THE USE OF RISK ANALYSIS AND RISK ASSESSMENT AT THE MASSACHUSETTS MILITARY RESERVATION

Resha M. Putzrath, Ph.D., DABT

INTRODUCTION

The frequently heard opinion, that risk was not and is not being considered during the decision-making process at the Massachusetts Military Reservation (MMR), is not accurate. The form of the risk analysis and the extent to which risk affected the actions taken, however, varies in time and by location (plume or source). The decision-making process, defined as risk management by the National Academy of Sciences, requires consideration of factors in addition to the risk assessment. Thus, the issue of whether risk has been appropriately used in any particular decision-making process at MMR is not solely dependent on the estimated risk. Other factors, such as community values, feasibility, and cost are appropriately considered in risk management decisions.

Any evaluation of risk analysis is complicated by the varied interpretations of what constitutes the process of evaluating risk. Risk analysis can range from the qualitative and subjective (e.g., any use or release of a man-made chemical into the environment must be reduced to the extent technically feasible); through screening procedures (designed to exaggerate exposures and toxicity to determine if a problem might exist); to complex, sophisticated, quantitative, site- and chemical-specific evaluation of current, future, or potential harm by methods developed within scientific disciplines (including toxicology, epidemiology, public health, and the ecological sciences) that result in a quantitative estimate as well as a qualitative assessment of risk.

A further distinction must be made between whether risks were considered in the decision-making process at MMR and whether risks were the driving or pivotal factor in specific decisions. As illustrated below, the decisions reviewed indicate that risk assessment, elements of risk assessment, or risk-based regulatory values were used in many if not all of the decision-making processes, but that risk was not necessarily the driving factor in those decisions.

The focus of this evaluation will be human health risks, as it is my area of expertise. In some cases, examples have been drawn from the evaluations of ecological risks, but review of these has been much more cursory. The review also focuses on activities since the arrival of the Air Force Center for Environmental Excellence (AFCEE), although previous activities and reports were also examined where necessary to understand AFCEE's actions.

This assessment was constrained by time, and it was not possible to review all risk assessments and all decisions. It was also not possible to speak with everyone who might

have been able to provide additional information. This evaluation therefore should not be considered comprehensive, but every effort was made to have the analysis be sufficiently representative to be useful. This report examines three major issues: (1) how risk analysis has been used at MMR, (2) how risk assessments have been conducted for MMR, and (3) potential utility of risk assessments for future decisions at MMR. The first two issues not only evaluate past practices, but also highlight issues that may become more important in the future, as risk assessment has been designated as a critical step in determining when remedial actions can be reduced or eliminated.

HOW RISK ANALYSIS HAS BEEN USED AT MMR

Various forms of risk analysis can all be useful, depending on the purpose of the evaluation. Quick and simple analyses, for example, may exclude situations considered to be of little or no concern from unnecessary further evaluation, or may highlight areas that warrant rapid assessment and/or action. Such screening assessments may have less utility in determining if actual harm has or is likely to occur. As will be discussed below, risk has been considered at MMR in various forms and by various protocols in the decision processes and documents.

The Groundwater Plumes Are Defined by the Maximum Contaminant Levels, which Are Considered Risk-based Values

A major use of risk-based analysis has been in defining chemical contamination in the groundwater. AFCEE has defined plumes as locations where the concentration of a chemical in the groundwater exceeds the drinking water standards as defined by the Maximum Contaminant Levels (MCLs). MCLs are defined as concentrations that are acceptably safe for drinking water; safety is the converse of risk. Since MCLs have been used to define contamination, this measure of acceptable risk is a *de facto* part of decisions for all groundwater remedies.

Note, however, that the process of determining MCLs, also consider factors in addition to risk, such as feasibility and cost for reducing the level of the chemical, as well as associated benefits, such as the contaminants that result from water chlorination. Thus, while MCLs are risk-based, they do not define a quantitative level of risk. If plumes were to be evaluated only on the basis of MCLs, chemicals that are present but which do not have an MCL would also not be evaluated.

Where Exposure Can Be Anticipated, i.e., a Completed Exposure Pathway Has Been Found, Action Was Taken

One component of risk assessment is an exposure assessment. For a risk to occur, there must be the possibility of exposure at a level that has the potential to cause harm. The oft-quoted phrase, “no exposure, no risk” reflects the requirement for a possible route for the

chemicals of concern to come into contact with the human, animal, or plant, which in risk assessment parlance is often called the receptor. If such a route or pathway exists, it is called a completed exposure pathway (as described, for example, in the Community Environmental Fact Sheet, "Understanding Exposure," April 1998). The presence of a chemical and a completed exposure pathway are necessary, but not sufficient, conditions for a risk to occur. Presence of a chemical below a level necessary to cause a toxic effect (i.e., below a threshold), for example, does not constitute a risk.

In some of the situations considered at MMR, a completed exposure pathway was discovered, e.g., concentrations above the MCLs in existing drinking water wells or chemicals in surface water that could pose a risk. In the cases that were examined where a completed exposure pathway existed, however, action was taken based on the potential for a real or perceived risk.

- When concentrations above MCLs were found in current drinking water wells, alternative water supplies were provided.

