Independent Technical Evaluation of Environment Canada's "Second National Assessment of Environmental Effects Monitoring (EEM) Data from Metal Mines Subjected to the *Metal Mining Effluent Regulations*" (MMER)

PREPARED FOR:
MINING ASSOCIATION OF CANADA

November 2012

Table of Contents

Table of Contents

List of Tables

Page

List of Figures

Page

1.0 Introduction

1.1 ENVIRONMENTAL EFFECTS MONITORING

The *Metal Mining Effluent Regulation* (MMER) requires metal mines in Canada to conduct Environmental Effects Monitoring (EEM) to assess the potential effects of mine effluent on receiving waters and their resident biota on a national basis. The EEM program (Environment Canada 2011) provides a framework for the consistent collection of data on five endpoints regarding fish health, and four endpoints regarding benthic invertebrate community structure. For fish health, the endpoints are condition, liver size, gonad size, weight-at-age, and age, and for benthic invertebrate-community structure the endpoints are density, taxon richness, the Bray-Curtis Index (BCI) and Simpson's evenness. The stated primary goal is to determine if the MMER effluent regulations, on a national basis, are (on average) protective of the receiving environment (Environment Canada 2012).

1.2 CONCERNS ABOUT EEM DESIGNS AND DATA INTERPRETATION

Environment Canada has interpreted the national EEM data and concluded that measurable environmental impacts are being caused by individual mines and the mining industry at large (Lowell et al. 2006, Environment Canada 2012). The Mining Association of Canada (MAC) is concerned about the technical validity and strength of Environment Canada's interpretation of the EEM data. Previously MAC retained independent technical expertise to review the scientific validity of Environment Canada's administration of the *Metal Mining Effluent Regulations* (MMER), especially on a site-specific basis. Findings of errors and systematic bias in the interpretation of the EEM data resulted in publications documenting these findings in respected aquatic toxicology journals (Huebert et al. 2011; Huebert 2012a, b). The findings also encouraged discussion between MAC and Environment Canada in December 2011 to attempt resolution of the identified issues. The findings included a pseudo-replicated basic study design, incorrect designation of experiment-wise error, and biased calculation of the Bray-Curtis Index.

1.3 SECOND NATIONAL EEM DATA ASSESSMENT

Environmental Effects Monitoring administered under the MMER has now been completed for two national assessment periods of monitoring (Cycle 1 and Cycle 2), and Environment Canada has just released its Second National Assessment of the EEM monitoring data (Environment Canada 2012). The second national assessment is a meta-analysis of the data generated from Cycle 1 and Cycle 2 of the metal mines EEM Program. The assessment uses the effect-size statistic, Hedges' d (=[Exposure Mean-Reference Mean]/Pooled Standard Deviation), and associated confidence intervals (CI), to summarize and interpret the EEM data for fish health and the benthic invertebrate community.

Hedges' d is a standardized, unitless statistical measure of the magnitude of biological effects. It was designed to compare and interpret results across multiple studies and endpoints (Nakagawa and Cuthill 2007). The Hedges' d test statistic, combined with presentation of the CI, is a powerful tool for assessing the relationships within data and their biological importance, and allows practitioners to view, summarize and interpret results in the context of previous and ongoing monitoring studies.

In their Second National Assessment report, EC concluded that effluent exposure was '…*more often associated with reductions rather than increases in the indicators (endpoints)*…', and that, '… *fish collected in areas exposed to effluent… showed significantly reduced condition, relative liver size and growth rate. Other effects included some reductions in gonad… size and a significantly increased age structure. In other words, fish collected in areas exposed to effluent were, on average, older, thinner and slower growing, with smaller livers and with more of a tendency to reduced gonad size*.' The second assessment also concluded for benthic invertebrates that there was '… *significantly reduced taxon richness*.', and that there were '… *different groupings of benthic invertebrates in exposure areas compared to reference areas*.' These conclusions were based on Hedges' d data summarized in Figure 4 (fish) and Figure 14 (benthic invertebrates) of the Second National Assessment report (Environment Canada 2012).

1.4 MAC INTEREST IN INDEPENDENT REVIEWS

With Environment Canada's recent release of its second review and interpretation of national EEM data, MAC considered it prudent to request two independent technical reviews of Environment Canada's second assessment. The first review was solicited from Dr. David Huebert of Stantec Consulting Ltd.'s Winnipeg office. The second was solicited from Dr. Shinichi Nakagawa of the University of Otago, New Zealand. This report documents Stantec's third-party review of Environment Canada's Second National Assessment.

The purpose of the following review is to determine if appropriate methods were used in the meta-analysis, and to determine if the conclusions within the second assessment, as outlined in Section 1 above, were supported by the meta-data (Hedges' d and CI values).

