A screening framework for oil spill impact assessment and research

Jon Moore
Coastal Assessment, Liaison & Monitoring Ltd.
jon@ticara.co.uk

Introduction

Laboratory based toxicological studies have long shown us that petrogenic oil can have toxic effects on a wide range of aquatic organisms; sometimes down to oil concentrations measured in parts per billion. In contrast, field-based studies, in all disciplines, often struggle to detect and describe impacts. There are many reasons for this, which this paper summarises. Understandably, in the aftermath of a spill, these difficulties are not necessarily highlighted, and some studies may be instigated with inadequate or no screening. When it becomes obvious that studies are not producing useful results they are usually dropped and forgotten, but sometimes they maintain a momentum that is not justified and may even reach white literature.

The PREMIAM guidelines (Kirby et al. 2018) provide a considerable resource for the design and implementation of an effective post-spill impact assessment. They include a section on selecting and prioritizing studies when developing an assessment programme. This paper provides an expansion of the guidance in that section and a personal viewpoint.

Screening framework

It is suggested that the following questions could be used when designing or assessing a proposed study or when reviewing its results. Many are focused on factors that can limit impact detection (i.e. potentially failing to detect an impact: often called a false negative result) or can increase the risk of wrongly concluding an impact (a false positive result). The term biological receptor is used herein as general term to include anything from an individual tissue biomarker to a broad community attribute.

<u>To what extent was the biological receptor exposed to the oil?</u>

Demonstrating exposure is a fundamental requirement of a damage assessment but is sometimes surprisingly difficult to achieve. Oil contamination, in all its forms, is usually patchy and transitory, with high levels of exposure typically affecting only a small proportion of the environment. This is the reason why broad scale randomised sampling designs rarely detect impacts – because the relatively small number of impacted samples are masked by the much larger number of apparently unimpacted samples.

Many impact studies don’t start until the main clean-up response has been completed, by which time the visual evidence of oiling may be much reduced and dispersed oil concentrations may be falling towards background levels. If there has been inadequate recording of oil exposure any assessment of impact is fundamentally incomplete.

Evaluating exposure to dispersed oil will always be challenging, even with the largest spills. Thus, with a lack of empirical data, some studies take what may be considered a cautious approach by assuming at least some likelihood of exposure, even when forecasts from spill modelling suggest very low oil in water concentrations. That assumption can be taken too far if a large amount of effort is required to provide some reassurance that a receptor was not impacted. Modern oil spill modelling is sometimes utilised only as a spill response planning tool, but its outputs have value for assessing the potential range of exposure to dispersed oil. While treating their detailed predictions with some caution, and preferably validating with some limited water sampling and hydrocarbon analysis, comparing predicted concentrations with typical background concentrations can provide a useful starting point for screening ongoing studies.

Does the design take sufficient account of the scale of natural variability in the biological receptor?
Both spatially and temporally, natural variability frequently masks the impacts of oil contamination, unless the sampling design includes sufficient replication. Replication is the bedrock of any manual on sampling design, but empirical data on natural variability in the specific biological receptor being studied is often unavailable until many samples have been analysed. Budgetary and other logistical constraints often decide how many samples and how often to take them; but does the design include a realistic assessment of the expected statistical power? Most do not, but many would have benefited from it. A conclusion of no detectable impact would be more informative if it was qualified with at least some explanation of the detectability of the method.

A frequent situation that may result in a false positive finding, is when damage assessment is based on comparison of post-spill data with inadequate pre-spill data. A single snapshot of pre-spill data can rarely be considered an adequate baseline.

Does the design take sufficient account of confounding factors?
Pristine environments are difficult to find, reference sites are not control sites and oil spills are not the only events that cause notable changes in biological receptors. Thus, while the selection of reference conditions may strive for ones that are as close as possible to those that were oiled, but without the oil, it is unrealistic to treat them as if we can control those conditions. Reference conditions will be influenced by other factors, natural or anthropogenic. Examples include other sources of contamination, storm events, proximity to freshwater inputs, differences in aspect, sediment silt content or other habitat characteristics. These factors are more of an issue in coastal areas, but some are also relevant in deeper water. So, does the design critically examine potential factors and explain how their influence has been accounted for in the design and analysis?

