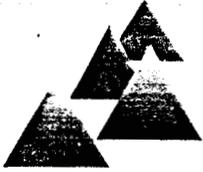


FINAL REPORT

**An Independent Review
of the Document**

**INCREMENTAL RISK ASSESSMENT
METHODOLOGY**

February 22, 1993



Rocky Mountain Universities Consortium

42 pgs.

FINAL REPORT

An Independent Review
of the Document

INCREMENTAL RISK ASSESSMENT METHODOLOGY

Prepared for:
EG&G ROCKY FLATS, INC.
Dan Shane

Prepared by:
The Rocky Mountain Universities Consortium

attention:
c/o C.S. Smith BE 224
Denver Research Institute
University of Denver
Denver, CO 80208

The Rocky Mountain Universities Consortium was formed to provide independent review and assessment of environmental restoration efforts. It is comprised of universities from several Western states. Reviews are conducted independently by selected teams of faculty and research staff.

A current list of Consortium reviewers is available on written request from the Consortium office.

The Consortium Operating Committee approved issuance of this Final Report on February 22, 1993.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Overview	1
Scope	1
Review Committee Structure	1
Critique of the IRAM Approach	1
Summary of Recommendations	2

SUMMARY OF REVIEWS 3

APPENDIX - Individual Reviewer Comments 6

Reviewer A	7
Reviewer B	11
Reviewer C	20
Reviewer D	25
Reviewer E	28
Reviewer F	34

EXECUTIVE SUMMARY

Overview

The Rocky Mountain Universities Consortium reviewed the report, " Final Incremental Risk Assessment Methodology for the Rocky Flats Plant" [IRAM], May 1992 (Project 22897-700), produced by Woodward-Clyde consultants. This report identified the potential methodologies for assessing incremental, cumulative risk. This is to support NEPA requirements for the site-wide environmental impact statement.

One of the components of this task will be to evaluate potential cumulative impacts associated with actions planned for restoration of the Rocky Flats site. Another may be to estimate the cost of reducing risk to achieve suitable levels of risk for different site end uses.

Scope

This review includes comments on specific items contained in the IRAM document as well as comments and suggestions which the reviewers believe will improve the underlying intent of that document.

Review Committee Structure

The review committee was comprised of six reviewers, and their individual comments are included. The reviewers have very diverse experience, academic and research backgrounds; consequently, their individual comments have quite different emphasis or focus. However, there are a number of common themes which appear in several of the individual reviews. Those common themes are embodied in the overall summary of recommendations given below.

Critique of the IRAM Approach

All of the reviewers expressed some degree of admiration for the general thoroughness of the IRAM document. But any attempt by a single contractor to develop an integrated approach to a task as complex and controversial as a general risk assessment of the Rocky Flats plant cleanup and restoration efforts is going to collect a lot of dissention.

If the IRAM document is considered as a "first-cut" to illuminate the principal problem areas, then it is a success. If its recommendations are taken as "the plan" for initiating the actual risk assessment studies, then the risk assessment will surely be a failure.

As indicated in recommendation nine below, a much larger, well coordinated effort with contributions from a wide assortment of experts will be necessary to develop an integrated risk assessment methodology that is consistent across the different cleanup alternatives, acceptable to government regulators and public interest groups, and compatible with the broad spectrum of scientific disciplines involved.

Summary of Recommendations

1. Inconsistent nomenclature and definitions should be remedied.
2. A coherent plan to verify and validate all conceptual (mathematical) models should be established. There must be special concern with spatial heterogeneity and temporal variability of processes, boundary conditions, and forcing functions.
3. An on-going effort to collect and improve site specific data used in the risk assessments should be maintained.
4. Consistent methods for validating all data values, whether measured or assumed, should be maintained.
5. "Additivity screening" should be expanded to consider or include uncertainties before any elements are eliminated.
6. A chart, logic diagram or similar pictorial representation summarizing the IRAM process should be prepared.
7. A program to up-date the integrated risk analysis and to reconsider its results throughout the clean-up and restoration work should be established.
8. The risk analysis and decision processes must be founded on consistent, probabilistic analysis.
9. Additional methodology studies focusing on principal portions of the integrated risk analysis are needed before specific risk analysis tasks are started.

Summary of Reviews

Reviewer A believes the recommended hybrid methodology to be appropriate but is especially concerned about the extreme generality of the document. Specific comments are given below.

- A standardized nomenclature is needed.
- There is a lack of specificity with respect to the Rocky Flats Plant.
- Treatment of key issues is insufficiently explicit.
- The IRAM proposals tacitly assume that the underlying science is "good" but this assumption is extremely restrictive. Scientific issues dealing with understanding and modeling of the transport processes and other phenomena related to environmental risk must still be resolved. In particular, the spatial and temporal heterogeneity of model inputs is treated too lightly.
- There is no provision for field validation of model predictions. The reliance on conceptual models is not considered adequate, especially in the face of so much uncertainty.
- Physical models, which have been verified with field data, are preferred, yet no plan for collection and validation of data is proposed. It was impossible to determine the adequacy of the existing data from the available information.
- There are several kinds of uncertainty: natural, model, parameter, economic, political, and technological. They have not all been addressed adequately.
- The methodologies are based on assumptions of independence and linearity; not all processes behave in this way.

Reviewer B offers a number of specific comments and cites the lack of consistent terminology, coherent risk assessment method, and a rational decision process as a principal failing of the report. Some comments follow:

- There is excessive preoccupation with detail rather than focusing on a unifying methodology.
- The reviewer suggests that the discipline of Safety Analysis may provide the unification needed. An overview of safety analysis is provided.
- Evaluation of the risk due to a treatment plan is not of much value without also evaluating the efficacy of the treatment plan.
- Apparent inconsistencies in ranking the alternatives in decision analysis are pointed out.
- There is need to plan for a "living" or "evolutionary" risk assessment activity.
- The reviewer suggests that a task force be formed to do a model IRAM, similar to that formed to do the Rasmussen Report on nuclear reactor safety.
- The inapplicability of the IRAM to CERCLA seems to imply wasted effort.
- Risks do not include explosions, falling equipment, fire, et. al.
- Several deficiencies in the proposed calculations are identified.
- Without assignment of values it will be impossible to compare human and environmental IRAM's.

Reviewer C commends the authors and reviewers for the great deal of work that has obviously gone into the document. Fifteen specific comments are offered.

- In discussing risk, insufficient emphasis is placed upon the fact that risks from low level environmental contaminants can be zero. Quotations from three referenced National Academy of Sciences reports are given to illustrate this point.
- The reviewer notes several internal inconsistencies which detract from the quality of the document. These include several different definitions of "slope factor," inconsistent use of DARA, mention of "proposed rule" with no indication of the nature of the proposed rule, and confusion of how IRAM and CERCLA interrelate.
- The document compounds inconsistent nomenclature.
- It is unclear why inconsistencies in chemical intakes are assumed. The description of chemicals of concern seems naive and incomplete.
- In refining the list of key receptors, direct contact with the U.S. Fish and Wildlife service is recommended in addition to reviewing the indicated information.
- The acronym list is incomplete; the definitions in the list do not always match the definitions in the text.

Reviewer D notes the encyclopedic coverage of the regulatory requirements but is concerned that the ascendancy of detail over focus makes it difficult to maintain a clear view of direction and purpose. The following suggestions and comments were provided:

- A chart, even a large fold-out chart, summarizing the features of the various approaches would have been welcome.
- Questions are raised about the combination of several methodologies for the recommended IRAM. Would execution of the combination of methodologies by different operators give the same results?
- The largest concern is the adequacy of existing methods for selecting receptors and exposure pathways. The physics of the various transport mechanisms are simply not understood.
- It is suggested that direct measurement of environmental concentrations, even with large measurement uncertainties, is preferable to modeling.

