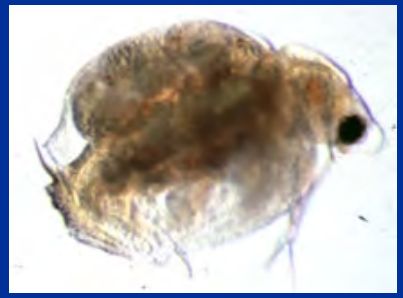
Interpreting Whole Effluent Aquatic Toxicity Tests: Avoidance of 'False Positives'

Stephen L. Clark

Pacific EcoRisk



BACWA Lab Committee
March 11, 2015

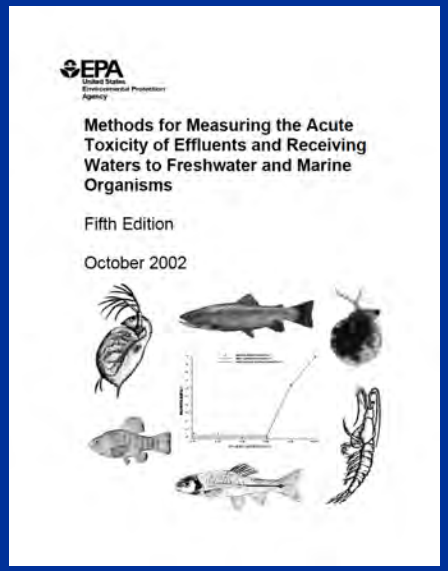
Presentation Overview

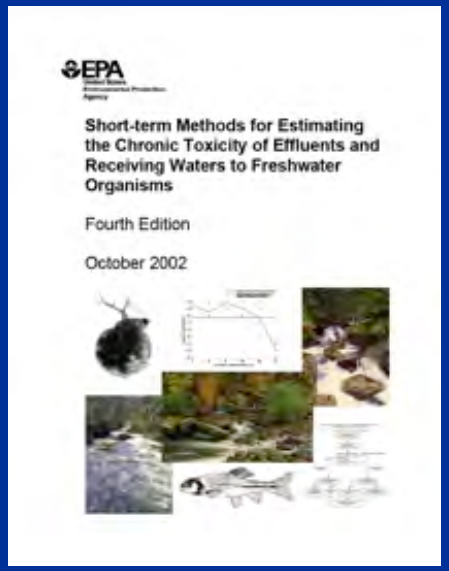
- Overview of Whole Effluent Testing
  - Cautionary Principles
- Confounding Factors Leading to 'False Positives'
  - Defined here as incorrectly identifying a sample as toxic when in fact it is not
- Conclusions

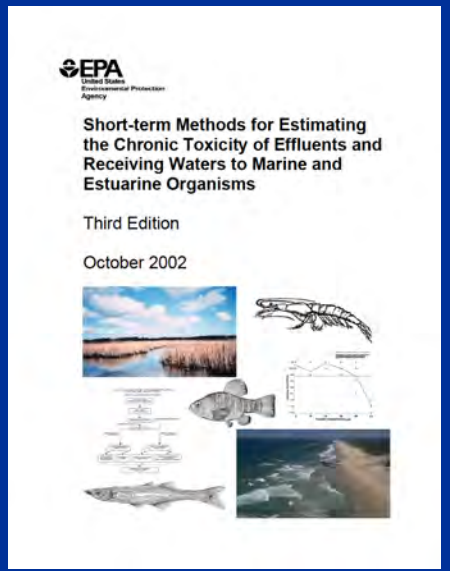
Overview of Whole Effluent Testing

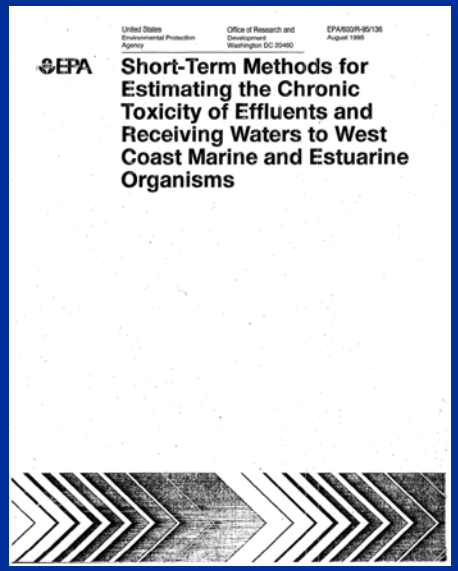
Whole Effluent Testing

- Whole effluent toxicity testing used in the National Pollutant Discharge Elimination System (NPDES) Permits Program is guided by various testing manuals.







