

## **COMMENTS ON DRAFT NUREG-1860**

Graham Wallis 07/07/2007

This document supersedes all earlier drafts. The first few pages provide an overview of my major conclusions. Pages 4-37 comprise a detailed review of the text, including the appendices. Pages 38-45 are an Appendix describing general features of ways to represent and manage risk.

### **OVERVIEW**

I fully support the April 2007 ACRS letter which stressed the importance and significance to the NRC of a general regulatory framework for future reactors. Development of such a framework represents a great opportunity to chart a regulatory course for coming decades.

In NUREG-1860 the staff describes a framework for regulations that could potentially apply to all future reactors, independent of their technology. A lot of work has been done and many useful ideas presented. However, I do not consider the result to be close to defining a mature framework. Several important features could be significantly improved and some may need to be changed. Far from describing a final design, the document is closer to the first step along what could be a fairly long road. Because of the great influence that the eventual framework is likely to have on the future of reactor technology, regulation, and public safety, it is essential that it be an authoritative, convincing document and close to optimum for practical use. Premature adoption of a process that later requires extensive revision could prejudice the success of the entire effort.

I developed this evaluation after studying the relevant sections of the document several times, in an attempt to develop a full understanding (which sometimes did not mature until reading the Appendices), but it is possible that I missed or misinterpreted some things, or am simply too poorly informed about regulatory thinking, in arriving at the following conclusions:

1. There is no development of a clear set of performance-based objectives, or top level design criteria, independent of the actual choice of design details, that can be used to guide the choices of approaches and to evaluate success at the end of the project. Many features are inserted into the proposed framework by description, without evaluating what function is being served and why the particular structure that is proposed has been chosen. These may unreasonably restrict choices available to a nuclear design engineer. A performance-based set of regulations should emphasize safety functions and metrics without overly specifying details of how to meet them.
2. Provision of containment/confinement appears to be a high level design criterion. It is hidden in the report as a brief reference in Section 6-4-3 to “controlled low leakage barrier” and in Table 8.1 as “provisions to establish a containment functional capability”.

This sort of key requirement should be given prominence early in the report, justified, and criteria for judging its adequacy developed.

3. There is no analysis and evaluation of various ways to describe and determine the impact on public safety of nuclear reactor operation. Apart from reference to the QHOs, choices are made with little justification or exploration of their impact on the achievement of safety, or the perception that adequate safety is being achieved.

4. Important criteria for the NRC and industry, as well as for public understanding and acceptance, such as economy, effectiveness, clarity, simplicity, enforcibility, and transparency, are not articulated or evaluated. They should be reflected in design specifications for any framework and used as part of the procedure for choosing between alternative approaches.

5. Though the QHOs are mentioned as the basis for future licensing, there is no explanation of how they are met by the framework, nor what is evaluated as the most effective and efficient means to do so. Surprisingly, they do not seem to be clearly quantified anywhere in the document in terms of dose and frequency, which are the outputs from the PRA that are intended to meet these criteria.

6. There is no comparison between the proposed framework and the existing system of regulation to show what improvements are being made.

7. No method is provided for adding up the risk represented by individual PRA sequences to obtain the risk of particular accidents, accident classes, or the total risk of the plant. This feature would appear to be an essential characteristic of a consistent regulatory framework that could be used for monitoring operation, changes, and appropriate regulatory actions in addition to initial licensing.

8. Though most, if not all, measures of public safety involve cumulative risk, the staff's "F-C curve" is designed to regulate at the detailed level of individual PRA sequences. This is far from the "top-down" philosophy mentioned by the ACRS in its April 2007 letter.

9. Since the QHOs, and other cumulative measures of risk, can be met individually without introducing an "F-C curve", its use and desirability needs to be justified. It simply appears out of the blue in the report.

10. Use of metrics resembling CDF or LERF, which have proved increasingly useful for evaluating risk of current plants, is dismissed summarily without adequate explanation. It is a proven method, is simpler and more economical to use than a two-dimensional F-C curve, and was recommended by the ACRS in their Sept 21, 2005 letter. How will the present functions of these metrics be performed? Can the QHO metrics substitute for them?