Reviewer E states that the document is strong and well put together. A shortcoming is that no indication was given as to how the four human and three environmental IRAM methodologies were selected for detailed examination. There is no listing of possible IRAMs from which the seven were chosen. Other comments include:

- It is inappropriate to assume that the data developed under CERCLA for the various OUs will be reliable, available, compatible and able to be integrated with other data to make an adequate NEPA analysis.
- Site characterization and data evaluation methods should be included in the IRAM.
- The document implies that remediation will lead to risk reduction; the opposite may be the case if, for example, altering subsurface conditions alters the pathways of contaminants from another OU.

- Field validation of the data used in the IRAM should be included in the document. Uncertainties might be huge, rendering estimates of potential change in risk meaningless.
- The IRAM will have to be iterated frequently to keep up with changing conditions at the Rocky Flats Plant.
- The cost of risk reduction is not considered.
- The time frame of the risk assessment is not clear although it should be.
- Two specific comments are offered expressing concern about screening out of "insignificant" contributors too early in the evaluation process.

Reviewer F dealt primarily with the ecological risk assessment. The authors have done a good job and the listing of advantages and disadvantages of each method was useful. The strongest criticism is that the document is very general with little information pertaining directly to the Rocky Flats Plant. Another strong criticism is the lack of field validation. Twenty specific comments are offered on the document and the recommended IRAM. A number of specific comments follow:

- There is too much emphasis on modeling and not enough on measuring.
- Much of the material is boilerplate directly from EPA documents, making a critique of the IRAM document a critique of EPA methods.
- The unique feature of the document is the hybrid IRAM.
- The use of biomarkers should be included in the methodology.
- Risks to one population have the potential to affect other populations.
- A more specific list of species and endpoints should be provided.
- Little insight is provided as to how effects of physical disturbance will be separated from those of the COCs.
- The DARA method would be more appropriate than the additivity screen for initial screening of key receptors.
- The cumulative risk methodology is the best approach to estimating environmental risk at the Rocky Flats Plant.
- Risk assessment verification may need to be done by comparing to a "reference area".
- Top predators are a poor choice for receptor organisms.
- Natural communities are better than individuals or populations for assessing ecosystem health. Species diversity, abundance and community composition are to be preferred over nutrient cycling and primary productivity as ecosystem indicators.
- Direct measurement is the best way to assess bioavailability and potential transfer of contaminants in natural ecosystems.

Appendix

Individual Reviewer Comments

Reviewer A

Reviewer A Comments

GENERAL COMMENTS

The document presents a detailed review of several human health IRAM's as well as several environmental IRAM's. My specific comments with respect to the recommended IRAM's follow. The general framework for the review is that the recommended IRAM must be appropriate with respect to the following general objectives:

- a) Characterize individual and cumulative risks and incremental changes to those risks from continuing and reasonable foreseeable future actions at the RFP site;
- b) Integrate site data from all OU's;
- c) Identify maximally exposed site-wide receptors based on potential cumulative exposure from multiple OU's;
- d) Identify important and unimportant combinations of OU's, chemicals/radionuclides, media, exposure pathways, and receptors with respect to incremental cumulative risk;
- e) Estimate exposure point concentrations based on cumulative site data, multiple sources, and exposure points of newly identified site-wide receptors;
- f) Evaluate uncertainties associated with combining pathways and population dynamics;
- g) Evaluate OU-specific uncertainties;
- h) Support risk management decision-making.

The objectives of my review below include the following issues: a) appropriateness of proposed incremental risk assessment methodology (IRAM) with respect to the objectives that must be achieved; b) the applicability of the methodology in the context of the specific situation under consideration; and c) availability of data. My assessment indicates that, if provisions are made so that some of the issues below are dealt with, the recommended hybrid methodology is generally appropriate.

The review assumes that all the science used in order to identify receptors, sources, pathways, risks of exposures, etc. is adequate and valid. This review will look only at the validity of the methodology given the science. However, it should be pointed out that this assumption is extremely restrictive. Clearly, scientific issues dealing with our ability to understand and model transport processes and other phenomena related to environmental risk in the presence of high spatial heterogeneity and temporal variability must still be resolved. These issues are at the leading edge of scientific research.

One of the major deficiencies of the document is its extreme generality. Perhaps, this is the most important flaw of the proposed IRAM, and one which may render the methodology inapplicable or not feasible of being monitored. The proposed methodology is not explicit enough in key issues. As a consequence, this review is also very general, because to do otherwise would imply a document beyond the scope of the review. This review does not intend to present the adequate IRAM but to point out potential problems with the proposed IRAM.

Another important flaw is the reliance of some aspects of the proposed methodology on conceptual models. Clearly, if the IRAM deals only with the assessment of risk once the

receptors, exposure pathways, etc. have been identified, then the conceptual model may be acceptable. However, when part of the procedure involves the identification of receptors, pathways, and so on, which are extremely dependent on the natural physical laws affecting transport and exposure processes, then conceptual models are inadequate and of very limited use. This is even more important in the presence of so much natural, parameter, and model uncertainty. Clearly then, what is required is the development and use of physically-based models which have been verified for the specific conditions of the RFP. Verification in the field should be a critical component of this model. It should be emphasized that the modeling exercise involves not only the conceptualization of process behavior but also carrying out model sensitivity analysis and the verification of model performance with field observations.

In addition, it is clear that the applicability of the proposed IRAM is dependent on the availability of sufficient data. With the available information it is impossible for me to determine the adequacy of the existing data. A document specifying what kinds of data are necessary to carry out the IRAM's should be an integral part of the proposed methodology. This document should also explicitly state whether these data sets are available, and if necessary, it should propose alternative methods and monitoring plans to obtain them.

In carrying out a risk assessment the most important issue to be resolved relates to the different kinds of uncertainty. Several kinds of uncertainty can be identified. *Natural uncertainty*, related to the inherent nature of the process under consideration. *Model uncertainty*, related to the uncertainty in the model chosen to represent a given process. For example, choosing a log-normal probability distribution function (pdf), instead of a normal pdf, to model the distribution of annual rainfall. *Parameter uncertainty*, related to the uncertainty present in estimating population parameters as a function of a limited sample. *Economic uncertainty*, related to the parameters of the cost functions, damage functions, benefit functions, utility functions, membership functions, etc. associated with the evaluation of risk. *Political uncertainty* and *Technological uncertainty* and others. For the case of the RFP these uncertainties can be expected to be very large. This is a result of the great natural heterogeneity of the hydrogeologic conditions to be expected, as well as of the intricacies and complexities of ecosystem and population dynamics and interdependencies. There is no provision in the recommended IRAM for an evaluation of the inherent uncertainty associated with the methodology. In addition, as new knowledge is incorporated and as new observations are made, and as remedial actions are taken, all of the above uncertainties will be modified. There is no objective definition of how these uncertainties will be affected. The concepts of expected value of sample information, expected opportunity loss, expected opportunity loss, etc. which are the basis of Bayesian Risk Analysis may be useful in refining the proposed IRAM.

As implied above, risk assessments presume development of functions associated with consequences of particular phenomena. Given so many different kinds of data, with different levels of associated uncertainty, which are part of both OU and site-wide IRAM's at RFP, a methodology must be developed to make all these different data types commensurate in terms of risk assessments. This is also lacking in the document.

However, as stated above, the recommendations have to be made specific for RFP. It is precisely the lack of specificity with respect to RFP conditions which makes this document so general. A few additional specific comments follow.

Recommended Human Health IRAM. The proposed methodology is a hybrid of the four methodologies reviewed in the IRAM document. Namely, additivity screen, cumulative risk assessment methodology (CRAM), decision analysis/risk analysis procedure (DARA), and a probabilistic analysis.

