FINAL PROJECT REPORT EVALUATION FORM

Referee ID: REF04

Project Number: SARF089	Completion Date: February 19, 2016
Project Title: A statistical analysis of sea-lice medicine use and benthic monitoring at Scottish marine salmon farms (2002 – 2014).	
1. In your view have the scientific objectives been achieved. If not, does this need to be addressed by SARF? We do not believe the SARF authors have adequately supported their conclusion that EMB is causing widespread effects to benthic crustaceans. There are multiple data and analytical inadequacies and inconsistencies that preclude the determination of a likely cause of the observed decline.	
2. Comment on the overall results of the project, including their significance for SARF. This project clearly demonstrates that additional data collection is necessary before a reliable and conclusive identification of the cause of observed declines in the benthic crustacean community can be made. The observation of effects at pre-determined reference sites is particularly concerning and should be taken into consideration when designing future sampling studies.	
3. Is there a need for further work? If so, explain. We strongly suggest that SARF authors pursue a broader, more methodical approach to determining the cause of the observed declines in benthic crustaceans. To this end, we propose the implementation of a causal analysis (see additional comments). Not only will this provide greater confidence and transparency in determining the plausibility of candidate causes of the decline, but the process should prove invaluable in identifying data gaps and guiding future data collections.	
Overall marking 1 - outstanding results 2 - results significantly above expectation 3 - satisfactory results 4 - results below expectation 5 - poor results	

Signature	Print Name	Date
February 19, 2016		
Please indicate whether you wish to receive payment (Yes/No) NO		

Additional Comments:
Please see attached document

EXTERNAL MEMORANDUM

SUBJECT: Response to SARF Responses to Commenters

In response to the SARF authors' "Reply to reviewer's comments", dated 17 Dec 2015, we reiterate the following original concerns that we believe have still been inadequately addressed by the revised report[1]:

1) Data quality issues
2) The problem of predicted impacts at reference stations
3) Inappropriate conclusions drawn from statistical assessments
4) The importance of Before/After treatment (or pre-EMB/post EMB) analysis

Based on authors' responses, these have still not been rectified and persist in the current version of the assessment (as do many issues identified by other reviewers). However, we believe that these are encompassed by a larger conceptual error in how the overarching problem is approached, which results in significant bias and cognitive errors. To rectify these, we strongly suggest the use of a Causal Analysis framework to sufficiently and accurately assess the cause of observed declines in benthic crustacean communities in Scotland's bays.

Causal Analysis (CA) is a systematic process of identifying and ranking the likelihood that a specific stressor or stressors are responsible for observed impacts to biological communities. Grounded in the widely-cited concepts of Sir Austin Bradford Hill[2], it is a weight-of-evidence

[1] SARF089: PAMP Refreshment Study- final report. A statistical analysis of sea-lice medicine use and benthic monitoring at Scottish marine salmon farms (2002 – 2014). Thomas A. Wilding and K.D. Black. Scottish Association for Marine Science, Scottish Marine Institute, Oban, Argyll, UK.

[2] Hill, A.B. 1965. The environment and disease: Association or causation. Proceedings of the Royal Society of Medicine, 58: 295-300.

approach that combines multiple sources of data and lines of evidence to determine causation[3][4]. This formal and inferential framework provides multiple benefits. First, a CA approach prevents cognitive errors that frequently arise in environmental injury assessments (e.g., focusing on a single hypothesis cause, overlooking data gaps or poor data quality). Next, the framework constitutes a formal and transparent process through which available data and lines of evidence may be assessed and ranked. Finally, this process improves confidence in results through implementation of a thorough and methodological framework. (For additional information on the CA approach, refer to Suter et al. 2010[5] and Cormier et al. 2010[6]). We believe that SARF's assessment of EMB will significantly benefit from such an approach.

Briefly, the process of identifying the stressor most likely to be causing biological impairment includes five components:

1) Definition of the biological impairment: identify the nature of the biological/ecological impact, area impacted, and potential reference sites;
2) Develop list of candidate causes: based on the identified biological/ecological changes, list all possible stressors, direct and indirect;
3) Evaluate all available field data: collate, evaluate, and analyze all relevant field data from the site and references areas, including environmental chemical, physical, biological, etc. data. Evaluate findings in terms of the strengths of different types of evidence (Table 1), while also identifying data gaps and insufficiencies;
4) Evaluate secondary data: incorporate data from the literature (i.e., toxicological data), and monitoring data from other areas. Evaluate findings in terms of the strengths of different types of evidence (Table 1), while also identifying data gaps and insufficiencies; and
5) Identify the most likely cause: compare strength of evidence for each candidate cause.