Is the design sufficiently robust to take account of inherent error and bias?
Few methods are so simple and well developed that potential errors and / or bias can be ignored. Quality assurance and quality control (QA/QC) procedures are a necessary requirement of any spill impact studies to provide confidence in the results, even with standardised methods and indicators. The PREMIAM guidelines highlight the importance of QA/QC procedures, but every methodology has its own weaknesses that may not be obvious to reviewers who are not very familiar with its application. Researchers should critically examine their own protocols and data for potential errors and bias during design, implementation, analysis and reporting. In some situations, the scale of such errors can be as great or greater than the scale of natural variability, greatly increasing the risk of a false negative conclusion. On the other hand, bias can increase the risk of false positives or false negatives.

Does the study design identify an impact mechanism that fits with the available evidence?
The study should endeavour to describe the causal relationship (often termed the *pathway*) for how the oil caused injury. Assumption that exposure of a biological receptor to a hydrocarbon with known toxic properties is sufficient explanation is not always enough. Thus, if the pattern of effects is not consistent, either across the study or with existing knowledge from previous studies, then the inconsistency requires reasonable explanation. If the correlation between exposure and injury is strong, but the mechanism for how one caused the other is not clear, then an interesting topic of research is waiting.

Examples of situations where causal relationships have been lacking include studies that neglect to explain the pathway from an external oil exposure to the internal biological tissue that was studied, or community studies that have described apparent toxic effects of dispersed oil on benthic species that are not typically sensitive while known sensitive species are not affected.

Other relevant questions
A review of potential studies might also consider the ecological, economic or conservation importance of the biological receptor, its known sensitivity to oil spills and the practicalities and logistics that affect project feasibility. These will be most relevant for damage assessment programmes focused on describing the most important and easily distinguishable impacts within a limited budget.

Future research

Failure to detect an impact does not mean that impact did not occur. Exposure of marine organisms to elevated concentrations of oil are likely to cause impacts, even if those impacts are within the range of natural background variation or less than effects of other anthropogenic activities. Optimistically it may mean that there is little cause for concern because the resource appears to be functioning and is likely to recover from any elusive impacts, particularly if concentrations have returned to background levels. Realistically it means that much more effort will be needed to improve our knowledge of the effects that oil spills have on biological receptors. If this requires increased sample replication, to increase statistical power, it will also require increased consideration for the damage caused by sampling.

Putting more effort into impact studies on individual receptors will only handle part of the challenge for modern spill science. Increasingly, concern is expressed for the impacts of spills at the ecosystem level, including productivity, nutrient cycling and larval supply. NRC (2013) highlights the concerns that large spills could impact the resilience of ecosystems to recover and starts to develop a framework for assessing that impact. The development of methodologies and novel attributes is underway but designing robust studies will need to take account of the factors described above. Reliably distinguishing the effects of a spill on, for example, productivity, from natural variability and other factors will be extremely challenging. Alternatively, it may now be appropriate to rekindle the power of field experiments to control these factors. It has been very difficult in recent years to get approval for experimental releases of oil into natural marine habitats, but the knowledge that has been gained from experiments has been enormous (Baker *et al.* 1993). Without them it will be difficult to further our understanding of spill impacts.

References

Baker, J. M., Little, D. I., and Owens, E. H. (1993) A review of experimental shoreline oil spills. International Oil Spill Conference Proceedings: March 1993, Vol. 1993, No. 1, pp. 583-590.

Kirby, M.F., Brant, J., Moore, J., Lincoln, S., (eds) 2018. *PREMIAM - Pollution Response in Emergencies - Marine Impact Assessment and Monitoring: Post-incident monitoring guidelines*. Second Edition. Science Series Technical Report. Cefas, Lowestoft. [in press at time of writing].

NRC 2013. *An Ecosystem Services Approach to Assessing the Impacts of the Deepwater Horizon Oil Spill in the Gulf of Mexico*. National Research Council of the National Academies. The National Academies Press, Washington DC. 235pp.