- a) The proposed hybrid methodology depends on developing a site-wide conceptual model in order to identify site-wide IRAM receptors likely to have impacts from multiple OU's and the OU/chemical/media pathways. As indicated below, physically-based models which have been verified with field data, and which have been subjected to a sensitivity analysis should be preferred.
- b) The methodology proposes that a standard set of exposure assessment equations and a standard nomenclature for input/output variables should be established. However, no specific recommendation is made on how to achieve this goal.
- c) In recommending that a Baseline Risk Assessment be performed for each OU the document proposes that combinations of OU/chemical/media/pathway/receptor whose risk to a given site-wide receptor are less than *some fraction* of a hazard index of 1, etc. What is this fraction and how is it going to be determined?

Recommended Environmental IRAM. A hybrid methodology based on additivity screen, DARA, and CRAM procedures. My specific comments emphasize those indicated above.

- a) This section, like all the others is extremely general in its recommendations, descriptions and statements. For example, in referring to CRAM it is stated that it will provide a ranking of ACTUAL RISK when *various endpoint measurement techniques* are implemented at the site. However, what these techniques are, and how the actual risk will be evaluated are left undefined.
- b) Data dependency is not addressed. No data collection plan, field verification, etc. are proposed.
- c) The issue of space-time variability of processes, boundary conditions, and forcing functions is not addressed. For the case of the environmental risk assessment, population dynamics and interdependencies only add to this problem.
- d) Underlying most of the methodologies proposed (both for human and environmental risk assessment methodologies proposed) are the assumptions of linearity and independence. Clearly, not all processes, exposure pathways, receptors, etc. behave linearly (that is, additivity can be used), nor are independent (that is, joint probability as a product). Procedures to allow for non linearity and interdependence should be investigated and included as part of the IRAM. This applies to both human and environmental IRAM's.

Reviewer B

Reviewer B Comments

I. GENERAL COMMENTS

The subject document contains a lengthy review of the several different risk definitions used by the government regulatory agencies concerned with the clean-up efforts at the Rocky Flats plant. The document authors state several times, all of Section 2.5.2, p.2-20 for example, that there are no consistent definitions of terms or methods between those government agencies. They seem to lament that fact but do not offer unifying suggestions. Apparently the authors assumed that this farrago of terms and methods will be sorted out later in the decision process. They were so devoted to examining the bushes, shrubs and trees that they missed the forrest. The forrest is, in my opinion, the certain futility of trying to satisfy those agencies unless the people doing the risk assessment employ consistent terminology, a coherent method and a rational decision process.

That failure is my major criticism of the report or of the task definition for the report. I believe that the report should have recommended common risk assessment terminology, steps for a coherent risk assessment method and principles for a rational decision process which incorporates the risk assessment results. These definitions and principles would then guide the development of specific risk assessment data, methods and recommendations. Those results could then be adapted to the idiosyncrasies of the different regulatory agencies.

Any effort to identify and develop risk assessment methodologies which satisfy different, even conflicting requirements based on different terminology and objectives is doomed to failure. Without clear, consistent definitions and principles, everyone will become trapped in a quagmire of conflicts. But, if the decision method and attendant risk assessment are developed from a rational, consistent basis, you will have a substantial base for seeking compromise with the regulators.

Because I feel so strongly about the necessity for a consistent terminology and a coherent risk assessment method, I will offer my views of the basics for meeting those necessities. Admittedly, my ideas are simple, perhaps too simple, but one should start from the simple and work towards the complex. Jumping into the middle of the complexities is usually a sure ticket to failure.

II. SAFETY ANALYSIS BASICS

Risk assessment is just one part of a larger process many of us refer to as safety analysis. The results of the safety analysis along with other considerations such as availability and dollar cost should be combined in a rational manner to arrive at a decision. McCormick, Ref. 1, provides an excellent introduction to performing and judging risk assessment. The decision is always which of several alternative activities to do; including the activity of doing nothing. I will come back to the decision process later but first I want to focus on risk assessment and its place in safety analysis.

Risk alone cannot form the basis of a rational decision; no one would ever choose to "take a risk" unless there is also some efficacy. Therefore, the decision maker must also have access

to an efficacy assessment on the proposed activity. Lowrance, Ref. 2, is perhaps the strongest proponent of this view. Efficacy is not as common a word as risk in our daily language but it should be. Efficacy and risk are essentially opposite sides of the same coin. Therefore, simultaneously stating the methodology for assessing both greatly assists in understanding their relationship and the assessment process.

Specific, simple definitions are needed for several words which are commonly used with widely different meanings.

- * Hazard- anything that has the potential of a bad result given an initiating event.
- * Salutory- anything that has the potential of a good result given an initiating event.
- * Consequence- the result of a bad thing happening.
- * Benefit- the result of a good thing happening.
- * Risk- the product of a consequence and its probability of occurrence summed for all possible hazards.
- * Efficacy- the product of a benefit and its probability of occurrence summed for all possible salutarities.

The sequential steps in the safety analysis and decision are:

1. Identify all hazards and salutarities.
2. Assess consequences and benefits.
3. Estimate the probabilities of occurrence.
4. Calculate the risk and efficacy.
5. Decide if the efficacy is sufficiently greater than the risk to make the proposed activity worthwhile.

The use of the mathematical terms product and summed in the above definitions implies that both consequences and benefits as well as their probabilities are quantitative terms. Ideally benefit and consequence are assigned dollar values so that a numerical efficacy/risk ratio can be calculated as a direct comparison. That ratio is the modern equivalent of the traditional benefit/cost ratio. The efficacy/risk ratio includes the probabilities for occurrence of the different initiating events and subject exposure whereas the traditional benefit/cost ratio assumes the events are certain. (The followers of Mr. Murphy are required to assume the costs have probabilities of one and the benefits have probabilities of zero; therefore, they are always against doing anything, including nothing.) Similar definitions of risk abound but most contain the same elements with different names. For example, McCormick, op. cit., defines risk as frequency times damage for each hazard and the sum for all hazards as composite risk.

In the ideal case the we will have both mean and variance values for the probabilities, benefits and consequences which can be combined into a mean and variance of the efficacy/risk ratio. If a statistical distribution is developed or assumed, a proper conclusion can be stated; for example, "we are 95 percent confident that the efficacy/risk ratio is greater than 1.2 for the proposed activity." The great conundrum of this whole process is the fact that the judgement as to whether the confidence level and the ratio are sufficiently high is an individual one. We will always be faced with the problem of obtaining concurrence on this judgement among different people in different government agencies and different citizen groups for any decision relating to clean-up of the Rocky Flats plant regardless of the safety analysis results.

The above definitions and descriptions are all given as if they are based on absolute values of benefits, consequences and their respective probabilities. But determining such absolute values is very difficult and expensive and really not necessary for decision making. Only relative values for the portions of the alternatives that are different are essential. Engineering cost estimating provides the best illustration I can think of for using relative values rather than absolute ones. Say we want to build a pickle factory and need to choose between two different pickling processes. The cost of all the equipment, construction, etc. for each process can be obtained provided the design is complete. These separate, absolute cost estimates would themselves cost a great deal of money. However, if we identify the differences between the two processes and obtain relative costs for those differences, then the choice between the two alternatives can be made with confidence, at least as far as cost is concerned. No doubt the owners will want an absolute cost estimate before proceeding with construction of the pickle factory but there will be a great deal more cost information with less uncertainty available after the design of the chosen process and its plant are complete. We can easily transform the concept of comparing cost differences to comparing risk differences. Moreover, we must recognize that alternative processes have different efficacies just as surely as they have different risks and include the efficacy differences in the decision process also.

