MANAGING ACCEPTABLE RISK RESEARCH:
WITH APPLICATION TO NUCLEAR WASTE[1]

Rex V. Brown
George Mason University,
rbrown@gmu.edu

Abstract

An approach to cost-effectively managing large-scale risk research for regulatory purposes is proposed. It includes a regulatory test of acceptable risk involving assessment uncertainty (2nd order probability), and corresponding research strategies for the applicant and others, using probabilistic goal allocation. It seeks to combine logic with practical and institutional realism, based on past consulting to NRC and DOE. Nuclear waste disposal, where $5 billion may have been spent ineffectively, provides illustration.

[1] This paper was supported by NSF/DRMS under grant ISI-8420732. Author thanks Lee Merkhofer, Jean Younker, Mike Voegele, Norm Eisenberg, Tom Isaacs and Vicki Bier, for helpful information and comment, but they are in no way responsible for his views

1 INTRODUCTION

When society decides whether to permit the operation of a potentially hazardous facility, such as a dam, reactor, pipeline or chemical plant, the risk involved is commonly evaluated by regulations that specify acceptable risk and how to verify it. Public concern may be enough to justify major risk assessment before a decision. The research effort needs to be in keeping with the social stakes involved and be efficiently spent. Sometimes great expense may be spent without commensurate progress toward protecting the public. Nuclear waste disposal may be such a case

1.1 Need for new approaches

1.1.1 A cautionary tale: nuclear waste

Nowhere is the research management issue more important to resolve than in the disposal of high-level nuclear waste, which is highly controversial and liable to searching political, regulatory, and ultimately judicial, review. By 2001, DOE had spent over $5 Billion on identifying a suitable site, notably at Yucca Mountain (YM) in Nevada.

Critics have charged that much of that effort has been misdirected (National Research Council 1990: NAS 2001) and much of the cost wasted (Keeney and von Winterfeldt 1994). I am concerned here specifically with the management of the risk assessment effort, by itself accounting for some five billion dollars. My views are based on over ten years experience as a consultant to both regulator (NRC) and applicant (DOE), early in the planning process. [2]

1.1.2 Other Applications

The YM project may be past help, but others of comparable scale could benefit from a viable methodology. For example, Congress has considered building a satellite solar power system that could cost up to $1 trillion (sic) over 20 years (Chinnis et al. 1986). Before it were actively pursued, massive research on the risk of a catastrophic risk accident would be called for.

Again, hundreds of millions of dollars a year are spent probabilistic risk assessments for regulatory purposes on some 100 US nuclear plants. A senior NRC regulator told me he made little use of these assessments when deciding whether to warn or close down reactors, partly because he did not find the research effort to be directed effectively at enhancing his judgment.

[2] From 1982 to 1992, I served variously as: consultant on nuclear waste rulemaking the Office of Nuclear Material Safety and Safeguards of NRC; as risk management advisor to the head of in the Office of Nuclear Reactor Regulation of NRC; and as peer reviewer, Site Characterization Plan Oversight Committee member and methodology advisor to successive heads of the Office of Radioactive Civilian Waste Management of DOE.

1.2 Scope of method

1.2.1 Objective

My hope is to contribute to applied decision aiding, rather than to science, other than to generate scientific ideas with practical value. A methodology for managing large risk research projects is to be developed, specifically a defensible rationale for allocating resources among often thousands of potential research activities.

The methodology should be useful in real cases, typified by YM, taking account of institutional, practical, cognitive and situational requirements. This requires *minimal treatment of many issues*, rather than technical closure on any single issue, and be broad enough to cover whatever it takes to produce usable methodology. Attention is focused on practical research-planning problem, but it is necessarily technically diffuse. The mixture should be rich, with its ingredients only refined enough to be digestible.

Any models should be realistic, without simplifying assumptions for tractability, and grounded in applied decision theory

1.2.2 Tasks to be aided

The primary client is the "applicant" (DOE) in the role of research planner, taking into account the chain of command above and below. Above, he (DOE) must comply with regulations (of NRC), which must conform to standards (from EPA) and to the needs of society (represented by "watchdog" bodies). Below, he directs research organizations (for-profit contractors). Each party in the chain has its own distinctive needs.

Applications are primarily risk regulation projects with the following features (typified by YM):

- Acceptable risk is a major test of compliance
- There is one major risk variable (e.g. cumulative release) and several lesser requirements.
- Stakes at risk and risk assessment effort are large
- Risk assessment is performed by the applicant himself, confirmed by a regulator
- Research activities are subcontracted to commercial firms

1.3 Background on nuclear waste case

In 1982, Congress directed DOE to identify a geologic nuclear waste site that would meet safety requirements, and to do the research necessary to assure that the requirements were met (NWPA, 1982). In 1987, DOE was directed to evaluate YM specifically (NWPA 1987)

Research was conducted over nearly two decades, at a rate of some $400 million a year, to determine a suitable site. (In parallel, research has been proceeding on a container and other engineered features of a repository). As of 2001, that process was drawing to a close, and DOE was preparing to apply to NRC for a license to build a repository at Yucca Mountain.

1.3.1 Setting at the time.

Much like business and law school case studies, I will describe the problem at a particular time[3], the late 80s, during early research planning, and consider how it could have been addressed. I will advance a generalizable approach to this case.

The plan then was for DOE's Office of Civilian Radioactive Waste Management (OCRWM) to gather data over a ten-year period, on the basis of which it would submit to NRC a license application to construct a repository at YM, if findings were favorable. NRC's approval would be contingent on demonstrating compliance with its regulations (10CFR60), derived from more general standards for radioactive waste disposal issued by EPA (40CFR191). NRC's determination would be subject to approval by Congress.

1.3.2 Regulatory regime

Regulations then specified a number of long-term isolation and other "safety performance" requirements for nuclear waste disposal. EPA's general standard for radioactive waste disposal (40CFR191) stated: "(There shall be) reasonable expectation based on performance assessment that the cumulative releases to the accessible environment for 10.000 years …shall have a likelihood of less than one chance in ten of exceeding (specified) quantities. … These estimates shall be incorporated into an overall probability of cumulative release to the extent practicable."

NRC made requirements for additional risk variables, for example dealing with "pre-closure" risks and certain determinants of the long-term risk (such as groundwater travel time). Specific research activities (such as on-site investigations) were prescribed, to provide "reasonable assurance". I am only concerned with site performance requirements (not engineered features).

1.3.3 Risk research program

DOE was responsible for conducting research onYM site performance. The site characterization budget was several $100 million per year, charged to the power industry. DOE's research planning depended, not only on its own discretion, but also on periodic direction from NRC and advice from a presidentially appointed Technical Review Board (TRB).

[Fig 1: HIERARCHY…]

DOE's initial site characterization plan (DOE 1985) identified a set of research issues, corresponding to part of a hierarchy of risk determinants, along the lines of figure 1. Research on these issues was subcontracted to specialized commercial engineering firms and a research organization (USGS). In principle, budget allocation among them was to be informed by analytic exercises, such as system performance assessment, probabilistic goal allocation and acceptance testing. However, it is not clear that these exercises had much impact on the research program (NWTRB 1998). I hope to develop an approach that is both usable and used.