- When ethylene dibromide (EDB) was found in surface water, a time-critical removal action was undertaken at Fuel Spill-28 (FS-28), which was evaluated by subsequent data collection and risk assessment ("Draft FS-28 Treatment System Extraction Well-1 (EW-1) Evaluation Report; Time Critical Response Action," Jacobs Engineering Group, Inc., (March 1998), document #11157). For non-time-critical actions, the decision tree, which outlines the paradigm to be used to determine what actions might be necessary, uses risk levels as a critical factor in the analysis (Figure 7.1, page 7-13 of "Draft Fuel Spill (FS-28) Plume Technical Decision Memorandum; Volume I," Jacobs Engineering Group, Inc., February 1997 document #10274). Similarly, risk assessment was used to evaluate potential risks from EDB in surface water resulting from Fuel Spill-1 ("Draft Letter Report. Surface Water Risk Evaluation at Fuel Spill Site No. 1 (FS-1). Quashnet River/Cranberry Bog Area," September 25, 1997).

- Concern about EDB reported to be on cranberries led to various actions. For example, cranberry crops that were grown in EDB-containing water were not sold.

- Concern about consuming fish from or swimming in the ponds that had MMR-related chemicals led to a risk assessment for those ponds. As reported in the June, 1999 Community Fact Sheet (issued by Massachusetts Department of Public Health, Massachusetts Department of Environmental Protection, the Agency for Toxic Substances and Disease Registry, and health agencies of the local communities) that Johns Pond, Snake Pond, and Mashpee-Wakeby Pond were deemed safe for recreational activities. The restrictions on consuming fish from these ponds is due to mercury, which has not been associated with MMR activities.

If Risk Is Potential (e.g., if Land Use Changes) or in the Range of Usually Acceptable, Other Considerations Tended to Drive Decisions

In many situations examined at MMR, risks were either potential or below a level of usual regulatory action, which for Superfund sites occurs: (1) for cancer risks somewhere between one-in-ten thousand and one-in-one million upper-bound incremental increase in incidence and (2) for non-cancer risks when a chemical exceeds a toxicity value specified by EPA. In these cases, other factors usually become more significant in the decision-making process. Thus, risk may not be a driving factor precisely because it is of less (immediate) concern.

Within the range of generally acceptable risks, most of the people interviewed indicated that MMR has been treated differently than most sites. Some thought it should be, with reasons including: (1) contamination of a sole source aquifer, (2) unique character of Cape Cod as a natural resource, and (3) finding of locally high cancer rates with unknown cause(s). Others expressed concern that at MMR more remediation was occurring at this site than would normally be justified based on: (1) the potential risks compared with other sites throughout the country, (2) other potential sources of higher risks at the Upper Cape, and (3) the presence of the same chemicals from other sources at the Upper Cape that were not treated with the same level of concern. These opinions are likely to have influenced, and to continue to influence, how risk assessments are used in the decision-making process.

Applicable or Relevant and Appropriate Requirements (ARARs) May Overrule Risk Assessments

The Superfund process requires consideration of applicable or relevant and appropriate requirements (ARARs) of federal and state environmental laws. ARARs often override risk considerations when selecting remedies. Many ARARs are chemical-specific and are, at least in part, risk-based. Other ARARs may be location- or site-specific, and may not be risk-based. For example, laws that protect archeological or historic sites have been considered ARARs.

Restoration of a Sole Source Aquifer as an ARAR

People may disagree as to whether the goal of restoration of a sole source aquifer is a risk-based standard, but it is an ARAR. Restoration may be considered risk-based, if considered in the context of current or future environmental damage to a limited natural resource. As such, it would be a qualitative rather than a quantitative risk analysis.

Restoration has been a consideration especially when risks are below a level that is generally of concern at Superfund sites. The letters in early 1999 concerning the Storm Drain 5 South Plume (SD-5S) between Jim Snyder (AFCEE) and Paul Marchessault (U.S. EPA, Region I) illustrate this point. Mr. Snyder's letter of 28 January 1999 states that

"findings of the risk evaluations should be the primary basis for determining the need for active groundwater remediation for SD-5 South." The February 11, 1999 letter of Mr. Marchessault states, EPA's decision is based in large part on the fact that Cape Cod is a sole source aquifer" even though the letter also states that potential uncertainties in the risk assessment are also of concern.

The goal of restoration as an ARAR does not guarantee that it will be achieved. As discussed later, other considerations, e.g., technical and economic feasibility, may be considered.

Decisions Based on Non-risk ARARs

Whether or not restoration is viewed as a risk-based ARAR, some of the decisions made at MMR were not based on risk assessment. The mere presence of EDB reported to be on cranberries, for example, does not present an unacceptable risk, as indicated by the statement that "MDPH has confirmed that eating the cranberries does not pose a risk to public health." (Andrew T. MacCabe, Estimation of Cancer Risk for Ethylene Dibromide Air Exposure Pathway, Ref MDPH Technical Memorandum, 5 March 97, Memorandum for AFCEE/MMR, 11 April, 1997).

According to one person interviewed, the EDB on cranberries might have been considered an acceptable level, if the EDB had been present as the result of pesticide application. The same amount of the chemical, however, was not acceptable when EDB was considered an adulterant because the EDB occurred via contact with contaminated surface water. Apparently, under Massachusetts law no EDB above the specified analytic limit of detection would be acceptable in this case. (For this and other references to Massachusetts laws, the statements are accepted as presented by the person interviewed. No review of these laws or their application at MMR was undertaken as part of this analysis.) According to another individual, however, the same standard (below a specified limit of detection) is required for EDB in infant or baby food. In either case, the law serves as an ARAR. Our laws often do not result in a consistent level of risk protection (risk assessment is neutral as to the origin of the chemical), but they must be obeyed. Massachusetts law rather than the estimated risk appears to have been a driving factor in determining the action to be taken.