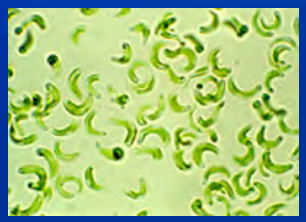


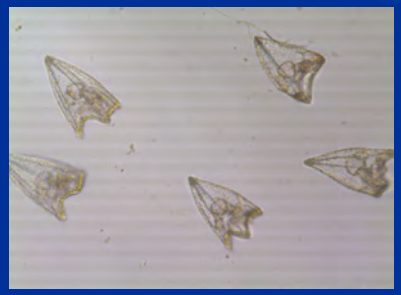
Typical NPDES Test Species

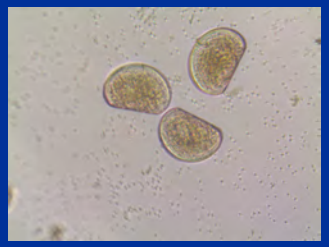
Algae

Invertebrates

Fish

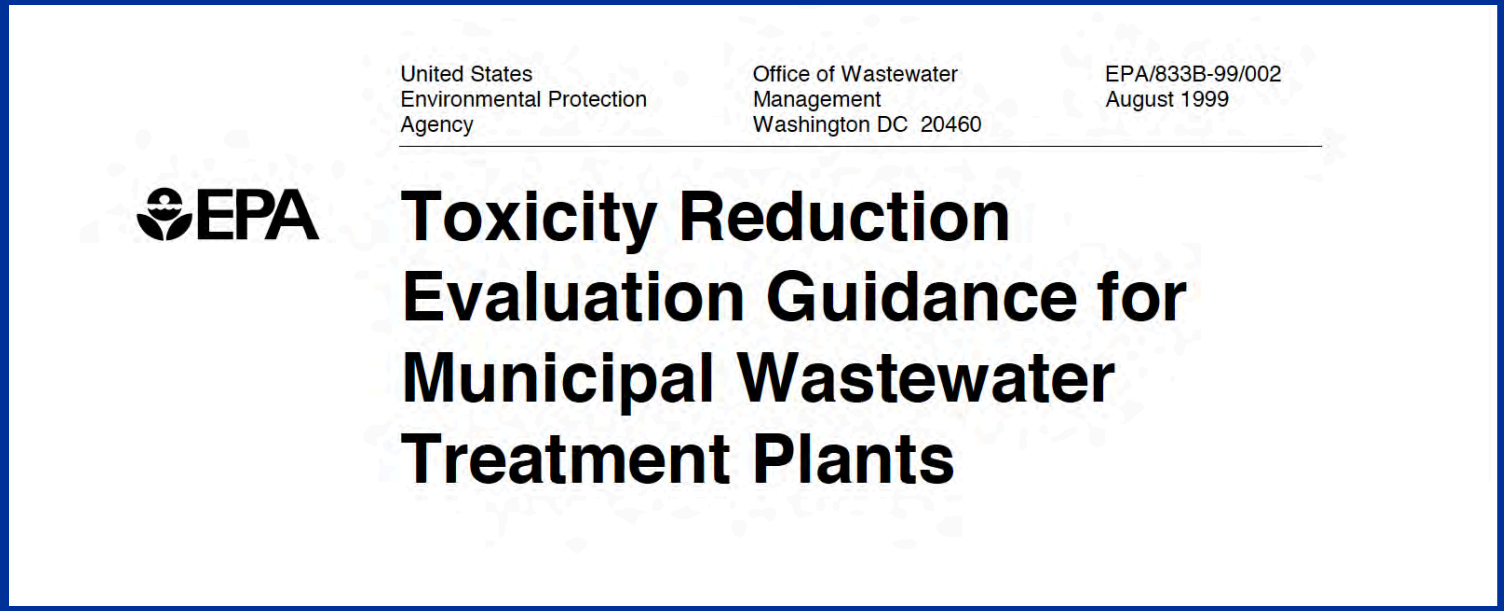






Whole Effluent Testing

- Once an effluent is identified as toxic, most permits require accelerated monitoring, and the implementation of a Toxicity Reduction Evaluation (TRE) if one of the accelerated monitoring tests exceeds the permit limit for toxicity.



United States Environmental Protection Agency | Office of Wastewater Management Washington DC 20460 | EPA/833B-99/002 August 1999

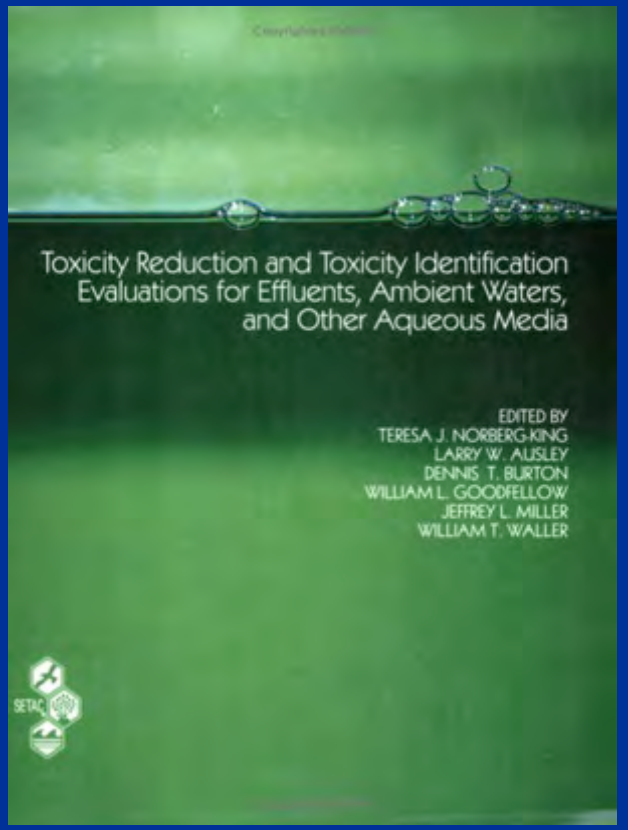
EPA

Toxicity Reduction Evaluation Guidance for Municipal Wastewater Treatment Plants

Whole Effluent Testing: Cautionary Principle

- " Permittee and the regulator should distinguish very early in the process whether an actual toxicity event has occurred or whether the effluent may appear to be toxic but may in fact be the result of an unusual or invalid test" (Ausley *et al*., 2005)

- A detailed review of the test data and laboratory methods will assist in making this decision



Cautionary Principle – Basic Test Review

- Prior to proceeding with reporting that toxicity is present, it is imperative that the laboratory comprehensively review the data to confirm that the testing is valid
  - Have the required test conditions been met?
  - Did the test meet the test acceptability criteria?
  - Were all water quality parameters within an acceptable range for the test species?
  - Is the test variability (PMSD) acceptable? Below the EPA 10th percentile?
  - Is the concentration response curve normal?