2.0 Methodological Evaluation

Stantec's evaluation of Environment Canada's methods and approaches used in analyzing the national EEM database identified four areas of methodological concern, including; lack of calibration of the meta-data, lack of effects confirmation, use of non-significant data in declaration of effects, and lack of evaluation of the statistical structure of the EEM program.

2.1 CALIBRATION

The use of Hedges' d and the associated CI is an appropriate and powerful tool for summarizing data in a meta-analysis (Nakagawa and Cuthill 2007). The Hedges' d statistic is considered, "…*statistically significant for the group as a whole if the 95% confidence interval does not overlap the zero effect line*. (Environment Canada 2012)". Environment Canada (2012) also states that "*Statistical significance therefore reflects repeated effects in the same direction over a large number of mines – a result that is also of biological significance.*" This latter statement is the foundation upon which the Second National Assessment is based. Unfortunately, it is not an acceptable approach within the context of the EEM statistical structure (Environment Canada 2011) and meta-analysis in general (Nakagawa and Cuthill 2007). Within the EEM program, alpha, beta and sample size are carefully considered and designated to discern significant effects at the designated Critical Effect Size for studies at the individual mine level (Environment Canada 2002, 2011). Within this context, significant differences are therefore designated as 'effects'. However, the basic EEM statistical structure is absent within calculation of the meta-data, and because the Hedges' d statistic is unitless, the derived Hedges' d values have no intrinsic meaning. Thus, they require calibration if the values are to be meaningfully interpreted (Nakagawa and Cuthill 2007). The purpose of the calibration is to determine the biological meaning associated with the Hedges' d values, similar to the calibration that is undertaken within the EEM statistical structure. Nakagawa and Cuthill (2007) state that, "…*it is important to know what magnitude of effect size constitutes something biologically important*.", and that biologists "…*should evaluate their effect sizes in light of their hypotheses and also of results from previous relevant studies*."

2.1.1 Calibration Options

There are four methods by which the Hedges' d meta-data could be calibrated, including;

- **Comparison of the Hedges' d values with established cutoff benchmarks**. Absent from any understanding of biological significance, and as a first approximation, it has long been considered (e.g. Cohen 1988, cited in Nakagawa and Cuthill 2007) that a Hedges' d value below 0.2 is indicative of 'no effect'. Values below this cutoff are usually considered to be biologically negligible and within the EEM context, indicate there is no discernible effect of effluent discharge on downstream receiving environments.

- **Comparison of temporal and spatial reference replicates to determine the average reference Hedges' d and 95% CI's for each of the nine endpoints.** These average values provide a cutoff benchmark for interpretation of the results. If the Hedges' d values for 'exposure' versus 'reference' sites were the same as, or less than, the 'reference' versus 'reference' sites, then it could be considered that there was a biologically unimportant effect of mine effluent on downstream fish health and benthic invertebrate communities. The 'reference' sites Hedges' d values therefore provide a basis for evaluation of the biological significance of the results. Calculation of 'reference' Hedges' d values is the optimal approach for interpretation of the meta-analytical data, and the required reference EEM data are available to Environment Canada staff. Unfortunately, the data are not publicly available (Section 2.5 herein), notwithstanding requests for publication of the raw data, so it is not possible to calculate the 'reference' Hedges' d values independently at this time. The authors of the 2nd national assessment report have declined to provide the raw data within the report.

  Despite the lack of public access to the comprehensive set of EEM data, the Mining Association of Canada (MAC) has been able to collect 26 benthic invertebrate datasets from nine mine sites across Canada. Five of these data sets used multiple reference areas and these data were then used for calculation of reference Hedges' d values for benthic invertebrate density and taxon richness. The mean reference Hedges' d value was approximately 0.8 for density and 1.2 for richness (Figure 3-1). These values are comparable to the toxicological benchmark of 0.9-1.0 (see below) and it can be considered that Hedges' d values below these benchmarks are biologically unimportant.

- **Comparison of Hedges' d values with the Critical Effect Sizes that are included within the EEM guidance document** (Environment Canada 2011), consistent with the hypothesis being tested within the EEM program. For the benthic invertebrate community-structure data, the Critical Effect Size (CES) is ±2 Standard Deviations, which corresponds to a Hedges' d of approximately 1.8-2.0. For the fish-health data, the CES is 25% for all parameters but fish condition, which has a stated CES of 10%. Assuming a coefficient of variation of 10% for fish condition, and 25% for the other endpoints (as stated in Environment Canada 2011), this corresponds to a Hedges' d of approximately 0.9-1.0. It can be considered that Hedges' d values below these benchmarks are not biologically meaningful in the context of the EEM program.

- **Comparison with toxicological data**, where it can be considered that the Lowest Observable Effect Level is a change of approximately 25%. Assuming a CV of 25% (Environment Canada 2011), this corresponds to a Hedges' d value of approximately 0.9-1.0, the same value as two of the three methods noted above. Hedges' d values below this benchmark can be considered biologically unimportant.