In decisions about EDB on cranberries, concerns about the requirements of the buyer regarding the marketability of the cranberries clearly were also a factor. The extent to which these concerns might have resulted in the same decisions absent the above-mentioned Massachusetts laws can not be easily determined.

Risk Estimates That Are Lower than Those Generally of Regulatory Concern May Not Be Believed and Therefore May Not Be Used

Finally, many people interviewed voiced a perception that estimated risks are only believed by some individuals to the extent that they find an unacceptable risk. Given the high cancer rates, some people questioned whether a risk assessment that does not indicate a danger is sufficiently comprehensive. Others felt that current knowledge about all chemicals is insufficient to declare any exposure safe. Some mentioned that any involuntary risk is unacceptable and should be reduced to the extent feasible. Others opined that, if risk estimates will not drive the clean up to background levels, the estimates will either be ignored or discredited, and another reason for cleaning to background will be found. These opinions suggest that current opinion of the risk assessment process may cause difficulties in the future, when risk assessment may become more important.

RECOMMENDATION: Hands-on experience with the risk assessment process may be useful to inform all participants about the strengths and limitations of risk estimates. Such knowledge should also allow more knowledgeable participation in the decision-making process.

Working with actual data from MMR, case studies could be developed that would illustrate the types of choices that must be made and their effect on the estimated risk. These case studies could be constructed for either group or individual use. While some people may justifiably still be skeptical about the risk assessment paradigm, such exercises would ensure that disagreements were based on common knowledge rather than lack of understanding or fear of the unknown. Such knowledge would also allow individuals to evaluate the reasons behind the differences when risk estimates from various agencies or individuals differ. (This process would have the additional benefit of educating individuals in techniques for evaluating risks reported in the media that chemicals, foods, or personal activities produce or reduce risks.)

HOW RISK ASSESSMENTS HAVE BEEN PERFORMED AT MMR

Criticizing specific risk assessments is easy, especially in retrospect. Choices must be made, and alternative decisions can often be justified, especially with the wisdom gained from hindsight. Individual opinions vary, and other risk assessors might make different choices or selections. A critical review of all individual risk assessments is beyond the scope of this review. Many of the procedures reviewed are common practice at Superfund sites. They are designed to overestimate risk. Thus, when the risks so estimated are below a level of concern, the true risk is below a level of concern.

The nature of the risk assessment process requires that each risk assessment include: (1) a clear statement of the purpose or goals of the assessment and (2) all the information (by inclusion or precise reference) necessary to repeat and/or modify the assessment to facilitate determining the effect of different choices. Without these elements, independent

review and interpretation of the results is difficult. Some or all of these elements were either lacking or difficult to find in the risk assessments reviewed.

Results of Screening Assessments Were Retained and Used as if They Were State-of-the-art Risk Assessments

Evaluating risk assessments at MMR requires understanding that many, if not most, were screening assessments, not actual risk assessments. Screening assessments based on risk assessment concepts are often used as a preliminary assessment of a situation to approximate the magnitude of the problem. Such screening assessments are designed to overestimate the potential risk to such a large extent that situations that pass the screen could not possibly pose an unacceptable risk, were an actual risk assessment to be conducted. An example of a screening assessment would be: if a chemical were spilled from a truck and if we find that if one individual were able to consume all of the chemical spilled he would not be harmed, then we can assume that the spill does not present an undue risk to human health. Situations that fail such a screening assessment may also not pose a risk, but they should be evaluated further. Screening assessments can also provide information as to the speed required for the actual risk assessment or for rapid action prior to the risk assessment. Using the same example, if the total amount of the spill were found to be 1,000 times more than the lethal dose for an individual, rapid action to contain the spill and limit exposure should be considered.

A potential problem is that the "risk estimates" produced by these screening assessments have been presented in MMR documents as if they resulted from state-of-the-art risk assessments, in particular in the summary documents such as Fact Sheets or Decision Criteria Matrix documents. When risks so estimated are below a level of regulatory concern, screening assessments are sufficient for determining the risk-based component of remedies. Even these assessments should be clearly labeled, however, otherwise people may wonder why the same concentrations result in different risk estimates if more accurate risk assessments are conducted later. If people don't understand the process, one of two reactions may occur. Some may assume the original risk is still present. Others may assume either one or both of the risk assessments are deliberately deceptive. Either situation is not desirable.

Screening Assessments Should be Clearly Identified

Screening assessments are often based on very limited, preliminary information. If only data on groundwater are available and if concern is raised about fish in nearby stream, several options exist. We can assume fish swim in the same concentration as was found in groundwater, possibly even the maximum concentration found in the groundwater. We can use models to estimate the concentrations that may get into stream, but use of models, especially with limited data, may introduce more uncertainty into the estimated risks. Finally, we can choose to ignore the fish, because we haven't found the chemicals in the stream to date.

Where examined, the analyses at MMR appear to use very conservative (i.e., use risk adverse) assumptions and models presumably to ensure that risks will not be underestimated in the screening process. (The examples discussed are largely based on interviews with people involved in the risk assessments; as mentioned below, this information is difficult to find in the documents.) The highest groundwater concentration, for example, was sometimes used as the concentration to which the fish would be exposed, i.e., the equivalent of assuming that the fish are swimming in the soil. Similarly, exposure to workers from chemicals in surface water, which sometimes is sprayed into the air, has been estimated by assuming that the cranberry bog workers were exposed while confined in a shower stall. Such assumptions will overestimate the risk, usually substantially.