Confounding Factors

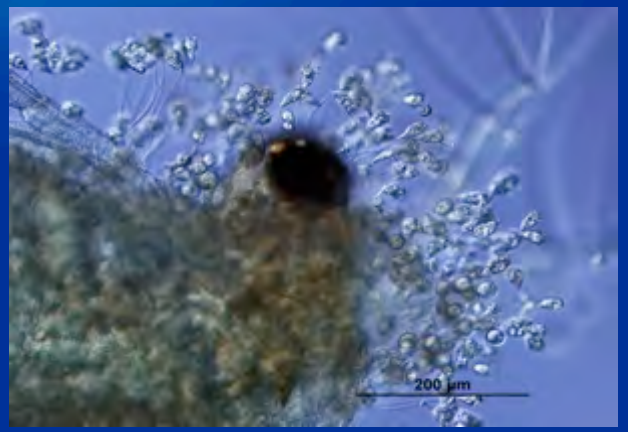
‘False Positives’ – Microbes

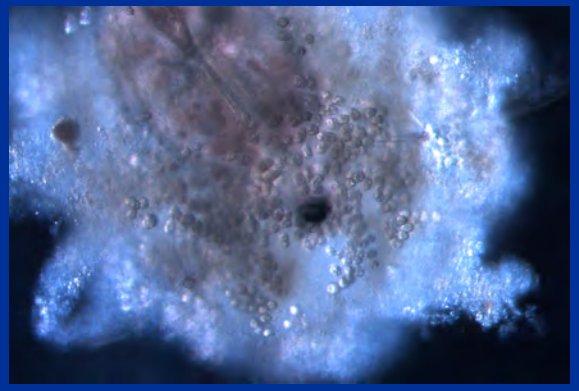
- ‘False Positive’ – defined here as identifying an effluent as toxic when in fact it is not!
- Fathead Minnows – Pathogen Related Mortality

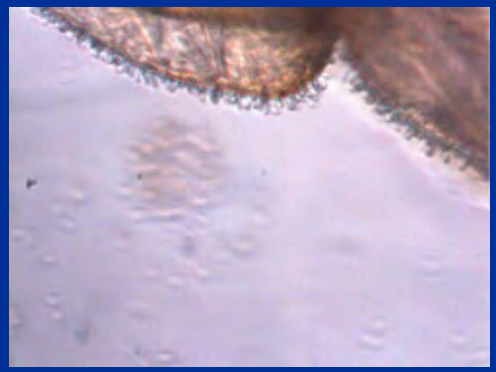
  - Pathogen – *a bacterium, virus, or other microorganism that can cause disease*
  - Easily observed
  - Mitigated by clean techniques, test modification (i.e., increase # replicates with 2 fish per replicate), and/or sample treatment (e.g., UV, filtration, chlorination/dechlorination, and antibiotics)

‘False Positives’ – Microbes

- *Ceriodaphnia dubia*
  - Epibionts – organisms that live on the external surface of another organism
  - Other microbial interferences
    - Reduced reproduction
  - Mitigation measures:
    - Assure that compositor tubing is replaced before each compliance monitoring event to avoid costly accelerated monitoring and TIEs