2.1.2 Environment Canada's Approach to Calibration

Unfortunately, as mentioned, no calibration was undertaken by Environment Canada for the national assessment; no reference Hedges' d values were calculated and no benchmark cutoff values were considered. Instead, the authors have relied solely on the use of 'statistical significance' to interpret the data. This approach has resulted in 'effects' being declared for differences between reference and exposure data that are negligible in comparison with natural variability, with the benchmarks discussed above, and with the Critical Effect Sizes (CES) established within the EEM program (Table 2-1).

For instance, it is stated that there is a significant reduction in liver size in fish collected from exposure areas (Environment Canada 2012). This is technically true, but the minimum Hedges' d value (i.e. the difference between the CI and Hedges' d = 0) for 'reference' vs 'exposure' fish is 0.01 in Cycle 1 and 0.04 in Cycle 2. These Hedges' d values are considerably less than the calibration benchmarks of 0.2 to 1.0 discussed above. Assuming a CV of 25%, these minimum Hedges' d values translate into a minimum difference between reference and exposure liver size (based on Figure 4 in Environment Canada 2012) of approximately 0.2% in Cycle 1 and 1.0% in Cycle 2 (Table 2-1). Stantec considers that this difference is negligible compared with both natural variability in liver size, and with the suggested biologically meaningful CES of 25% that is discussed within the EEM guidance document (Environment Canada 2012).

Table 2-1: Calculated Minimum Effect Size for the Nine EEM Endpoints

Organism	Endpoint	EEM Critical Effect Size	Minimum Effect Size (%)[1]	
			EEM Cycle 1	EEM Cycle 2
Fish	Condition	±10%	-1.2	0.0
	Liver Weight	±25%	-0.2	-1.0
	Gonad Weight	±25%	0.0	0.0
	Weight-at-Age	±25%	0.0	-0.2
	Age	±25%	0.0	+0.5
Invertebrates	Density	±2STD (~±50%)	-0.8	+2.8
	Richness	±2STD (~±50%)	-7.5	-6.5
	Evenness	±2STD (~±50%)	+3.0	0.0
	Bray Curtis Index	±2STD (~±50%)	+32.8	+39.5

[1] Calculated as the smallest difference between the CI and Hedges' d =0. Assumes a Coefficient of Variation of 25% for all but fish condition, which is assumed to have a Coefficient of Variation of 10% (Environment Canada 2011).
Source: Figure 4 and Figure 14 in Environment Canada (2012)

FINDING: Sole reliance on 'statistical significance' of the meta-data is an inappropriate approach (Nakagawa and Cuthill 2007) that has ultimately resulted in a finding of 'effect' at differences that are typically negligible in comparison with natural variability, calibration benchmarks and EEM Critical Effect Sizes.

2.2 CONFIRMATION OF "EFFECT"

One of the fundamental guiding principles within the EEM is that if an 'effect' is found, it must be confirmed in a second round of monitoring (Environment Canada 2011). To be confirmed, the same effect must occur and it must occur in the same direction (e.g., both significantly inhibitory or both stimulatory). If the results are not consistent, then a follow-up monitoring cycle is required. Although not explained, the guiding principle of consistency of effect has not been considered within the national assessment. The result is an attempt to infer significant 'effect' for data that show a positive Hedges' d in one round and a negative Hedges' d in the second round, or a significant difference in one round and none in the second, or vice versa. Examination of the meta-data indicates that for only three endpoints was a statistically significant 'effect' seen in the first monitoring cycle and subsequently confirmed in the second cycle (Table 2-2).

Table 2-2: Results of the Hedges' d Meta-Analysis from the Second National Assessment

Organism	Endpoint	Cycle 1 Effect	Cycle 2 Effect	Confirmed as per EEM protocol?
Fish	Condition	Negative	NS	No
	Liver Weight	Negative	Negative	Yes
	Gonad Weight	NS	NS	Yes - NS
	Weight-at-Age	NS	Negative	No
	Age	NS	Positive	No
Invertebrates	Density	Negative	Positive	No
	Richness	Negative	Negative	Yes
	Evenness	Positive	Negative	No
	Bray Curtis Index	Positive	Positive	Yes

Source: Figure 4 and Figure 14 in Environment Canada (2012)

FINDING: Within the structure of the EEM, it is inappropriate to interpret data that have not been confirmed, except to state that a further round of confirmatory investigation is required. The lack of consideration for confirmation of effect has resulted in the finding of 'effect' for endpoints for which results are inconsistent between the Cycle 1 and Cycle 2 monitoring data.