---

[3] I have taken minor liberties with historical accuracy for clarity of exposition

1.3.4 Later case developments

After the period of this case, there were several developments. The timeline for completing the site approval process[4] doubled from 10 to 20 years and the funding more than doubled, from $2B to $5B. The original EPA and NRC regulations were put ”in remand”, while changes (such as replacing release by dosage) were being considered[5].

1.4 Structure of paper

Section 2, characterizes and diagnoses what ails common risk research practice. Section 3 proposes new regulatory ground rules in the form of a modification to EPA standards. Section 4 proposes procedures to help the applicant meet these requirements. Section 5 suggests how the distinct interests of the society can be safeguarded. Section 6 discusses general issues and conclusions.

2 CAUSES OF WASTED RESEARCH RESOURCES

Resources may be wasted if those who control them are not rewarded by the public interest, and if there is nothing to oblige them to be cost-effective. Wastage is often due to three related impediments:

- Conflicts of interest
- Regulatory requirements
- Lack of control

2.1 Conflicts of interest

It is common practice for risk research responsibility to be given to the regulated party itself (e.g. power companies assess risk at their own nuclear plants), presumably to save public resources. There is obvious scope for conflict of interest that harms both the cost and effectiveness of the research.

2.1.1 Misdirected research

If the applicant wishes to get his project (the waste site) approved, neither he nor his research contractors have an interest in aggressively seeking out evidence of unacceptable risk, unless obliged. The misdirection may also stem from researchers' having priorities beyond the public interest. I observed that the YM effort was guided more by how it resolved technical or scientific issues, than by decision needs[6].

2.1.2 Over-spending

Overspending is likely if the cost does not come out of the applicant's budget. YM research was funded by a fixed charge on power used by consumers. This does not

[4] Reported at www.ymp.gov.
[5] See 40CFR197, 10CFR63. The changes are not directly relevant to historical case analysis. One of them, however, would have simplified application of the propose methodology—changing EPA's hazard control parameter from 90 percentile to expectation (see below).
[6] A USGS source acknowledged that DOE had agreed to research projects that claimed no relevance to the repository but were of scientific interest to the Survey.

motivate DOE to economize, nor to finish early. Indeed, DOE's budget was invariably fully used ("use it or lose it"), and the ten years originally planned more than doubled.

Conversely, in the more common case where if applicants pay for evaluation research themselves, the pressure is to economize *too much*, without regard for quality[7].

2.1.3 Misallocation

Commercial research contractors have a financial interest in making their pieces of the effort as large as possible (as well as in keeping the client happy). They may resist reallocation of resources to another contractor, in response to evolving evidence on which issues are most critical (e.g. at YM, on radioactive retardation vs. gaseous release). Indeed the rare reallocations seemed to depend more on bureaucratic politics than on public interest.

2.2 Ill-formulated regulatory requirements

Conflicts of interest can most readily be indulged if demonstrating compliance with regulatory requirements allows the applicant great discretion or inappropriate.

2.2.1 Common 1st order AR requirements

It is becoming common for regulation to specify an acceptable probability of mishap. This is unquestionably a major advance over previous deterministic requirements.[8] . Though reasonable as one objective of research, it does not specify when enough research has been done. Though there may be (as for YM), a vague requirement of "reasonable assurance", it does not specify how "firm" the assessed probability should be

Sometimes regulation addresses this issue by prescribing minimal research activity to be done. In the YM case this includes a broadly specified program of on-site exploration and testing for "potentially adverse and disqualifying conditions" (PACDs) (10CFR960).

2.2.2 Problems with 1st order requirement

However within the discretion allowed by these directives, there can be a serious motivational problem. If a site is already in first order compliance, based on preliminary evidence (commonly the case), there is no assurance that the site will *still* pass after prescribed additional research. This test gives no motivation to the interested parties to "look for bad news" (i.e. to vigorously search out PACDs). Moreover, the research that is prescribed in regulation has to be based on outdated knowledge available when the regulation was set, and so cannot be too specific or limiting.

---

[7] After an unsuccessful bid for an FAA risk assessment contract, the government client told me: "Executive order 12291 requires us to do a risk assessment before we can go ahead on this project". The cost comes out of my budget and yours was not the cheapest proposal."

[8] Notably the much vilified Delaney Amendment, which demanded zero carcinogenic effect from foodstuffs.

2.3 Lack of effective control procedures

Lack of a responsible, enforceable and reviewable rationale makes it easier for special interests to dominate the research management process. In the YM case, few procedural controls were in place to guide and discipline the allocation of research resources. Available management science theory was not developed into methodology that was used in this case. A review board (NWTRB 1998) recently reported that comprehensive risk assessment, for example, appears to have played little role in guiding the DOE effort.

2.4 Possible remedies

Three promising avenues to mitigate these problems are:

- for the regulators to specify acceptable risk more appropriately
- for the applicant to devise a procedure to help him comply
- for another party to enforce cost-effectiveness in the public interest

3 PROPOSED TEST OF ACCEPTABLE RISK

Environmental regulation commonly requires that hazardous activity can be demonstrated to have "acceptable risk" (AR). Other questionable areas of regulatory policy not relevant to research management methodology will not be addressed here.[9].

3.1 Is acceptable risk (AR) an appropriate test?

Decision analysts often advise against "acceptable risk" as a basis for regulation, since there are criteria other than safety (like economics), and the appropriate degree of risk thus varies with context (Fischhoff 1994)[10] Nevertheless, a fixed acceptable risk can be an appropriate guide for a class of risky choices that is homogeneous in respects other than risk (provided its context dependence is recognized). Nuclear waste siting may be such a case.

Moreover, society is not a unitary actor. Multiple players, including Congress and business, are involved in setting and enforcing compliance with regulations. Acceptable risk has bureaucratic appeal, in that it is compact, easily understood and lends itself administratively to division of labor between safety regulators (like NRC) and agencies (like OMB) with complementary concerns. In any case, acceptable risk is a persistent feature of regulatory policy, and needs to be accommodated.

3.2 Assessment uncertainty (AU) as a feature of acceptable risk

1st order risk assessments may shift as research uncovers new evidence, giving rise to "assessment uncertainty" (AU). I interpret this to mean uncertainty about what the assessment would become with unlimited research (Brown 1991, 1993).

3.2.1 Incorporating AU into AR

A reasonable interpretation of the EPA requirement of a "reasonable expectation" (of

[9] In the YM case, one might, for example, consider discounting the future.

[10] A given risk, say from drilling for oil on ANWR, may be acceptable in oil crises, but not otherwise Russians accept higher risks from oil than us because a shortage there could threaten economic breakdown.

meeting 1st order test of AR) would be that, in addition to first-order compliance, an acceptable risk assessment should be "firm," i.e. have low assessment uncertainty. (Reducing AU is presumably a reason for characterization). I call this second-order compliance, which requires that no further research could "plausibly" overturn 1st order acceptability.

3.2.2 Proposed second order test of acceptable risk

I suggest a form of words in safety regulation along the following lines:

For a facility to be in compliance with safety requirements, it must pass two tests:
a) First order test: the probability of (specified mishap) must not exceed x%[11].
b) Second order test: the probability that the first order test would be failed given ideal evidence must itself not exceed y%.