Conclusions within the SARF report and responses to comments indicate that a causal analysis approach would be an ideal tool for determining the cause of observed declines in benthic communities. In their response to comments, authors state "we believe that the EMB linkage is the most plausible explanation for the effects". However, determining that EMB is the most plausible cause of biological impairment requires an equally rigorous assessment of alternate candidate causes. As pointed out in Suter et al (2010), "it is not possible to confidently identify the cause of an environmental impairment" but "it is possible to apply abductive inference to determine which of a set of causal hypotheses is best supported by the evidence". When a systematic framework is employed, elimination of candidate causes through assessment of available data improves the plausibility of remaining candidate cause(s). As such, plausibility

[3] Norton, SB, Cormier, SM, and Suter II, GW. 2014. Ecological Causal Assessment. CRC Press. 513 pp.

[4] 3 .003-6702

[5] Suter, G.W., S.B. Norton, and S.M. Cormier. 2010. The science and philosophy of a method for assessing environmental causes. Human Ecological Risk Assessment, 16: 19-34.

[6] Cormier, S.M., G.W. Suter, and S.B. Norton. 2010. Causal characteristics for ecoepidemiology. Hum. Ecol. Risk Assess. 16: 53-73.

of a candidate cause cannot be determined by only assessing a single candidate cause, which means other candidate causes *must* be assessed before EMB use can be determined to be “the most plausible explanation for the effects”.

> The following aspects of the current version of the SARF EMB assessment are highly troubling, and hence necessitate a more comprehensive, scientifically rigorous and systematic analysis of *all* candidate stressors, in order to support the authors' conclusions:

1) The lack of specificity of the observed response (decreased health of benthic crustacean communities)
2) The failure to consider alternate stressors or multiple stressors acting in concert, and/or the less rigorous assessment of other candidate stressors.
3) The wide spatial scale over which effects were reported by SARF authors
4) Biological impairment noted at reference stations
5) The conclusion that adverse effects to benthic crustacean communities occur at sites where EMB cannot be reliability quantified
6) The fact that the BACI (before after control impact) analysis suggests a lack of EMB impact

These have previously been brought to authors' attention, and as of yet are still not resolved. As such, we will address how each could be approached within a casual analysis framework, thereby improving the assessment process.

Lack of crustacean response specificity

A highly specific response to a candidate stressor, such as the induction of a specific gene or the presence of certain tumors, provides strong evidence that biological impairment is due to that particular stressor (see "Symptoms" category in Table 1). However, as we previously commented, changes in benthic crustacean communities are largely *non-specific* responses, which render it a poor line of evidence. Benthic crustaceans are notably sensitive organisms that demonstrate adverse responses to metals, PAHs/oil, nutrient enrichment, pesticides, and other stressors. The SARF authors make two claims in response to this. First, they assert an association between EMB use and declines in crustacean communities based on the fact that "EMB is specifically designed to be toxic to crustaceans and is known to remain for extended periods in sediments". This is an example of general causation (i.e., exposure to EMB causes adverse effects in crustaceans). However, the purpose of this assessment is to determine *specific* causation—that EMB is the primary cause of observed declines in crustacean communities in Scotland's bays. Confusing the two types of causation has led to a considerable overestimation of the strength of this line of evidence.

Second, and more critically, authors claim that "alternative sources of contaminants and confounding were discussed at length with SEPA . . . none were identified." However, no evidence is provided as to the adequacy of assessment and availability of supporting data. This constitutes a profound flaw in the assessment of the causes of observed biological impairment. In fact, Suter et al (2010) states that the most critical error of many stressor identification exercises "is that they are *ad hoc* and opaque . . . that is, the assessors integrate the evidence in the privacy of their own minds and provide *post hoc* justifications for the results." If the SARF authors intend to conduct a scientifically rigorous analysis of the cause of benthic community impairment, it is imperative that they address potential stressors in a consistent, transparent, well-documented, and organized manner.