Assigning a dollar value to each consequence or benefit is assailed by many as being too crass and materialistic. I believe that this resistance arises because people want to weight some of those items much more heavily than any reasonable dollar cost estimate can justify. Such desire must arise from factors such as voluntary vs. involuntary participation as described by Chauncy Star, Ref. 3, or from outrage elements as outlined by Peter Sandman, Ref. 4. Consequently, many partial definitions of efficacy and risk just employ the probabilities of occurrence and subject exposure. No attempt is made to incorporate the magnitude of the benefit and consequence values directly into a single measure such as the efficacy/risk ratio. That simplification is OK as long as only a single benefit or consequence is considered. But ultimately the relative value of the several benefits and consequences and their respective probabilities must somehow be combined in the decision process. Therefore, if we want a rational decision process which is broadly acceptable, we must incorporate some form of importance weighing in the definition of risk and efficacy.

III. DECISION PROCESS

A decision process which is rational and broadly acceptable must incorporate a number of features:

- * There must be two or more alternatives which meet some basic requirements in order to qualify for consideration.
- * There must be one or more attributes which are common to all alternatives.
- * These attributes must have relative importance values or weights which remain constant throughout the decision process.
- * A relative score must be assigned to each alternative on an attribute by attribute basis.
- * The alternative having the highest total weighted score will be the alternative desired.
- * The total weighted score isn't as important as the process of assigning the weights and scores.

The Kepner-Tregoe method, Ref. 5, or a similar matrix based decision process best embodies these characteristics.

Even if a efficacy/risk ratio, as defined above, has been calculated we still need to employ a matrix based decision process in order to include additional factors such as availability or the dollar cost of completing each alternative. There are likely several other important considerations which wouldn't normally be included in the safety analysis but should be included in the decision process as additional attributes.

Uncertainty in the quantified attributes is best accommodated in the decision process by expressing each at say a 95 percentile confidence level; they should all be expressed at the same confidence level whatever it is. The uncertainty of non-quantified attributes, say operability, is best accommodated by judgement in assigning the weighing factor for that attribute or the score for each alternative with respect to that attribute.

IV. DOCUMENT SPECIFIC COMMENTS

Section 1.0, 6th paragraph, p. 1-2: Why are the IRAM methodologies to be independent of CERCLA/RCRA risk assessments and not applicable to the individual or cumulative risks under the IAG? That sounds as if a lot of effort is going to be wasted.

Section 2.1, 1st paragraph, p. 2-3: The four listed risk assessment components correspond more or less to the four I used in part II of this review: site characterization and data evaluation to hazard identification; exposure assessment to probability estimation; toxicity assessment to consequence assessment; and risk characterization to risk calculation. This comparison illustrates the extreme breadth of the descriptions of risk assessment methods. However, since the last sentence in the paragraph omits any mention of the need to include efficacy in the decision process there are clearly differences in my method and the methods proposed in this document.

Section 2.2,p. 2-4: Shouldn't this section be expanded to include explosions, fire, falling equipment, etc. Although these hazards primarily impact workers, they shouldn't be neglected in an overall risk assessment.

Section 2.2.2.1, 1st paragraph, p. 2-6: Worst case estimates are difficult enough to interpret without being allowed to extend into an impossible range. The use of optimistic, most probable and pessimistic values or the use of mean values and error factors are much less likely to lead into an impossible range.

Section 2.2.4, 1st paragraph, p. 2-8: Something is wrong with the sentence starting "To characterize potential. . ." It seems to mix noncarcinogenic and carcinogenic statements but doesn't offer any hint as to their relative importance.

Section 2.2.4.1, last paragraph, p. 2-9: The document is quite correct in warning that the Hazard Index should not be used as a probability let alone a risk assessment. The HI is essentially a consequence quantification; perhaps suitable for additivity screening as described

later. Something needs to be done to get a probability of exposure above the Hazard Quotient for each chemical of concern. Surely there are a number of ways to estimate such probabilities or at least relative probabilities for the different chemicals.

Section 2.2.4, first paragraph, p. 2-10: The SF*CDI product is a good approximation to a true probability as long as the product is below 0.1; at least it is for an exponential probability density function. A dimensionless number is not necessarily a probability or a risk. Moreover, using the 95 percentile confidence limit value for SF as described ignores any uncertainty in the CDI value which is surely just as uncertain as the SF value. And using the product of 95 percentile values for both SF and CDI will not yield a 95 percentile value for the product. To a first approximation, mean values of the SF and CDI must be used to calculate the mean value of the probability. Then the mean value and variance of each term must be used to calculate the variance of the probability. Finally, the 95 percentile value for the probability can be determined for an assumed distribution. The process may appear tough but it is easy and can't be effectively short-cut.

Section 2.2.5, 2nd paragraph, p. 2-11: I certainly agree that collecting and analyzing the data necessary to develop absolutely correct probability distributions is probably too expensive. But there are ways to approximate probability distributions that need not be so expensive. What shouldn't be acceptable is to grossly oversimplify a risk assessment to satisfy some regulatory check-off and then use that gross simplification as a basis for significant decisions.

Section 2.4, 1st paragraph, p. 2-12: The lead sentence says that ecological risk assessment predicts the probability of adverse effects. Without any assessment of value, how can such ecological "risk" be compared to say human health risk or worker risk or anything else? We could chose between alternative processes on the basis of such ecological "risk" if it were the only attribute to be considered.

Section 3.1.1, p. 3-7: The Additivity Screen method described proposes the direct addition or summing of carcinogenic risks and noncarcinogenic hazard indexes. Those items were previously defined on p. 2-9 and 2-18 respectively. I believe typical values for such carcinogenic "risks" would range around 10^{-6} while typical hazard index values would range around 10^{-2} ; obviously the sum of such values would be completely dominated by the hazard indexes. This magnitude dominance is one reason that risk should be defined as the product of probability and consequence. Risk should have units; dimensionlessness is not always a desirable attribute for a measurement concept. When one dimensioned value is much larger than another with the same dimensions, we may neglect the smaller without concern. But that freedom cannot be extended to dimensionless quantities. I doubt if anyone can get a general agreement on a single risk dimension whether it be dollars, number of cancers per 100,000 human population, or fatalities per thousand golden eagles. Combining such diverse risks for a decision requires assigning a weight factor to each risk that, in effect, sets its relative value with respect to the other risks.

Section 3.1.2, p. 3-13: The Comprehensive (Site-wide) Risk Assessment procedure outlined here sounds good provided the risk values are comparable, i.e. not dimensionless numbers with completely different references. But summing mean values and "reasonable maximum values" as suggested in Section 3.1.1 will yield a result for which no one can assign a probability. I

want to reiterate that basing a decision on the relative risks between two alternatives implies that you are assuming the efficacies of each alternative are equal. Likewise, if you use only probabilities for risks or efficacies, you are assuming the respective consequences or benefits are equal.

Section 3.1.2.1, p. 3-15: I certainly support the idea of establishing a systematic evaluation of the parameter values and, consequently, the data on which those parameter values are based as an integral part of the risk assessment process. Whether the parameter values are based on measurements or opinion surveys made at the Rocky Flats plant or whether they are generic values from world-wide experience, they need to be warranted somehow.

This is a good place to suggest that a system also be established to maintain the risk analysis as a "living" process. As the remedial actions are implemented at Rocky Flats plant there will be many changes in the values used in the risk assessment, in the clean-up processes, in the attitudes of the decision makers, in every facet of the on-going activities. These changes need to be recorded and their combined effect on the risk assessment evaluated periodically. As the parameter values change, whether the change is to the mean or just to the variance, changes may be needed in the remedial action plans in order to maintain the same or a new level of overall risk.

Section 3.1.3, p. 3-25: The DARA model described is really just the simple matrix decision tool I mentioned previously. Labeling it as "computer based" is an unfair slur or maybe it is supposed to be a compliment. In any case, only the human decision maker can assign weighing factors and scores and write the rationale for those choices. Those activities are the important parts of the decision process not calculating a final score or other computer magic.

