Considerations for the Study from the Expert Panel

- Finding: Insights about lab performance have been gleaned from historical review and from recent lab testing, but important sources of variability remain to be identified

- Recommendation: Additional time is necessary to study the data further before a complete set of final recommendations can be provided

Culture Health and Procedures Are Key

- Finding: Method guidance exists for an acceptable brood board culturing procedures
- Recommendation: Need better documentation, reporting and evaluation of cultures and brood boards
- Recommendation: Providing specifics for culturing and the brood board characteristics should be the focus in the additional intercalibration.
  - Focus should be on communication about culturing and collaborative efforts for continuity and consistency

Focus on Ongoing Culture Health and Performance

- Finding: Method guidance exists for an acceptable brood board culturing procedures
- Finding: most likely source of the variability in mean neonate production among labs
- Recommendation: Laboratories need to develop clear, step-wise operating procedures (OP's), documentation and evaluation of brood board health and do not initiate tests when cultures do not meet minimum health standards
- Recommendation: Additional method refinement or optimization should focus on brood boards, particularly variability in age of the female used to start the brood board

Agenda

1. Opening Remarks and Review of the Agenda (5 min)

2. Baseline Intercalibration Study summary of results (30 min)

3. Expert Science Panel conclusions and recommendations for the scope of the next tasks (15 min)

4. Schedule and Next Steps (5 min)

Project Timeline

Milestone	Timeframe/ Deadlines
~~Task 2- Baseline testing (preliminary test by one lab and 3 round robins with 11 labs)~~	~~August 15 – October 6 , 2022~~
~~Task 2- Baseline testing: Deadline for data submission~~	~~October 18, 2022~~
Task 2- Meetings with Expert Science Panel for review and analyses of baseline testing data	November 7, December 2 and 21, 2022
Task 2- Public report out of baseline testing results and next steps	December 21, 2022
Task 3- Develop and execute plan to optimize select test parameters	January 2023
Task 3- Public report out of optimization testing results	By March 15, 2023
Task 4- Discussions among stakeholders and ESP to develop key recommendations	April 2023
Task 4- Oral report and technical memo of results	By May 30, 2023
Task 5- Draft recommendations report	By June 30, 2023
Task 5- Final recommendations report	July 15, 2023