Appendix 36

Meadowbank Addendum to Cycle 4 Study Design

Addendum to Meadowbank Cycle 4 Study Design

The following are Agnico Eagle's responses to comments and recommendations provided by the Technical Advisory Panel (TAP), consisting of representatives from Environment and Climate Change Canada (ECCC), Nunavut Water Board (NWB) and Crown Indigenous Relations and Northern Affairs Canada (CIRNAC), in the letter entitled '*Meadowbank Mine 4th EEM Study Design – action items identified*" submitted May 5, 2020.

Action Items

1. **p. 33 (ECCC): The proposed near-field exposure area identified in Figure 2-2 is within the 1% effluent plume based on the modelled summer maximum extent scenario. Will efforts be made, prior to or during benthic invertebrate sampling, to delineate the plume in the field and to verify whether the selected sampling location is within the 1% effluent plume?**

   Response
   If effluent is being discharged when the benthic invertebrate sampling is conducted, delineation of the plume will be undertaken using conductivity.

2. **p. 66 (ECCC): The study design indicates that two sub-samples of the benthic community will be composited for each sampling station. Note that the Metal Mining Technical Guidance for EEM (EC 2012, section 4.3.3) recommends that each composite sample consist of three or more benthic invertebrate field sub-samples. Moreover, the previous interpretative report for Meadowbank (p. 79) indicated that the results of the sampling precision analysis showed that four sub-samples would be required to estimate evenness within 20% of the true value at the station level in the exposure area. Although the exposure area has changed, please consider this information for the current benthic invertebrate community study design and evaluate whether it would be prudent to increase the number of sub-samples per station. Moreover, you could consider assessing within-station variability in Second Portage Lake, given that this is a new exposure area; see section 4.4.2 of the technical guidance (EC 2012) for details on determining sampling effort for field sub-sampling.**

   Response
   The issue of sub-sampling has been raised in every EEM Cycle for the Meadowbank Mine. The rationale for continuing to use two-grab samples is presented below.

   The Technical Guidance document (ECCC, 2012) recommends a minimum of three subsamples (grabs) be composited for benthic samples. The Agnico Core Receiving Environment Monitoring Program (CREMP), which began in 2006 and was approved by regulatory agencies, uses 2-grab composites. The CREMP sampling program was designed with before-after-control-impact (BACI) comparisons in mind and has collected data annually from multiple exposure areas and multiple reference areas to support those ideal designs.

   If three-grab samples are used in the Meadowbank EEM, the two-grab benthic samples collected for CREMP cannot be used in the analyses. The baseline data that are available from CREMP sampling for Second Portage Lake and the reference lakes data are very valuable for understanding natural variations in effect endpoints. The ANOVA designs supported by these data are also far more relevant (have much more 'logical power') for identifying true effects related to perturbation than the basic EEM reference-exposure design (Green, 1979; Underwood, 1991, 1992, 1993, 1994; Kilgour et al., 2007; Somers et al., 2018).

The ratio of within-station variance relative to the among-station variance is the factor that determines power to detect differences among stations (Kilgour and Rosaasen, 2008). A simulation conducted on the effect of the ratio the within-station variance relative to the among-station variance is presented in Appendix A. The design proposed in this Cycle 4 EEM for the Meadowbank Mine (3-Areas, 2-Grabs), has better power than a conventional EEM design (2-Areas, 3-Grabs) across the entire range of within-station variances explored. When the among-station noise is less than or equal to the among-area noise (i.e., when $\sigma_W:\sigma_A$ is ≤1:1) in 3-Area designs there is little difference in power between designs with 2-Grab and 3-Grab samples. When the among-station noise is greater than the among-area noise (i.e., when $\sigma_W:\sigma_A$ > 1:1) the ability to detect differences diminishes for all designs but the 3-Grab design has more power. When the $\sigma_W:\sigma_A$ is 3:1, a design with 3 grabs modestly outperforms a design with 2 grabs, with about 8% greater likelihood of detecting a difference.