Such practices may be prudent for protecting public health when data are limited. If no unacceptable risk is found, no further refinement of the risk assessment is necessary. Nevertheless, such screening assessments should be carefully qualified and the major screening assumption(s) appropriately identified, least they be interpreted as actual estimates of risk from consumption of fish or working in the bogs. Not only were these screening assessments not identified, but it was often difficult to determine which estimated risks contained such assumptions that, while justifiable for screening assessments, would not be reasonable for an assessment of actual risk. Based on several discussions, almost all of the people who would be using such estimated risks appear to be unaware of the extent to which the risks are sometimes overestimated beyond that normally encountered in regulatory risk assessments at Superfund sites.

Potential Problems with Screening Assessments

Screening assessments have resulted and will continue to result in assessments that are not only less sophisticated than state of the art but also may produce values that can not be interpreted. The ecological risk assessment in the "Remedial Investigation Report. Area of Contamination SE-4. Volume I - Text" (Final. April 1996, Table 8-102) includes hazard quotients exceeding 11 million and hazard indices exceeding 28 million. (A hazard quotient is the ratio of the expected exposure to the acceptable exposure for an individual chemical; a hazard index is the sum of these ratios for a mixture of chemicals. If either the hazard quotient or hazard index exceed one, further assessment of risk is warranted.) Prior to this document, the highest hazard index I had seen, even in a screening assessment, was about 200. A hazard quotient of the magnitude reported in the MMR document can not be reasonably interpreted unless one is, for example, finding dead bodies, in this case, of birds. Moreover, the values from this document were carried forward into another, recent document by another consulting firm ("Record of Decision. Areas of Contamination FTA-2/LF-2, PFSA/FS-10/FS-11, SD-2/FS-6/FS-8, SD-3/FTA-3/CY-4, SD-4, and SD-5/FS-5 Source Areas," September 1998).

It is well known that worst-case, screening assessments can use assumptions and models that lead to ridiculous values. It is incumbent on someone to: (1) recognize that such

values are either nonsense or verify that the risk actually exists, (2) rectify the error(s) in the analysis, (3) see that such nonsense values are not included in any documents, and (4) ensure that such values are not perpetuated in future documents by the absence of a reasonable review. It is not sufficient to state, "Although HIs [hazard indices] for the black-crowned night heron and ospreys were 3 to 6 [sic] orders of magnitude greater than 1, field observations during the RI [remedial investigation] suggest that the area would not be used extensively by either the black-crowned night heron on [sic] the osprey. For these reasons, response objectives were not identified for wetland sediment and surface waters." (page 7-62, "Record of Decision. Areas of Contamination FTA-2/LF-2, PFSA/FS-10/FS-11, SD-2/FS-6/FS-8, SD-3/FTA-3/CY-4, SD-4, and SD-5/FS-5 Source Areas, September 1998"). Other implications of the absence of an appropriate, independent, and overarching review of risk assessments conducted at MMR is discussed in more detail later in this analysis.

RECOMMENDATION: All screening assessments should be identified, and major assumptions that greatly overestimate risk should be highlighted.

Numbers such as risk estimates are frequently considered more certain than is warranted by the process used to generate such values. Special care, therefore, should be taken to identify values that are even more uncertain and likely to overestimate risks. If screening assessments are not properly characterized, more sophisticated and accurate risk assessments prepared in the future may not be believed.

Various Protocols with Different Risk Assessment Methods Have Been and Are Being Used, Depending on When the Assessment Process Began

At least three distinct processes have been used and are being used to conduct risk assessments at MMR: (1) the Risk Assessment Handbook (RAH); (2) the Decision Criteria Matrix (DCM), and (3) risk assessment protocols developed for Superfund sites, primarily described in the volumes of EPA's Risk Assessment Guidance for Superfund (RAGS) and associated documents, such as EPA's Exposure Factors Handbook and various directives that update or supplement RAGS. (For brevity, risk assessments conducted under the Superfund guidance will be designated as following RAGS protocols.) The first two were developed for MMR, while the third is used throughout the country. In addition, at least one formal public health evaluation was conducted, and several epidemiology studies have been performed on the Upper Cape. Some of these protocols result in more sophisticated and accurate estimation of risk than others.
The use of several different protocols for risk assessment has recently been codified. The Remedial Program Managers Consensus Agreement # 99-01(March 1999) states, "Past decisions and recommendations <u>already made</u> in Feasibility Studies, Decisions Documents, Proposed Plans, and Records of Decision submitted for public comment, and other documents based upon the 1994 Massachusetts Military Reservation Risk Assessment Handbook will not be revisited in regard to their risk assessment guidance." [emphasis original]

The decision to continue to use several protocols for assessing risk concurrently is likely to increase confusion and misinterpretation of the risk estimates. The established policy has and will continue to result in contemporaneous documents using different risk assessment protocols. Results may be confusing, e.g., the same chemical concentration in the same medium assessed by different methods can result in different risk estimates. Risk assessments in documents with the same submission date can contain risk assessments that differ substantially in their quality and degree of accuracy. For example, two documents containing risk assessments are dated September 1998. The first, "Southwest Operable Unit. Baseline Risk Assessment Memorandum," has risk estimates calculated based on RAGS protocols. The second, "Record of Decision. Areas of Contamination FTA-2/LF-2, PFSA/FS-10/FS-11, SD-2/FS-6/FS-8, SD-3/FTA-3/CY-4, SD-4, and SD-5/FS-5 Source Areas," contains assessments performed under the RAH, which are less sophisticated. As these documents bear the same date, most people would assume, incorrectly, that these documents used the same protocols, and therefore, can be directly compared.