Failure to consider alternate stressors

In response to questions regarding potential alternate causes of benthic crustacean declines, the SARF authors stated "to our knowledge, there is no confounding in relation to e.g. point-sources of pollution, temperature or salinity, for example, there is no basis for thinking that Reference stations, that are associated with farms using EMB, will be commensurately impacted by any other source of impact". First, this statement is vague and thus troubling. The phrase "to our knowledge" may indicate that there is insufficient field data regarding environmental concentrations of other stressors known to impact benthic crustaceans, or that the authors simply have not fully pursued this line of thinking. In any regards, the plausibility of alternate stressors must be investigated as thoroughly as EMB usage was to support exclusion as candidate causes; further, this investigation must be completely documented within a consistent and transparent framework. Otherwise, this appears to comprise a profound source of bias within the assessment, as only one candidate cause is investigated and then determined to be "most plausible".

It may be that the SARF authors have conducted this assessment, and merely failed to adequately present the process and results. However, this needs to be scientifically and logically documented within the report rather than asserted. Further, if there are no data or insufficient data concerning presence or concentrations of other stressors (physical, chemical or biological, e.g. example organic enrichment), this comprises a significant data gap and source of uncertainty in the assessment. This, in turn, will reduce the strength of evidence of the association between EMB use and observed benthic community impacts, as the lack of data for other candidate causes does not reduce their plausibility.

Effects observed over a wide spatial scale

In both the first and revised drafts of the SARF report, authors assert that "repeated, large-scale EMB use may adversely affect crustacean populations at the scale of an entire sea-loch" and that it may be necessary to "consider the likely ecosystem consequences of large-scale reductions in crustacean richness and abundance (at the scale of sea-lochs)." However, it should be noted that large areas of impacts are also indicative of multiple and/or non-point stressors. In fact, a causal analysis guidance document authored by the State of California[7] states that assessments conducted over "large areas, such as a watershed or sub-watershed, may complicate the process since more than one stressor, or single stressors at various magnitudes, can be acting

[7] Schiff, K., Gillett, D. J., Rehn, A., & Paul, M. (2013). Causal Assessment Evaluation and Guidance for California. SCCWRP Technical Report 750.

in different portions of the case." This, coupled with the lack of transparency (or lack of thorough assessment) concerning analysis of alternate candidate stressors, increases the likelihood that biological impairment could logically result from other stressors or combinations of stressors. However, instead of addressing this possibility, the SARF authors presume that all observed effects, including those at observed at reference stations (discussed below), are due to EMB use.

Effects observed at reference stations

The SARF authors report significant declines in benthic crustaceans at reference stations in both the original and revised versions of their report. Many commenters were very concerned by this finding and the fact that authors presumed that this was simply a result of a wider distribution of EMB than previously suspected. As discussed above, a wide-scale observation of effects is indicative of both multiple and/or non-point source stressors; however, this possibility is not addressed by the SARF authors. Instead, they state that "we have acknowledged the deficiencies in observational research, but we maintain (and justify in the report) that EMB could, logically, be having detrimental effects at the Ref stations." Again, this conclusion is based on general not specific causation, which is inadequate for identifying the cause of the observed benthic decline. The fact that impacts to benthic crustaceans were observed at sites considered to be reference sites (and did not contain detectable levels of EMB) is very concerning and indicates significant errors in the original understanding of the sites in question. It is inappropriate to adopt the observation of effects at reference stations as evidence that EMB is causing widespread effects on crustaceans. It should be noted that the authors again acknowledge that deficiencies in the underlying data, which significantly reduce the certainty with which the cause of the biological impact can be identified.

Observation of biological impacts in sediments with non-detectable concentrations of EMB