3.2.3 Implications of 2nd order test

The "specified mishap" could be any undesirable event, typically a discrete occurrence (like a core melt). In the YM case, the event relates to a variable. The event is "more radioactive emission than (specified limit) reaches the accessible environment over the next 10,000 years".

An alternate form of the 1st order test in the "variable" case would refer to *expectation* (of radioactive release), rather than 90 percentile. This permits simpler application of the research planning methodology, and is easier for the potential user to understand, which may be critical (see below).

Where to set the acceptable 2nd order probability y% is a question of regulatory policy. For of expository purposes I will again use 90% (as in EPA's 1st order test).

[FIG 2: ACCEPTABLE…]

3.2.4 YM example of unacceptable 2nd order risk

Non-compliance with 2nd order AU is illustrated in Figure 2. At the top, Figure 2a shows a plausible 1st order assessment, i.e. a probability distribution on total release, based on information available at the outset of the research program. The dashed vertical line indicates the EPA limit (L), requiring at least a 90% probability that it will be met. The solid vertical line is the 90th percentile of the current assessment. Since it is well below L, the site complies with 1st order AR on this issue.

Two possible shifts in the simple assessment of release in Figure 2a after such research are shown in Figure 2b. One meets the 90% goal, as a result of a favorable characterization and the other does not. Figure 2c represents a complete AU distribution, summarized by a solid horizontal 80% credible interval bar[12].

---

[11] Or a comparable expectation requirement. which NRC has been considering
[12] For an expectation rdquirement, vertical lines and the horizontal bar would simply move to the left

It shows   a 15% chance of failing to meet 2nd order AR, thereby failing to comply with regulation.

3.2.5 Prescribed research as safeguard

Even with a 2nd order AR test, it makes perfectly good sense for regulation to prescribe *some* research as a kind of "defense in depth" (as in the YM case).

3.3 Nature of probabilistic assessments

3.3.1 Personalist interpretation

For regulatory purposes, I take a "personalist" position on the basis of the assessed risk (1st or 2nd order). Thus, probability is a neutral statement of uncertainty (not, say, conservative), elicited from qualified assessors, and based on all evidence, hard, soft and judgmental, available at the time of assessment.[13]

3.3.2 Assessment method

Determining current 1st order risk is a familiar (if controversial) exercise in Probabilistic Safety Assessment and requires no special methodology. It can be derived by propagating any lower level of risk assessments, taking account of functional relationships[14].

The derivation of 2nd order assessments is more problematic, but it could be propagated from lower level 2nd order assessments (see below), say, by adapting the theory of the distribution of sampling statistics (Kendall and Stuart, 1961).   The default assessment would be directly or indirectly elicited professional judgment. However, this subjectivity constrains, if not fatally[15], the defensibilility of assessments for compliance purposes. ("Garbage in, garbage out")

4 RESEARCH MANAGEMENT: APPLICANT'S PERSPECTIVE

This section presents a research management tool for use by the applicant as research manager, seeking to comply with 2nd order AR.

4.1 Task formulation

4.1.1 Allocate resources among research activities

The task is essentially to allocate research resources among a hierarchy of research issues (see figure 1) and to determine when to stop. I seek to develop a tool useful to the applicant.   Incorporating "society's" and other perspectives is addressed in section 5.

---

[13] This does not coincide with much common PRA practice (Brown 1994), including DOE's past interpretations, reflected in Total System Performance Assessments for YM (TRW 2000)

[14] From theory of distribution of functions of random variables.

[15] Lawyer Stephen Breyer (personal communication, 1980) says that the judicial system will generally accept as "non-arbitrary" an *indirect* (i.e. modeled) assessment based on subjective judgment.

4.1.2 Applicant/Research manager's distinctive objectives

The regulator and applicant have distinct objectives. The applicant wants to demonstrate that the facility (site) *does* comply with regulation. The regulator, on the other hand, wants to learn it, if the facility *does not* comply.

4.1.3 Projection, not optimization

The end product of the method is only to project the consequences of research plans in a convenient form. The gap between current and acceptable risk assessment is integrated informally with other considerations into research planning. No formal optimization is attempted, to avoid the impractical burden of making sure that all critical decision considerations are realistically modeled.

4.1.4 Single primary risk measure

A single risk measure is addressed (cumulative release). If it appears that any other measures (such as pre-closure performance) could significantly change the research plan, they are to be factored in formally or informally.

4.2 Illustrative planning choice: allocation among modes of release

As an example of method in use I will take a particular planning problem at YM: how to allocate resources among the main modes of radioactive release: gas, water and human intrusion.

4.2.1 History

The initial YM budget was devoted largely to water-borne release (which had traditionally dominated comparable siting studies and had the most influential research team). Early findings suggested that gaseous release was a serious concern and much less was known about it, but funding was not reallocated to gas (which had a smaller and less influential team). Human intrusion (which had fewest advocates) was the least well understood, but was assigned the least research effort.

4.3 Informal reasoning on research planning

The essence of the proposed approach is contained in the following line of informal reasoning (not necessarily authoritative). A research manager might usefully adopt it as it stands, without further modeling.

4.3.1 Top-level appraisal

"I will periodically appraise the gap between the current and acceptable risk assessments and orient my research activities to bridging it. It looks as though the probability of release exceeding the limit is less than the required 10%, i.e. 1^{st} order acceptable. However, additional research could easily indicate an unacceptable higher probability. This is too much assessment uncertainty to satisfy NRC's requirement of 'reasonable assurance'."

4.3.2 Resource allocation

"Waterborne release has been extensively studied and much research effort would need to be expended to reduce AU. Gaseous release, on the other hand, has been much less studied and modest research effort is likely to reduce AU more. So research on gaseous release would be more productive, up to a point[16]."

"Current uncertainty about human intrusion is immense, but no feasible research is going to affect it much, so research on human intrusion is least cost-effective. Moreover, there is a quite separate reason to downplay human intrusion. If comparing alternative geologic sites were to become relevant, since they are unlikely to differ much on human intrusion, which becomes less important (compared to water and gas release). However, if non-geologic waste disposal were ever to be considered, human intrusion risk might well discriminate among options on the grounds of relative accessibility. Total release is unlikely to differ much from the sum of these three modes, due to other release modes (e.g. earthquake) or non-additivity."

4.3.3 Halting research

"If it becomes clear that no further research could produce either an acceptable or unacceptable risk assessment we will stop work. If neither happens within a reasonable time, say, because human intrusion uncertainties prove irreducible, we will explore with NRC an alternative stopping criterion (e.g. one that only considers risks other than human intrusion)."

4.4 Modeling the top-level risk

This informal reasoning might be enough for a research manager to make his own planning decisions. However, a formal methodology like the following may be needed to convince others. Figure 3 illustrates the form such modeling might take[17].

[FIG 3: PARTITIONED…]

4.4.1 First order assessment

Figure 3a addresses how far from being acceptable the current risk assessment is, both 1st and 2nd order. The bell-curve is a current 1st order assessment (corresponding to Figure 2a). The vertical solid line is its 90 percentile. Since that is below the regulatory release limit L (indicated by a vertical dashed line), the site's risk is 1st order acceptable.