Section 3.1.3.1, p. 3-28: The discussion of weighing and scoring for the decision matrix in this section puts too much into the significance of the values themselves. I think the authors were trying to incorporate the missing consequences into them. The weighing factors should, in my opinion, reflect only the relative importance of each attribute (criterion) and therefore be fractional values adding to one for all attributes. The scores should be simple integers, say 1-5 or at most 1-10, which reflect the ranking of each alternative taken one attribute at a time. In addition the scores should always be assigned from the top down, that is the best alternative receives a 5 or a 10 score. That technique prevents the score values from making a defacto change in the weighing factors assigned to each attribute.

Section 3.1.4.1, p. 3-36: There has been quite a bit of research on canvassing experts, not "academic" experts but "field" experts, in relation to PERT analysis as described in Ref. 6. The preferred technique is to seek judgements of optimistic, pessimistic and most probable values. People seem to respond to questions on those values without feeling too "pressured" and their responses are good estimates of one in twenty events or 5 and 95 percentile values. But as stated previously, when calculations are made the parameters should be valued at their mean and the uncertainty propagated as a variance.

Section 3.1.4.3, p. 3-44: The calculation done here appears to AND four 95 percent confidence level probabilities of 0.05 and then take the one's complement of the result as a new confidence

level. That's so completely the wrong calculation to make that I can't decide just how to point out its errors. I will try to reconstruct what I think was the intent and do the appropriate calculations.

Assume the mean of each exposure event probability is say 0.030 and the standard deviation is 0.012, then 0.05 is close to the 95 percent confidence upper limit for a normal distribution.

We can OR the four mean probabilities together by using De Morgan's Theorem to get:

$$P = 1-(1-0.03)^4 = 0.1147$$

The result is the mean value of the probability of one or more exposures.

Apply the Taylor series approximation to calculate the variance of the probability:

$$V(P) = 4*((1-.03)^3)^2*(0.012)^2 = 4.798*10(-4)$$

The 4 is not squared because the calculation is for four statistically independent events not four identical events.

The standard deviation of the probability is the square root of the variance:

$$SD = (4.798*10(-4))^{1/2} = 0.022$$

The 95 percent confidence upper limit of the probability, assuming a normal distribution, is:

$$UL = 0.1147+1.645*0.022 = 0.151$$

A proper statement of the result is "we are 95 percent confident that the probability of one or more exposures is 0.151 or less.

If I have missed the whole point of what the authors intended by the calculation in this section of the report, I apologize now. However, my calculations demonstrate how the mean values and variances have to be combined to get a result at a desired confidence level.

III. CLOSING COMMENTS

I have run out of allotted time for review and comment and I have begun to repeat myself. However, I do want to close by saying that I believe the IRAM document does an excellent job of addressing the complications inherent in trying to develop a coherent overall risk assessment process to aid decision making about how to clean and restore the environment in and around the Rocky Flats plant. I obviously believe the risk assessment process needs to be approached much more carefully than just trying to shoe-horn the various risk definitions developed for much more narrowly focused activities into an integrated process for the Rocky Flats plant site restoration.

I want to finish by suggesting that DOD, DOE and EPA consider a special task force, similar to that assembled to produce the famous Rasmussen Report on nuclear reactor safety, Ref. 7, to address the myriad problems of an integrated risk assessment for a plant such as Rocky Flats. Such risk assessments will be, perhaps, more site specific than nuclear power plant PRAs but assembling a team of experts to work out the basic difficulties in a common, concerted effort would be a much more efficacious approach than that being taken now.

IV. REFERENCES

1. McCormick, N. J., Reliability and Risk Analysis, Chapters 11, and 16-18, Academic Press, 1981.
2. Lowrance, W. W., Of Acceptable Risk, William Kaufmann, 1976.
3. Starr, C., "Social Benefit Versus Technological Risk", Science, 165, p. 1232, 1965.
4. Sandman, P., Keynote Lecture to 1989 AIChE Meeting summarized in ChExtra, Chemical Engineering Progress, Nov., 1989.
5. Kepner, C. H. and Tregoe, B. B., The Rational Manager, McGraw-Hill, 1965.
6. Moder, J. J. and Phillips C. R., Project Management with CPM and PERT, 2nd Ed., Van Nostrand Reinhold, 1970.
7. Reactor Safety Study - An Assessment of Accident Risks in U. S. Commercial Nuclear Power Plants, WASH-1400, NUREG-75/014, 1975.

Reviewer C

Reviewer C Comments

1. GENERAL COMMENTS

This document is remarkable because in most cases, as I read the document, questions would come to mind, only to be answered as I read further. It is apparent that a great deal of work has gone into the preparation and review of this document. The authors and reviewers should be commended for a job well done.

1.1 Risk from Very Low Exposures

Although I have difficulty dealing with the enormously large uncertainties and often extremely small estimated risks that arise from the risk assessments, I have come to accept the need for risk assessment methodology. As documents such as this are prepared I believe it is important to stress not only the uncertainties but the fact that in many, if not all, cases the carcinogenic risk from low level exposures could very well be zero. This is mentioned in passing on page 2-10 but I believe it deserves more emphasis. The National Research Council-National Academy of Sciences has issued a series of reports over the years on the risks from low levels of ionizing radiation. These reports provide risk rates (Slope Factors in the EPA jargon) to be applied to for exposure to low levels of ionizing radiation. However the reports also qualify their numerical risk rate estimates and conclusions. Some examples:

"Moreover, epidemiologic data cannot rigorously exclude the existence of a threshold in the milliseivert (100's of mrem) range. Thus the possibility that there may be no risks from exposures comparable to external natural background radiation cannot be ruled out. At such low doses and dose rates it must be acknowledged that the lower limit of the range of uncertainty in the risk estimates extends to zero." (NAS, 1990)

"The Committee does not know whether dose rates of gamma or x rays of about 100 mrads/yr are detrimental to man. Any somatic effects at these dose rates would be masked by environmental or other factors that produce the same type of health effects as does ionizing radiation." (NAS, 1980)

"Hence expectations based on linear extrapolation from the known effects in man of larger doses delivered at higher dose rates in the range of rising dose-incidence relationship may well overestimate the risks of low-LET radiation at low dose rates and may, therefore be regarded as upper limits of risk for low-level, low-LET irradiation. The lower limit, depending on the shape of the dose-incidence curve for low-LET radiation and the efficiency of the repair process, could be appreciably smaller (the possibility of zero is not excluded by the data." (NAS, 1972)

I don't have sufficient familiarity with the chemical carcinogenicity literature to quickly find similar caveats but I suspect they exist. It is important for consumers of risk assessments to realize that the risk may well be zero. Unquestioning acceptance and propagation of the silly notion that there is "no safe level" of substances only serves to perpetuate inappropriate fear of technology.

1.2 Consistency

The document has some internal inconsistencies which, if cleared up, would improve the readability and utility of the work. Examples follow.

1.2.1 Slope Factor (SF)

On page 2-10 slope factor is described as

"...the upper 95% percentile confidence limit of the slope describing the probability of carcinogenic response versus the unit intake of a given carcinogen over a lifetime."

However if one reads page 3-20, one finds that

"...radionuclide SFs represent best estimate values whereas chemical SFs are usually based on the upper 95 percent confidence interval on the slope of the dose response curve for that chemical."

Should one look up slope factor in the definitions at the end of the document, one finds Slope Factor to be

"A plausible upper bound estimate of the response per unit intake of a chemical over a lifetime."

It is confusing when a term as central to risk assessment as Slope Factor is defined variously as "95% confidence level," "best estimate" and "plausible upper bound."