In both versions of the SARF report, the authors note that EMB residues were mostly below detection limits, especially when collected at sites more than 25 m from cage edge. Yet, in absence of quantitative EMB residue data, the authors still maintain that EMB is the main cause of observed declines in crustacean communities. This constitutes a leap in logic that is not supported by empirical evidence, nor supported by the causal analysis approach. Determination of an exposure-response relationship between candidate stressors and biological impairment is a key component of several types of evidence utilized within the causal analysis framework. The strength of evidence for several lines of evidence (spatial/temporal co-occurrence, evidence of exposure or biological mechanism, causal pathway, and field stressor-response relationships; see Table 1) depends on a quantitative demonstration of exposure in order to logically link the stressor to the observed impairment. In their response to comments, the authors state that the "analysis indicates an association between EMB use (not EMB residue) and declines in crustacean metrics. . . given the toxicity of EMB to crustacea, it is fair to speculate that the association is caused by the toxicity of EMB to crustacea." First, EMB *use*, in absence of measured sediment residues, is simply insufficient to support spatial/temporal co-occurrence, evidence of exposure or biological mechanism, causal pathway, and field stressor-response relationships; these lines of evidence require quantitative data that demonstrate that biological receptor(s) are exposed to critical concentrations of a chemical stressor As such, these lines of evidence would be weak or negative in support of EMB as the candidate cause of biological impairment. Further, it calls into question the validity of the authors' speculation of an association. Also of concern is the authors' statement in their response to comments that they believe that "a no-effect position doesn't exist, [and] impacts will occur on a continuum." This statement is concerning, in that, given basic tenets of toxicology, there should be a point along this "continuum" at which EMB concentrations are sufficiently low so as to not elicit effect. Toxicity has long been shown to be a threshold response. The authors also speculate that this no-effect concentration would be below the analytical detection limit (hence the observations of biological impairment at sites showing no quantifiable EMB concentrations), but provide no evidence that this is occurring. As such, this comprises a key data gap and uncertainty that prevents the conclusion that EMB is the primary cause of benthic crustacean declines.

Evidence from BACI analysis

Temporal sequence (i.e., the stressor precedes the effect) is another key line of evidence used to judge the relative plausibility of candidate stressors within the causal analysis approach (Table 1). This line of evidence was addressed by the authors as part of their examination of sites both pre- and post-EMB usage. Unfortunately, the data set is too small to conduct a robust analysis. Yet, the lack of clear crustacean response to implementation of EMB use at these sites provides at minimum a weak line of evidence that EMB use may not be causing observed declines in benthic crustaceans. Obviously, alone, this finding cannot support or refute EMB as the primary candidate cause of crustacean decline. However, it does provide an instance in which casual analysis can be used to clarify the SARF analysis and support a specific causation.

Conclusions

In light of the SARF authors' responses to commenters, we urge the authors to consider a broader and more scientifically rigorous approach, along the lines of a causal analysis. It is inappropriate to make a conclusion of plausibility when only one potential cause of apparent declines is assessed in depth (EMB). The conclusion of "most plausible" insinuates that other causes have been investigated and determined to be a less plausible cause than EMB. Further, the fact that bay-wide declines (and declines at "reference" sites) are somehow concluded to occur as a result of concentrations of EMB that are below the analytical detection limit indicates that the plausibility of different potential causes has not been fully investigated and ranked. If the authors' hypothesis would be true, then the reference sites are of no investigative use and the study has no negative controls to test the hypothesis, which would also be of concern. We note that other reviewers had similar concerns, and that a broader causal analysis approach might better present a more complete and transparent assessment of the cause of observed benthic crustacean declines.

Table 1: Lines of evidence for determining the plausbility of candidate causes (from Suter et al. 2010)

Type of evidence	**Concept**
Spatial/temporal cooccurrence	The biological effect is observed where and when the cause is observed, and must not be observed where and when the cause is absent.
Evidence of exposure or biological mechanism	Measurements of the biota show that relevant exposure to the cause has occurred, or that other biological mechanisms linking the cause to the effect have occurred.
Causal pathway	Steps in the pathways linking sources to the cause can serve as supplementary or surrogate indicators that the cause and the biological effect are likely to have co-occurred.
Stressor-response relationships from the field	As exposure to the cause increases, intensity or frequency of the biological effect increases; as exposure to the cause decreases, intensity or frequency of the biological effect decreases.
Manipulation of exposure	Field experiments or management actions that increase or decrease exposure to a cause must increase or decrease the biological effect.
Laboratory tests of site media	Controlled exposure in laboratory tests to causes (usually toxic substances) present in site media should induce biological effects consistent with the effects observed in the field.
Temporal sequence	The cause must precede the biological effect.
Verified predictions	Knowledge of a cause’s mode of action permits prediction and subsequent confirmation of previously unobserved effects.
Symptoms	Biological measurements (often at lower levels of biological organization than the effect) can be characteristic of one or a few specific causes.