Statistical and Biological Interpretation of Test Endpoints Relative to Unacceptable Risk Levels

Frank S Dillon/CH2M HILL

Chris A. Mebane/U.S. Geological Survey

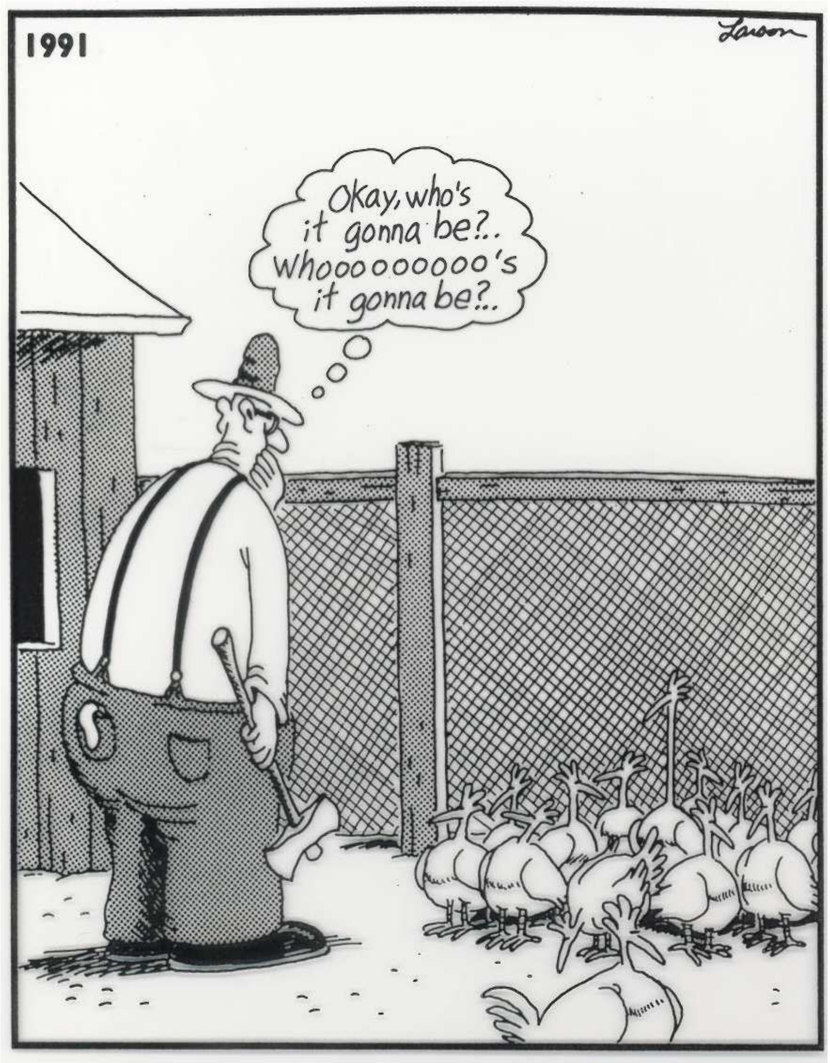
Dan P. Hennessy/Anchor Environmental, L.L.C.

How Should We Interpret Toxicity Tests?

What endpoint should we select?

How should we estimate adverse effects?

What to Do?



Estimating Chronic Effects: Problem Statement

Current guidance specifies that chronic criteria can be derived through:

- A series of chronic tests with a range of species,
- or through the development of acute-to-chronic ratios (ACRs).

There is the potential for significant variation in ACRs depending on a number of test factors and statistical analysis approaches.

