CONFIRMING RESPONSE EFFECTIVENESS: AN OVERVIEW AND GUIDE TO OPERATIONAL MONITORING

Alan J. Mearns
Biological Assessment Team
Hazardous Materials Response and Assessment Division
National Oceanic and Atmospheric Administration
7600 Sand Point Way NE
Seattle, Washington 98115

ABSTRACT: *Monitoring is a common, if generally implicit, part of spill response. However, there are no general guidelines for designing and carrying out spill response monitoring. With a few notable exceptions, as in shoreline assessments, monitoring rules are made up as the response progresses. Important rules of monitoring, such as those concerning adequate controls, are often left out or remain ill-defined, leading to unjustified product claims. As a result, there is often considerable debate and confusion about the effectiveness of protection and cleanup actions as well as the need to move on or terminate the response.*

A framework has been developed to help make monitoring an explicit process in spill response management. This framework forces (1) clear identification of the objectives of prevention and cleanup strategies, (2) selection of meaningful and useful end points, to quickly document effectiveness, (3) an appropriate strategy (statistical or otherwise) for collecting and reporting needed feedback information, (4) an often-overlooked quality assurance/quality control plan, (5) a data or information management plan, and (6) confirmation that a decision was (or was not) made based on monitoring results. The most important aspect of an explicit monitoring plan is that it forces not only clear definition of "how clean is clean enough," but also how this measure will be operationally documented. Elements of this process are indeed covered in many past responses and in some recent documents. Nevertheless, there is a critical need to develop additional rapid response tools for chemistry, biology, and toxicology and for a more thorough operational monitoring guide for spill response.

This paper reviews the roles of monitoring in spill response and the development of monitoring strategies that make a difference. It is based largely on the author's direct experience in designing and carrying out monitoring programs not only for spill response,[5] but also other kinds of marine pollution monitoring. General approaches to marine environmental monitoring have been debated on many occasions,[8] while, guidelines for monitoring specific oil spill response activities exist, but are scattered among various documents.[3,5,7,11]

The primary purpose of this report is to provide an overall context for operational monitoring and to ensure that specific monitoring activities adhere to basic monitoring principles. This report does not provide detailed step-by-step methods for specific kinds of monitoring. Instead, it defines some basic rules for developing operational monitoring approaches and provides some examples. It shows how to think about monitoring as an integral part of the day-to-day and long-term response decision process. It is meant to provoke responders to include monitoring explicitly in strategic planning for better spill response management.

Generally, those who carry out monitoring functions are experts far removed from direct communication with the on-scene coordinator (OSC) and OSC staff. An underlying objective of this paper, then, is to provide the OSC staff with some insight into the considerations, problems, and challenges that monitoring personnel must address during the course of a response. Likewise, members of the scientific monitoring community must understand the needs of the OSC and the constraints imposed on monitoring by the spill response. Finally, demonstrations or tests of new products are often done at spills of opportunity. Often, these tests are missing an important component of monitoring, such as adequate controls, which may lead to false or unverified product claims. It is thus incumbent on vendors as well to understand basic monitoring strategy and the need for it.

What is monitoring? "Monitor" simply means "to watch or check on." In this sense, monitoring involves repeated observations to document trends (are things getting better or worse?) or to confirm that an intended effect actually occurred. But there is more: "to monitor" also means "to keep track of systematically, with a view to collecting information; to test or sample on a regular or ongoing basis; to keep close watch over; to supervise; to direct." In this sense, monitoring is a purposeful, structured process that leads to specific decisions and actions. In effect, monitoring addresses uncertainties about alternatives and effectiveness: if you are certain of the outcome, you don't have to monitor—a decision that should be made explicit.

Why is monitoring important? Spill responders need timely, accurate information about the effectiveness and progress of response actions. Monitoring provides that information. Thus, monitoring is not an isolated activity but part of a management process that provides feedback confirming that the intended action not only took place but also resulted in the claimed or desired benefits. Monitoring is necessary for our expectations and perceptions to become facts. Some perceptions will be supported, reinforced; others will not. If done properly, according to specific guidelines (see below), monitoring will provide the basis for comparing alternative actions, deciding which courses of action to take next, and determining that a response is over. Monitoring will not resolve all uncertainties about response actions, but it will reduce uncertainty about what to expect from the actions chosen.

What is operational monitoring? Operational monitoring is any measurement or observation activity required to ensure the success of a response, and in particular, to direct or redirect the response decision process. Operational monitoring includes not only environmental trend monitoring to document oil removal effectiveness or shoreline recovery, but also compliance monitoring to ensure operations are kept within predetermined limits, product and equipment monitoring to track performance, and hypothesis-testing to resolve uncertainties about alternative outcomes of actions. All these types of monitoring are already evident in responses. Examples include collecting weather and sea state data and forecasts, observing and mapping oil or pollutant movements, confirming forecasts of pollutant fate, documenting the effectiveness of dispersant operations, recording amounts of hazardous waste recovered, and confirming that operations comply with incident-specific guidelines, such as those relating to wildlife conservation areas. What makes any kind of monitoring operational monitoring is timely reporting in a manner that can influence day-to-day as well as longer term strategic response decisions.