In summary, the 3-areas, 2-grab program proposed by Agnico is more statistically powerful than a conventional 2-areas, 3-grab program and is approximately equivalent in terms of statistical power to the 3-grab survey recommended by Environment Canada. The program proposed by Agnico has the advantage (with the 2-grab program) that it can compare the benthic data collected in the present program to data from the baseline period, and from other reference lakes in the CREMP, and thus has greater logical power than a 3-grab design which would not permit such comparison.

3. **pp. 73-74 (ECCC): The report lists the various analytes that will be included in the water quality monitoring program; however, the list does not include chromium, cobalt, thallium, uranium, phosphorus, or manganese. The MDMER require that water quality monitoring includes the concentrations of substances set out in Schedule 5, paragraphs 4(1)(a) to (p). Please ensure that water quality monitoring is conducted as per the requirements of Schedule 5, section 7, including recording the concentration of substances set out in paragraphs 4(1)(a) to (p).**

Response
Agnico will make sure to include all parameters in the analyte list.

4. **Throughout (ECCC): Study designs shall contain a description of the exposure and reference areas where biological monitoring studies will be conducted, and an explanation of how the study will provide the information necessary to determine if the effluent has an effect (Schedule 5, section 10). The report identifies sampling areas in two reference lakes, as well as near-field and far-field areas within Second Portage Lake. Overall, the study design indicates that the near-field and the reference lakes will be compared to identify any effects of effluent on the benthic community, and that data from the far-field site will be used to evaluate possible confounding effects of the dike. However, there are some inconsistencies that warrant additional clarification; please address the following points:**

    a. **The CREMP benthic sampling locations within Second Portage Lake are outside of the modelled 1% effluent plumes (Figure 2-2, p. 33). Within the report, this area is sometimes identified as the far-field control area (e.g., p. 60), and sometimes as the far- field exposure area (e.g., p. 68). Similarly, the report sometimes distinguishes between near-field and far-field areas within Second Portage Lake, but sometimes refers more generally to the exposure area within Second Portage Lake. Please be explicit as to whether the exposure area to identify any effects of**

effluent on the benthic community is only the near-field area, or if it includes the far-field area as well.

Response
Agnico will be consistent with terminology going forward. The inconsistent language is partly a function of the CREMP terminology which considers a lake to be 'exposed' once it experiences a mine-related effect. In Second Portage Lake, the first mine-related effect was due to dike construction. The far-field (CREMP) sampling area is exposed to dilute concentrations of effluent and in that sense the sample area is a 'far-field'. The far-field sampling area can be anticipated to have been historically exposed to suspended sediments when the dyke was constructed. In that sense the far-field exposure area may also be a control (reference) for the effects of dyke construction. There is some uncertainty here, since the CREMP sampling area is more removed from the dyke than the near-field exposure area; nonetheless, use of the CREMP data (which will be available) has some potential to partition effects related to dyke construction, if any, from effects related to mine-effluent exposure, should those various influences be significant.

b. **Table 3-2 (p. 46) presents the coordinates for the approximate centre of the sampling area for the three lakes included in this EEM program. However, it does not distinguish between the near-field and far-field sites within Second Portage Lake. Please provide the required information for both the near-field and the far-field areas within Second Portage Lake.**

Response
The coordinates of the sampling areas in Second Portage Lake are provided in Table 1, below.

Table 1. Latitude and longitude of benthic invertebrate sampling areas in Second Portage Lake.

Area	Latitude	Longitude
Near-field	65°1.33'N	96°2.326'W
Far-field	65°1.197'N	96°1.959'W

c. **Please confirm that the SP referred to in ANOVA 1 (p. 68) and in assessment of Bray- Curtis distances (pp. 71-72) relates to the near-field exposure area and that the far-field area within Second Portage Lake will not be included in this analysis. This interpretation is consistent with Table 5-3, for ANOVA 1, but the text is unclear.**

Response
Yes, SP in Hypothesis 1 refers to Second Portage Near-field, as in Table 5-3 on page 69. For that first hypothesis, the data from the far-field will not be included, thus the absence of contrast coefficients for the far-field for the first null hypothesis.

d. **The report indicates that the far-field sampling area was influenced by dike construction but that it is much less likely to have been influenced by mine effluent since it is outside the 1% mixing zone. In order for ANOVA 2 and 3 (pp. 68-69) to be useful in parsing out the possible effects of the berm and effluent, one would need to understand to what extent the far-field area is / was exposed to effluent. Do you have any information as to whether the far-field area was within the 1% effluent plume in the past, when effluent discharges were larger?**

AGNICO EAGLE

Response

The far-field site was not within the 1% mixing zone in the past.