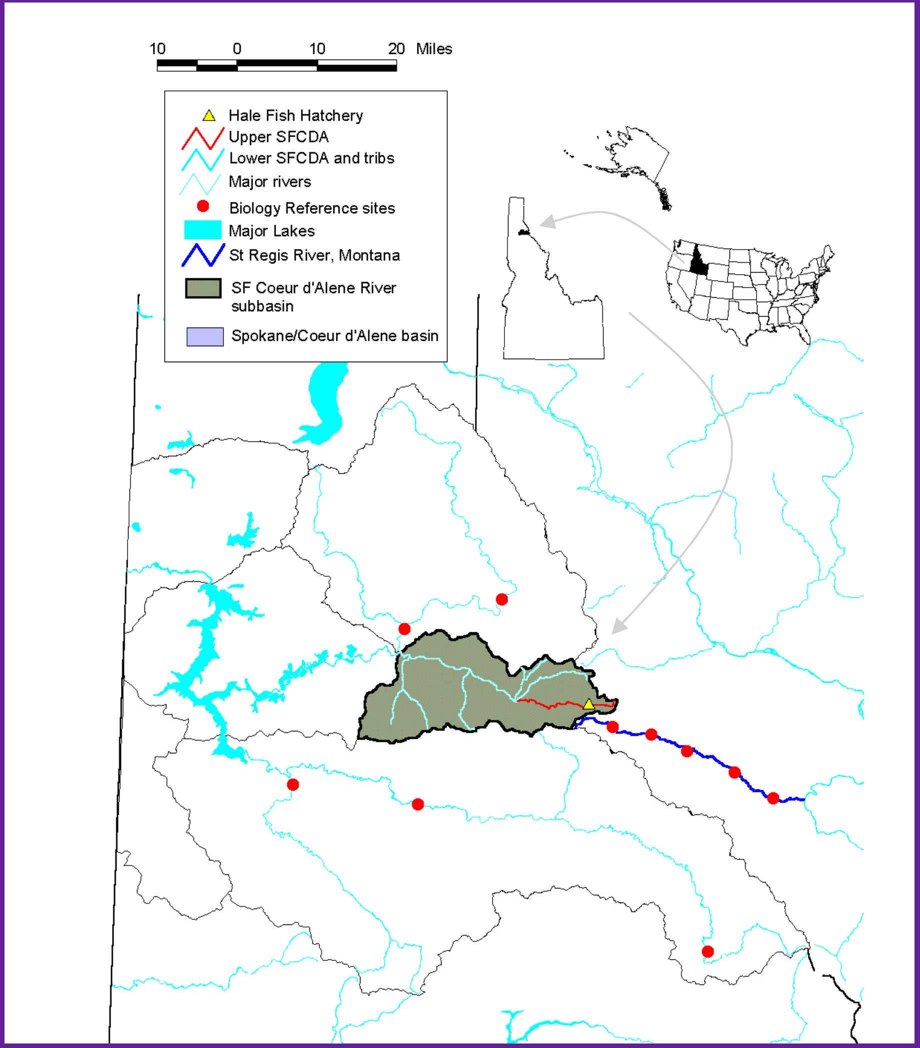
Variability of ACRs

- We found that the choice of test endpoint and statistical analysis influenced ACR estimates by up to a factor of four.

Acute Value	ACR	Chronic Value
100 ug/L	8	12.5 ug/L
100 ug/L	2	50 ug/L

Project Background

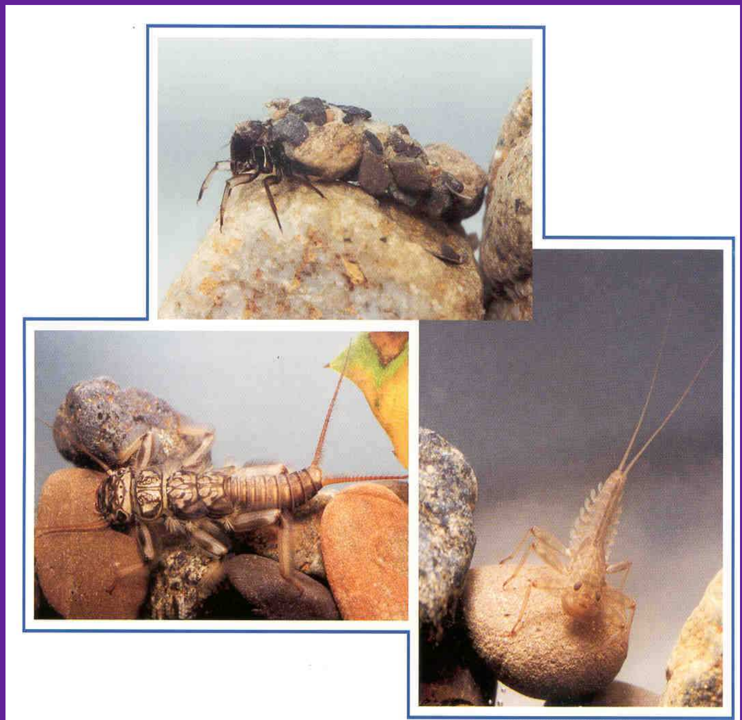
Development of Water Quality Criteria for Metals for the South Fork Coeur d'Alene River Basin, Idaho