4.4.2 Second order assessment

The two bars below the curve in 3a summarize 2nd order assessments--solid for current, dashed for target--as 80% credible intervals, with the upper end of the bar marking the 90 percentile. Since the solid bar extends to the right of L, the current 2nd order assessment

---

[16] This finding was generally rejected later in the research program, as the test of acceptable risk shifted from release to dosage, which is less serious for gas. As perspective and knowledge evolve, the preferred research strategy can change.

[17] The scales and assessments in figure 3 are no more realistic than needed to make a point For example, log scales would be more realistic, but would not show equal areas for equal probabilities.

is unacceptable (if 10% is the 2nd order threshold). The dashed target bar corresponds to one possible acceptable 2nd order assessment, since it just reaches the limit line.

4.5 Modeling contributors to total risk

Figure 1 illustrated how total risk (release) can be partitioned into a hierarchy of contributing risk variables. Modeling the risk variables at lower levels in the hierarchy performs two functions:

- To help make the top level assessments
- To help allocate resources at that lower level

4.5.1 Partitioning current assessments

As far as the first of these is concerned, assessments at any level (1st and 2nd order) are implied by assessments at the level below, provided:

- all variables at that level are accounted for,
- the functional relationship among them is specified, and
- an error term makes the function an identity.

Figure 3b illustrates this propagation, for just the second tier in the hierarchy (three release modes). For each variable and residual error, current 1st and 2nd order assessments are shown by a curve and solid bar, as before.

4.5.2 Judgments reflected

As drawn, these current assessments are consistent with the informal reasoning above. They show widest curve (1st order assessment) is for human intrusion, then gas, then water, then error.

However, modes are ranked differently on AU (solid bars)—gas has most (i.e. is most shiftable), then water, then error, and lastly human intrusion. This reflects a judgment that there is most still to learn about gaseous release and a good deal less about water-borne release. Error uncertainty may be already small, but almost all of it could be resolved with enough research[18]. Intrusion may be extremely uncertain now, but there is very little that research can do to resolve it. This difference between 1st and 2nd order rankings has *a major effect on research planning*.

4.6 Target assessments

Allocating resources among research activities in order to close the gap between current and acceptable 2nd order assessed risk is largely determined by 2nd order assessments of risk variables that these activities contribute to. (Any 1st order gap will automatically be closed if the 2nd order gap is closed).

[18] On the other hand, predicting ground water travel time from a given set of geologic variables and any given mathematical flow model may be quite doubtful as witnessed by disagreements between geologists on the relative merits of certain fracture flow and equivalent-porous-media flow models. Thus the top-level assessment cannot be routinely inferred from lower level assessments (e.g., by simulation), no matter how ambitious the intervening physical process modeling has been. In particular, the top uncertainty will be seriously understated if no allowance is made for modeling and residual error. Nevertheless, even without error terms, the assessments are bounded by those below them, permitting a consistency check on the hierarchy of assessments.

4.6.1 Goal allocation as planning guide

A research plan can be specified by allocating resources to variables at any level of the hierarchy of risk variables in Figure 1. Ultimately the allocation among variables must translate into allocation among specific research activities that directly influence the lowest level variables and through them to higher levels. Allocation among release modes is at the highest level.

The allocation procedure is illustrated in Figure 3. It is basically a 2nd order adaptation of a well-known procedure, probabilistic goal allocation, that is normally directed at 1st order assessments[19].

For each risk variable at a given level (release mode in this case), a 2nd order "target assessment" is specified, shown as a dashed bar in figure 3. The complete set of target assessment is chosen so as to produce an acceptable top-level target assessment, as shown by the dashed bar in figure 3a. Note that its 90 percentile just reaches limit L. Each contributing target assessment can be characterized by its 90 percentile, which can be interpreted as a target, comparable to L at the top level. Thus, target T(g) is the gaseous release with just 10% probability of being exceeded in the target assessment[20].

4.6.2 Target Assessment combinations

There are many sets of possible target assessments, and therefore many research plans, consistent with a single acceptable top-level 2nd order risk. Among them a set is chosen to be technically feasible at reasonable cost and delay. The choice takes into account how far each target assessment falls short of its target, and what it will take to achieve it. The latter requires considering

- The gap between 2nd order current and target assessments (the difference between solid and dashed bars), and
- The cost-effectiveness of research activities needed to close the gap.

This permits deducing (formally or informally) a research program projected to produce and acceptable top-level risk assessment cost-effectively.

I suggest the cost of achieving any given reduction in 2nd order uncertainty be derived by informal argument, along the following lines. In this case, effort studying gaseous release appears to be very cost-effective, since a large reduction in 2nd order risk (shortening solid bar to produce dashed bar) can be achieved at modest cost. On the other hand, reducing 2nd order risk of water-borne release by enough to reduce total 2nd order risk equally (same shortening of solid bar at top of figure 3) would be more costly. Human intrusion permits even less cost-effective reduction in 2nd order uncertainty (negligible shortening of solid bar). Modeling error is already tiny and already thoroughly studied, so not worth much effort to reduce.

---

[19] See Apostolakis (1985); Sung and Cho, (1989); Hunt and Modarres (1984); OECD (1989). Under the name "performance allocation", this approach was, in fact, attempted for this very problem at DOE in the late 80s (DOE 1988). However, it was not, I believe, successfully completed or followed through, at least to the point of significantly impacting research resource allocation (Brown 1988).

[20] The sum of the contributing Ts is greater than L, according to statistical theory.

4.6.3 Lower level targeting

The same logic could be used to address lower level risk levels, down to the very bottom of the hierarchy in figure 1[21]. This would be an indirect way of both making higher level current and target assessments at a lower level[22].

If competing tasks address the same sub-tree in the hierarchy of variables, one need only consider impact on the lowest higher order variable they share[23].

4.7 Other considerations in prioritizing research issues or activities

The cost-effectiveness of a research plan in producing acceptable risk for this one variable (post-closure release) is not the only consideration in choosing the :best" plan—though it may be the main consideration.
Others include:

- Risk requirements other than the one (s) modeled by the above approach (e.g. preclosure release).
- Delay in disposing of nuclear waste
- Sequencing: the research plan is subject to change as assessments change in response to evolving evidence and this may affect the initial plan.

4.7.1 Optimization?

These considerations could all be modeled explicitly in a way that permits explicit optimization, and therefore calculation of the preferred plan. However, this would be unmanageably burdensome to do realistically, and, I suggest, better handled by informal adjustments.

4.7.2 Trial and error

I recommend iteratively evaluating plausible target allocations in terms of what they do to cost and these other consequences, as well as the modeled top-level risk assessment. The effectiveness of informal adjustments will be greatest if the above target assessment exercise accounts for most of a comprehensive "bottom line" plan evaluation. How close to an appropriate criterion is maximizing probability of complying AR for a single variable (release) at minimal cost?

This approach has drawbacks, even to advance just the applicant's interest. The output will not explicitly prescribe what research managers and their sub-contractors should do, since it involves no optimizing algorithm—only a clearer representation of the research consequences of options, from which persuasive informal argument can be built. A numerically more explicit refinement and implementation of the approach would carry

[21] Engineered features would not be included for the present exercise.