Discharge to Second Portage Lake began in 2014. As indicated in the Study Design report, the 1% plume was modelled at the minimum, mean, and maximum daily discharge rates for the period 2017-2019 (i.e. since the last study design was submitted) and at the diffuser design discharge rate of 12.0 L/sec.

Summer effluent discharges were not larger in the past. During the summer period, the maximum daily discharge in the past (2014-2016) was 9.1 L/sec (785 m^3 day). The maximum daily effluent discharge volume during the summer for the period 2017-2019, was 10.8 L/sec (935 m^3/day). The maximum extent of the 1% effluent plume during the summer in the past would have been less than the maximum extent during the period 2017-2019, which is shown in Figure 2-2 of the Study Design Report.

During the winter, the maximum daily discharge in the past (2014-2016) was 9.1 L/sec (785 m^3 day). This is greater than the maximum daily effluent discharge volume during the winter for the period 2017-2019, which was 7.9 L/sec (680 m^3/day). The model, however, predicted the extent of the 1% plume at the diffuser design capacity of 12.0 L/sec (1000 m^3/day) to be 119 m, which is approximately 100 m less than the distance of the maximum summer extent for the period 2017-2019. Therefore, the maximum summer extent shown in Figure 2-2 of the Study Design Report represents the maximum extent since discharge began and it can be concluded that the far-field site was not within the 1% mixing zone in the past.

e. **Please confirm whether ANOVA 3 relies on the assumption that the far-field area was influenced by dike construction but that it is much less likely to have been influenced by mine effluent. If so, can you provide additional explanation for why you are only comparing baseline period 1 (2006-2007) to the exposure period (2014-2020) and not also to baseline period 2 (2008-2013)? As well, please articulate what information would be gained if the hypothesis is rejected (e.g. evidence of some effect of the dike).**

Response

It is correct that if ANOVA 3a identified a difference between effluent exposed and baseline 1 then understanding if the dyke construction caused a change would be relevant. In that case, a contrast to test that effect specifically would be developed using the data from the far-field sampling area.

5. **Appendix C (ECCC): The effluent plume modelling study indicated that the plume would attach to the shoreline in all cases. Can you provide some additional explanation of this point, given that Figure 2-2 shows concentric circles with no influence of the shoreline?**

Response

The plume ‘attaching to’ (i.e. intersecting) the shoreline will influence the movement of the plume. Lower current speeds near shore would limit the advection of a plume compared to stronger currents further from shore. Therefore, the concentric circles represent a conservative prediction.

Other Items

6. **p. 43-44 (ECCC): The study design provides estimates of the concentration of effluent at 100 m and 250 m from the point at which effluent enters the exposure area from a final discharge point. These estimates are used to assert that a study respecting fish population is not required (Schedule 5, paragraph 9(1)(a)).**

   The study design also indicates that a study respecting fish population would not be required based on the results of the previous two EEM fish population studies. In regards to this second point, please note that a fish population study would be required unless the results of the previous two biological monitoring studies indicate no effect or an effect less than the critical effect size for all effect indicators (Schedule 5, subparagraphs 9(1)(a)(i) and (ii)). If one or more effect indicator (e.g. reproduction) was not assessed in the previous two EEM fish population studies, then this exemption could not apply.

   Response
   To date, the situation where the need for a fish survey was dependent upon the results of previous fish surveys has not arisen for Agnico, but this could be of relevance in the future. This interpretation appears to mean that if a fish survey study design cannot examine all of the effect indicators (for example if a non-lethal survey is the accepted design) then the fish survey must be conducted during every cycle, regardless of the results of previous cycles. This seems counter-intuitive from a science perspective.

7. **pp. 66, 74, and 75 (ECCC): There are a few instances where cited references are not displaying properly.**

   Response
   Based on a search of the document there are three locations where this occurs.

   In Section 5.2.2 the reference should be to Figure 5.1.