Resident Species Approach

1. Range-finding tests to estimate relative sensitivity among indigenous species
2. Definitive testing to establish acute values
3. <u>Chronic testing with trout and invertebrates to estimate acute to chronic ratios</u>
4. Establish hardness-toxicity relationships
5. Metal mixtures toxicity testing

Acute Testing Overview



- Tested Westslope Cutthroat trout and Shorthead sculpin
- Tested numerous macroinvertebrates, especially mayflies
- Keystone species and potentially sensitive species
- Cadmium (Cd), lead (Pb), and zinc (Zn) in 96-h acute exposures

Chronic Testing Overview

Rainbow trout Cd, Pb, and Zn 60+ day early-life stage (ELS) exposures.

We also tested the acute and sublethal responses of a mayfly (*Baetis tricaudatus*) and a midge (*Chironomus dilutus*) with Pb.

Results of Chronic Testing

Analysis Approach

- Acute tests
  - Exposure/Dose concentrations for use in statistical analysis were calculated as geometric means of concentrations at test initiation and at 96 h.
  - LC_{50} estimated by regression analysis of the geometric means for 96 hr tests.
- Chronic Test
  - test concentrations for effects analyses were calculated as time-weighted arithmetic averages (TWA) of measured concentrations.
  - Test results were analyzed both by hypothesis testing and regression analyses.
    - Treatment responses were compared to control responses using Dunnett's multiple comparison test using α of 0.05 (one tailed) to define "significant" results (EPA 1993; Zar 1984).
    - If Bartlett's test indicated heterogeneity of variance, we used a nonparametric version of Dunnett's test with rankit-transformed data.
    - Since the rainbow trout early life stage (ESL) tests with the different metals were run side-by-side in the hatchery with the same water source and using eggs from a common cohort, we pooled the control replicates in order to increase the statistical power of Dunnett's test to detect changes from controls (Zar 1984:195).