1.2.2 DARA

Another item related to consistency is the term Decision Analysis/Risk Tree (DARA). This approach is central to the recommended environmental IRAM but on page 3-25 and in the list of acronyms and definitions, DARA is "Computer based approach to evaluation of human health risk."

1.2.3 Relationship to CERCLA

In the introduction (1-1 and following) it is made clear that the IRAM is not intended to satisfy any CERCLA requirements. However in Section 4, it is pointed out that the recommended IRAMs, and in particular the recommended Human Health IRAM, is consistent with CERCLA guidelines. Perhaps the interrelationships among the CERCLA, RCRA, NEPA, and the IRAM could be explained more clearly.

1.2.4 Proposed Rule

On page 3-94 reference is made to "this proposed rule," but I could find no prior mention of a proposed rule. This seems like a very important concept because OU's would be excused from NEPA compliance if it were adopted. What is the proposed rule?

1.3 Time Period

The time period over which the risks are to be estimated is not stated. Perhaps this omission is deliberate but a discussion of time periods would seem to be crucial to the usefulness of the document.

2. SPECIFIC COMMENTS

Page Comment

- 1-2 I suspect that the ERP also looks at waste disposal sites with a potential to release to the environment, not just the sites which are releasing.
- 1-2 Define OU (first time used).
- 2-1 The HEAST tables would be another useful reference here.
- 2-2 "instructional" should probably be "institutional."
- 2-4 The description of chemicals of concern seems naive and incomplete. Why the concern for "compounds," but not elements? What about toxic elements such as mercury and arsenic. Why the concern for "highly toxic chemicals found in low concentrations?" Why not be concerned about high concentrations of highly toxic chemicals? Why should there be concern for "tentatively identified chemicals that are historically associated with the site" unless they are also hazardous?
- 2-7 Add "for" between information and COCs.
- 2-8 The term "weight-of-evidence determination" should be explained.
- 2-11 I believe that radiation slope factors are published in the HEAST tables as well as the IRIS database.
- 2-12 Why should it be assumed that assumptions used to estimate chemical intakes and risks may be inconsistently applied? Isn't the ERP under a single management scheme?
- 3-85 combined instead of combine
- 3-85 I doubt if one could find "implementability" in Webster.

Page Comment

- 3-86 set of rather than site of.
- 3-95 In refining the list of key receptors, direct contact with the U.S. Fish and Wildlife service is recommended in addition to reviewing the indicated information.
- 3-96 Are the key receptors referred to in the text the same as the keystone species mentioned in the referenced table?
- 7 The acronym list is incomplete. For example RPM (page 4-5), OM and RFV (Table 3-10) are not included in the list. Also the definitions in the list do not always match the definitions in the text (see section 1.2.1 in this review).

REFERENCES

NAS (1990). Committee on the Biological Effects of Ionizing Radiation. Health effects of exposure to low levels of ionizing radiation. National Academy Press, Washington, D.C.

NAS (1980). Committee on the Biological Effects of Ionizing Radiation. The effects on populations of exposure to low levels of ionizing radiation. National Academy Press, Washington, D.C.

NAS (1972). Advisory Committee on the Biological Effects of Ionizing Radiation. The effects on populations of exposure to low levels of ionizing radiation. National Academy of Sciences, Washington, D.C.

Reviewer D

Reviewer D Comments

GENERAL COMMENTS

This review presents the viewpoint of one whose relevant experience is in atmospheric transport, diffusion, and deposition, one of the required areas of expertise specified in the guidance for the review of this report. The reviewer has had experience evaluating the results of risk assessment efforts, but has not participated in their formulation or execution.

This report told me more about incremental risk assessment methodology (IRAM) than I wanted to know. The encyclopedic coverage of the regulatory guidance and the available methodology, coupled with the high density of acronyms and abbreviations, had a soporific effect. It was difficult to maintain a clear view of direction and purpose while grinding through the report. Nevertheless, the coverage of the regulatory guidance and alternative methodologies for risk assessment is thorough and detailed. The recommendations for the appropriate methodologies to be applied to the human health and environmental IRAM's are well presented in outline and in detail.

My primary reaction to the report as a whole is the ascendancy of detail over focus. The authors are obviously thoroughly familiar with the subject, and the report is written for readers who are equally familiar. Each option for the human health and environmental IRAMs is presented in detail, and its advantages and disadvantages are listed. However, as an example, this reviewer would have appreciated a chart summarizing these points on a single page or fold-out. By the time we got to the Recommended IRAM (Section 3.3), it was difficult to recall the previously listed advantages and disadvantages for the risk assessment methodologies recommended at each step.

To one who has not worked directly in this field, the proposed combination of pieces of different methodologies into the final recommended IRAM's seems surprising. Are the stated combinations of approaches clearly understandable to one skilled in this art? Could execution of these mixed methodologies by different operators give substantially different results?

Aside from these comments, we will leave the discussion of risk assessment methodologies to reviewers who are more familiar with the subject. Our largest concern is with the adequacy of existing methods for selection of receptors and exposure pathways (Section 3.0).

Many of the essential data elements or mechanisms required for assessment of risks from such spatially and temporally diverse activities as those at the RFP are simply not yet understood. Example: wind-related suspension and transport of soil particles containing radioactive materials, and their subsequent inhalation by humans. How can we assess uncertainty when we are not even sure we know the basic physics?

There are basically two ways to characterize the exposure of a receptor at a given point to a threat originating in a given operational unit: by direct measurement (air or water sampling networks), and by modelling. The variance in estimates of exposure obtained using either approach may be large.

In a recent consortium review of an air modeling effort for the RFP, concerning transport and deposition of wind-generated dust containing radioactive particles, it was shown that the state of the science for re-suspension of plutonium-carrying soil particles by the wind is totally inadequate for even order-of-magnitude estimates of the quantity of dust generated. Current models of transport, diffusion, and re-deposition of dust have comparable problems, especially when applied over complex terrain such as that around the RFP. A risk assessment based on model results will likely be a classic exercise in GIGO (garbage in, garbage out). We suspect that comparable shortcomings exist for models of ground-water transport.

At the RFP, the data from air sampling networks have large error bars because measured levels of radioactive substances are generally near the detection limits. Nevertheless, they give a meaningful measure of uncertainty which can be used for risk assessments. This is much preferable to the use of model results, where we are not even sure that we have the physics right.

Reviewer E

Reviewer E Comments

I. GENERAL COMMENTS

A. What is the Basis for Selection of Risk Characterization Methodologies?

The IRAM document does an excellent job of reviewing and evaluating selective procedures and general guidelines for assessing cumulative risk. There is only methodological discussion; sources, chemicals pathways, and receptors specific to Rocky Flats Plant are not identified. Four approaches for Human Health Risk Characterization (Additivity, Comprehensive, DARA, and Probabilistic) and three Environmental Risk Characterization methods (Additivity, Cumulative, and DARA) were identified as suitable for more detailed evaluation as part of Task 720.

The review of these seven approaches was quite thorough and rigorous. However, there is a potential for criticism of the selection methodology of the seven approaches which were deemed suitable for review. There is no indication of the process for selection of these seven "suitable" methods. There is no list of other, suitable or unsuitable methods which were considered, but rejected for discussion. The basis for inclusion in the list of seven is not at all clear, nor is the basis for exclusion from the list. For example, why is the probabilistic approach not suitable for a more detailed evaluation under Environmental IRAMs?

If this document is to be truly defensible, surely some discussion is appropriate on why these seven are the most suitable approaches for assessing cumulative risk. The selection criteria are given, but it is left to the reader's imagination to determine why these are the best approaches for Rocky Flats Plant.