   In Section 6.2 the "table below" referred to is Table 6-1.

   In Section 6.3 the "table below" referred to is Table 6-2.

8. **p. 67 (ECCC): The report refers to calculating total abundance for the benthic invertebrate community. Note that the MDMER include total benthic invertebrate density as an effect endpoint. Please confirm whether abundance is equivalent to density, and consider using the terminology from the MDMER in the interpretative report.**

   Response
   The term 'abundance' in the report is equivalent to the use of 'density' in Environment Canada's (2012) guidance document. Agnico will use Environment Canada's preferred terminology in subsequent documents.

9. **p. 69 (ECCC): There is a possible typo of "exposure" instead of "reference". The sentence is unclear and it makes it difficult to understand the interpretation of the various outcomes of ANOVA 3.**

   Response
   The reviewer is correct. There is an error and "exposure" should be "reference".

10. **p. 69 (ECCC): The Metal Mining Technical Guidance for EEM (EC 2012) indicates that one could do an ANOVA to compare among reference, near-field and far-field areas. You could consider this as an alternative to the separate ANOVAs 1 and 2b that are currently proposed.**

    Response
    ANOVA 1 considers both reference areas and the exposure (nearfield) area in a linear contrast, testing specifically that the mean of the reference areas differs from the mean of the nearfield area, and in Agnico's opinion is the most efficient use of the data.

11. **p. 70 (ECCC): The study design presents an approach for the assessment of covariable effects. Please note that the Metal Mining Technical Guidance for EEM (EC 2012, section 8.5.1) suggests that ANCOVA can be used for control-impact or multiple control - impact designs to factor out covariates that may create noise that makes it difficult to make simple ANOVA comparisons of reference to exposure areas.**

    Response
    Discussion will provide clarity, but this appears to be two different ways of saying the same thing.

12. **pp.73, 76, and 77 (ECCC): In Section 3, the report authors present their rationale for why a fish population study is not required as part of the current EEM biological monitoring study.**

    However, there are a few instances later in the text (e.g. on pp. 73, 76, and 77) that mention work to be done on fish or in fish sampling areas. Please clarify this discrepancy and confirm whether the current EEM biological monitoring study will include any fish sampling.

    Response
    The references to work done on fish or fish sampling in Section 7 are not relevant and should not be present. As indicated in Section 3.1, a fish survey is not proposed because the 1% effluent concentration does not extend for 250 m from the discharge point.

References

C. Portt and Associates and Kilgour & Associates Ltd. 2014. Environmental Effects Monitoring: Agnico- Eagle Mines Ltd.- Meadowbank Division Cycle 2 Study Design. Prepared for Agnico-Eagle Mines Ltd., Regional Office - 93, Rue Arseneault, suite 202, Val-d'Or, Québec, J9P 0E9. 55 p. + 4 appendices.

Environment Canada (EC) 2012. Metal Mining Technical Guidance for Environmental Effects Monitoring.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley, New York.

Green RH. 1989. Power analysis and practical strategies for environmental monitoring. Environ Res 50:195-205.

Kilgour, B.W., M.G. Dubé, K. Hedley, C.B. Portt, K.R. Munkittrick. 2007. Aquatic environmental effects monitoring guidance for environmental assessment practitioners. Environmental Monitoring and Assessment, 130:423-436.

Somers, K.M., B.W. Kilgour, K.R. Munkittrick and T.J. Arciszewski. 2018. An adaptive environmental effects monitoring framework for assessing the influences of liquid effluents on benthos, water and sediments in aquatic receiving environments. Integrated Environmental Assessment and Management, 14:552-566.

Underwood, A.J. 1989. The analysis of stress in natural systems. Biological Journal of the Linnean Society, 37:51-78.

Underwood, A.J. 1991. Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Aust. J. Mar. Freshwat. Res. 42:569-587.

Underwood, A.J. 1993. The mechanics of spatially replicated sampling programmes to detect environmental impacts in a variable world. Aust. J. Ecol. 18:99-116.

Underwood, A.J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications, 4:3-15.