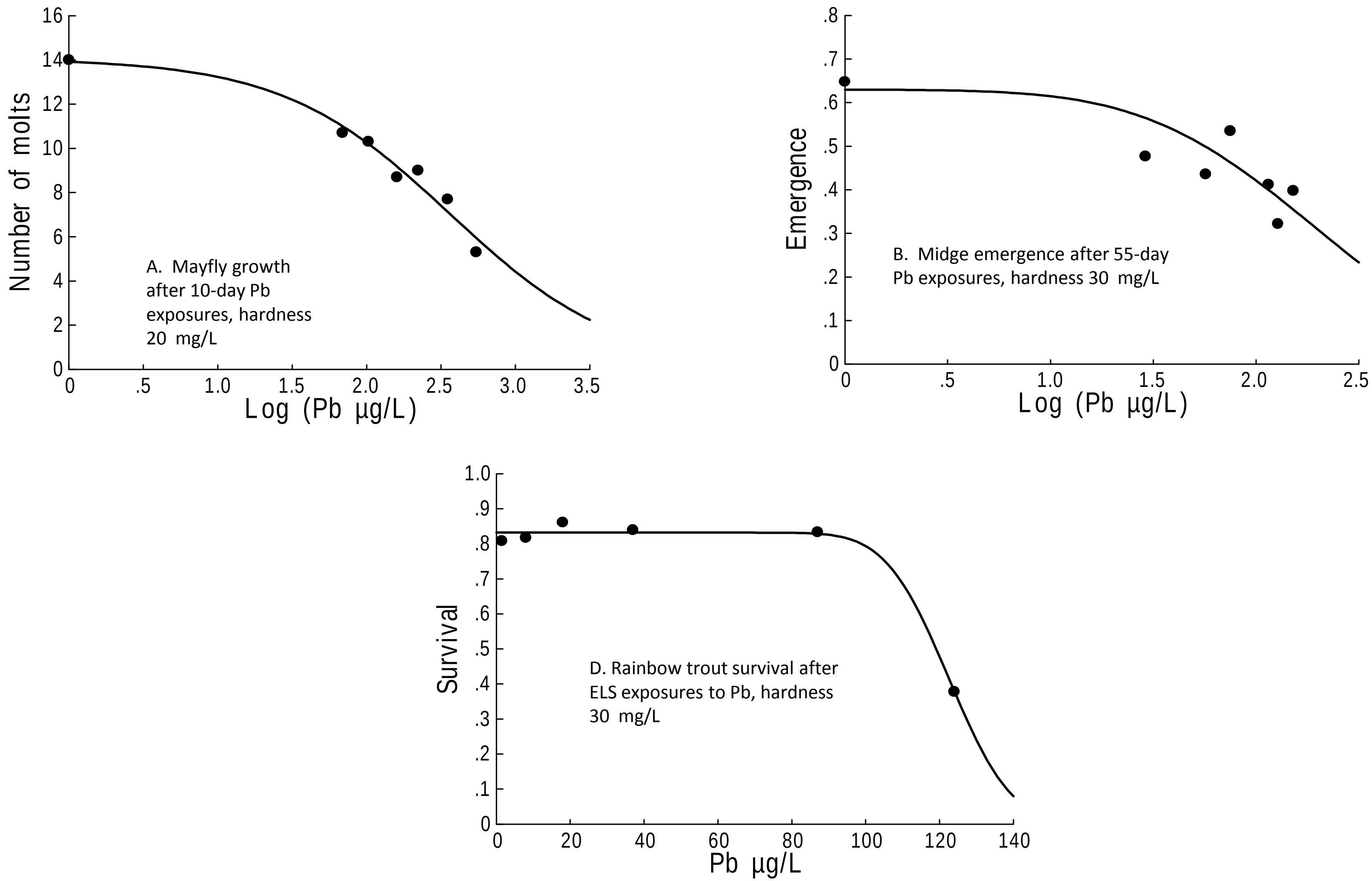
Analysis Approach

- Two values were derived from Dunnett's test,
  - No-observed-effect concentration (NOEC) = highest concentration with response = not statistically different from controls
  - Lowest-observed-effect concentration (LOEC) = lowest concentration with response statistically different from controls
  - NOEC and LOEC commonly used as "chronic limits;"
- The hypothetical "maximum acceptable threshold concentration" (MATC) or "chronic value" is an assumed threshold for toxic effects between the NOEC and LOEC.
- The MATC is a point estimate set as the geometric mean of the NOEC and the LOEC.
- We report NOEC, LOEC, and MATC values for our tests. We also report estimates of the minimum statistically detectable difference (MDD) of the hypothesis testing to detect minimum responses when setting α of 0.05.

Analysis Approach

- The 10th and 20th percentile effects concentrations (e.g. EC10, EC20, ECp) from distribution or regression based dose–response curves were also calculated.
- Mortality ECp values were calculated using a three factor probit model (Caux and Moore 1997; Erickson 2002).
- Growth or reproduction inhibition effects concentrations were estimated by least-squares nonlinear regression analysis (Erickson 2002) or nonparametric linear interpolation (Norberg-King 1993).
- We calculated EC10s and EC20s for the following reasons.
  - EC10s are assumed to be near the low limits of an ECp value that could be calculated with confidence using usual test procedures and EC10s have been suggested as a regression-based replacement for the NOEC. (Moore and Caux 1997; van der Hoeven et al. 1997).
  - We also calculated EC20s since they have been used recently in lieu of MATC values to derive criteria or to report sublethal effects (e.g., EPA 1999; Hansen 2002a).