2.3 INFERRED MEANING FROM "NON-SIGNIFICANT" DATA

The second assessment concluded that effluent exposure was '…*more often associated with reductions rather than increases in the indicators (endpoints)*…', and that, '… *fish collected in areas exposed to effluent… showed significantly reduced condition, relative liver size and growth rate. Other effects included some reductions in gonad… size and a significantly increased age structure. In other words, fish collected in areas exposed to effluent were, on average, older, thinner and slower growing, with smaller livers and with more of a tendency to reduced gonad size*.' The second assessment also concluded that there was '… *significantly reduced taxon richness*.', and that there were '… *different groupings of benthic invertebrates in exposure areas compared to reference areas*.' These conclusions, however, infer meaning from statistically non-significant data without any persuasive explanation or justification (Table 2-2).

Even though sole reliance on 'statistical significance' is inappropriate in the context of a meta-analysis, at a minimum endpoints should at least be consistently and significantly different than Hedges' d = 0 if they are to be discussed in terms of effluent 'effects'. This is not what occurred in the analysis of the data for fish condition, gonad weight, age, and weight-at-age (Table 2-2).

FINDING: Within the national assessment, the basic concept of statistical significance has apparently been ignored, rendering many of the conclusions both statistically invalid and methodologically inappropriate.

2.4 ASSESSMENT OF THE STATISTICAL STRUCTURE OF THE EEM PROGRAM

Environment Canada's guidance document notes that the statistical structure within the EEM carefully considers the risk to both industry and the environment (Environment Canada 2011). It is considered within the guidance document that the risk should be equal. The statistical method outlined in the guidance document therefore sets alpha and beta equal. The prescription designates a sample size of 20 for fish and 5 for benthic invertebrates. Assuming a known coefficient of variation, this sample size theoretically results in a finding of 'significance' at the designated Critical Effect Size. For fish, this is stated as a 10% difference for fish condition, and a 25% difference for all other fish endpoints, between fish collected in a 'reference' area versus those collected in an 'exposure' area. For benthic invertebrates, this is stated as a difference of ±2 standard deviations.

Within the assessment document, it would be useful to impact-assessment professionals if the success of the above-stated EEM statistical structure had been evaluated and declared. Unfortunately, while the document provides summary figures, they are not suitable for assessing the success of the EEM program.

Further, key questions cannot be answered, such as:

- Which mines (with identifiers removed) specifically showed significant effects, for which endpoints?
- For which mines were observed effects in Cycle 1 confirmed in Cycle 2?
- How many times were there differences between reference and exposure sites >CES that were determined not to be statistically significant?
- Which 'significant effects' for which specific mines (with identifiers removed) were >CES?
- Which 'significant effects' for which specific mines (with identifiers removed) were <CES?

These data are important for understanding the prevalence and distribution of biologically significant effects, the nature of the effects (e.g., whether the BCI is driving the regulatory process), and whether the statistical structure of the EEM is performing as designed.

FINDING: Without answers to the questions above, it is difficult to fully address the objectives of the national assessment as stated within the EEM guidance document, particularly given recent concerns regarding the statistical structure within the EEM program (Huebert et al. 2010; Huebert et al. 2011; Huebert 2012a, b).

2.5 AVAILABILITY OF DATA

Despite several requests, Environment Canada has, to date, declined to publish the national dataset upon which its assessment is built. While the absence of the data relied upon in deriving a conclusion is inconsistent with basic science reporting, it is also inconsistent with federal treatment of data from other national pollutant-release and associated monitoring programs (e.g., National Pollutant Release Inventory [NPRI]) and with other high-profile large-scale multimedia environmental-monitoring programs (e.g., Alberta RAMP). Unwillingness to publish the dataset can create the potential for challenge to the legitimacy of the entire program, especially Environment Canada's assessment findings. At the very least, a summary table containing the effect sizes for each endpoint for each mine would be useful.

3.0 Reinterpretation of the National Metadata

The following section is a re-interpretation of the meta-data contained within the Second National Assessment, using first principles of EEM statistical and meta-data analysis. This includes the principles of 'significance' and 'effects confirmation' at a minimum, and the application of four calibration techniques for inference of biological meaning. Each of the nine endpoints is evaluated separately, although they are summarized graphically in one figure (Figure 3-1).

3.1 FISH CONDITION

The Hedge 'd' values for fish condition were inconsistent between the two national cycles and were both below the cutoff benchmark of 0.2 (Figure 3-1; Table 3-1). Environment Canada's statement that "…*during the second national assessment period, exposure area fish again showed significantly reduced condition*…" is therefore not supported by the meta-data.

FINDING: The meta-data indicate that, on average, there has been no measurable effect of mine effluent on fish condition in exposed fish collected from downstream 'exposure' monitoring locations.