[22] For example, it may demonstrate that spreading effort evenly over several geologic strata representing serial barriers is a more cost-effective way to produce the water target assessment in fig 3b than devoting all resources to one primary barrier (taking due account of differing accessibility of strata and therefore cost-effectiveness of effort).

[23] One research task may, however, address several different variables. For example, sinking an exploratory shaft, might cast light on more variables throughout the hierarchy and so deserve high priority.)

more weight, and enable a research manager to take more active control of research planning.

Although full numerical application, even without formal op0timization, involves ambitious analytic effort and new technical development, it can be—and has been—used as an informal knowledge-structuring tool. In particular, in cases with lower stakes than YM, it may assure that relevant considerations are taken into account. However, explicit quantification may be needed to be fully effective in *enforcing* sound risk research planning on the recalcitrant.

4.8 Halting research

A research program is complete, and can be halted, when it establishes that the facility (site) either does or does not comply with regulation, however specified. A reasonable criterion for a good research strategy is to minimize the expected cost of such completion (not necessarily derived explicitly).

4.8.1 Successful compliance

Applicant would reasonably halt research if 2nd order acceptability is achieved, i.e. when the dashed bar is entirely to the left of limit L, as shown in figure 3a. This would be achieved if all target 2nd order assessments, represented by dashed bars lower in the tree were met.

4.8.2 Failed compliance

Presumable applicant would also stop when all hope of success had gone, i.e. if the dashed bar in figure 3a settled entirely to the right of L.

5 RESEARCH MANAGEMENT: SOCIETY'S PERSPECTIVE

5.1 Distinctive interests of applicant and society

The applicant's interests differ from those of "society". His primary objective is presumably to have his application succeed, and so has an interest in maximizing the *unconditional* probability of a "positive" finding (see below). Society's interest is to properly balance safety, economics and other considerations. Sec 4 has sought an appropriate research strategy from the applicant's point of view.

5.1.1 Enlightened self-interest in an adversarial process.

However, interests may converge under the right conditions. In our society, contentious issues are often resolved by participants pursuing their own interests (within rules to safeguard public interest), sometimes subject to an arbiter[24]. Appropriate regulations discipline the environmental protection process (see section 3). "Enlightened self-

[24] In the courts, counselors contend and jury decides. Economic processes foster enlightened self-interest among competing businesses. The regulator in this case could be treated, not as an arbiter, as an *opponent* of license approval, rather like a counsel for the prosecution. (TRB, reporting to the President, would be judge/jury). Regulators may justify their existence by "ratcheting", i.e. progressively piling on requirements. Experts on advisory bodies may urge research on topics they are interested in

interest," argues that the better the parties advance their own interests, the better public interest is served. Thus, helping the applicant may help society, but we cannot take that for granted.

5.1.2 A tool to serve society directly

Alternatively, a research-planning tool can be developed to serve public interest explicitly.

The applicant's (DOE's) discretion may be limited by direct intervention of others representing society's interests more closely. In particular, NRC and TRB make suggestions (typically to do additional research) that are often interpreted as directives by DOE. Without some systematic research planning rationale, even these bodies may be diverted from best serving the public interest[25].

There are two promising approaches, one based on the acceptance-testing paradigm, the other on adapting the target assessment approach.

5.2 Decision theoretic acceptance testing

There are well-established decision theoretic approaches to designing information-gathering options. The basic decision theoretic logic from the perspective of the ultimate client, society in this case, is the "preposterior analysis" paradigm of statistical decision theory (Pratt et al. 1993).[26]

Mattson et al. (1991) have applied this approach to YM. They prioritized tests for "potential concerns" about radioactive release, and concluded that "The tests of highest priority are those for gas flow (carbon-14) above the repository, and possibly (other tests)". Interestingly, this is consistent with the results of target assessment applied hypothetically to the release mode issue above[27]. Both approaches were predicated on the knowledge and regulatory requirements in 1990

Society has an interest in reducing two types of error: low probability of either finding a facility acceptable when it is not, or finding a facility unacceptable when it is in fact acceptable. The appropriate research plan depends on the relative cost of the errors: false positive and false negative, respectively). However, the complexity of this analytic task (especially the inclusion of value judgments) makes it difficult to do without making simplifying assumptions that sacrifice equivalence of the model to realism. Accommodating multiple constituencies adds complexity.

---

[25] A research-planning tool might be developed for the regulator that paralleled the approach put forward in section 4, but where the target assessment shows a *high* probability of exceeding the limit. .

[26] Non-decision theoretic approaches to the "acceptance testing" version of the problem are familiar in engineering (Oliver and Smith 1940, Diamond 1989).

[27] Then, radioactive *release* was treated as the primary risk variable to be controlled. However, when regulatory attention later (and I believe appropriately) switched to *dosage*, the case for gaseous study significantly weakened. A given gaseous release produces less dosage due partly to dispersion into the atmosphere. Reapplication of either approach, with appropriately changed inputs, would probably have confirmed this shift, and possibly made it more convincing.

The model used departs from being an "equivalent substitute" for the real problem by unrealistically treating as discrete both the options—either to test or not on each issue--and the potential test outcomes--false-positive, etc. The optimizing feature of acceptance testing produces unambiguous direction for research managers to follow, thereby restricting their and others' ability to serve private agendas. However, the very absence of optimizing in the target assessment approach permits informal accommodation of considerations inevitably omitted from the model or unrealistically simplified.

5.3 Target assessment as element of acceptance testing

The target assessment approach can be adapted as a kind of partial acceptance testing. An important difference is that, unlike section 4, we would not set *unconditional* target assessments aimed at demonstrating compliance (for applicant) or non-compliance (for regulator or environmental intervener). Instead, we specify two sets of *conditional* target assessments. The condition is whether the facility is or is not acceptably safe (2^{nd} order). These correspond to true-positive and true-negative in the acceptance testing model

Both target sets are compared with current assessments, along the lines of figure 3. In the true-negative case, the counterpart of figure 2 would then show a top-level dashed target bar entirely to the *right* of the release limit (rather than to the left). As before professional judgment is used (as before) to make the actual resource allocation decision. The resolution would depend *implicitly* on the relative cost of a false-positive and false-negative, but without formal optimization. I am not clear how exactly this idea would be implemented or how practical.

5.3.1 Combining both approaches

These two research-planning approaches might be pooled with plural evaluation techniques[28]. They have the complementary strengths of realism for target assessment and optimization for acceptance testing.

Optimization produces a publicly reviewable decision procedure and reduces irresponsible discretion. That discretion can also be beneficial in giving a responsible risk manager wiggle room to combat defects in regulation. Calculating an optimum allocation under simplifying assumptions may be insightful, provided it is not treated as an equivalent substitute for realistic, comprehensive reasoning.

6 DISCUSSION

6.1 Status of research planning tool

I have attempted, at a meta level, to cover all technical and behavioral bases that are critical to usefulness (Brown 1991). This inevitably leaves many loose ends to be tied. The methodology presented here is certainly still primitive and largely conceptual.

[28] E.g. with plural evaluation techniques (Brown and Lindley, 1986).

However, I did use its logical structure to formulate tentative research planning suggestions to my DOE consulting client. They implied significant redirection of research resources among research teams. I am not sure this had much practical impact. However, it got the nervous attention of a contractor who got wind of it and whose research budget it seriously threatened[29].