Selected Chronic Toxicity Relationships



Mayfly subchronic responses to dissolved Pb

Measured Pb, µg/l	**Control (<3)**	**69**	**103**	**160**	**222**	**350**	**546**	**Minimum detectable difference from control (% of control)**
Percent mortality	31.3 (20.8)	24.2 (6.1)	31.3 (13.1)	33.3 (10.5)	45.5 (9.1)	72.7[a] (0)	93.9[a] (5.2)	23% (72%)
Number of molts	14.0 (3)	10.7 (2.3)	10.3 (2.3)	8.7[a] (2.5)	9.0[a] (1.7)	7.7[a] (1.2)	5.3[a] (2.3)	4.6 (32%)

[a] – Significantly different from control at $P < 0.05$ using Dunnett's procedure.

Mean (± SD, n) hardness 20.7 mg/l as $CaCO_3$ (± 0.58, 3); alkalinity 19.8 mg/l as $CaCO_3$ (± 1.04, 3); pH 6.64 (± 0.18, 16); conductivity 47.7 µS (± 1.72, 20); DO 10.1 mg/l (± 0.45, 22) temperature 9.3°C (± 0.67, 22).

Midge Responses to Dissolved Pb

Nominal Pb µg/l	**Control**	**50**	**79**	**126**	**199**	**315**	**500**	**Minimum detectable difference from control (% of control)**
Measured Pb (± SD)	***<3***	***29 (9.8)***	***57 (35)***	***75 (16)***	***115 (40)***	***128 (32)***	***152 (38)***	
Mortality –day 20 (%)	38.9 (18.9)	66.7 (26.2)	41.7 (19.5)	41.7 (22.1)	38.9 (31.2)	38.9 (22.9)	33.3 (25.0)	48% (79%)
Weight – day 20 (mg)	2.95 (0.12)	2.24[a] (0.37)	2.4 (0.35)	1.98[a] (0.39)	1.47[a] (0.05)	1.56[a] (0.14)	2.1[a] (0.29)	-0.57 mg (19.6%)
Eggs/case – day 20	690 (144)	746 (181)	792 (242)	762 (210)	646 (169)	751 (190)	798 (202)	-240 (35%)
Percent emergence – day 55	64.8 (10)	47.7 (19)	43.6 (10)	53.5 (19)	41.2 (5.1)	32.2[a] (7.8)	39.8[a] (4.7)	-25% (38%)

[a] – Significantly different from control at P <0.05 using Dunnett's procedure.

Average (± S.D. n) hardness 32 mg/l as $CaCO_3$ (± 3.2, 8); alkalinity 31 mg/l as $CaCO_3$ (± 3.0, 7); pH 7.26 (± 0.21, 18); conductivity 76 (± 4.9, 18) µs; dissolved oxygen 7.8 (± 0.8, 112) mg/l; temperature 22.2°C (± 1.0, 112). Average (maximum, n) dissolved metals in control samples: Cd <0.2 µg/l (<0.2, 9), Pb <3.0 µg/l (5, 9), Zn 12 µg/l (62, 9). nc – not calculable

Mortality and Growth of Rainbow Trout Exposed to Cd

Nominal Cd	**Control**	**0.25**	**0.57**	**1.3**	**2.9**	**6.6**	**Minimum detectable difference from control (% of control)**
Measured Cd (±SD)	***<0.2***	***0.3 (0.05)***	***0.6 (0.12)***	***1.3 (0.20)***	***2.9 (0.47)***	***6.9 (0.77)***	
Mortality (%) (53days, note a)	14.1 (5.9)	14.8 (4.6)	15.6 (6.7)	26.7[a] (11.8)	70.4[a] (6.8)	85.2[a] (8.4)	7.6% (8.9%)
Mortality (%)	18.7 (4.8)	b	b	33.3[a] (9.7)	86[a] (7.1)	96[a] (6.4)	10.3% (12.7%)
Wet weight (g) (a)	0.218 (0.04)	0.208 (0.01)	0.219 (0.01)	nm	nm	nm	- 0.048 g (22%)
Wet weight (g)	0.369 (0.02)	b	b	0.342 (0.04)	0.222[a] (0.04)	0.146[a]	- 0.031 g (8.5%)
Biomass (g)	13.47 (0.72)	b	b	10.2[a] (0.65)	1.3[a] (0.36)	0.049[a]	-0.93 g (7.0%)
Length (mm)	31.6 (1.7)	b	b	31.7 (1.9)	28.4 (5.9)	25[a]	-3.0 mm (9.5%)