The framework: elements of monitoring

Although it may not always be apparent, any monitoring strategy has a number of basic elements: clear objectives, measurable or meaningful end points, a survey approach, timely conduct and reporting, logistical considerations, and quality assurance/quality control (QA/QC) procedures.

Clear objectives. It is critical to state and document the objectives of both response and monitoring actions carefully, no matter what level of monitoring is undertaken. What is the goal of the response in terms of oil removal? In terms of resource protection? How clean is "clean enough," in terms of amounts of visible or chemically detectable oil? When we say "remove the most oil with the least amount of additional impact on resources," what does this mean quantitatively?

The point of monitoring is to resolve uncertainty. The best way to define clear monitoring objectives is to carefully define what decisions will be made based on the monitoring results. What are the management or treatment alternatives under consideration, and what is it we are uncertain about? For example, can dispersants continue to be used? What are the specific benefits and consequences of increasing the temperature or volume of hot-water washing? What is the most critical information needed to decide if dispersants or in situ burning can be used? Is a specific proposed shoreline cleaning agent more effective and less harmful to resources than hot-water washing? Is bioremediation being considered, and specifically, do we have enough information about the factors limiting degradation?

Trends are important. For example, is the viscosity of the oil changing rapidly enough to prompt changes in response strategies (a clear consideration in the *Exxon Valdez* spill? Is oil already dispersing or sinking?

Degree of difference is important. In comparing response alternatives, what level of difference in effectiveness is being sought—two-fold, five-fold? Or, what amount of variation—10 percent, 0 percent, 100 percent?—is acceptable?

It is necessary to be quantitative about expectations. What will constitute a deciding dispersant run and what will not? A certain percent loss in cover of the slick? On one pass? Observable loss in two of three runs? In three of five? What specific weather conditions merit continuing or shutting down an in situ burn, and as measured by what? What specific recovery rate or biological impact condition(s) merit continuing or terminating shoreline washing? Shoreline cleaners? Bioremediation? How, specifically, should oil removal effectiveness be compared with resource damage?

Defining both response and monitoring objectives is an iterative process. It must be done frequently, with team members reminded of objectives daily. Moreover, the objectives may be refined as other specific monitoring considerations are reviewed (meaningful measures, survey design, timeliness, and so forth, as discussed below).

Meaningful measures (end points). What measures of effectiveness and effects are most appropriate and logistically feasible to monitor for expected outcomes or decisions? For slick dispersal, will the measure be percent cover estimated visually during overflights? Of what—sheen, slick, mousse? For in situ burning, is it feasible to measure particulate concentrations at a certain point to meet predetermined concentration criteria? Should effectiveness be measured in terms of oil recovered, or oil remaining, or both? By what kind of observation method? How do you measure oil recovery during shoreline operations? Or, should the measure be of residual oil, regardless of recovered amounts? How should remaining oil be measured? Or, should effectiveness be judged by the abundance or variety of shoreline marine life remaining, instead of by oil recovery? Will an easily observed index do in place of a more complicated measurement? Will the index be calibrated against the measured variable?

Although it may be assumed that operators monitor equipment to maintain peak performance, it should not be assumed that such "internal" monitoring will provide the information needed to decide about the effects or effectiveness of response methods. For example, shoreline washing equipment is typically regulated in terms of pressures and temperatures experienced within the equipment itself:[2] actual water velocities, impact pressures, and temperatures experienced by shoreline surfaces and marine life may be entirely different and even unrelated to equipment performance.

Survey design and sampling strategy. At a minimum, basic monitoring should replicate observations or sampling at *both* treated and untreated areas before and after treatment (or any other specified or anticipated action). In developing the survey design, it is necessary to consider at least five basic elements: controls, replication, stratification, equalization, and randomization.

Controls. A "control" is an untreated reference condition or location. Several kinds of controls must be considered. Nature plays many tricks. Effectiveness, in terms of percent improvement or removal, can only be convincingly demonstrated by comparison with an untreated or unaffected control. It may not be sufficient, and may even be misleading, to use conditions at a site before treatment as its own control. A slick can coincidentally disperse because of changes in wind. The next high tide can remove oil regardless of the treatment.

Often, new products or methods are tested at spills of opportunity. Shoreline cleaners and bioremediation agents are often applied in conjunction with high-pressure washing (as at a recent demonstration with a shoreline cleaner in Alaska). Without separating the two components—the cleaner and the washing—how can we know if effectiveness is due primarily to the chemical or to the water washing? Of course, if the system works, who cares? We should care, especially if a product vendor will subsequently claim that, without additional information, it was the chemical that made the difference. Could the same or a similar result be achieved without using the chemical?