6.2 Work to be done

To make the target assessment approach to research planning more useful, defensible and applicable to a wider class of risk research planning cases, several types of methodological development are needed.

6.2.1 Scope of method

The scope of the approach presented could be enlarged to address balance of applicant's research effort between *evaluating* a fixed risk (such as a repository site with given engineered features) and *reducing* the risk (e.g. through more effective containment and other engineering). This would permit reviewable validation of recent shifts in research planning at Yucca Mountain that favor engineering enhancement, as confidence in the ability of the site on its own to isolate release has receded.

6.2.2 Algorithms

Operational algorithms need to be developed for propagating contributing first and second order assessments of risk variables into total risk. The derivation of the assessments themselves will be somewhat issue-specific and require extension of familiar PRA techniques to incorporate informed judgment (NASA 2001 [?], Brown 2001).

To turn a logical rationale into an operational methodology and extend it to include the false-negative/positive issue will require work. Work is needed to develop the specific procedures and software to implement the rationale in a practical context. What input can be cognitively supplied, what output can be institutionally used and what logical algorithms are appropriate to link the two? .

6.2.3 Case development

Most important, any evolving methodology will need to be exercised, developed and tested in the context of live research management cases.

6.3 Conclusion

This paper reports on work that attempts to be prescriptive—in the sense that it produces usable and useful methodology. In particular, it has proposed a regulatory initiative to motivate responsible research direction by the applicant, and a resource management

[29] The potential of this work to save DOE money, and the power of contractors to thwart it, are illustrated by the following incident. My consulting contract was administered through one of the major research contractors. When it became clear by 1991 that my interim reports to the DOE Office head and to the Technical Review Board could lead to selective cuts in the research budget that could cost that company many millions of dollars, I was warned by a vice-president "We expect you to be more docile (*sic)*. We are not going to pay for the bullets that will kill us, and can get your contract cancelled". It was.

procedure to help the applicant to control research sub-contractors. If successfully implemented, it may provide an enforceable rationale for research programs that are more cost-effective in the public interest.

It also provides a convenient framework within which to present competing arguments about what should be (or should have been) done. The approach can be used to validate whatever research management strategy is proposed (even if it was not used in the first place to derive that strategy). It should help the risk manager, DOE in this case, to withstand, not only regulatory and legal challenge, but also the resistance of researchers to having their projects terminated for programmatic (rather than scientific) reasons.

Although the large scale of the nuclear siting problem is exceptional, the issues and proposed resolutions typify a wide range of risk management problems, especially where large funds and public controversy are involved. This is true, whether we are talking about genetic engineering, industrial safety, the health risks of consumer products or fatal errors in medical facilities.

REFERENCES

Apostolakis G, Kafka P. Advances in Prob Safety Assess. Nucl eng des 134 (1) 141-148 May 1992.

Apostolakis, G. Some issues related to goal allocation and performance criteria," Paper M2 4/3, 8th International Conference on Structural Mechanics in Reactor Technology, Brussels, Belgium, August 19-23, 1985

Breeding RJ, Helton JC, Gorham ED Summary description of the methods used in the probabilistic risk assessments for NUREG-1150. Nucl eng des 135 (1) 1-27 Jun 1992

Brown, R.V. Characterizing a Nuclear Waste Site: A Logical Framework for Allocating Resources, Proceedings 1988 Las Vegas Annual Meeting. Decision Sciences Institute. 1988.

Brown, R.V., and Ulvila, J.W. Does a reactor need a safety backfit? Case study on communicating decision and risk analysis information to managers. Risk Analysis, 1988, 8(2), 271-282.

Brown RV. Some Reflections on DOE Nuclear Waste Program. Consulting report to Dr. John Bartlett, Head of DOE Office of Civilian Radioactive Waste Management. Decision Science Consortium, Inc. February 7, 1990.

Brown, R.V. Assessment uncertainty technology for making and defending risky decisions. Journal of Behavioral Decision Making, 1991.

Brown, R.V. Impersonal Probability as an Ideal Assessment based on Accessible Evidence: A Viable Construct? Journal of Risk and Uncertainty, 7:215-235, 1993.

Brown RV. "Extending Probabilistic Risk Assessment to include Soft Data". Conference on Probabilistic Safety Assessment Methods of IEEE San Diego, Spring, 1994.

Brown RV "Should a regulator require specific action, acceptable risk or just the public interest?" Working Paper Series, School of Public Policy, George Mason University. Fairfax, Va. 1996.

Brown RV. How Decision Aids can Mislead Policy Makers. TIPP working paper series. GMU. 1997.

Calabria R, Pulcini G, Rapone M. Optimal reliability and maintainability allocation in manufacturing systems. Qul reliab eng int 11 (3) 183-188. May June 1995

Chinnis JO. *Evaluation of a proposed Satellite Power System.* Technical report to DOE. Decision Science Consortium, Inc. 1986 [?]

Cho NZ. Multiobjective programming approach to reliability allocation for nuclear power plants. Nuclear Science engineering. 95 (3) 165-188 Mr. 1987

Congress of the US. The 1987 Nuclear Waste Policy Amendment Act, 1987.

Covello, Vincent T. and Merkhofer, Miley W, Risk Assessment: Methods and Approaches for Quantifying Health and Environmental Risks, N.Y., Plenum Press, 1994 [?].

CRWMS (Civilian Radioactive Waste Management System) and M&O (Management and Operating Contractor). Total System Performance Assessment—Viability Assessment (TSPA-VA). Las Vegas, Nevada. CRWMS M&O. 1998

Diamond, William J. Practical Experimental Designs for Engineers and Scientists, 2nd Edition, N.Y. Van Nostrand Reinhold, 1989.

DOE Viability Assessment of a Repository at Yucca Mountain. DOE/RW-0508/V4. DC: OCRM Dec. 1998.

DOE, US. Site Characterization Plan, Yucca Mountain Site, Nevada Research and Development Area. DOE/RW-0199. Washington DC: OCRWM Dec. 1988.

Federal Register 64 FR 8640. 1999a. Disposal of High Level Radioactive Waste in a Proposed Geologic Repository at Yucca Mountain. Proposed Rule: 10CFR Part 63.

Federal Register 64 FR. 1999b. OCRWM/NWR: Yucca Mountain Site suitability Guidelines. Proposed rule making. 10CFR963.

Federal Register 64 FR 46976. Environmental Radiation Protection Standards for Yucca

Mountain, Nevada. 40CFR Part 197.

Fischhoff B. Acceptable Risk: a conceptual proposal. Risk: Health, safety a1and the Environment. Winter1994.

Hunt, R.N.M. and Modarres, M. Integrated economic risk management in a nuclear power plant. In Proceedings of the Annual Meeting of the Society for Risk Analysis, October 1984, Knoxville, TN, Plenum Press, New York.

Keeney RL. An analysis of the Portfolio of Sites to Characterize a Nuclear Repository. Risk Analysis 7. 195-218. 1987.

Keeney, R.L. and Merkhofer, M. A multiattribute utility analysis of alternative sites for the disposal of nuclear waste. Risk Analysis, 2(7), 1987.

Keeney RL and von Winterfeldt D. Managing Nuclear Waste from Power Plants, Risk Analysis, 14, 107-130. 1994.