11. It is not clear that any method of regulation of individual PRA sequences, which is the basis of the processes described in Chapter 6, can succeed because the results of these sequences can be arbitrarily changed to meet the criteria by manipulating the structure of the PRA, particularly by subdividing the sequences, without changing any feature of the actual design.

12. The “F-C curve” in the report is not the usual F-C curve. Where individual steps in the curve are given a rationale, there appears to be a use of cumulative dose or frequency as criteria for acceptable dose and frequency of individual PRA sequences, which seems inappropriate.

13. There is no indication of awareness of the additional complexity and difficulties introduced by having the PRA output expressed as the continuous variable, dose, rather than in terms of the binary parameter, CDF. Several significant complicating features need to be explained and analyzed.

14. The need for LBEs, their definition, function and use, should be better explained. There are conflicting, ambivalent, incomplete and undefined terms used in describing their development and use. Neither the text nor Appendix E shows how they are needed to make the described decisions, which could perhaps just as well be based on all the PRA results. The statement in the Executive Summary: “The purpose of LBEs is to demonstrate the conservatism of the PRA analysis”, is unsubstantiated.

15. The proposed LBEs appear to be quite different from the traditional DBAs. Apart from their technical modeling in the PRA, no more detailed technical analysis of them is performed, unlike the practice in the present Chapter 15 of SARs. The statement in the Executive Summary that “certain event sequences are chosen for more conservative deterministic analysis” appears unsubstantiated by anything in the report. It confuses the reader until he reads as far as Appendix F.

16. There appears to be a basic departure from what constitutes a traditional “deterministic” analysis and how it is related to the corresponding PRA sequence which as also analyzed deterministically. There are some additional constraints imposed on the PRA sequences in the name of defense in depth. These are called “deterministic”. The traditional “deterministic” analysis of the type found in Chapter 15 of current SARs appears to have been abandoned completely, though it has been the basis of reactor regulation for decades.

17. The criterion of ignoring all sequences with 95<sup>th</sup> percentile probability less than 1E-7/yr appears to cut out major contributors to violating the early QHO fatality criterion by a factor of five in the example in Appendix E. This appears unacceptable.

18. It appears from Appendix F that the expectation is for computer codes, of the sort that are presently used to analyze DBAs, to be incorporated into the PRA. This is presently infeasible and would require considerable research and development in order to be made practical.

19. The promising approach to evaluating defense in depth by analyzing the uncertainties in the PRA is mentioned several times but is not developed as a main element in the structure of the framework.
20. Safety margin is not defined in a useful form for performance-based regulation.

## **DETAILED COMMENTS ON THE TEXT**

### **SECTION 1. INTRODUCTION**

#### *1.2 OBJECTIVES*

The objectives should be a clear statement of what the authors set out to achieve. Since this project presents a great opportunity to rethink and improve the regulations, the objectives should be cast in terms of general needs of the agency. They should not anticipate specific details of solutions to these needs.

The authors probably needed to spend more time with the Commission, spelling out the objectives and the basis for regulations that they were constructing. The first stage in a top-down design process is understanding and approving the main objectives, accepting what are to be used as criteria for meeting them, clarifying goals for new plant risk levels, whether to include societal risk, deciding whether or not all risks should be combined on a dollar scale, whether the QHOs needed revisiting, whether a measure of overall plant risk status is required, whether to require a containment, and so on. A new set of regulations must rest upon clear policy decisions, and to some extent upon guidance about the details of use.

A list of objectives is also a chance to sell the project to the customers, including the informed public, for whom it is intended.

An item that is listed in the objectives makes the authors accountable for demonstrating how that objective is met. For example, statements such as “minimize complexity in the regulatory process”(p.1-1), “enhance efficiency and predictability” (p1-1), “defensibility in the development of the requirements”(p.1-2), “ensure that NRC actions are effective, efficient, realistic and timely” (p1-4), should not be hidden in the text but be condensed into equivalent objectives. Then the text of the document should demonstrate how these are met.

The suggested framework is more complex than the present process in that it requires a PRA with broader scope (p6-2), comprehensive evaluation of all uncertainties (p4-14), and a measure of consequence, dose, which is continuous and not binary. Many of the present regulatory requirements are carried over with the new framework. Therefore it appears that the proposed scheme will be more burdensome than the present regulations.