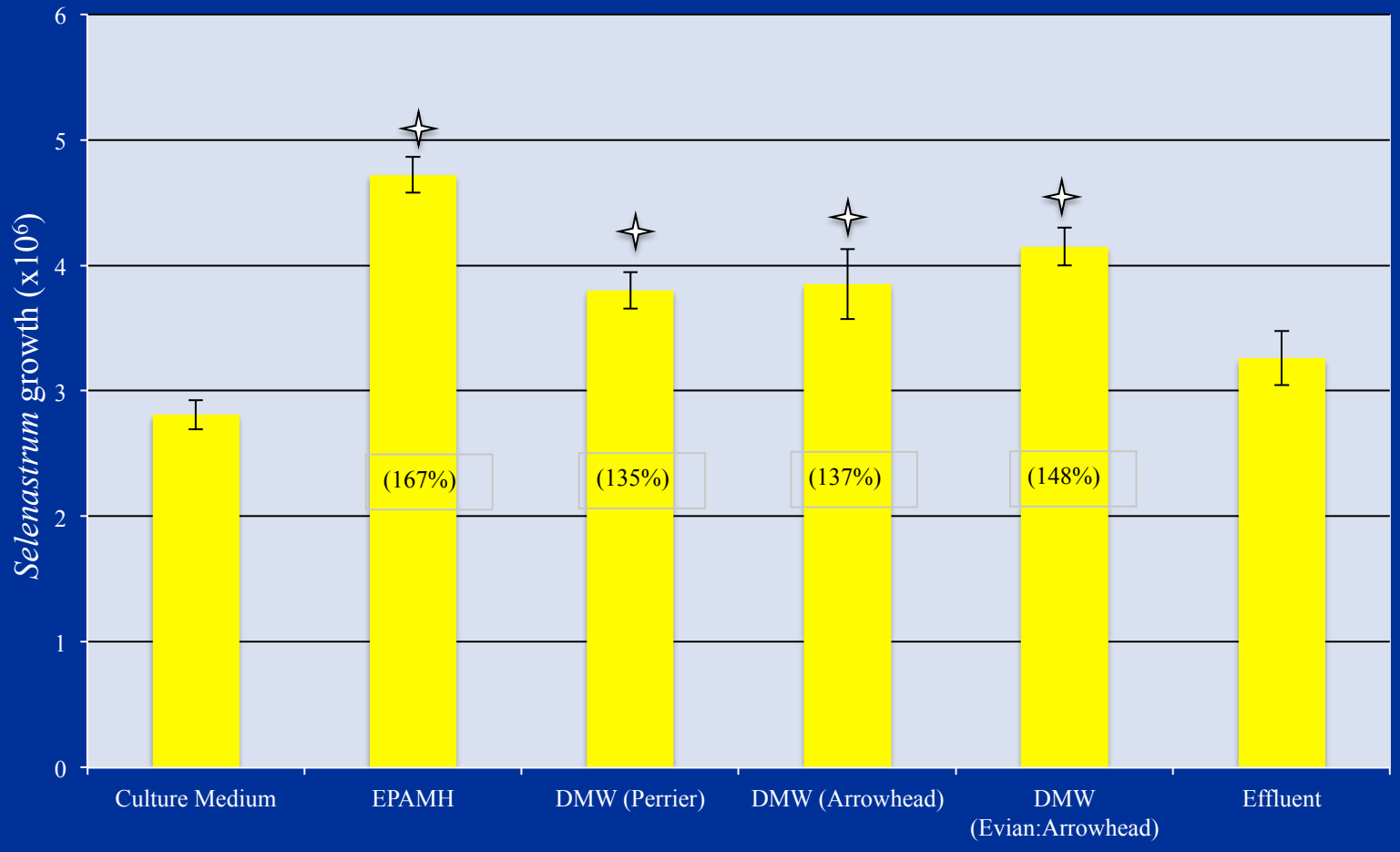






‘False Positives’ – Lab Control Media

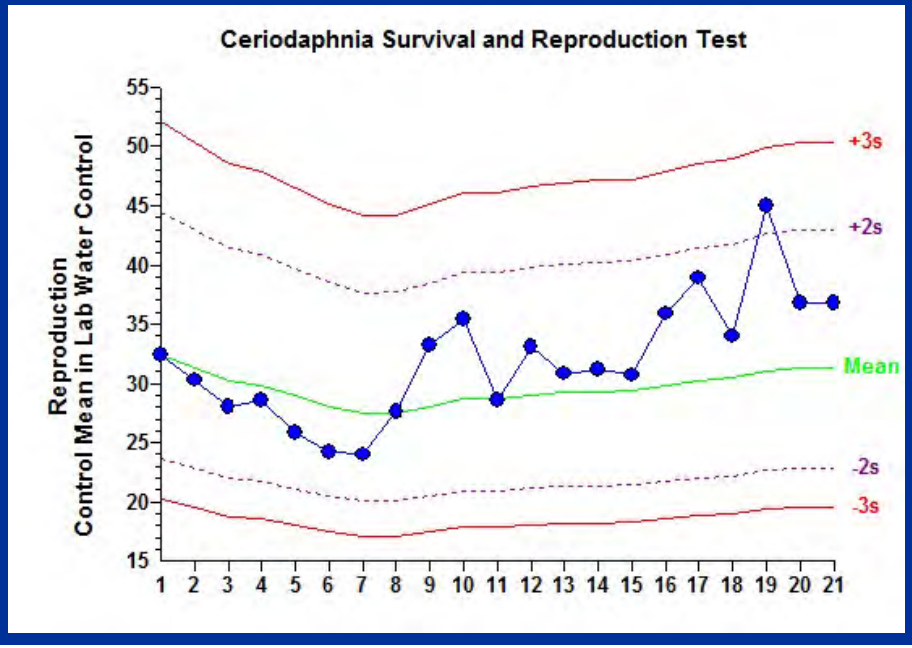
- Lab control treatment media selection can affect the determination of toxicity for the *Selenastrum* test



‘False Positives’ – Lab Control Well Above Historical Mean

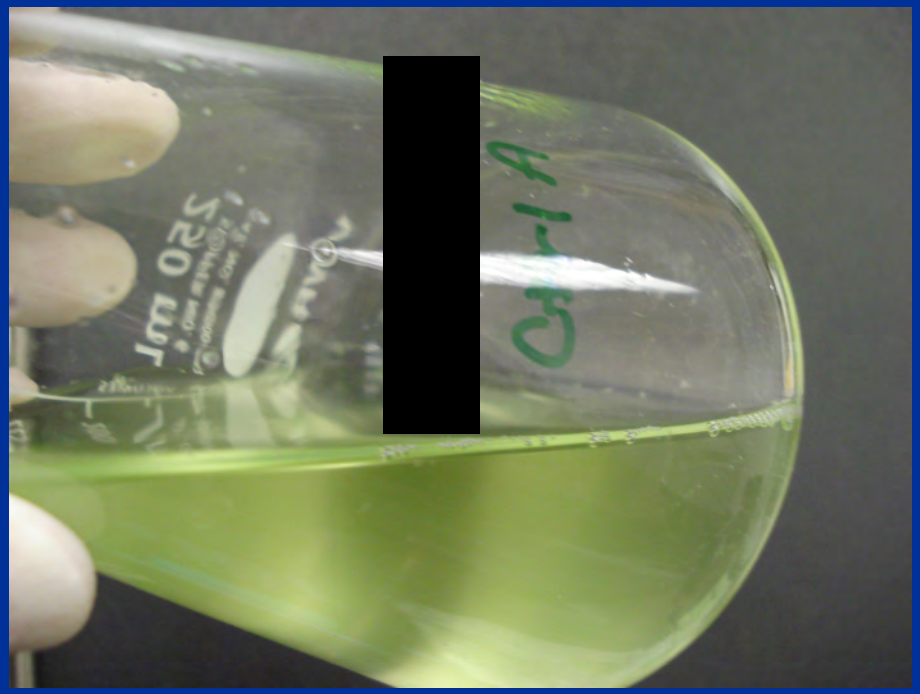
- Is the effluent identified as toxic due to a stimulated Lab Control treatment (i.e., significant effects at all test concentrations)?