B. Lack of Specificity in Site Characterization Procedures

Section 2.1 of the IRAM document states that Risk Assessment has the four components:

1. site characterization and data evaluation
2. exposure assessment
3. toxicity assessment
4. risk characterization

The IRAM document does not treat these four areas with equal specificity, particularly with regard to site characterization. For example, the IRAM makes references to EPA guidance documents for site characterization procedures, and in Section 2.2.1 the authors generically discuss site characterization. In Section 2.2.1, data validation is discussed, but in Section 3 there are no recommendations made for data validation, or for that matter, little if any quality assurance procedures specific to site characterization at the Rocky Flats Plant.

Perhaps at issue are several of the underlying assumptions made by the authors of the IRAM document in section 3.0, pages 3-2 and 3-3. It was assumed by the authors that

the risk assessments for each OU would be performed under CERCLA, and further, it was assumed that site characterization data and corresponding risk information would be available for each OU. These assumptions reduce the need to recommend, in the IRAM document, any specific site characterization or data methodologies, assuming this would be carried out under CERCLA guidance.

The IRAM document is meant to identify potential methodologies for assessing incremental, cumulative risk to support NEPA requirements. NEPA was intended to be innovative and flexible, and attendant guidance for NEPA even allows for consideration of alternative actions which could be outside current regulatory guidelines, (including CERCLA). It is inappropriate to make a blind assumption that the risk information developed under CERCLA for each OU will be reliable, available, compatible, able to be integrated with other data, and appropriate to support decisions pertaining to NEPA analysis for evaluation of potential cumulative impacts at the Rocky Flats Plant.

The one thing that could make or break the proposed risk methodologies in the IRAM document is the goodness and timeliness of data supplied to them. Specific site characterization strategies, and data evaluation procedures which support them, should be specified in the IRAM document.

C. IRAM assumptions on the nature of "Changes" in risk attributable to remediation of individual OUs.

One of the stated criteria of used to identify favorable method of risk evaluation the potential for the method to "Assess changes in risks as OUs are remediated" (page 3-1). At several points in the text, however this is restated as "assess incremental reductions in risk due to phased implementation of remedial actions."

It is important to recognize from the outset that remedial actions to reduce risk at one OU have the potential to increase risk from other OUs. This means that the term "incremental reductions" is the hoped for case, but not the only possibility. Some remedial techniques could significantly change the pathway component of risk for some chemicals with time. One example might be changing pH or redox conditions (during vapor extraction and/or air sparging) which could speciate some metals from any nearby OU into a much more mobile fraction. Remedial techniques using surfactants to flush soils could mobilize contaminants. Likewise groundwater pumping or drain systems (some drain systems have already been installed at RFP) could dramatically effect contaminant pathways, and therefore, risk.

Risk could vary with time for any OU. This is particularly true at Rocky Flats, where mixtures of contaminants could create cosolvent effects, daughter products of varying toxicity could be created through radioactive decay or biotransformation, and other changes could occur.

It is crucial that, for the IRAM to be useful, it must be implemented correctly. The proposed risk assessment methodologies should not blindly assume that across the board

reduction of risk will occur from remediation of a single OU. The difference between the model (theory) and the field realities of site remediation must be incorporated into the IRAM.

- D. Field validation of the proposed IRAMs is essential, but is not currently a specific part of the proposed IRAMs. It should be.
- E. Concern over implementation of the IRAMs

This reviewer believes that the proposed IRAMs are acceptable, but is greatly concerned over the potentially overwhelming problems with their implementation. There are several potential difficulties with the implementation of remedial action under the guidance and constraints of the proposed models.

1. Uncertainty. Actual field uncertainties might be huge for OU/chemical/pathway/receptor risk combinations. Essentially, any true incorporation of real world uncertainties might render any estimate of the potential for change in risk to be meaningless. There is an underlying assumption in the IRAM document that there is a good degree of confidence in the ability to quantify the efficacy of any given remedial approach. This may not be true.

For example, the very science of contaminant transport is changing as more is learned about transport processes. Recently, it has been shown that sorption of some contaminants onto soils is hysteretic; that a recalcitrant phase may hang on to soils for a long time, rendering many current contaminant transport models inaccurate. It may be impossible to incorporate OU-specific uncertainties into the proposed IRAM.

2. Need for iteration. Interim remedial action or temporally staggered cleanup can change the pathway component of risk analysis significantly for other, non-remediated OUs. The timing of cleanups at the Rocky Flats Plant is likely to conflict with site characterization schedules. Decision makers will be faced with the choice of whether or not to begin remediation without complete site characterization and associated risk information from all the OUs. Indeed, even some interim site remediation has already been undertaken at RFP. Will the IRAM be able to keep up with all the changes?
3. The proposed IRAMs are meant to assist in risk management decision making. Part of that decision making process is achieving cost effective solutions and balancing those solutions with expected risk reduction. The cost of risk reduction is not considered in the proposed methods.
4. The scope of risk reduction is also not well addressed. For example, the contaminant mass might be shipped to Beatty, Nevada for disposal, but will the risks in shipping and the risks to receptors in Nevada be considered?

5. It may be extremely difficult to choose model tolerances, and endpoints (acceptable risk) involved.
6. The time frame involved for risk estimates is somewhat unclear and not defined in the IRAM. NEPA requires that both short and long term considerations be appraised. Risk appraisal under long time frames, such as the 10,000 years time frame associated with many radionuclides, would pose a difficult constraint on any risk methodology, including the proposed IRAMs.

II. SPECIFIC COMMENTS

A. Problems with the screening methodology.

Page 3-88 and 3-91. While the need for screening is recognized as a measure to keep the total OU/chemical/pathway/receptor combinations reduced to a workable number, I am concerned that proposed "hybrid" IRAM human health risk assessment model proposes to screen out "insignificant" contributors to cumulative risk (step 4) before uncertainty is quantitated (steps 7 and 8). The removal of these "insignificant" contributors from a master list of OU/chemical/pathway/receptor combinations presupposes not only that the combinations have a level of acceptable uncertainty, but that the risk associated with these combinations will be time invariant.

Exacerbating this problem is a mistake in the text. On page 3-86 the text reads that steps 4 through 9 will be repeated, but on page 3-94 the text reads that steps 5 through 9 will be repeated. Will the screen (Step 4) be repeated? The merits of iteration are supported earlier in the document, but does your iteration go this far (to step 4)?

If great uncertainty exist in a OU/chemical/pathway/receptor combination to be screened, under the proposed system, it will not be defined and is potentially nonconservative to delete that combination from the master list. Additionally, some remedial techniques could significantly change the pathway component of risk for some chemicals with time. One example might be changing pH or redox conditions (during vapor extraction and/or air sparging) which could speciate some metals from any nearby OU into a much more mobile fraction. Likewise groundwater pumping or drain systems could dramatically effect pathway, and therefore, risk.

B. Typographical Errors

1. In the Table of Contents page 3-94 should be changed to 3-84
2. The title which appears on page 3-84 is incomplete and should be changed to the same title that appears in the Table of Contents.
3. Page 2-8, section 2.2.4. The 4th sentence should be 2 sentences, after the word "values".
4. Page 3-86, last sentence "set" not "site of sitewide receptors"
5. Page 3-88, first paragraph, second sentence. Capitalize first word.
6. Page 3-91, paragraph two, sentence two. Capitalize first word.

Reviewer F

Reviewer F Comments

GENERAL COMMENTS

The stated purpose of this document was to evaluate procedures and guidelines for assessing cumulative human health and environmental risks at the Rocky Flats Plant (RFP). The document will recommend a method for assessing the reduction in cumulative risks (=combined site wide risks from all OUs) following remediation of the operational units (OUs).

I feel that the authors have done a pretty good job of reviewing relevant approaches and methodologies for performing risk assessments. Discussion of the advantages and disadvantages of each approach was especially useful. I feel that one of the primary strengths of this document is that it provides a procedure for estimating cumulative risk from each operational unit (OU) at RFP in order to estimate total risk.