There is a clear need to convince users that the framework is truly effective and efficient by evaluating these attributes and comparing with alternative regulatory schemes.

Since it will be useful to refer to specific objectives throughout the report, for example in order to show that design decisions are being made logically and create a result that meets these objectives, it is useful to list them, identifying each with a number.

The first paragraph in the NUREG section 1.2, with the heading “Objectives”, contains two objectives. There is also a somewhat hidden objective later on about meeting the Commission’s expectations for future reactors.

The rest of the section does not describe objectives but is concerned with means to achieve them, such as protective strategies and defense in depth, and ways to make them practical, such as classifying risk significant components. These belong in later parts of the report, if such a level of detail is appropriate in a framework document. If stated as part of the objectives they may inhibit the creation and evaluation of alternative ways to meet the functional objectives.

The first objective is clear. In order to refer to it later, it should be numbered:

Objective 1. Develop a framework that provides the technical basis, including guidance and criteria, for writing risk-informed, performance-based requirements for future reactors.

In view of my later discussion, I’d add the words “internally consistent” before the word “framework” above.

The second objective, “demonstrate that the NRC mission..is met”, seems to describe the entire operation of the NRC and might be something the GAO would set out to do. I suggest a slightly different wording which is more functional:

Objective 2. The framework should be derived from, and support, the NRC mission of protecting public health and safety, as expressed by Commission papers, existing generic requirements, and other appropriate measures of performance.

I suspect that more objectives are needed in order to guide later developments in the report, reflect the bulleted items in the Executive summary, and gather key objectives that are hidden in the text. Some (suggestions only and not a complete list) might look like this:

Objective 3. Provide a structure for constructing regulations that are clearly derived from a set of basic safety measures, are selected on a basis of efficiency and effectiveness for the agency and the industry, and will be convincing to the informed public.

Objective 4. Reduce future effort and expense by establishing the basis of regulations with minimal complexity that can be applied to a variety of new designs, which may involve technologies significantly different from those in use today.

Objective 5. Demonstrate the utility of the framework by applying it to the example of an existing LWR design.

Objective 6. Demonstrate the utility of the framework for licensing new reactors by applying it to the PBMR design.

Some of the “Desired Principles” listed in Section 1-4 might be candidates for inclusion at the level of primary objectives.

Besides giving inadequate treatment to objectives, Section 1.2 contains several assertions which appear to be too vague or incongruous to be useful to guide creation of the framework. When a statement appears in a key section such as this, it is there to be used later in evaluating features of the design and must be usable in that context.

For example: “defense in depth (DID) structure...will ensure safety limits are met”. Is there any means of comparing DID with safety limits? Safety limits are usually numerical criteria set by the regulator and met by the designer through analysis and experiment. DID is an add-on imposed for qualitative reasons when one is unsure if safety limits that do not include it are adequate. For example, a safety limit of frequency of release of radioactive isotopes can be met with one barrier, if it is strong enough, and vendors may advocate such a design. Why have more barriers? How does one determine the need for DID and when it is enough?

Another statement is “safety fundamentals defined in terms of protective strategies”. Safety fundamentals are concerned with measures of safety to determine if an adequate level has been reached, i.e. they are performance-based. Protective strategies are means of achieving this performance in various ways; they need to be compared and selected on the basis of criteria such as effectiveness, efficiency and simplicity of implementation.

A third statement is “enough safety margin to withstand unanticipated events”. It is hard to design to withstand something totally unanticipated, such as a visit of Martians or a swarm of bees in the control room. Safety margin is usually imposed as a result of experience with several unquantified influences and imagining possible scenarios, that are qualitatively recognized but are not included in the formal analysis.

Is it clear what “safety margin” is? Is there an unequivocal working definition of it? In pipe design it is derived from experience and because of an anticipated event, the pipe breaking. One might use an established code to design for 2000psi when the expected pressure is 1000psi because past experience shows that that was a good idea. Would safety margin be similar, designing for 1E-5 CDF in order to achieve 1E-4? What units

are used to measure safety margin? If the PRA includes the probability of the pipe breaking, is this already a measure of safety margin of the pipe against rupture?

### 1.5 RELATIONSHIP TO CURRENT LICENSING PROCESS

p1-5 “the DBAs...and PRA are important components of the safety analyses, but there is no direct link between these components”.