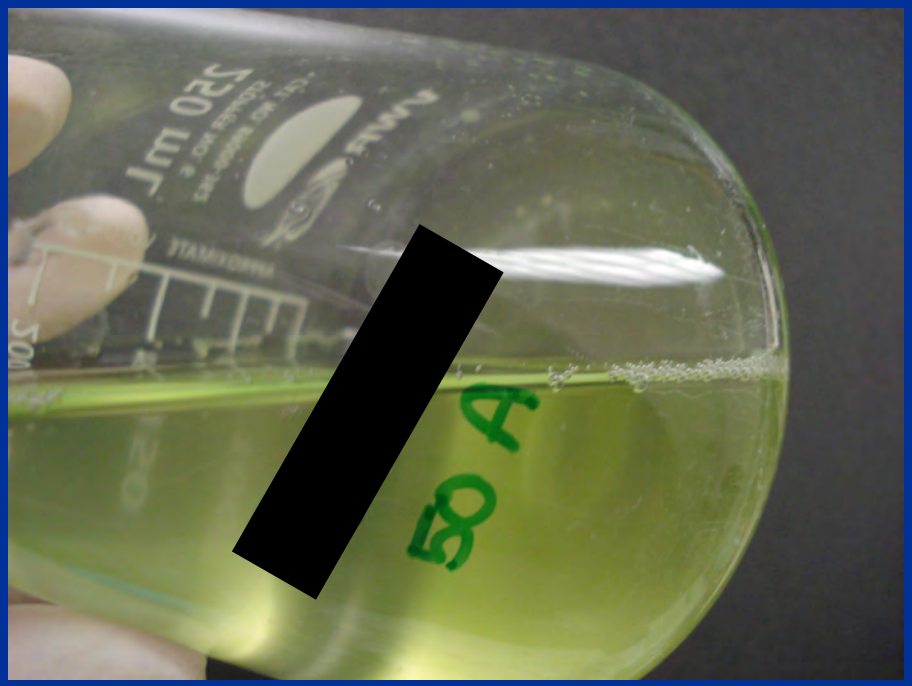
Test Effluent Treatment	% Survival	Reproduction (# neonates/female)
Lab Water Control	100	42.7
Culture Control[a]	100	31.9
12.5%	90	31.1
25%	100	36.5
50%	100	38.3
75%	100	33.5
100%	80	30.4[b]



‘False Positive’ – *Selenastrum*

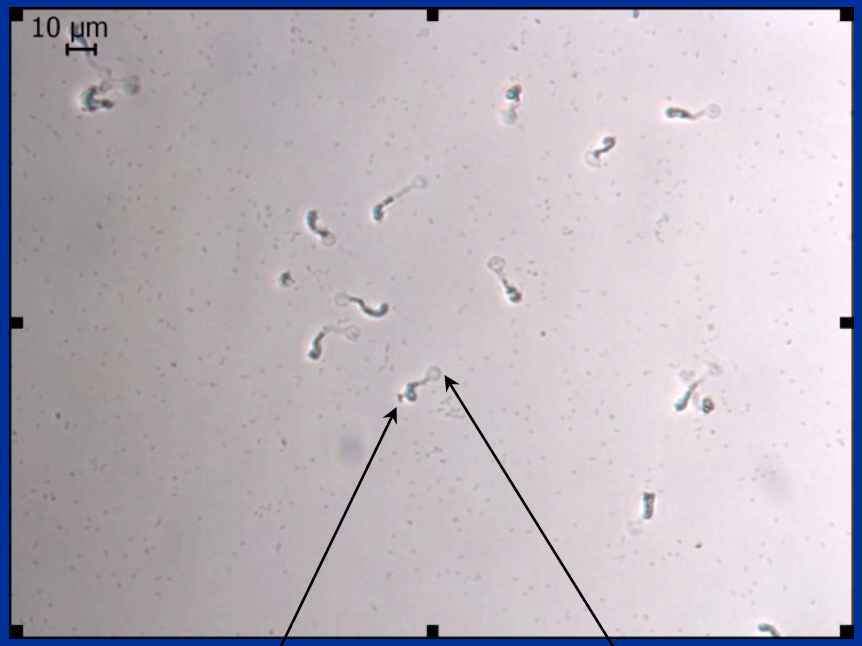
- ‘Plating’ of algae can occur, reducing the algal count
- Need to re-suspend the algae to obtain a count in the flask