3.2 FISH LIVER SIZE

Reliance on 'statistical significance' in the meta-analysis has resulted in unsupportable interpretations. For instance, there is a stated 'significant effect' on liver size because the CI is 0.04 Hedges' d units from the 0 line (Figure 4; Environment Canada 2012). This is a reduction in liver size in exposure fish of approximately 1%, assuming a CV of approximately 25% (as in Environment Canada 2011). If the standard deviation assumption is increased to 40% (an upper level of variability for EEM data), this results in a difference of approximately 1.6%. To infer an 'effect' at this level of difference would typically be considered unsupportable, having regard for the natural variability of liver size and considering the carefully developed statistical structure outlined within the EEM guidance document (Environment Canada 2011), where the sample size, alpha, and beta are set to result in statistical significance at the stated CES (25% effect). It appears that the authors have abandoned the statistical structure and experimental design established within the EEM guidance document. The Hedges' d values are also below the cutoff benchmark of 0.2 (Figure 3-1: Table 3-1).

FINDING: The meta-analysis indicates that, on average, there is no measurable effect of mine effluent on relative liver size in exposed fish collected from downstream 'exposure' monitoring locations.

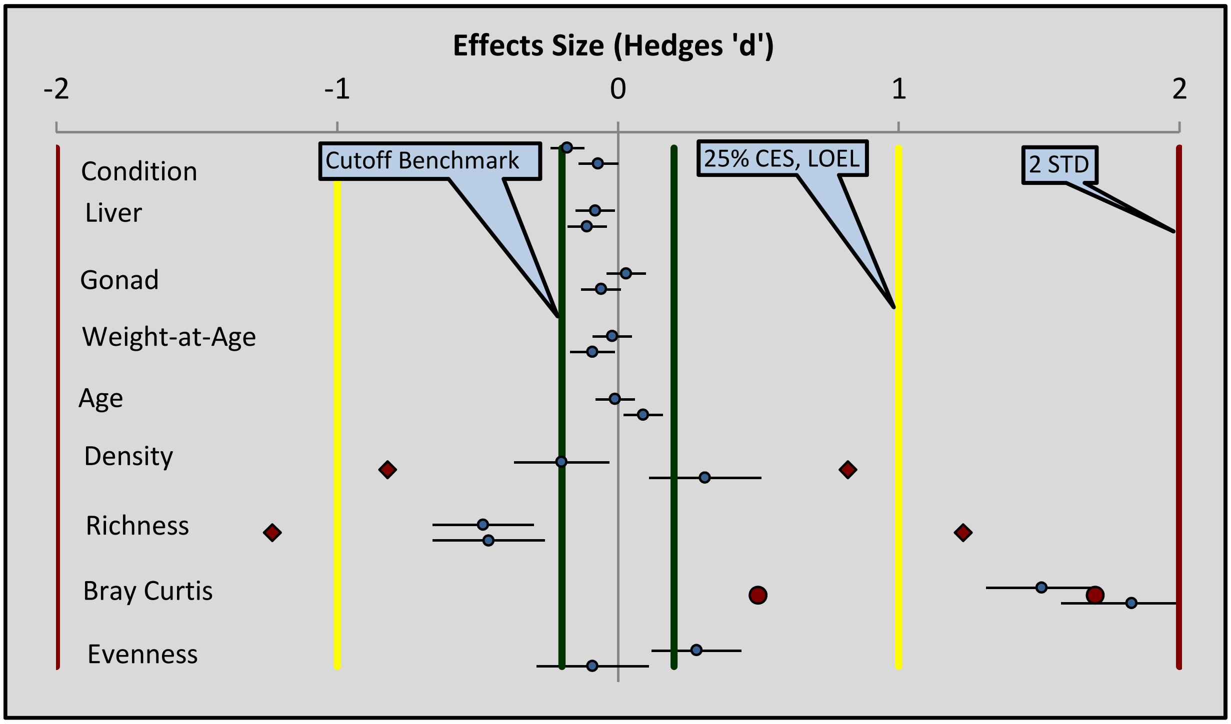


Figure 3-1: Meta-analysis of Cycle 1 and Cycle 2 Environmental Effects Monitoring Data

The green line is the cutoff benchmark, the yellow line is the toxicological Lowest Observable Effect Level (LOEL) and the fish Critical Effect Size (CES) benchmark, and the red line is the invertebrate Critical Effect Size benchmark of two standard deviations (STD). The red diamonds are the reference benchmarks and the red circles are the independently calculated Bray-Curtis Index values.

3.3 FISH GONAD SIZE

The average Hedges' d for gonad size in the first national Environment Canada assessment was approximately +0.03 (<1% increase in gonad size), but was not considered to be statistically significant. The average value in the second assessment was approximately -0.06 (<1.5% decrease in gonad size) and was also not statistically significant (Environment Canada 2012). Without persuasive explanation, the authors conclude that the meta-analysis indicated a, "…*tendency towards reduced gonad size*." This statement is unsupported by the data contained within the report. Firstly: the results are not statistically significant, and secondly: the results are inconsistent from one assessment to the next, and thirdly: the Hedges' d values are less than 0.2 (Figure 3-1: Table 3-1).