Kendall, M.G. and Stuart, A. The advanced theory of Statistics. London: Hafner, 1961.

Mattson SR, Judd BR, Sinnock SR, Hoxie DT. Testing Priorities at Yucca Mountain: Recommended Early Tests to Detect Potentially Unsuitable Conditions for a Nuclear Waste Repository. USDOE/OCRWM. March 1991.

NAS 2001[Report on disposal of HLW]

National Energy Association. Proceedings of an NEA Workshop on Uncertainty Analysis for Performance Assessments of Radioactive Waste Disposal Systems, Seattle, February 24-26, 1989, OECD, Paris.

National Research Council (Board on Radioactive Waste Management). Rethinking high-level radioactive waste disposal. Washington, D.C.: National Academy Press, 1990.

NWTRB “Report to US Congress and the Secretary of Energy”. Jan-Dec 1998.

[Oliver and Smith 1940]

Sung, S.K. and Cho, N.Z. Determination of performance criteria at hierarchical levels in a nuclear power plant, Reliability Engineering and System Safety, 24:231-256, 1989.

Whipple C, Budnitz B Ewing R, Moeller D, Payer J Witherspoon P. 1999. Final Report. TSPA Peer Review Panel. Las Vegas, Nevada.

Yang XP, Kastenberg WE, Okrent D. Optimal safety goal allocation for nuclear power plants. Reliab eng syst safe 25: (3) 257-278 1989.

Yucca Mountain Project. Timeline/Milestones. www.ymp.gov/timeline. Oct. 2000

10CFR60. Code of Federal Regulations. Nuclear Regulatory Commission Title 10; Disposal of High-Level Radioactive Wastes in Geologic Repositories. Part 60. Washington D.C. Government Printing Office, 1992.

10CFR960. Code of Federal Regulations. General guidelines for the recommendation of sites for nuclear waste repositories. USGPO. 1987.

33CFR404. Code of Federal Regulations. Navigation and navigable waters: Title 33. Part 404 [?] Washington D.C. Government Printing Office, 1986.

40CFR191. Code of Federal Regulations. Environmental Protection Agency: Title 40; Environmental standards for management and disposal of spent nuclear fuel, high-level and transuranic radioactive wastes: Part 191. Washington D.C. Government Printing Office, 1985.

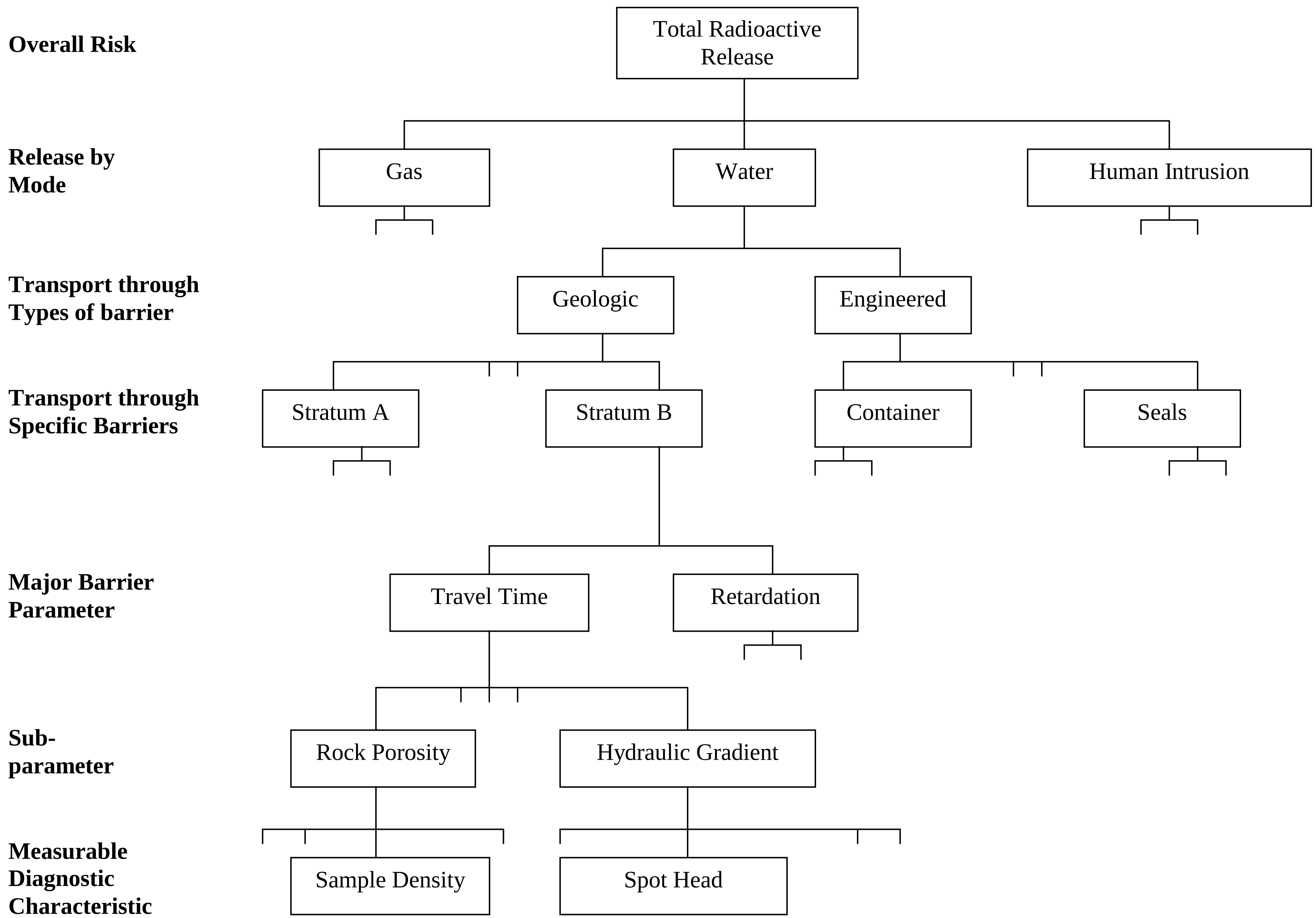

FIG. 1: HIERARCHY OF RADIOACTIVE RISK VARIABLES
(Partial)