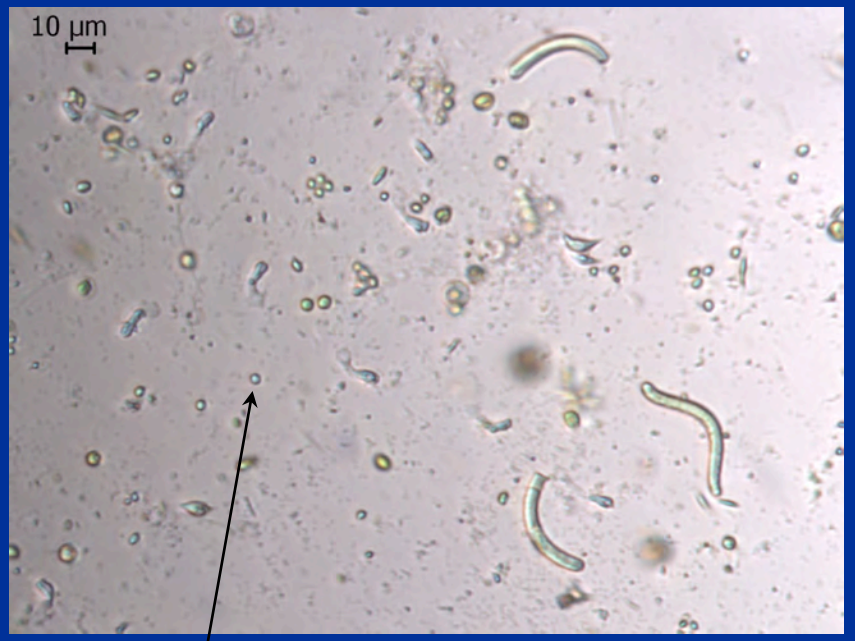


‘False Positive: Giant Kelp

- Invalid test due to elevated ocean temperatures (≥72 ºF).
- Method-required 60 μm filtration of samples allows for presence of resident organisms that can confound the test.



Embryonic gametophyte Kelp spore

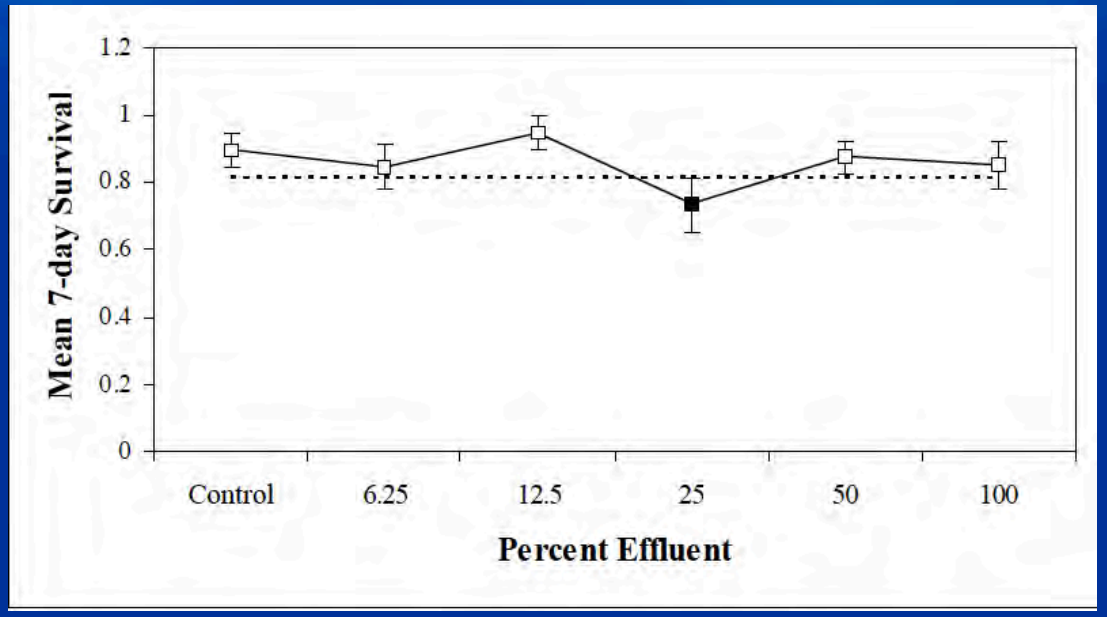


Kelp spore or resident organism from sample?

Concentration Response Relationships Must Be Assessed

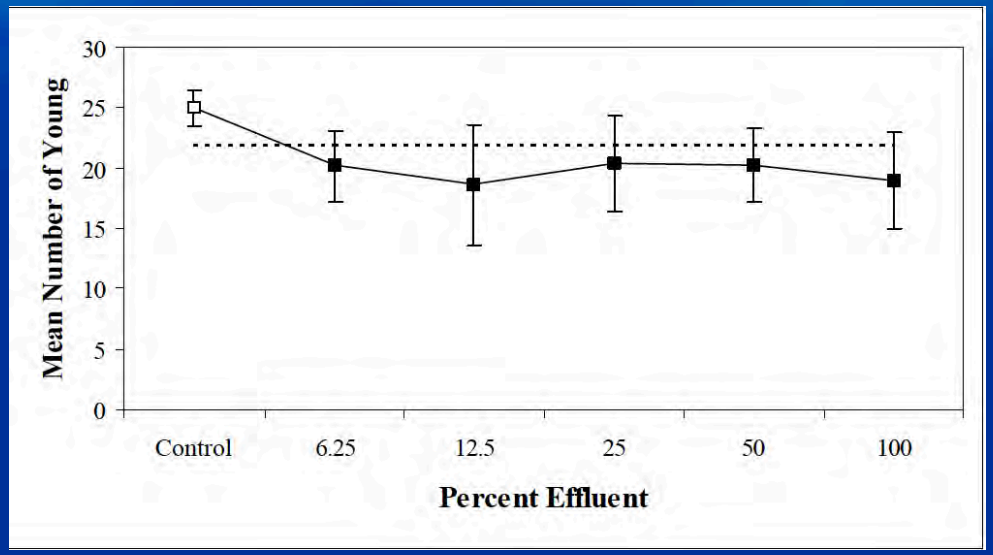
- All or nothing
- Stimulatory at low concentrations and detrimental at higher concentrations
- Stimulatory at low concentrations but no effect at higher concentrations
- Interrupted concentration response - significant effect bracketed by non-significant effect
- Interrupted concentration response - non-significant effect bracketed by non-significant effect
- Significant effects only at highest concentration
- Significant effects at all test concentrations but flat concentration response curve
- Significant effects at all test concentrations with a sloped concentration response curve
- Inverse concentration response relationship