FINDING: The meta-analysis indicates that, on average, there is no measurable effect of mine effluent on relative gonad size in exposed fish collected from downstream 'exposure' monitoring locations.

3.4 FISH WEIGHT-AT-AGE

The average Hedges' d value for weight-at-age in the first national Environment Canada assessment was approximately -0.02 (<1% decrease in weight-at-age), but was not statistically significant. The average value in the second national assessment was approximately -0.09 (<2.5% decrease in gonad size) and was statistically significant by 0.02 Hedges' d units, which is an approximately 0.5% difference between the 'reference' and 'exposure' data (Environment Canada 2012). Without adequate explanation,, the authors conclude that the meta-analysis indicates fish were "…*thinner and slower growing*." This statement is unsupported by the data contained within the report. Firstly: the results were not significant in the first cycle, secondly: the results were inconsistent from one assessment to the next, and thirdly: the Hedges' d values were less than 0.2 (Figure 3-1: Table 3-1), with the stated difference being less than 2.5% in all cases. This would typically be considered a negligible difference, having regard to natural variability and the stated CES of 25% within the EEM guidance document.

FINDING: The meta-analysis indicates that, on average, there is no measurable effect of mine effluent on weight-at-age in exposed fish collected from downstream 'exposure' monitoring locations.

3.5 FISH AGE

The second assessment concludes that there is a '*significantly increased age structure'* in '*at least one of the national assessment periods*' (Environment Canada 2012). While this statement is technically correct, it misinterprets and overstates the biological meaning of the difference in the age structure. Firstly: the results were not statistically significant in the first cycle, secondly: the results were inconsistent from one assessment to the next, and therefore at a minimum require confirmation, and thirdly: the Hedges' d values were less than 0.2 (Figure 3-1: Table 3-1), with the stated 'significant effect' dependent on a 0.5% difference in age (Hedges' d = 0.02, 25% CV) between the 'reference' and 'exposure' fish (Figure 3-1: Table 3-1).

FINDING: Considering that the CES for age is a 25% change, and considering that the coefficient of variation is assumed to be approximately 25% (Environment Canada 2011), a difference of 0.5% would typically be considered biologically negligible and not consistently distinguishable from site-specific differences. The meta-analysis indicates that, on average, there is no measurable effect of mine effluent on fish age in exposed fish collected from downstream 'exposure' monitoring locations.

Table 3-1: Evaluation of the Assessment Data Using a Variety of Benchmarks

Benchmark	Fish					Benthic Invertebrates			
	Condition	Liver	Gonad	Weight	Age	Density	Richness	BCI	Evenness
Are the results of the first and second Assessment consistent with each other?	N	Y	Y-NS	N	N	N	Y	Y	N
Is the Hedges' d value for reference versus exposure sites greater than the temporal and spatial reference replicates?	U	U	U	U	U	N	N	U	U
Is the Hedges' d value > 0.2, which is the lower cutoff benchmark for environmental significance (Cohen 1988, in Nakagawa & Cuthill 2007)?	N	N	N	N	N	Y/N	Y	Y	Y /N
Is the Hedges' d value > 1.0, which is the approximate value at the fish Critical Effect Size (25% diff assuming a 25%CV), and consistent with the Lowest Observable Effect Level in toxicology?	N	N	N	N	N	N	N	Y*	N
Is the Hedges' d value > 2.0, which is the approximate value at the benthic invertebrate Critical Effect Size of 2 SD?	N	N	N	N	N	N	N	N	N
What is the magnitude of the mean difference between the reference and exposure means for those endpoints with a consistent difference (assuming a CV of 25%)?		-2 to -3%					-12%	+40 to+45%*	

*Calculation of the BCI within the EEM program is incorrect. When the correct calculation is used, significant differences disappear and the Hedges' d becomes consistent with other invertebrate endpoints and less than the cutoff benchmark of 1.0.

N = No, Y = Yes, U = Unknown, NS = Not Significant

3.6 BENTHIC INVERTEBRATE DENSITY

The Hedges' d values for density were inconsistent between cycles and therefore at a minimum require confirmation. Furthermore, the values were both below the stated invertebrate CES of 2 standard deviations and the toxicological benchmark of a 25% effect (Hedges' d = 0.9-1.0), and both were below the mean reference Hedges' d value of 0.8 (Figure 3-1: Table 3-1).

FINDING: The data indicate that, on average, there has been no measurable effect of metal mine effluent on benthic invertebrate density.

3.7 BENTHIC INVERTEBRATE RICHNESS

The average Hedges' d value for the two national assessments was approximately 0.5 (Environment Canada 2012). Assuming a CV of 25%, this is a difference between 'reference' and 'exposure' sites of approximately 10-15% in species richness. However, the Hedges' d values for taxon richness were both below the toxicological benchmark of 0.9-1.0 and were both below the reference Hedges' d value of approximately 1.2 (Figure 3-1: Table 3-1).