[a] – Significantly different from control at P <0.05 using Dunnett's procedure. b – Results as of day 53; the two lowest Cd treatments were overdosed on day 54, causing complete mortality in those treatments. nm – not measured.

Mean (±SD, n) hardness 19.7 mg/l as $CaCO_3$ (1.5, 17); alkalinity 19.6 mg/l as $CaCO_3$ (2.2, 17); pH 6.75 (0.41, 570); conductivity 45.8 µs (2.2, 475); DO 10.2 (0.70, 982) mg/l; temperature 9.8°C (0.6, 1474). Maximum metals in control samples (n=12) – Cd <0.2 µg/l, Pb <3.0 µg l^-, Zn <10 µg/l.

Mortality and Growth of Rainbow Trout Exposed to Pb

Nominal Pb	**Control**	**10**	**22**	**48**	**100**	**230**	**Minimum detectable difference from control (% of control)**
Measured Pb (±SD)	***<3***	***12 (2)***	***24 (3)***	***24 (9)***	***54 (25)***	***143 (20)***	
Mortality (%)	18.7 (4.8)	16.3 (6.7)	19.3 (4.6)	20.0 (5.9)	60.7[a] (20)	94.1[a] (4.6)	9.9% (12.2%)
Wet weight (g)	0.369 (0.02)	0.324 (0.05)	0.357 (0.07)	0.360 (0.01)	0.282[a] (0.03)	0.108[a] (0.04)	-0.033 g (8.9%)
Biomass (g)	13.47 (0.72)	12.27 (2.7)	12.98 (0.64)	12.96 (0.91)	4.82[a] (2.2)	0.31[a] (0.26)	-1.6 g (11.8%)
Length (mm)	31.6 (1.7)	31.36 (1.9)	31.37 (1.3)	32.61 (0.63)	29.07 (1.4)	21.30[a] (2.9)	-2.1 mm (6.7%)

[a] – Significantly different from control at P <0.05 using Dunnett's procedure. b – Results as of day 53; the two lowest Cd treatments were overdosed on day 54, causing complete mortality in those treatments. nm – not measured.

Mean (±SD, n) hardness 19.7 mg/l as $CaCO_3$ (1.5, 17); alkalinity 19.6 mg/l as $CaCO_3$ (2.2, 17); pH 6.75 (0.41, 570); conductivity 45.8 µs (2.2, 475); DO 10.2 (0.70, 982) mg/l; temperature 9.8°C (0.6, 1474). Maximum metals in control samples (n=12) – Cd <0.2 µg/l, Pb <3.0 µg l^-, Zn <10 µg/l.

Mortality and Growth of Rainbow Trout Exposed to Zn

Nominal Zn	Control	100	170	300	500	870	Minimum detectable difference from control (% of control)
Measured Zn (± SD)	***<10***	***113 (16)***	***208(28)***	***362 (27)***	***607(200)***	***900 (414)***	
Mortality (%)	18.7 (4.8)	36.3[a] (11.1)	47.4[a] (10.2)	48.1[a] (12.2)	80.7[a] (12.2)	100[a]	10.3% (12.7%)
Wet weight (g)	0.369 (0.02)	0.365 (0.05)	0.317 (0.01)	0.307[a] (0.05)	0.350 (0.03)		-0.037 g (10.1%)
Biomass (g)	13.47 (0.72)	10.28[a] (0.39)	7.52[a] (1.34)	7.00[a] (0.73)	2.90[a] (1.68)		-1.12 g (8.3%)
Length (mm)	31.6 (1.7)	31.38 (1.3)	29.68 (0.40)	27.84[a] (4.73)	29.43 (0.91)		-2.5 mm (7.9%)

[a] – Significantly different from control at P <0.05 using Dunnett's procedure. b – Results as of day 53; the two lowest Cd treatments were overdosed on day 54, causing complete mortality in those treatments. nm – not measured.