RECOMMENDATION: The critical differences in the three protocols used to evaluate risks should be summarized in a manner that allows everyone to understand them.

Documents containing risk estimates derived from different protocols are still being produced. In order to properly interpret these values as well as those from previous risk assessments, knowledge about the key differences among the three protocols is necessary. This information may be especially important if these risks are reexamined as part of an analysis of changes in remedial action, as described in the final section of this report.

Decisions Within the Three Protocols that Will Also Affect Risk Estimates

Once the protocol to be used for the risk assessment has been determined, other decisions may affect the risk assessment process and the resulting risk estimate. Decisions that substantially affect the estimated risk should be discussed in documents where the risk estimates are characterized for those who have not participated in the risk assessment process. Such decisions will affect the appropriate interpretation of the estimated risk.

Regulatory Risk Assessment Procedures Differ from Unconstrained Risk Assessments

In order to understand and interpret the risk assessments conducted at MMR, we should distinguish between regulatory and other risk assessments. In this context, regulatory risk assessments are defined as those that follow default guidance assumptions and models without exception. Regulatory risk assessments generally use worst-case assumptions. Specifically, RAGS and other guidance for Superfund risk assessments encourage that practice. Even when exposure values are adjusted to provide a "central tendency" estimate, the norm is to adjust a person's activities, e.g., how often a person might contact contaminated soil. Other parts of the exposure assessment, such as the rate of penetration

of the chemical through the skin, are not adjusted from worst-case assumptions. Moreover, the toxicity values provided by EPA are worst-case estimates, and are not adjusted for the central tendency estimate. The ability of some chemicals that cause cancer at high level of exposure to cause cancer at low levels of exposure, for example, has been challenged by many toxicologists, especially when cell or organ damage appears to be required for the carcinogenic effect to occur. Thus, even when two calculations, a reasonable maximally exposed individual (RME) and a central tendency (CT), are provided, they are both likely to overestimate the risk. While such conservative risk assessments are often appropriate for many reasons, they must be considered when interpreting the estimated risks.

Risk Assessment Decisions Derived from Consensus Procedures

One of the situations examined in greater depth was the presence of ethylene dibromide (EDB) in surface waters, including cranberry bogs. This situation was selected because cranberry bogs do not exist near most Superfund sites, and thus it would illustrate how risk was used in a non-standard situation. Many decisions regarding cranberries and bog workers were made by consensus through discussions that usually included representatives of AFCEE, EPA Region I, Massachusetts Department of Environmental Protection, Massachusetts Department of Public Health, and when appropriate, organizations such as the cranberry growers and Ocean Spray. Such a consensus process can resolve disagreements, especially with regard to differences in general mandates or biases resulting from their agency's or group's general practices.

In general, this process appears to have worked reasonably well. Participants who were interviewed appeared relatively sanguine about general outcomes, although all had problems with one or more specific decisions. Sometimes one agency had a method or parameter that was essentially non-negotiable. Sometimes agency representatives believe their views were not given appropriate credence, especially when they view themselves as having fewer resources to allocate to MMR than some other groups.

Risk assessment by consensus can resolve differences early, often before the risk estimates are calculated. Thus, it is less likely that disagreements will result solely because a group does not like the outcome. The process has the possibility of producing a less biased, more generally accepted risk assessment. Existence of a forum for discussing alternative approaches for risk assessment can also provide a mechanism to change assumptions or models when more data become available. This appears to have been done on two issues regarding EDB: (1) the number of days a worker would be in the bogs and (2) the bioaccumulation factor for EDB in fish.

Risk assessment by consensus presents at least two potential dangers. First, if one party in the negotiation has undue real or perceived power, it may influence the results regardless of merit. Second, some data required for risk assessments are not subject to negotiation. One can negotiate how conservatively to model how many days an individual may wade in

a stream. One can even debate how best to interpret and model toxicity data from rodents to potential for effects in people. One can not negotiate certain biological, physical, or chemical constants, which would be the equivalent of voting on the law of gravity.

This consensus process does not appear to be well documented. While not all of the discussions need be recorded, the final decisions and rationale behind these decisions are important and currently are not easy to locate. Thus, it is not possible to evaluate many of these decisions, or the decision-making process, independent of discussions with participants. The rationale for many of the decisions is in meeting notes that are not generally available and may become inaccessible as people archive or discard old project information, or when they move to new positions.

RECOMMENDATION: When negotiated assumptions, models, or values are used in a risk assessment, the rationale and supporting documentation should be provided in the document containing the risk assessment.

In order to evaluate or update a risk assessment, the original rationale must be known. Some decisions may be based on site-specific data, in which case more recent data could be substituted. Others may involve decisions as to which assumptions (values or models) were evaluated concurrently. For example, a decision might be made to leave one parameter at the highest value and use average values for other parameters in the same risk calculation in order to fairly represent a high but reasonable risk. If these decisions are reexamined, these parameters should again be considered together. If the high value were examined alone, it might be considered unreasonable, because it was not in conjunction with the other assumptions. Such documentation should be part of future risk assessments, and to the extent possible, should be reconstructed for critical decisions made previously.