Replication. One observation can be misleading. Differences among treatment areas or actions can only be demonstrated by comparison to differences within areas or actions. Failure to visibly disperse oil after one pass may lead to the false conclusion that the oil is not dispersible: what if it did disperse in two out of three passes? In three out of four? The same questions can be asked of any other treatment action. There are several rules for determining the amount of replication needed: all deal with the amount of variation expected and the amount of difference one wishes to detect between treated and untreated situations.[4]

Stratification. Large areas should be broken up for analysis into relatively homogeneous subareas. Shoreline types vary, including, for example, rocky and boulder-cobble shores, sand beaches, and marshes. Further, some shorelines can be covered with surviving marine plants and animals, while others are bare. What works in one area may be ineffective or even counterproductive in another. Observations to document treatment effectiveness must be made by comparing similar geomorphological and biological environments.

Equalization. Similar numbers of samples or observations should be taken from both treated (or affected) and untreated (unaffected) areas, or from the same area before and after treatment. What if two of three treated slicks dispersed, but, because of changing wind conditions, so would two out of three untreated slicks? Was the dispersant effective? Extending an overflight a few more minutes to observe the fate of untreated slicks might make a big difference in the final decision of whether to continue.

Randomization. Randomization is the final act of objectivity. Once replicate conditions are identified, treatments can be assigned randomly so that no bias is involved. Such bias might represent geographic or temporal gradients, as in alongshore salinity (which can alter dispersant effectiveness) or marsh cover density (which can alter nutrient dynamics during bioremediation).

The statement that basic monitoring should involve replicate observations or sampling at *both* treated and untreated areas before and after treatment (or any other specified or anticipated action) will cause considerable debate, as it should. It implies invoking a statistical approach, which many may not feel is appropriate in spill response. However, when sampling, there is no such thing as a nonstatistical approach: samples are just that—samples—and may or may not be representative of the larger area of concern. Statistics is simply the science of making larger inferences from samples. Thus, any sampling or observation activity represents a statistical approach, whether or not one is declared. For example, the decision to take a single sample or observation is in fact a decision not to replicate.

Logistical constraints. Operational monitoring is encumbered with the same logistical constraints as the response itself, and more. Monitoring requires logistical support from the response, in personnel, transportation, and time and attention from the OSC staff, and thus may be seen as adding to the response burden. However, if it reduces the response effort required, or helps set priorities, monitoring can reduce the load. Monitoring equipment can and should be simple, but there is need for further development of simple tools for chemistry and biology measurements (see below). Proper labeling and shipping is time-consuming and critical for chemical and toxicological samples and should be acknowledged as a critical activity.

Timeliness. The monitoring strategy and methods must be designed to provide information in a timely manner. If results are needed on a daily basis (as for the evening report) then monitoring personnel must undertake measurements and observations that are simple, yet that do not seriously compromise the tenets of monitoring strategy described above. Visual, olfactory, and audio observations and simple physical measurements, such as depth of oil penetration in shoreline sediments, salinity assessed with a refractometer, phone contacts with others who have taken recent measurements, are the only feasible tools under this time constraint. Measurements involving detailed oil chemistry analysis with gas chromatography/mass spectrometry (GC/MS) involve sample collection, labeling, shipping, extraction, and analysis at distant locations. They yield extraordinarily important information on weathering and biodegradation, but cannot be expected to yield such information in less than one to two days. However, this information is highly relevant to decisions regarding seafood contamination and bioremediation. Thus, responders should weigh the costs as well as the benefits of postponing some decisions to obtain the best information possible.

Quality assurance/quality control. Sampling or observation methods as implemented by different observers or types of equipment may not be comparable. Training is essential. Observers, samplers, and monitoring equipment should be periodically calibrated against each other for precision and against an accepted reference standard for accuracy. For visual observations, multiple observers are better than one: they can discuss their observations and come to some agreement (or not, which is just as important). Samples sent to laboratories for chemical analysis should include material from similar untreated sites as well as the source oil.

Implementing a strategy

Setting objectives. There is plenty of room for bringing monitoring into the planning phase of spill response. For example, response organizations can adopt and include in planning documents one of several shoreline assessment guides, such as NOAA's.[7] Plans are also available for some districts to monitor dispersant effectiveness, in situ burn plumes, and bioremediation. These plans bring decision-related objectives into planning and may represent templates for prespill design of other operational monitoring programs.

During a response, and in the absence of an operational monitoring plan, it is incumbent that the OSC and monitoring personnel meet directly and outline both spill response and operational monitoring objectives on scene. Appointment of a dedicated monitoring or QA/QC Officer should be considered. This officer's primary job is to ensure that both response and monitoring objectives are clearly understood and met.