This only appears to be so because these two approaches grew up separately and different success criteria were developed for them. However, they both describe the same scenario. In principle, a DBA is no more than a PRA sequence with limiting assumptions made about some of the probabilities (and perhaps a different thermal/hydraulic analysis of the same event).

This is the point in the report where it could be made clear that the traditional role of DBAs is being subsumed into the PRA, which models all events with full technical sophistication. There appears to be no technical evaluation required except what is in the PRA. LBEs are defined, but they are not DBAs and fulfill none of the functions of the traditional DBAs. It is not clear to me what function they do perform.

This section should be expanded significantly to make clear what innovations are being introduced in the new framework, why they are there, and how they relate to and improve on existing methods. This would help to orient the reader, who now may have to wait till reading an Appendix (he may not get that far) to understand some of the new processes.

### STRUCTURING A FRAMEWORK TO MEET THE OBJECTIVES: SECTIONS 2 TO 5

After defining objectives, the next step in the design process is usually to analyze the objectives and determine, first in general terms, what is needed in order to meet them. This usually leads to a set of specifications and functional requirements that any design should meet, without specifying any particular regulatory design at this point. This step enables alternative ways of achieving such features to be imagined and evaluated. It appears to be missing from the report. The authors jump right into describing a particular structure.

I expected the staff in NUREG-1860 to start with a blank page and develop a more creative and optimum set of regulations that could be justified in a top-down sense from basic goals for public health and safety. I expected them to explain and justify the purpose of each piece of their framework, not simply describe it, and to rationalize why some other way of performing the desired function had been rejected.

The several objectives that I have suggested require:

1. a. Definition of *risk*, measures for it, assessment of applicability,

- b. Ways to evaluate *performance*, both at the design and operational stage,
  - c. Use of 1a and 1b for choosing *criteria* for evaluation.
2. a. Determination of what the NRC, particularly the Commission, wishes to use as *measures of "public health and safety"*.
- b. Review of *alternative measures*, particularly those used previously and those that appear suited for new designs.
  - c. *Coordination* with 1a, b, c so that the chosen measures and criteria fit these measures of safety.
3. a. Ways to assess *effectiveness and efficiency* of the new framework, at least qualitatively. Since there is often a conflict between these two figures of merit, it has to be worked out how to make this tradeoff.
- b. Ways to assess the *uses* of the framework and how *convincing* it appears,
  - c. *Logical threads* that derive the framework from the analysis of 1 and 2 above.
4. a. *Assessment of the effort and expense incurred by use of the framework* versus alternatives,
- b. *Assessment of other pros and cons* of using this framework or using other possible alternative regulations for new designs.
5. a. "Blind" *application* of the framework to some existing design.
- b. Determination of the *bases for comparison* with existing methods,
  - c. Comparison with existing methods, using some *figures of merit*.
6. Like 5, using available documentation for the PBMR.

In the above, I have indicated by italics some of the things that appear to need to be worked out, decided upon, and plans made for their resolution, before proceeding very far with the details of the design. In the NUREG these are incorporated into the several sections, when they appear at all, and have to be found there.

## *RISK*

Items 1a and 2a in the above list are the elements that provide rationale for the main structure of the proposed framework. Being risk-informed, or developing risk-derived results, implies having some measure of risk and ways to use it. It would be useful at this stage in the document to evaluate various ways of representing risk and determine which will meet the objectives most effectively.

### *Present uses of risk.*

“Risk analysis” today usually means the use of probabilistic risk analysis (PRA) to develop values of core damage frequency (CDF) and large early release frequency (LERF). They are obtained by adding up the values of frequency associated with the end-states on a PRA tree. Because there is a single binary outcome, core damage or not, the frequencies of individual outcomes can be added up without worrying about the extent of core damage. This makes for efficiency, but may lack effectiveness, since all core damage events are not comparable. The approach also has the virtue that CDFs are additive. If one has them for several events, addition gives the value for a class of events. Adding the CDFs from all events gives the net CDF for the plant.