Mean (±SD, n) hardness 19.7 mg/l as $CaCO_3$ (1.5, 17); alkalinity 19.6 mg/l as $CaCO_3$ (2.2, 17); pH 6.75 (0.41, 570); conductivity 45.8 µs (2.2, 475); DO 10.2 (0.70, 982) mg/l; temperature 9.8°C (0.6, 1474). Maximum metals in control samples (n=12) – Cd <0.2 µg/l, Pb <3.0 µg l^-, Zn <10 µg/l.

Comparison of Different Statistical Test Endpoints - Cd

Species	Metal	Hard.	Days	Endpoint	NOEC	LOEC	MATC	EC10	EC20
Rainbow trout	Cd	20	69	Mortality	<1.3	1.3	nc	0.89	1.4
Rainbow trout	Cd	20	53	Mortality	0.6	1.3	0.88	0.87	1.4
Rainbow trout	Cd	20	69	Weight	1.3	2.9	1.9	nc	nc
Rainbow trout	Cd	20	69	Biomass	<1.3	1.3	<1.3	nc	nc
Rainbow trout	Cd	20	69	Length	2.9	6.9	4.4	nc	nc
Rainbow trout	Cd	29	62	Mortality	1.0	2.5	1.6	1.6	3.0
Rainbow trout	Cd	29	62	Weight	<0.2	0.2	<0.2	0.18	>2.5
Rainbow trout	Cd	29	62	Biomass	1.0	2.5	1.6	ndr	1.6
Rainbow trout	Cd	29	62	Length	<0.2	0.2	<0.2	>2.5	>2.5

Table notes: nc – not calculable; ndr – no dose response, ECx value was lower than higher concentrations with higher values; w- Calculated by Weibull regression model, insufficient response for Probit model.

Comparison of Different Statistical Test Endpoints - Pb

Species	Metal	Hard.	Days	Endpoint	NOEC	LOEC	MATC	EC10	EC20
Mayfly	Pb	19	10	Mortality	222	350	279	177	213
Mayfly	Pb	19	10	Molting	103	160	130	30	58
Midge	Pb	32	20	Mortality	152	>152	>152	nc	nc
Midge	Pb	32	20	Weight	57	75	65	15	28
Midge	Pb	32	20	Fecundity	152	>152	>152	>152	>152
Midge	Pb	32	55	Emergence	152	>152	>152	10	20
Rainbow trout	Pb	20	69	Mortality	24	54	36	36	47
Rainbow trout	Pb	20	69	Weight	24	54	36	10	39
Rainbow trout	Pb	20	69	Biomass	24	54	36	25	30
Rainbow trout	Pb	20	69	Length	54	143	88	45	83
Rainbow trout	Pb	29	62	Mortality	87	125	104	93^{w}	110^{w}
Rainbow trout	Pb	29	62	Weight	37	87	57	7	>87
Rainbow trout	Pb	29	62	Biomass	87	125	104	88	97
Rainbow trout	Pb	29	62	Length	8	18	12	>87	>87
Rainbow trout	Zn	20	69	Length	208	362	274	300	>362

Comparison of Different Statistical Test Endpoints - Zn

Species	Metal	Hard.	Days	Endpoint	NOEC	LOEC	MATC	EC10	EC20
Rainbow trout	Zn	20	69	Mortality	<113	113	<113	75	123
Rainbow trout	Zn	20	69	Weight	208	362	274	187	>362
Rainbow trout	Zn	20	69	Biomass	<113	113	nc	54	106
Rainbow trout	Zn	20	69	Length	208	362	274	300	>362

Table notes: nc – not calculable; ndr – no dose response, ECx value was lower than higher concentrations with higher values; w- Calculated by Weibull regression model, insufficient response for Probit model.