Risk Assessment by Computer Spreadsheet

Increasingly detailed regulatory risk assessment guidance and default values coupled with accessible computer software has led to a disturbing increase in risk assessment by computer spreadsheet. In this process, the default assumptions and equations are used in an uncritical manner to generate “risk assessments.” Most such computer programs have few, if any, judgment nodes. Thus, the risk estimates so generated incorporate neither judgment nor thought beyond that initially used to write the most generic guidance which, by its nature, can not include site-specific information beyond the concentrations of the chemicals present. Unless these evaluations are carefully reviewed and, if necessary, modified by someone with experience with risk assessment and knowledge about the site, they are unlikely to be the most accurate risk assessment possible with the available data.

All experienced risk assessors have an unfortunate abundance of examples of absurd “risks” that can be generated by the unthinking use of equations and default assumptions. A computer spreadsheet is not capable of determining when default assumptions and equations lead to either a misrepresentation of the situation or un-interpretable results,

e.g., the ecological risk hazard index, discussed previously. Fortunately, guidance is guidance, not law or regulation. Most (if not all) of EPA's guidance documents indicate that the assumptions contained therein should not overrule judgment that would lead to better risk assessments when more situation-specific information is available.

Many of the risk assessments reviewed appear to have used default equations and guidance uncritically. Such guidance errs on the side of overestimating risks, and the risks so generated have usually not been sufficiently large to alter remediation decisions. This less sophisticated method of analysis may therefore be sufficiently accurate for the current purposes, although they may not be sufficiently accurate for future purposes. As with screening assessments, such generic risk assessments should be clearly labeled.

RECOMMENDATION: AFCEE should consider conducting an audit that would determine and evaluate the major decisions regarding risk assessments, including procedures, models selected, exposure assumptions, etc.

Currently, many of the stakeholders appear to be unaware of a number of the decisions that have been made regarding risk assessment assumptions that could substantially overestimate the risk. The information that is available in existing compendia such as MMR's Risk Assessment Handbook or EPA's RAGS and Exposure Factors Handbook need not be repeated in an audit. Many risk estimates used at MMR, however, require information not in these manuals, such as parameters for exposure to workers in cranberry bogs. Furthermore, even working within any written protocol, choices must be made. For example, air dispersion models must be selected. Which key choices were made for each risk assessment needs to be documented, if individuals not personally involved in the risk assessment process are to understand and interpret the risk estimates.

To be most useful, this audit should be accessible to the layman. Thus, models should not merely be listed by name, but also a short description of the assumptions (e.g., of the model and its use at MMR) should be included. For example, does the model used to determine the concentration a person might breathe from gases released from surface water assume: (1) that all of the air is evenly distributed in all directions, (2) that the wind direction is always toward the person of concern, or (3) that the air is confined in the equivalent of a shower stall? The choice of data used in the model should also be described in a non-technical as well as a technically precise manner.

Such an audit would be useful for interpreting current risk estimates, as it appears that everyone is not equally aware of the overly protective nature of many of the assumptions used in the estimations of risk. The audit would help all stakeholders understand what has been done to date in risk assessments. Also, the results of an audit could serve as a basis for beginning discussions with the community as well as regulators as to whether these assumptions, which may have been appropriate for deciding if actions should begin, are still appropriate for deciding when activities may be reduced or stopped.

Dealing with Uncertainty When Assessing Risks

The risk assessment paradigm assumes we can predict potential effects on health by building upward from knowledge about the effects of individual chemicals (often in laboratory animals), models of fate and transport of chemicals in the environment, and assumptions concerning activity patterns. In contrast, the public health paradigm starts with observed effects in people and attempts to attribute causes for the observations. Both techniques are useful and have been successful in some cases. Neither is infallible. For a population such as that near MMR that is looking for causes and solutions to elevated cancer rates, neither process may be able provide answers with sufficient certainty. The potential for both procedures to be unable to answer the questions of most concern to some of the population should be discussed, both before studies are conducted as well as after the findings are available.

Need for a Clear Statement of Purpose for Each Risk Assessment

Before any risk assessment is conducted, the question to be answered or the issue to be resolved should be determined. Not determining the purpose of a risk assessment before it is done is equivalent to digging a hole in the ground without deciding whether the hole is to look for precious metals, to produce a drinking water well, or to become a swimming pool. A hole dug for one purpose may not be useful for another. Similarly, the purpose for risk assessments may vary. Many risk assessments are possible with the same data, and risks estimated for one purpose may not be useful for another. All may be accurately conducted, but the utility of the resulting assessment will depend on how accurately the assessment assists in resolving an issue under consideration.

One illustration of the effect of different goals for different risk assessments is the baseline risk assessment for the Southwest Operable Unit (SWOU). For the purposes of this assessment, areas within the SWOU were segregated by land use rather than by groundwater plume, as appears to have been done for other risk assessments at MMR. Thus, each land-use area could have portions of more than one plume. For example, Area 1 "includes part of the LF-1, CS-4, CS-10, and FS-13 plumes, as well as approximately 20 identified potential source areas." ("Southwest Operable Unit Baseline Risk Assessment Memorandum." September 1998. document # AFC-J23-35Q86101-M14-0001, page 4). There are several consequences that result directly from the decision to conduct the risk assessment in this manner.

- People are often concerned with the risks of a particular area such as the Crane Wildlife Management Area, which is one of the areas of the SWOU. The land use classification also segregates residential from non-residential areas. By focusing attention on areas of interest to different segments of the population, questions concerning those areas can be addressed.