FINDING: The results indicate that, on average, there has been no biologically meaningful effect of metal mine effluent on benthic invertebrate taxon richness. Based on the reference benchmark, the observed effect is just as likely to be from random error and/or site-specific differences as from exposure to mine effluent.

3.8 BRAY CURTIS INDEX (BCI)

The Hedges' d values for the BCI are inconsistent with the Hedges' d values for density and richness. This inconsistency is of importance because the BCI is calculated from density and richness data and it would be expected that meta-analysis would result in similar effect sizes for all three endpoints. However, the inconsistency of the BCI metadata is perhaps not surprising, since within the prescribed EEM methods calculation of the BCI is contentious, with a strong bias towards increasing the difference between 'reference' and 'exposure' values (Huebert et al. 2010; Huebert et al. 2011; Huebert et al. 2012b).

To test this hypothesis, Hedges' d values were calculated for five datasets (obtained independently by MAC) for which there were multiple reference sites using both the EEM procedure and a standard pairwise comparison (e.g., Anderson et al. 2011, Huebert et al. 2011). Using the biased EEM procedure, the Hedges' d for the five datasets was calculated as 1.7, which was similar to the values calculated for the national assessment (Figure 3-1: Table 3-1). Using an unbiased pairwise comparison, the Hedges' d was calculated as 0.5, considerably lower and consistent with the magnitude of effect size calculated for richness and density (Figure 3-1: Table 3-1). This suggests that the Bray-Curtis values calculated for the second assessment are incorrect and should be recalculated. Unfortunately, this is not possible for the larger data set at the present time for two reasons; the complete national EEM data are not publicly available (in spite of requests for publication of these data), and few EEM studies have used multiple reference sites.

3.8.1 National Implications of BCI Metric

Based on the marked inconsistency between the Hedges' d value for the BCI and the other benthic invertebrate endpoints (and particularly in comparison with invertebrate density and richness), it would be valuable to know for how many of the mine sites used in the national assessment the BCI was the only endpoint showing an 'effect'. Unfortunately, there is no summary table of mine site versus endpoint result within the assessment report. There are many summary figures that illustrate how many mines show particular 'effects', but they do not address important questions such as:

- How many mines specifically showed only one or two effects, and how many times was at least one of these effects the BCI?
- How many mines specifically showed no effects for any of the benthic invertebrate or fish endpoints?

FINDING: Until these questions are answered and the BCI is recalculated, concern remains regarding the current interpretation of the benthic invertebrate data. Re-interpretation of the national meta-data indicates that, on average, there has likely been no biologically meaningful effect of metal mine effluent on invertebrate community structure. Based on the recalculation, the observed effect is just as likely to be from random error and/or site-specific differences as from exposure to mine effluent.

3.9 BENTHIC INVERTEBRATE EVENNESS

The Hedges' d values were inconsistent between studies and therefore at a minimum require confirmation, one of the values was statistically insignificant and below the cutoff benchmark of 0.2, and both were below the toxicological benchmark of 0.9-1.0 (Figure 3-1: Table 3-1).

FINDING: The national meta-data indicate that, on average, there has been no measurable effect of metal mine effluent on invertebrate community evenness.

4.0 Conclusions

4.1 METHODS

The following methodological deficiencies were noted within the second National Assessment:

- Hedges' d values were not calibrated and were interpreted solely on the basis of 'statistical significance.' This resulted in determination of 'effect' for differences that would typically be considered negligible in comparison with natural variability, cutoff benchmarks, EEM Critical Effect Sizes, and are therefore of not likely to be of any consistent biological or ecological significance.
- The EEM guideline principle of 'effects confirmation' was not adhered to with the result that inconsistent results were routinely reported as 'effects.'
- Statistically non-significant data were routinely reported as indications of 'effects' or trends.
- The performance adequacy of the EEM statistical structure was not evaluated.

4.2 REINTERPRETATION OF THE DATA

Examination of the Cycle 1 and Cycle 2 fish data indicates that all five fish endpoints are below the minimal cutoff of 0.2, that four of the endpoints show inconsistent results, and that for four of the endpoints the differences are not significant in one or both cycles (Figure 4; p. 18 in Environment Canada 2012). Further, none of the five fish endpoints showed consistent effects that were above what is typically considered to be appropriate (i.e. , biologically meaningful) effects benchmarks.

Stantec considers that, based on the foregoing, The defensible conclusion from examination of Figure 4 (Environment Canada 2012) is that, on average, there is no measureable biological effect of mine-effluent discharge on fish health in downstream receiving environments.