CDF does not appear in most (perhaps not in any) of the regulations because they predate its development. It appears in guidance, such as RG 1.174, and in several recent staff and Commission statements. It is clearly useful and understood. It must be a candidate for evaluation as at least part of the basis of new regulations. If it is to be discarded, there must be arguments given, or perhaps some substitute identified. These could be developed by trying to define an effective measure of “core damage” for anticipated new designs, showing how it can be done, and evaluating how good it would be for regulatory purposes. The arguments given briefly on p3-3 require expansion to explain what is gained by the added complexity of using dose, a continuous parameter, as the measure of consequence, and why the added complexity is justified.

Looking at Objective 2, one would like to relate the chosen measure of risk to agency policy. The Commission has not usually been specific about using CDF this way, though one can find staff and Commission statements that suggest policies based on CDF. EPRI and the European Utility Requirements Documents (ACRS letter, Sept 21, 2005) have been more specific.

The Commission issued a safety goal policy statement some time ago (1986) that describes quantitative health objectives (QHOs). These have to do with the probability of death or latent cancers to particular individuals who spend time around power plants. They do not include total deaths, which seems strange from a public cost point of view, as one could then put a reactor in Central Park if it met individual risk criteria. Nor do they include additional ill effects, such as loss of property and environmental degradation. “Protection of public health and safety *and the environment*” is mentioned on p1-4. Consideration of the environment, and perhaps other societal risks, may be too important to be relegated to an Appendix.

Though there is a (rather dated) Policy Statement, it seems to have been little used for regulating, which presumably implies a reason. Current plants are not required to meet QHOs and there does not seem to be a requirement to assess whether they do, as part of the risk analysis, and to publish it. Currently, risk analysis develops values of CDF and LERF but does not assess probabilities of individual deaths or latent cancers.

The accident at TMI2 in 1979 was terrible from a core damage point of view but had little effect on public health and safety. It was the most traumatic setback to occur to nuclear power in the US and had great effect on public opinion about safety and NRC performance. Perhaps this indicates that some measure of core damage should appear in the new regulations, even if there is no Commission statement about it.

Regarding Commission statements about the level of safety of new plants, there is some uncertainty. One Commission expressed a desire for enhanced margins of safety (what does that mean?). Another, more recently, seems to have required a comparable level of safety. This may indicate that the expectations for future reactors need to be clarified.

The other Commission statement, quoted on page 2-2, is that “advanced” reactor designs will comply with the Commission’s safety goal policy. This would indicate that the QHOs are to be enforced for such plants, which represents a change in policy. On p3-3 it is stated that “current reactors ...in many cases achieve a level of safety comparable with the QHOs”, which might actually allow future reactors to be somewhat less safe than the current designs, depending on what “comparable” means.

Do these “advanced” designs cover all anticipated applications for the proposed framework? Would the framework also provide a more effective and efficient way to regulate existing designs, particularly if more of these are built?

The safety goals have to be cast in terms of *some appropriate overall cumulative measures of safety*, mathematically related to the outputs of the PRA, in order to be practically useful as the basis for a new regulatory framework that is risk-derived. The proposed framework describes criteria limiting frequency and consequence of individual PRA sequences but supplies no overall measure of “plant risk” so that designs can be compared and changes evaluated, as in RG1.174, or for power uprates, or during maintenance and so on, as is now done using CDF and LERF. This gap in the proposed framework can be rectified if risk is defined as dose times frequency for each PRA sequence. When this is added up on a plant-wide basis, it provides a metric of the total plant risk,  $R_q$ , for comparison with the latent cancer QHO. The fatality QHO can be related to the cumulative probability,  $F_q$ , of exceeding the fatal dose.

I’m not sure how the expressed desire to integrate security into the design framework fits with a risk-derived basis, since security events are not usually analyzed on this basis.

### *Ways to represent risk.*

PRA produces outcomes of a very large number of sequences. Each outcome has associated with it a frequency and consequences. In reality there are many consequences to society. If all of them are represented, the PRA is gargantuan and unwieldy. If too few are represented, it may be too crude for some significant regulatory purposes.

Generally a unique outcome  $i$  from a PRA has a frequency  $f_i$  and consequences  $C_{ij}$ , if there are  $j$  measures of consequence. CDF is an example of a simple binary outcome that involves no particular measure  $C$ .