Appendix A

Examination of the ability to detect differences among areas by three sampling designs

Sample sizes for EEM benthic invertebrate designs are constructed around the premise that differences between reference and exposure areas that exceed about two times the among-station standard deviation (i.e., ± 2 σ_A) are important to detect. Anything that results in an increase in within-area variance will not change the reported power of the study but will change the magnitude of the true effect that can be detected. Further, a study that minimizes the within-station variance will generally have greater likelihood of detecting effects equal to 2 σ_A than a study that has a higher within-station variance. In the Guidance Document, Environment Canada recommends that the number of within-station sub-samples be the number that ensures that means of effect variables will be estimated with a precision (D) of 0.2, or to within ± 20% of the true value. A better precision, or smaller D would reduce within-station variance, while a poorer precision and larger D would increase within-station variance. Not meeting Environment Canada's precision guideline will increase the within-station variance and would have some impact on the ability of a sampling program to detect differences of ± 2 σ_A. However, it is not so much precision (relative to the station means), but the magnitude of the within-station variance relative to the among-station variance that affects the likelihood of a sampling program to detect differences of ± 2 σ_A (Kilgour and Rosaasen, 2008).

The statistical power of a nested design with sub-samples within stations, and stations within areas, is not handled by any of the statistical power analysis packages we are familiar with (e.g., SYSTAT, PASS2020, GPower 3.1, piface). Agnico therefore explored the statistical power of 2-grab and 3-grab survey designs through a simulation exercise in Excel. The exercise used a synthetic variable (say density) for which the following conditions applied:

- The reference and exposure areas were assigned the same among-station variance (σ_A^2).
- The reference areas had the same mean, but the exposure area had a mean that was truly two standard deviations (+2 σ_A) from the mean of the reference areas.
- The within-station standard deviation (σ_W) was varied in the calculations from 0.1 to 1.0 x σ_A to explore the influence of the ratio of within- to among-station noise (i.e., σ_W:σ_A).

The following design variations were compared in their ability to reject the null hypothesis that there was 'no difference' between reference and exposure area means:

- Design 1: Conventional EEM design with 2 Areas, 5 stations per area, and 3 grabs per station (i.e., 2-Area, 3-Grab);
- Design 2: Proposed Meadowbank design with 3 Areas (2 reference, 1 exposure), 5 stations per area, and 2 grabs per station (i.e., 3-Area, 2-Grab); and,
- Design 3: Modified Meadowbank design with 3 Areas (2 reference, 1 exposure), 5 stations per area, but with 3 grabs per station (i.e., 3-Area, 3-Grab).

In Excel, station means were determined from normal populations with the defined Area means and standard deviations. Sub-samples were subsequently drawn for each station given the sampled station mean and assigned among-station standard deviations. Data from sub-samples were then 'pooled', and analysis of variance was used to test for a difference between reference and exposure areas. For the 3-Area designs (i.e., Designs 2 and 3), linear orthogonal contrasts were used to test for the difference between reference and exposure areas. For the simpler Design 1, a simple one-way ANOVA (or t test) was carried out. From theses ANOVA's, F ratios with p values < 0.1 were used to 'reject' the null hypothesis. These analyses were 'run' 1000 times per design. The fraction of the 1000 runs that produced a significant difference was used as the estimate of power, or the likelihood of rejecting the null hypothesis of 'no difference'.

The range of within- and among-station variances in this simulation is similar to what was observed by Azimuth in the preparation of the study design for the first EEM study Meadowbank Mine. The data presented in Table 1 shows that the standard deviation of log of density within stations (i.e., σ_W) is about 0.27 while the standard deviation of station means (i.e., σ_A) is about 0.09, or a ratio of about 3:1 ($\sigma_{W:}\sigma_A$). For family richness σ_A is about 0.76, while σ_W is about 0.73, for a ratio of about 1:1 ($\sigma_{W:}\sigma_A$). The variance ratios explored in the analysis (see Figure 1) therefore covered those that have been described for density and richness in for lakes in the Meadowbank study area.

Table 1. Within-station standard deviations (SD) for abundance and richness, from the 1st Cycle EEM program for the Meadowbank Mine.