Examination of the Cycle 1 and Cycle 2 benthic invertebrate data indicates that all four benthic invertebrate endpoints are below the Hedges' d effects benchmark of 2.0, that all four of the endpoints are below the 'effects' benchmark and Lowest Observable Effect Level of 1.0 (when the BCI is recalculated), that two of the endpoints are below the reference effect benchmark, and that two of the endpoints show inconsistent effects between Cycle 1 and 2 (Figure 14; p. 33 in Environment Canada 2012). Further, none of the four invertebrate endpoints showed consistent effects that were above the appropriate biologically meaningful effects benchmarks.

The defensible conclusion from examination of Figure 14 (Environment Canada 2012) is that, on average, there is no measurable and/or meaningful biological effect of mine-effluent discharge on benthic invertebrate community structure in downstream receiving environments.

Re-analysis of the meta-data indicates that the findings of the second national EEM data analysis are not supported by the data. Methods prescribed by Environment Canada in its own EEM Guidance Document have been ignored without explanation. Statistical-analysis first principles have been ignored without explanation. Development of appropriate benchmarks has not been undertaken. An assessment of the performance of the statistical structure within the EEM program has not been considered.

5.0 References

Anderson, M.J., T.O. Crist, J.M. Chase, M. Vellend, B.D. Inouye, A.L. Freestone, N.J. Sanders, H.V. Cornell, L.S. Comita, K.F. Davies, S.P. Harrison, N.J.B. Kraft, J.C. Stegen, N.G. Swenson. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology Letters 14: 19-28.

Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. 2nd Ed. Lawrence Elbaum Associates. NJ. 567pp.

Environment Canada. 2012. Second National Assessment of Environmental Effects Monitoring Data from Metal Mines Subjected to the Metal Mining Effluent Regulations. National Environmental Effects Monitoring Office, Forestry, Agriculture and Aquaculture Division, Environment Canada. Cat. No.: En14-64/2012E-PDF. 68pp.

Environment Canada. 2011. 2011 Metal Mining Environmental Effects Monitoring (EEM) Technical Guidance Document. Accessed at; http://www.ec.gc.ca/Publications/default.asp?lang=En&xml=D175537B-24E3-46E8-9BB4-C3B0D0DA806D.

Environment Canada. 2002. Metal Mining Guidance Document for Aquatic Environmental Effects Monitoring. June 2002.

Huebert, D.B., M. McKernan and C. Paquette. 2010. EEM; An outsider's perspective. Proceedings of the 37th Annual Aquatic Toxicity Workshop: October 3 to 6, 2010, Toronto Ontario, Can. Tech. Rep. Fish. Aquat. Sci. 2949: xxv + 145 p.

Huebert, D.B., M. McKernan and C. Paquette. 2011. Canada's environmental effects monitoring program; areas for improvement. Integr. Environ. Assess. Manag. 7(1): 143-144.

Huebert, D.B. 2012a. Rebuttal to the response of Bosker et al. (2012) to "Canada's Environmental Effects Monitoring Program: areas for improvement." Integr. Environ. Assess. Manag. 8(2): 382-383.

Huebert, D.B. 2012b. Rebuttal to Lowell et al. (2012) regarding Canada's environmental effects monitoring (EEM) program. Integr. Environ. Assess. Manag. 8: 568-569.

Lowell, R.B., A. Willsie, M.E. Bowerman, C. Tessier, S.L. Walker, D. Gautron and K. Hedley. 2006. National assessment of magnitudes and patterns of effects on fish and invertebrates exposed to metal mining effluent. Proceedings of the 33rd Annual Aquatic Toxicity Workshop: October 1-4, 2006, Jasper, Alberta.

Nakagawa, S. and I.C. Cuthill. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol. Rev. 82: 591-605.

6.0 Closure

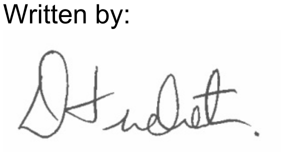
This report was prepared for the Mining Association of Canada. The report may not be relied upon by any other person or entity without the express written consent of Stantec Consulting Ltd. and the Mining Association of Canada.

Any use which a third party makes of this report, or any reliance on decisions made based on it, is the responsibility of such third parties. Stantec Consulting Ltd. accepts no responsibility for damages, if any, suffered by any third party as a result of decisions made or actions based on this report.

The information and conclusions contained in this report are based upon work undertaken by trained professional and technical staff in accordance with accepted scientific practices current at the time the work was performed. The conclusions and recommendations presented represent the best judgment of Stantec Consulting Ltd. based on the data obtained from the work and on the site conditions encountered at the time the work was performed at the specific sampling, testing, and/or observation locations.

STANTEC CONSULTING LTD.

Written by:



Dave Huebert, Ph.D., P.Biol.
Environmental Scientist
Tel: (204) 928-7613
Fax: (204) 942-2548
dave.huebert@stantec.com

Senior Reviewer:

Mike McKernan, M.Sc., MES, P.Biol.
Principal and Practice Leader
Tel: (204) 924-5734
Fax: (204) 942-2548
mike.mckernan@stantec.com