My strongest criticism of this document is that it is very general, with little specific information pertaining directly to RFP. To be fair, the document states that the scope of the current project did not include actual development of the recommended IRAM to the point of implementation. One must assume that a more detailed plan will be forthcoming. Most of the information contained in the document was obtained from recent EPA publications on risk assessment and the document offers little more than a review and summary of current EPA guidelines on these procedures. As a consequence, review and critique of this document becomes a critique of the current EPA protocols. I feel that the only unique aspect of this document is the recommended "hybrid" IRAM, which consolidates aspects of each of the methodologies reviewed. A large portion of this document is simply "boilerplate" material, which could be used to develop risk assessments for almost any hazardous waste site.

My second general comment pertains to the overall lack of emphasis on measuring (not modelling!) contaminant transport at RFP. I am a critic of most "armchair" risk assessments, which rely exclusively on literature values and computer models to predict risks to unique, complex ecosystems. While developing conceptual models of exposure pathways and predicting contaminant concentrations in ecosystem compartments may be completed through literature reviews and appropriate software, I feel that this approach is useless without field validation. This important step deserves additional consideration in the document. Once potential contaminant exposure pathways have been identified, I would like to see a plan that validates predictions of these exposure models.

Third, I feel that the document should give more consideration to the use of biomarkers for assessing contaminant exposure. There have been numerous recent publications describing the use of biomarkers to assess exposure to both specific contaminants as well as general classes of compounds. Although these methods cannot be used to measure effects directly, they are quite useful for determining if exposure has occurred.

SPECIFIC COMMENTS

Page

1. 2-14 The suggestion that population endpoints are superior to community and ecosystem measures of stress is not supported. There exists a large body of literature indicating that responses at higher levels of biological organization (community) are at least as good as population responses. At the very least, a balanced approach that integrates endpoints at each level of organization is necessary.
2. 2-17 The document notes that there is little EPA guidance for determining risk of simultaneous exposures. I agree that evaluating multiple sources of exposure represents the major challenge at RFP. The statement "if two pathways do not affect the same individual or subpopulation, neither pathway's individual risk estimate or hazard index affects the other..." requires some qualification. With respect to ecosystem responses, risks to one population certainly have great potential to affect other populations.
3. 2-20 Although treatment of each of the OU's as discrete units may be practical from a management perspective, most species at RFP ignore these artificial boundaries and readily move between areas. Assessing risk to these highly mobile organisms (e.g., mule deer, raptors) will require that each species' habitat, home range, foraging area, etc. be included in the evaluation. I strongly agree with the idea expressed in the document (pp. 2-22) that assessment of sitewide risks will require detailed analysis of spatial and temporal patterns of habitat utilization.
4. 3-4 The selection of receptors should be more specific. The document states that the "objective of this stage of the IRAM is to identify receptors most likely to receive the greatest degree of exposure..". I agree and was expecting to find details of exposure pathways and receptors. I was disappointed to find only general lists of potential receptors, that are found in recently published EPA guidelines. Certainly there is enough information on the general ecology of this system that at least a conceptual model and a list of likely receptors could be provided. With a little additional research the authors could provide a more specific list of species and endpoints.
5. 3-59 The document notes the necessity of distinguishing effects of chemicals of concern (COC's) from "other effects, such as competition, or natural community variation." This will be one of the major obstacles for developing a useful risk assessment at RFP. Evaluating potential effects of COC's on populations, communities and ecosystems at RFP will require 1) adequate characterization of background or reference conditions; and 2) isolating effects of COC's from other anthropogenic impacts (e.g., physical disturbance).
6. 3-59 The document offers little insight as to how effects of physical disturbance, which are very important at RFP, will be separated from effects of COC's. Physical disturbance at RFP will also complicate the selection of potential reference sites.

I question whether the "natural state" of this system can be determined and recommend that the authors consider using both on-site and off-site reference areas to characterize this system.

7. 3-62 The "additivity screen" approach is at best a limited screening method. If neither the amount of time an organism spends at an OU nor the concentration of COC's at the OU is included in the approach, I doubt that it can be useful even for screening.
8. 3-73 Of the three methods reviewed, I feel that the "cumulative risk analysis methodology is the best approach for estimating environmental risk at RFP. The major advantage of this approach is the establishment of spatial and temporal boundaries based on natural history of the specific receptors at RFP.
9. 3-74 I agree that key receptors should be selected based on a sitewide basis. This will greatly simplify comparison among OU's. The list of criteria for selection of these key receptors is reasonable.
10. 3-75 The statement that organisms at the top of the food chain are at greater risk to contaminants that bioaccumulate is true for only a limited number of materials (e.g., organochlorines, PCB's). First, almost any contaminant may bioaccumulate, while only a few will biomagnify. Even classic examples of biomagnification (DDT, dieldrin) are currently being reexamined in light of recent research on the importance of lipids, age, size, etc. For several reasons, top predators are a poor choice for receptor organism. I argue that because of their higher mobility it would be more difficult to assess and compare effects of individual OU's on these organisms.
11. 3-75 Although there is no specific definition of ecosystem health, Rapport et al. (American Naturalist, 1984, 125:617-640) have provided a general list of ecosystem responses to perturbations that are common among several ecosystem types. The most important consideration, as noted in the document, is comparison to a suitable reference area.

More importantly, assessment of ecosystem indicators such as nutrient cycling and primary productivity are problematic because of high variability and functional redundancy. I am in favor of approaches that consider community structure (e.g., species diversity, abundance, community composition) over these functional endpoints. Several researchers have demonstrated that these approaches are usually less variable and more sensitive to perturbation.

12. 3-80 The document correctly states that no single approach will be appropriate for assessing ecological risk at RFP. I feel that this point deserves emphasis. The limitations of using a single approach, especially in situations where perturbations are subtle. An integrated approach, using multiple receptors at several levels of biological organization, will be essential for successful completion of a risk assessment at RFP.

COMMENTS ON RECOMMENDED IRAM

1. 3-94 I agree that a "hybrid" IRAM methodology is appropriate at RFP. However, I doubt that the Additivity Screen, because of its inherent weaknesses noted above (comment # 7), will be of much use in providing even a preliminary screening tool.
2. 3-95 The DARA methodology would be more appropriate for initial screening of key receptors.
3. 3-95 Despite some of the limitations of the Cumulative Method listed in section 3.2.3.3, I feel that this approach will be especially useful at RFP, particularly after preliminary screening of receptors and pathways using the DARA model.
4. 3-95 Selecting key receptors is a critical step in this process. If key receptors are missing or are at low abundances at RFP because of previous operations, how will this be assessed? Again, it seems that examination of off-site references areas will be necessary to identify potential key receptors.
5. 3-98 I feel strongly that Exposure-Pathways Analysis should be an integral component of the IRAM at RFP. As stated above, contaminant transport from abiotic compartments to various biotic compartments must be incorporated into decisions regarding potential remediation.
6. 3-98 The observation that population responses are an inadequate surrogate for community and ecosystem responses to contaminants is quite true. Regulatory decisions based on simplistic and inexpensive approaches, such as water quality criteria or single species thresholds, risk being grossly over- or underprotective.
7. 3-99 I strongly support the use of natural communities for assessing biological integrity and ecosystem health. This approach has numerous advantages over individual or population responses.
8. 3-99 Despite sources of variability, direct measurement of contaminants in different biological compartments is the best way to assess bioavailability and potential transfer of contaminants in natural systems. This approach should be an integral component of the RFP IRAM.

DISTRIBUTION

Team Leader, Tal Neill (1)
University of Denver (2)
Colorado State University (2)
Idaho State University (2)
University of Nevada, Las Vegas (2)
EG&G Rocky Flats Plant:
 Lou Medal (4)
 Dan Shane (1)
 Records (1)
File (1)