NUREG-1860 uses dose, a continuous variable, as the measure of consequence. This would make the PRA more complicated and more expensive to produce than it is today, as the level of dose would have to be calculated for each sequence. What justifies this increased level of detail? Presumably it is because a measure of consequence is needed to relate the PRA outputs to the QHOs, involving deaths and cancer. The QHOs can be met by computing these measures directly without use of any sort of F-C curve.

There are many consequences to society besides people receiving doses of radioactivity, some of which lead to individual deaths and cancer. These include total deaths, total health effects from all causes, loss of property, environmental degradation, evacuation costs, and loss of faith in nuclear power (even with very small actual health effects, e.g. TMI2). Therefore the outcome of a given PRA sequence could be represented on a  $j+1$  dimensional diagram, with one axis being frequency. This would be a huge step away from the binary CDF and LERF measures. It would be a more effective representation of real societal risk, but might not be at all efficient in use.

All of the consequences, though they have different measures, can be combined into a single parameter by using weighting parameters,  $w_j$ , so that the net consequence is  $C_i = \sum w_j C_{ij}$ . The most obvious weighting parameter is equivalent dollar cost. There is perhaps much to be said for developing such parameters and not focusing regulations entirely on one or two measures of individual health risk.

The choice of which description to use for consequences is key. In NUREG-1860 the authors do not evaluate alternative possibilities but simply state that they use “dose”, which introduces a single particular measure of consequence,  $C$ .

### *Use of a single consequence measure, $C$ .*

Suppose that one wishes to take a single step towards a more complete description of risk by introducing one continuous measure of it, rather than using a binary outcome like CDF.

Then each PRA sequence has a frequency  $f_i$  and a consequence  $C_i$ . How can they be used in regulation? The designer has to make selections in terms of how to formulate composite measures in order to satisfy his objectives. This is not a trivial matter. Using

a continuous variable as a measure of consequence changes the PRA fundamentally. Since each end-state has frequency and consequence paired in a two-dimensional continuum, one can no longer add up frequencies and consequences independently, but must make suitable definitions in order to represent cumulative effects that describe the safety impact of the plant or of certain classes of accident.

One could try to develop regulations based directly on the frequency and consequences of each PRA sequence. This appears to be what the authors are trying to do in the NUREG because page 6-6 says “with the kind of acceptance criterion for individual sequences described above”. This implies regulating each one of many outcomes that have any consequences. It depends on the structure of the PRA tree, which can be manipulated to change the frequencies, as desired. It is not directly compared with the QHOs, provides no indication of the overall risk status of the plant, and is not evidently either efficient or effective.

If one were to use all the  $f_i$  and  $C_i$  as a basis of regulation, this would be a very fine-grained approach. Since the intent can hardly be to regulate every PRA outcome, there has to be some way of combining these for more practical purposes. One cannot simply add up the frequencies and consequences to get an effective  $f$  and  $C$  for an event, accident, or accident class. As explained in the Appendix (The NUREG could do with an Appendix, or even a section in the text of this nature, describing the various ways to aggregate frequency and consequence), additive properties can be developed in terms of F-C curves, or alternatively by aggregating “risk”, defined as frequency times consequence. One then has measures of the probability  $F$  of exceeding some specified consequence  $C$  in the entire group of outcomes under consideration, or alternatively of the total risk involved from any aggregation of the PRA outcomes.

The F-C curve can also be used to evaluate the total plant risk and the cumulative frequency of consequences within some specified range of interest, such as “small” and “large”. The aggregated risk can also be used to evaluate subdivisions of events or to assign allowable risk over specific ranges of consequences. These are practical measures that can be used at any level of subdivision for each plant risk profile. They could provide simple and clear ways of explaining to the informed public the rationale behind regulatory decisions and for comparing different nuclear power system designs, as well as for regulating day-to-day operation.

I expected the staff to devote an early and key part of the document to Top Level Regulatory Criteria (TLRC). They are the basis for everything that follows. In principle, a design that meets them is acceptable. Subsidiary requirements merely serve to reinforce the confidence that the staff and public have in how well they are met. Before designing any regulatory framework there has to be agreement that the TLRC are an adequate set, are consistent, are truly technology neutral, and reflect actual public safety concerns. (For example, why are some on a “per event” basis while others are on a cumulative dose basis?) Then, in a top-down approach, everything in the framework of more detailed requirements, such as LBEs, needs to be justified as the optimum way to enforce the basic TLRC.