- The approach used ignores the plumes and looks only at the analytical results of the groundwater samples in the area. Thus, all chemicals are evaluated, regardless of source.

- By not conducting risk assessments for individual plumes, however, the linkage between the presence of chemicals associated with MMR and any estimated risk is diminished. While the risk of an area will be designated, determining what portion of that risk is due to contamination from MMR would require either another analysis or at least a separate examination of the risks posed by individual chemicals.

- Remedies are usually addressed by source or the results of a source such as a plume. This is apparently the case for the plumes in the SWOU (as described in the MMR Installation Restoration Program Fact Sheet, "Proposed Plan For the Southwest Operable Unit (CS-4, CS-20, CS-21 and FS-13 plumes only), June 1999). In the SWOU risk assessment, however, it appears that data from several plumes are analyzed as though they are from one unknown origin. Thus, the maximum contamination appears to be the maximum detected in any plume in the land-use area, and the central tendency appears to be based on statistical analyses of all of the data considered from within an area, regardless of source. Thus, it is not possible to determine the effect of remediating one or more of the plumes on the risks evaluated for the land-use areas.

Currently, No Person (or Team) Is Responsible for Reviewing All of the Risk Assessments

Confusion resulting from the concurrent use of various risk assessment protocols is compounded by the lack of one person who is solely responsible for reviewing the quality of the risk assessments produced by MMR. No one person determines or reviews the potential utility or the purpose of a risk assessment before it is conducted. No one person knows the major decisions that have been made for all of the recent or past risk assessments. Many of the issues raised in the discussions above might have been resolved if such a person (or team) were assigned this responsibility.

Apparently, technical review of some of the risk assessments was initially provided by the Technical Review and Evaluation Team (TRET). In some cases, their comments were used to improve the risk estimates. Recently, however, it appears that the TRET only reviews documents on specific request, in part due to the large number of documents generated and in part due to the observation that most of the risk estimated are below a level that is generally of regulatory concern.

RECOMMENDATION: The goals for and results from risk assessments should be reviewed by an experienced risk assessor independent from the entity conducting the risk assessment. This person should be familiar with MMR and be responsible for ensuring that the risk assessments are posing and answering the appropriate questions to achieve the goals set by AFCEE.

All people and firms get bound by habits. Within any standardized risk assessment protocol, some choices must be made. An experienced reviewer with a different set of experiences and habits may discover methods to improve risk assessments. Furthermore, an independent reviewer can more easily determine if the risk assessments are meeting the purposes for which the assessment was conducted, as well as the overall goals of remediation of MMR, i.e., to see if the risk assessments are answering useful questions in useful ways.

Determining the Most Recent Risk Assessment for Each Plume or Other Evaluation is Difficult

One of the basic difficulties in reviewing the risk assessments conducted for MMR is with document control. It may be difficult to worry about document control when trying to act quickly, especially if some of the information is not being used at the time it is generated, e.g., risk assessments were sometimes completed after the action was taken based on finding a completed exposure pathway. Document control appears to be getting better, but it is still difficult to find some information quickly.

Risk Assessments are Contained in MMR Documents with Various Titles

Some risk assessments can be in separate documents, e.g., the SWOU baseline risk assessment mentioned previously. Other risk assessments are part of a larger document such as a remedial investigation (RI). Knowledge about the Superfund process can assist in locating the risk assessments for various sites, but as many of the evaluations were performed outside the formal Superfund process, risks are calculated in documents with other titles that give no indication that they contain risk assessments.

AFCEE Does not Generate or Control All of the Data Regarding Risk at MMR or the Upper Cape

Some of the factors in improving document control are within the authority of AFCEE, while other are not. People have tested their own wells. Regulatory agencies (both federal and state) have different mandates that may require data collection and analyses. Clearly, not all of these data are under the control of AFCEE.

Data sharing, which evidently has sometimes been a problem in the past, is improving according to all who were interviewed. Agreement for how and when data will be released to all interested parties is an important issue that has become more structured. Accessibility and transparency of use of the data are important for risk assessments to be understood; they are essential if risk assessments are to be accepted. Toward this goal, all data which are the basis of decisions should be available to all interested parties.

RECOMMENDATION: AFCEE should continue to improve document control and should attempt to obtain copies of all data used for analyses that are critical for decisions.

If risk assessments are to be part of a decision-making process, the documentation of the methods and assumptions underlying the risk estimates must be readily accessible. Participants in the decisions should be able to determine, for example, if screening assumptions that generate large overestimates of risk were used and to approximate the effect of substituting alternative assumptions. Otherwise, risk estimates must be taken on trust and as absolutely certain values, rather than one factor in the process.

All of the Data Necessary to Evaluate and Replicate a Risk Estimate may not be Presented nor Referenced in the Document where the Risk is Calculated

> Some of the decisions and data relevant to understanding the risk assessments currently reside in personal files or otherwise not readily accessible through only public documents. Some of the necessary information has been included in documents produced subsequent to the risk assessment, which can make it difficult to locate. In one case, for example, the source of the model used to generate dermal permeability constants was always cited, but the values used to calculate this parameter, as well as a table of the final parameters were not in the risk assessment document, but in a later document. Summary documents, such as the Fact Sheets, do not indicate which of the multitude of other documents provide the supporting information for the risk values cited.

RECOMMENDATION: All future documents should contain references to the documents where the supporting data, models, or calculations can be found.