Variable	Replicate Station	Mean	Within Station SD
Log of Density	1	1.42	0.24
	2	1.49	0.46
	3	1.38	0.10
	4	1.27	0.28
	Among station SD	0.09	
	Average within station SD		0.27
Richness	1	4.20	0.28
	2	3.60	0.58
	3	2.75	0.83
	4	2.50	1.33
	Among station SD	0.73	
	Average within station SD		0.76

Table Note: Copied from C. Portt and Associates and Kilgour & Associates Ltd. (2014), study design for the 2nd EEM program for the Meadowbank Mine. SD = standard deviation.

The results of the simulation are presented in Figure 1. The design proposed in this current Phase 4 program for the Meadowbank Mine (i.e., Design 2 with 3-Areas, 2-Grabs), has better power than a conventional EEM design (2-Area, 3-Grab) across the entire range of within-station variances explored (Figure 1).

When the among-station noise is less than the among-area noise (i.e., when σ_W:σ_A is ≤1:1, such as for family richness in Table 1) the power of Designs 2 and 3 are about equal (Figure 1). When the among-station noise is greater than the among-area noise (i.e., when σ_W:σ_A > 1:1 such as for Density in Table 1) Design 3 the ability to detect differences diminishes for all Designs. When the σ_W:σ_A is 3:1, Design 3 (with 3 grabs) modestly outperforms Meadowbank's Design 2 (with 2 grabs), with the difference in detection likelihood being about 8%.

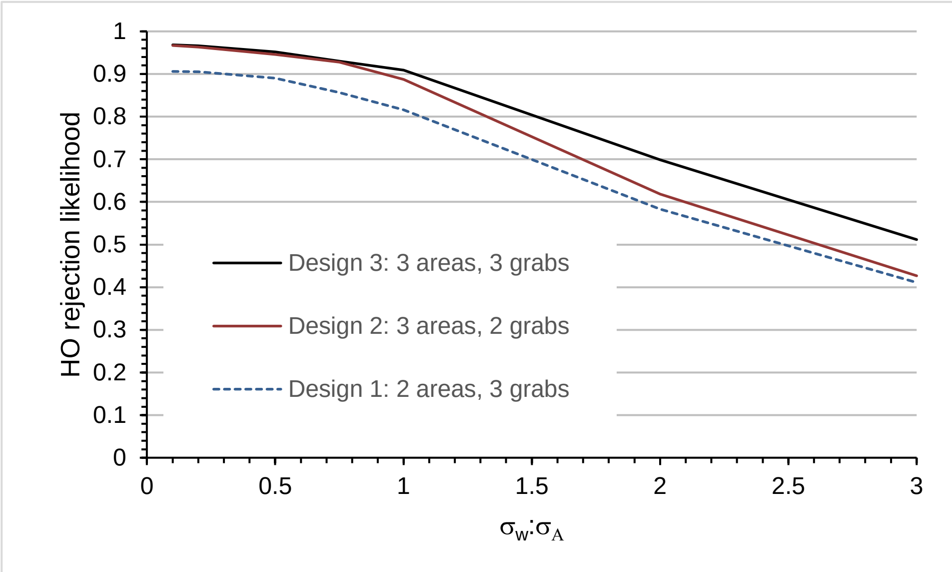

Figure 1. Power curve for CREMP and EEM design options

The Meadowbank Mine, through the CREMP, has been implementing an even more powerful design, collecting 2-grab samples every year from each of 5 stations from the Second Portage far-field exposure area and each of the reference areas, and has been doing so since 2006 for several of the sampling areas. So, not only is the proposed 2 Reference Area, 1 Exposure Area, 2-grab surveys proposed by Agnico Eagle for the Meadowbank Mine Phase 4 EEM approximately equivalent in terms of statistical power to the 3-grab survey recommended by Environment Canada, but the ability to incorporate Before-After-Control-Impact comparisons makes the proposed program with 2-grabs far superior to the changing to a 3-grab design.

References

Kilgour, B.W. and A. Rosaasen. 2008. The influence of subsampling on the ability to detect effects in surveys of benthic macroinvertebrates. K. Liber, D.M. Janz and L.E. Burridge (eds). Proceedings of the 35th Annual Aquatic Toxicity Workshop: October 5 to 8, 2008, Saskatoon, Saskatchewan.