ACRs Resulting from the Most Sensitive Endpoints of Each Chronic Test

Species and test endpoint	**Metal**	**96-hour LC50 (95% confidence interval)**	**ACR by MATC**	**ACR by NOEC**	**ACR by EC10**	**ACR by EC20**
Mayfly-10d molts	Pb	664 (368 – 1315)	5.2	6.4	22	11
Midge-20 day weight	Pb	3323 (1983 – 8971)	51	58	221	118
Rainbow trout- 69 day weight	Pb	120 (104 – 138)	3.3	5	12	3.1
Rainbow trout- 62 day weight	Pb	133 (a)	11	17	19	<1.5
Rainbow trout- 53 day mortality	Cd	0.84 (0.62-0.95)	0.95	1.4	0.96	0.60
Rainbow trout- 62 day mortality	Cd	0.89 (0.8 – 0.98)	0.62	0.9	0.55	0.29
Rainbow trout- 69 day biomass	Zn	123 (b)	nc	nc	2.3	1.2

Conclusions

Conclusions

- Increasing the number of control replicates by 2 to 3× decreased the minimum detectable differences by almost half.
- Pb ACR estimates mostly increased with increasing acute resistance of the organisms (rainbow trout ACRs <≈ mayfly < Chironomus).
- The choice of test endpoint and statistical analysis influenced ACR estimates by up to a factor of four.
- ACRs based on MATCs with rainbow trout and Cd were 0.6 and 0.95; Zn about 1.0; and for Pb 3.3 and 11. The comparable Pb ACRs for the mayfly and Chironomus were 5.2 and 51 respectively.
- Our rainbow trout ACRs with Pb were about 5–20× lower than earlier reports with salmonids.

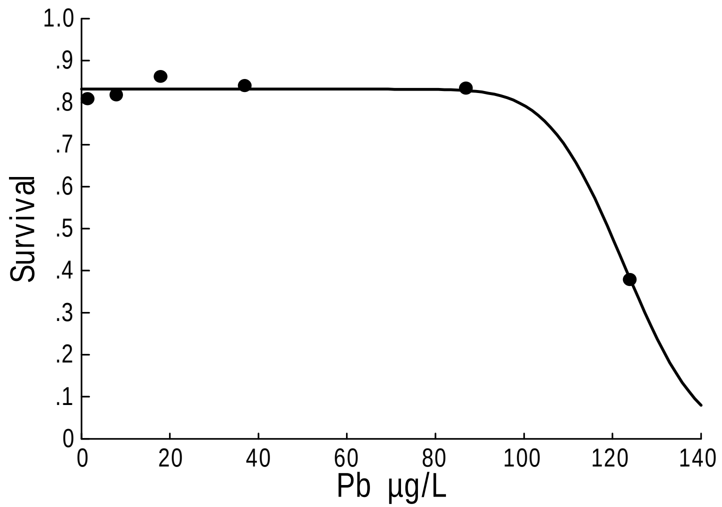
Considerations

What Did We Learn?

- In the tests we reviewed for comparison with our results, hypothesis testing was by far the most commonly used statistical analysis approach to establish chronic values and a few studies used no statistical analyses whatsoever.
- However, regression techniques have a number of advantages over hypothesis testing for making point estimates of low toxic effects, and so where possible we calculated comparable ECp values from the previous reports.
- In our tests and in most of the tests we reviewed and calculated ECp values from, NOECs were about as high or higher than EC_{10} values. In other words, NOECs corresponded with low-effect concentrations, not no-effect concentrations.
- Increasing the number of control replicates by 2 to 3× decreased the minimum detectable differences by almost half.
- Pb ACR estimates mostly increased with increasing acute resistance of the organisms (rainbow trout ACRs <≈ mayfly < Chironomus).

Implications for Water Quality Criteria Development

- Continue to standardize test and analysis methods.
- Consider tests designed to optimize the minimal detectable differences and increase confidence in affects estimate.
- Continue discussions on ecologically meaningful effects thresholds. For example MATC, EC_{10} or EC_{20}.



Thank You

Frank S Dillon

Principal Technologist

CH2M HILL

7927 Nemco Way

Suite 120

Brighton, MI 48446

1-517-488-4901

frank.dillon@ch2m.com