@fig4 bare tree
1/22/01 5:37 PM

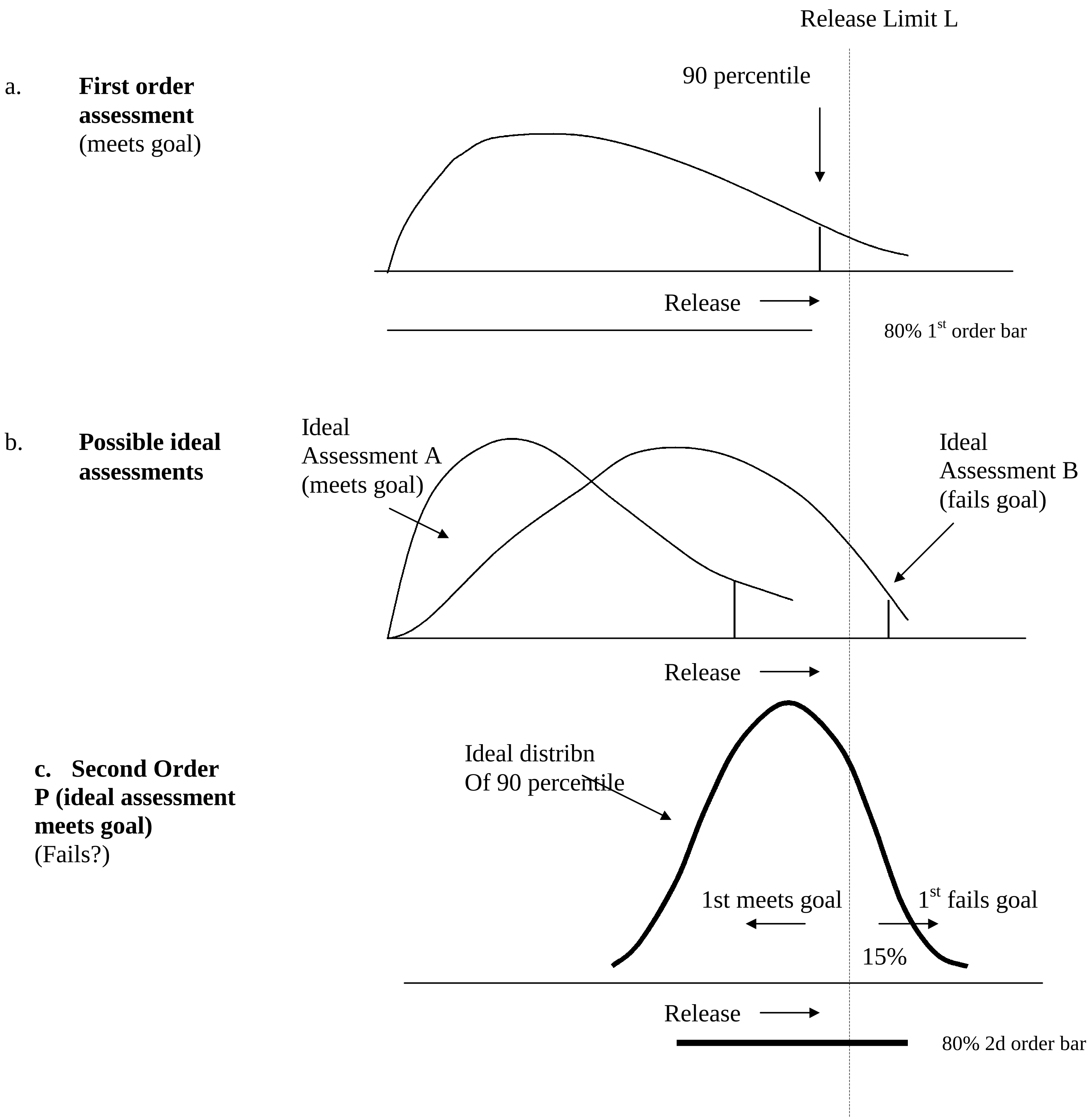

FIGURE 2
ACCEPTABLE 1ST AND UNACCEPTABLE 2ND ORDER ASSESSMENTS

@@fig2 1st 2d bells
01/22/2001 5:36 PM

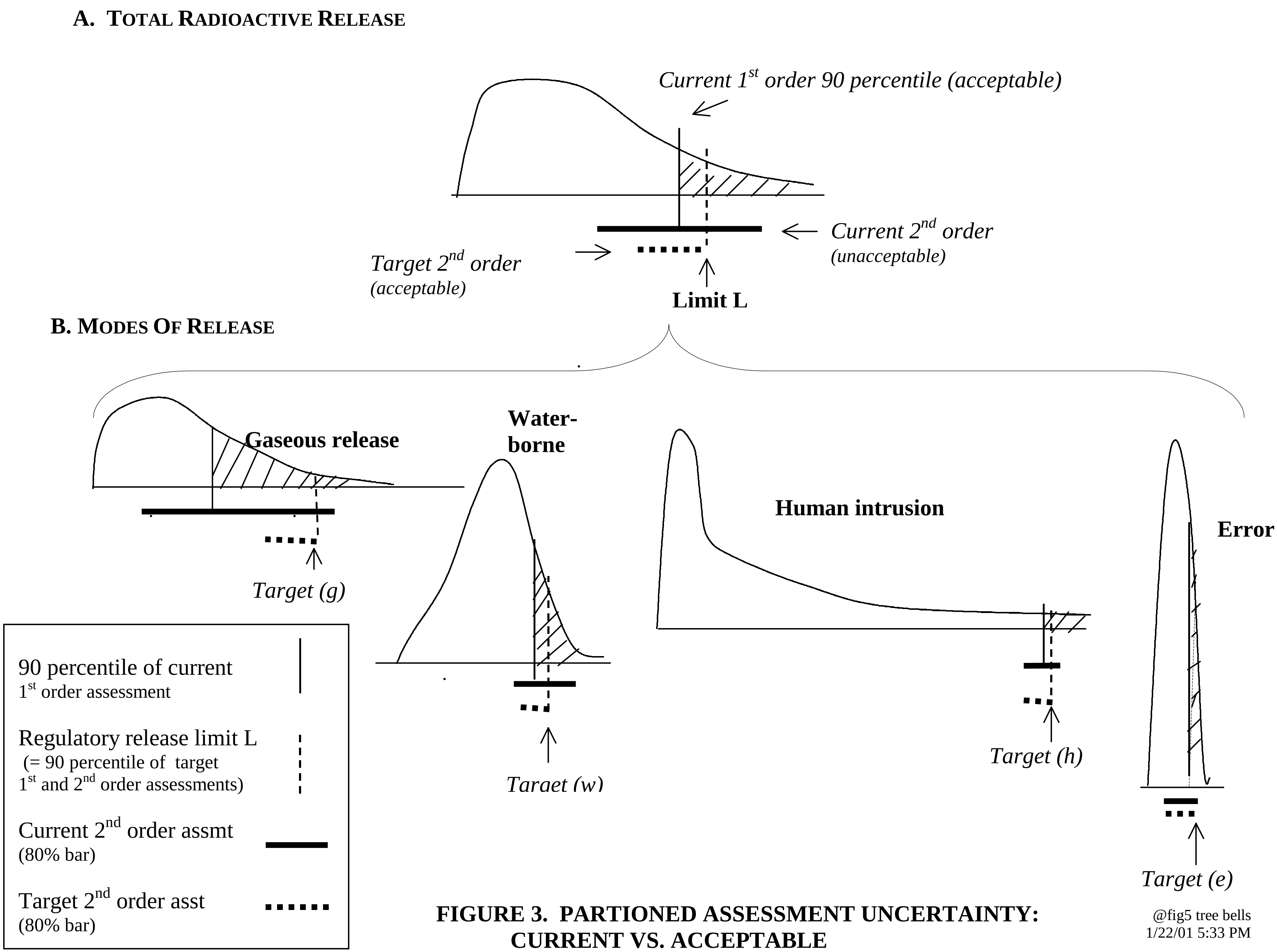

FIGURE 3. PARTIONED ASSESSMENT UNCERTAINTY: CURRENT VS. ACCEPTABLE