Unusual Response Curve Example



- Must evaluate for procedural errors (e.g., D.O.), within treatment variability, and test sensitivity (PMSD).
- Outcome could be that the 25% effluent treatment in the example above is an outlier and that the NOEC should be 100% effluent (i.e., not toxic) versus the 12.5% effluent (i.e., VERY toxic).

Unusual Response Curve Example



- Must evaluate test sensitivity (PMSD), unusually high control response, dilution water (lab vs. receiving water), and ***consider*** pathogen interference
- Be cautious to not jump to conclusions that pathogens are the only driver. If weight of evidence leads to pathogens, perform appropriate treatments (e.g., filtration, UV, chlor/dechlor, and antibiotics) for conclusion.

Other Potential 'False Positives'

- Matrix interferences:
  - Low hardness waters
  - High hardness waters
  - Can cause cell lysing in *Selenastrum*
  - Can cause low reproduction to mortalities in *Ceriodaphnia*
- Basic testing errors and reporting errors

Conclusions

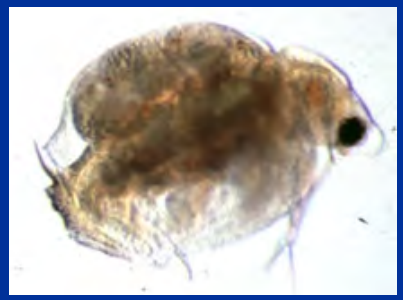
Conclusions

- It is imperative that a critical evaluation of test data is performed to assure that an effluent preliminarily identified as toxic in fact is toxic (avoid false positives) and use appropriate ‘off ramps’ to avoid unnecessary accelerated monitoring and TIEs
  - Assure that test data have been comprehensively reviewed and that the results are acceptable
  - Assure that test interferences have been eliminated as causative factors for a “toxic” response
  - Address plant operations (TRE)

Key Points of the State Board Toxicity Policy and the TST Statistic Related to the Laboratory

Stephen L. Clark

Pacific EcoRisk



BACWA Lab Committee
March 11, 2015

Toxicity Objective

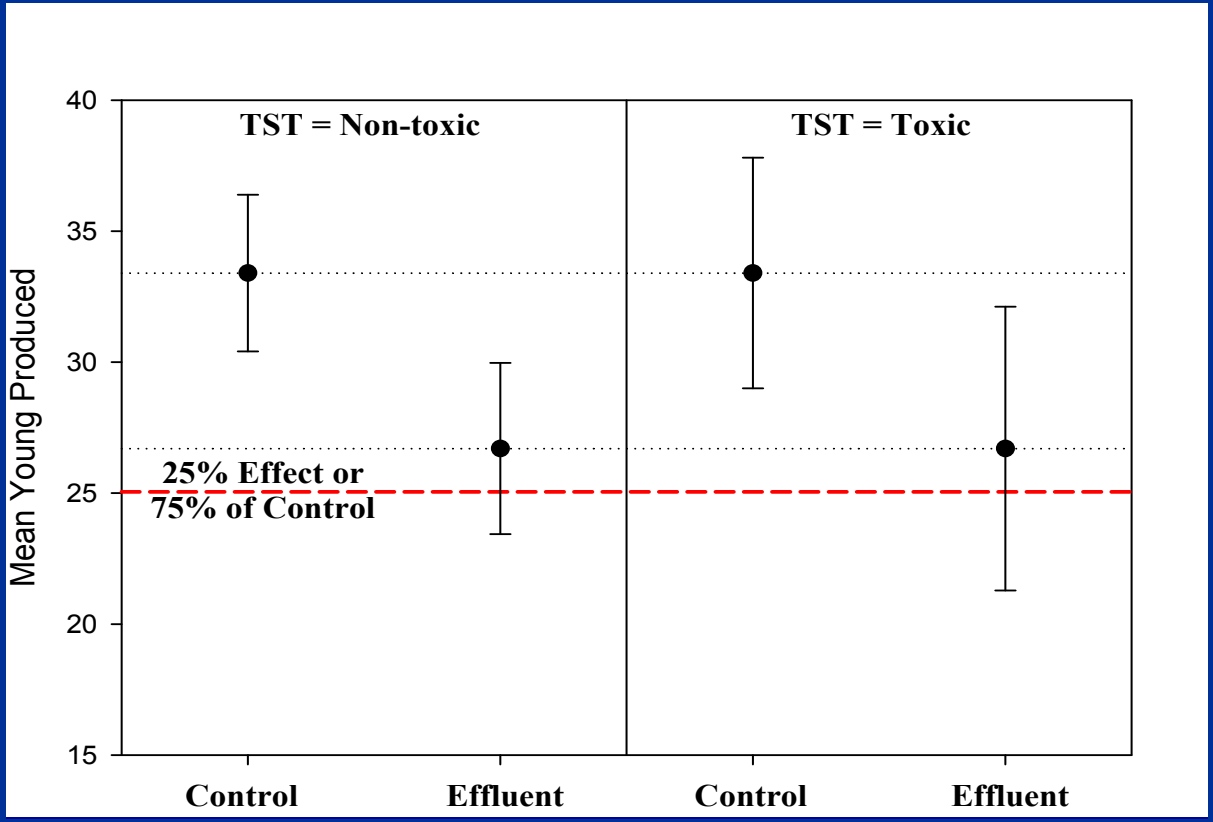
- Current objective is narrative – "no toxics in toxic amounts"
- Transition to new numeric objective applied to monthly toxicity testing requirement
- Regulated in permits as maximum daily effluent limit (MDEL) and monthly median effluent limit (MMEL)
  - Chronic Toxicity
    - a ≥25% effect at the *instream waste concentration (IWC)* would be deemed an unacceptable level of toxicity (= "fail").