This is not a recommendation for a long list of documents in a separate reference section. Rather, it is hoped that the appropriate references will be interspersed where they will be most useful. For example, in Fact Sheets that contain risk estimates, the document that contains those calculations would be cited with the table or paragraph containing those values. Similarly, if a table contains a value (such as chemical-specific skin permeability constants), both the overall guidance (which is usually mentioned) as well as the MMR document (including page or table number) in which each constant was calculated (which is generally absent) would be cited.

THE FUTURE UTILITY OF RISK ASSESSMENT AT MMR

Risk assessments are likely to become more important in future decisions at MMR. Initially, people are often less interested in the quantitative risk estimate and more concerned about whether an activity (eating the fish, swimming, breathing the air) is safe. Yet, with a little reflection, most people understand that none of these activities is completely safe and therefore "safe" usually becomes "sufficiently safe" or "incrementally safe" relative to the activity without the incursion of concern. Risk assessment alone can not answer the questions of how safe is sufficiently safe, as that answer involves issues beyond risk. Risk assessment, however, can be extremely useful in providing information helpful for an informed decision-making process.

Risk Assessment as a Method for Defining "Mess"

It has been said in several venues that MMR should clean up the mess it made. Risk assessment may be useful for determining what constitutes "the mess."

Consider the following illustrative example. A person washes his vehicle with tap water every month. Is the run-off from that washing a mess that must be cleaned up? What if there were 100 such vehicles? What if the washing were done every day? Now consider the same series of questions, but assume that soap were added to the water. Consider these actions again, assuming an organic solvent, such as trichloroethylene (TCE), were used as part of the process.

Most people would not consider that washing vehicles with tap water constitutes a mess that needs remediation, but somewhere along the changing situations, most people would consider a mess to have been created. That changing the chemical and changing the amount of the chemical changes the potential for harm are key elements of risk assessment.

Now consider washing one vehicle once a month with tap water right next to a sensitive dessert environment. Clearly, even tap water has the potential to affect the dessert environment more than if the site is MMR, but will the effect be harmful, or in risk parlance, adverse? Would one bucket of water cause harm? one hundred buckets? Differences in response by various people, animals, plants, and ecosystems are part of the risk assessment process. Similarly, in the assessment of toxicity, the distinction between effects and adverse effects must be made. Eating a meal will have an effect on the body, but usually it is not an adverse effect. This exercise can be stretched beyond its utility, but it serves to illustrate how the concepts of risk assessment might be useful for converting the intuitive concept of "mess" into a defined concept that can be discussed and serve as the basis for decisions.

Risk Assessment as a Procedure for Organizing Information

To conduct a quantitative risk assessment, information that most people would want to know about a situation must be collected: which chemicals, how much of them, what effects they might cause at those levels, how people (or animals or plants) might get exposed, and what is the chance that this might occur. Risk assessments, therefore, could provide a method of organizing information in a manner that was relatively simple to understand and, if properly documented, could facilitate access to more information as desired.

Risk Assessment may be Useful for Determining When Remedial Actions can be Reduced or Eliminated

In all cases of environmental contamination, everyone would prefer if every molecule could be recaptured or otherwise eliminated from the environment. This goal is rarely technically or economically feasible. At some point, therefore, we usually have to decide if sufficient remediation has occurred. Such decisions are rarely based solely on the results of risk assessment, but risk assessment can provide information that is useful when making such decisions.

A decision-making process has been determined for groundwater clean-up at MMR. As reported in the MMR Installation Restoration Program Fact Sheet, "Proposed Plan For the Southwest Operable Unit (CS-4, CS-20, CS-21 and FS-13 plumes only)," June 1999, it consists of a three-step process:

1) "Remediate the aquifer to federal and state drinking water standards (MCLs).

2) "When MCLs are achieved and before the system is shut off, perform a risk assessment to determine if unacceptable risks are present, continue system operations and/or pursue additional measures as required to achieve acceptable risks.

3) "Once acceptable risks have been achieved, evaluate the technical and economic feasibility of additional remediation to approach or achieve background concentrations."

The second step, a risk assessment, can have various purposes. A risk assessment would presumably include chemicals for which no MCL is available, but for which toxicity information is available so that a risk assessment can be performed. The assessment of residual risk after MCLs have been achieved would also determine if the residual risk is below a level that is normally of regulatory concern, as MCLs are not determined solely by risk and each MCL may represent a different level of risk.

As discussed above, current assessments concerning MMR follow various protocols and are not of the same accuracy. The acceptability of the risk is likely to be affected by the

magnitude of the estimate, and the method used may substantially affect that value. Thus, the issues raised in the previous section regarding different methods for evaluating risks will become more important. Furthermore, risk assessment is an evolving discipline, and advances made in risk assessment subsequent to any previous assessment should be used.

RECOMMENDATION: AFCEE should contact stakeholders to begin the process of determining how such residual risk assessments will be conducted.

To the extent that agreement can be reached before any one remediation is examined, the result is likely to be viewed as less biased. Current procedures that would be used if the assessment were conducted today can be identified. Selection of procedures should identify choices that must be made with regard to models and assumptions. Processes for updating these procedures can be determined. Furthermore, ways of including alternatives to the regulatory default practices could be identified. For example, a limited number of risk assessments based on alternative assumptions championed by certain groups or individuals could be conducted in parallel to determine the effect of such assumptions on the estimated risk. Such alternative risk estimates could provide confidence in the selected course of action or could highlight uncertainties in the standard practice. Either outcome would provide more and useful information upon which to base decisions.