Controls: the set-aside controversy. Securing and using controls—untreated areas—may be one of the most critical and hotly debated aspects of a response, because such an action implies delay in cleanup, at least for those sites. However, the value of maintaining, respecting, and monitoring untreated "set-asides" may outweigh the costs, which should be an explicit point of discussion. After scorching debate, set-asides were established at the *Exxon Valdez* spill in Prince William Sound and were effectively used to separate the effects of hot-water washing from those of the oil itself.[6]

Training and drills. Some monitoring scenarios are implicit in the early stages of most drills—collection and reporting of weather, and estimating the suitability of early countermeasures, such as dispersant use and in situ burning. However, later stages, particularly involving shoreline impacts, are often neglected. Extending drills to include monitoring, assessment, and decision making concerning the recovery of stranded oil could provide needed focus on the monitoring issues that extend real spills.

Do we need all this? It may be that some or most of the conditions described cannot be met during response monitoring. What should be discussed openly, however, are the consequences and uncertainties of not meeting them all. Those consequences may be acceptable or the risks of ignoring unmet conditions may be worth taking. These should be explicit decisions.

Stick with the results. It should be kept in mind that an unexpected or undesirable result is not necessarily a valid reason for rejecting the method.

Where to get more information

Any basic statistics book will reveal the considerations that must be taken into account in developing monitoring programs. As noted in the introduction, general approaches to marine environmental monitoring have been debated on many occasions, including a dedicated Oceans '86 Conference.[1,9,10] These papers deal less with operational monitoring and more with long-term environmental monitoring, but they do document the strategies and approaches that need to be considered in operational monitoring. Finally, guidelines for monitoring specific oil spill response activities are scattered among various documents.[3,5,7] Indeed, there is a critical need to develop a more thorough operational monitoring guide for spill response.

Acknowledgments

Special thanks are due to Robert Pavia and Jean Snider, Hazardous Materials Response and Assessment Division, NOAA, for their persistence and encouragement in helping develop this paper. I also appreciate the support of several NOAA scientific support coordinators. There remain important omissions in this discussion, which will be rectified in due time.

References

1. Bernstein, B. B., 1986. A philosophy for effective monitoring. *Proceedings, Oceans 86, Vol. 3, Monitoring Strategies Symposium*, Marine Technology Society, Washington D.C., pp1024–1029
2. Exxon Production Research Company, 1990. Valdez Oil Spill Technology: 1989 Operations. Houston, Texas, 100pp
3. Hayes, M. O., R. Hoff, J. Michel, D. Scholz, and G. Shigenaka, 1992. An Introduction to Coastal Habitats and Biological Resources for Oil Spill Response. Report No. HMRAD 92-4. Hazardous Materials Response and Assessment Division, National Oceanic and Atmospheric Administration, Seattle, Washington
4. Jaquez, G. M. and F. J. Rohlf, 1986. Problems in the variance of nine environmental monitoring variables: determining the number of samples needed to detect a change in mean of 50%. *Proceedings, Oceans 86, Vol. 3, Monitoring Strategies Symposium*, Marine Technology Society, Washington, D.C., pp974–979
5. Mearns, A. J., 1991. Observations of an Oil Spill Bioremediation Activity in Galveston Bay, Texas. National Oceanic and Atmospheric Administration Technical Memorandum NOS OMA 57. June, 38pp
6. Mearns, A. J., 1993. Recovery of shoreline ecosystems following the *Exxon Valdez* oil spill and subsequent treatment. *Coastal Zone '93, Proceedings, 8th Symposium on Coastal and Ocean Management*, American Society of Chemical Engineers, pp466–479
7. National Oceanic and Atmospheric Administration (NOAA), 1992. Shoreline Countermeasures Manual: Temperate Coastal Environments. Hazardous Materials Response and Assessment Division, National Oceanic and Atmospheric Administration, December
8. National Research Council (NRC), 1990. Managing Troubled Waters: The Role of Marine Environmental Monitoring. National Academy Press, Washington, D.C.
9. Segar, D. A. and E. Stamman, 1986. Monitoring in support of estuarine pollution management needs. *Proceedings, Oceans 86, Vol. 3, Monitoring Strategies Symposium*, Marine Technology Society, Washington D.C., pp874–877

10. Sullivan, P. K. and H-J. Krock, 1986. An objective water quality monitoring strategy. *Proceedings, Oceans 86, Vol. 3, Monitoring Strategies Symposium*, Marine Technology Society, Washington, D.C., pp1021–1023

11. United States Coast Guard (USCG), 1994. Gulf Strike Team Special Response Operations Monitoring Program (SROMP). Gulf Strike Team, U.S. Coast Guard, New Orleans. April 26, 28pp