TST Statistical Analysis

- Both the NOEC and the TST use statistical hypothesis testing
- San Francisco Bay region – EC/IC point estimates used for permit compliance
- All are fundamental statistical concepts
- TST assumes guilty until proven innocent. It 'flips' the hypothesis testing question
- Therefore, implications associated with 'within test' variability and replication are very important

TST Statistical Analysis

- 'Within test' variability:
  - If test variability is high, it is possible to have less than a 25% effect (chronic) and still result in a "fail" by the TST



High variability tests will have a greater likelihood of "failing" via the TST

TST Statistical Analysis

- 'Within test' variability:
  - Toxicity labs should work to minimize variability
    - Staff training, QC testing, etc.
    - Dischargers should consider adopting more rigid QA requirements for their contract labs
      - May consider adopting NELAP approach for QA evaluation
  - Parameters out of the control of the lab
    - Stress to organisms during shipping
    - Precision of different tests may vary
  - Implement strategies to reduce variability – may include test design modification

Does Increasing Replication Increase the Probability of Passing?

Test Endpoint	% Reduction	S.D.	Number of Replicates		
Americamysis bahia growth			8 replicates	12 replicates	16 replicates
	15%	25th percentile	Mostly pass	Mostly pass	Always pass
		50th percentile	Mostly pass	Mostly pass	Mostly pass
		75th percentile	50/50	Mostly pass	Mostly pass
	20%	25th percentile	**50/50**	Mostly pass	Mostly pass
		50th percentile	--	50/50	50/50
		75th percentile	--	--	50/50

- Mostly pass – *in general*, >75% probability of passing
- Always pass - >95% probability of passing

Does Increasing Replication Increase the Probability of Passing?

Test Endpoint	% Reduction	S.D.	Number of Replicates		
Ceriodaphnia dubia reproduction			10 replicates	15 replicates	20 replicates
	15%	25th percentile	Mostly pass	Always pass	Always pass
		50th percentile	Mostly pass	Mostly pass	Mostly pass
		75th percentile	50/50	Mostly pass	Mostly pass
	20%	25th percentile	Mostly pass	Mostly pass	Mostly pass
		50th percentile	50/50	50/50	Mostly pass
		75th percentile	--	50/50	50/50

- Mostly pass – *in general*, >75% probability of passing
- Always pass - >95% probability of passing

Does Increasing Replication Increase the Probability of Passing?

Test Endpoint	% Reduction	S.D.	Number of Replicates		
Oyster, Mussel, Urchin, and Sand Dollar Development tests			5 replicates	7 replicates	10 replicates
	15%	25th percentile	Always Pass	Always Pass	Always pass
		50th percentile	Always Pass	Always Pass	Always Pass
		75th percentile	Mostly Pass	Always Pass	Always Pass
	20%	25th percentile	Always Pass	Always Pass	Always Pass
		50th percentile	Mostly Pass	Always Pass	Always Pass
		75th percentile	50/50	Mostly Pass	Mostly Pass

- Mostly pass – *in general*, >75% probability of passing
- Always pass - >95% probability of passing

EPA/Tetra Tech Guidance

- Where beneficial, increase the number of test replicates in each test treatment (e.g., Control and IWC) above the minimum required
  - Two factors affect this decision:
    - How small of an effect is typically observed when testing the sample?
    - How low is the lab's within-test CV for the test endpoint of interest? *Happy if at the EPA 50th percentile.*
  - Further recommending doubling the number of replicates should be sufficient with not much additional benefit beyond doubling
- Select laboratories with QA/QC practices that contribute to increased within-test precision

Monitoring and Reporting Requirements

- What are the implications of a "failed" routine compliance test?
  - Accelerated Monitoring Requirements for exceeding the MDEL or MMEL
    - At a minimum, discharger is to initiate two, full dilution chronic toxicity tests conducted within the the month that the compliance test was initiated
    - Dilution series test will include the IWC and four concentrations
    - A chronic test that fails at the IWC and exhibits ≥ 25% reduction requires the initiation of a TRE
      - Discharger must submit a TRE Work Plan to the Water Board

Conclusions

- State Board Toxicity Policy
  - TST statistical analysis of the data will be required
    - Automatically fails if >25% reduction in organism response compared to Control
    - Rewards high precision testing (and high precision test methods)
    - May benefit from increased replication
  - Minimum – test at IWC and Control
  - If contracting out work, select a lab with a stringent QA program (e.g., NELAP)