In setting the stage for their framework for 10 CFR 53 in Section 1 of NUREG-1860 the staff mention that it will interface with other parts of 10 CFR. They present a diagram, Figure 1-1, showing the connections but do not go into details of aspects which might influence the design of the framework. At no point do they review the present 10 CFR to extract and establish high level regulatory criteria which are appropriate for carryover to the new part 53. This is a major defect. Up until the introduction of Figure 6-2, the only quantitative TLRC that are mentioned are the QHOs.

I looked over 10 CFR and found it singularly devoid of generic quantitative high level design criteria. There is a 10mrem/yr ALARA limit in Part 50.34, which exceeds the 5mrem/yr quoted in Figure 6-2 of NUREG-1860 from 10 CFR 50 App 1. There is also a 100mrem/year public dose limit in 10 CFR 20. Both of these exceed the cumulative dose limit of 4mrem/yr which results from dividing the QHO latent cancer risk of  $2e-6$ /yr by the latent cancer fatality risk coefficient of  $5e-4$ /rem. If the QHOs are indeed the only TLRC, there may need to be an evaluation of whether they are a sufficient set to form the basis of an entire framework of new regulations.

It would help NUREG-1860 substantially to state in Section 1 that the framework is designed to meet a set of explicit TLRC. These should be spelled out and critiqued to assess if the set is adequate. If they are the only legal requirements that establish a quantitative performance base for future reactors, it needs to be explained why the framework follows logically in response to them.

## SECTION 2 FRAMEWORK OVERVIEW

This section outlines the structure of the report. It might be expected to explain how the elements of the framework meet the expressed objectives, particularly numbers 1 and 2.

P2.2 "...integrates the NRCs expectations for safety, security and preparedness to achieve the desired level of safety". The only safety expectations that are actually mentioned here specifically are the QHOs. Are there no other functional specifications for "desired levels" to be met?

For the security expectations the only design specification seems to be that "the overall level of safety should be consistent with the Commission's expectations for safety from non-security related events". Other items listed are qualitative or describe the means to achieve expectations, such as the DBTs, without defining their function. Are security requirements then supposed to be derived from the QHOs?

For preparedness expectations the discussion seems to indicate that they are not part of the framework, though the first statement above said they were "integrated" into it. "Making emergency preparedness more risk-informed and performance-based is a possibility" does not indicate that how to do it was considered and recommendations made.

P2.3 Defense-in-depth is said to be a “safety philosophy” which makes it difficult to define performance bases for it.

P2.4 “..incorporating successful past practices and lessons learned”. Can these lessons be made more explicit? The purpose of the framework is to rethink and devise an improved structure for regulation, rather than to perpetuate past practices.

Figure 2-2 may be misleading, giving the impression that the two legs “deterministic” and “probabilistic” are given the independent complimentary treatment typical of present regulatory practice. It appears evident from Chapter 6 and Appendix E that the framework is based entirely on the PRA and contains nothing equivalent to the traditional “deterministic” analyses, such as those found in Chapter 15 of LWR SARs, which formed the basis of regulatory decisions for decades, until being risk informed by also considering PRA results, as appropriate.

P2.5 “top-down hierarchical approach, starts with a desired outcome and identifies protective strategies to ensure this outcome is achieved even if some strategies should fail”. These five strategies are not really performance-based, but describe various means to achieve a desired end. The top-down “desired outcome” is to “ensure public health and safety”, therefore it needs to be explained how this is to be demonstrated in a performance-based way.

P2.6 The five strategies are called objectives, which seems to mix terminologies.

Section 2.5 describes design criteria and design objectives which “provide overall goals that the protective strategies are intended to meet”. The only specific objective that is cited here is to meet the QHOs.

P2.7 The F-C curve, LBEs and SCCs are means to an end, not really “design objectives”. They might be justified in terms of suitable more fundamental performance-based objectives, or in terms of defined needs for enforcement, links that appear tenuous in the present report.

In discussing PRAs it is explained how they are related to several elements of the framework but, strangely, no demonstration is provided of how they help meet the only high level objective that has so far been explicitly developed in the report, the QHOs.

What are the “licensing risk criteria”?

P2.11 Figure 2-5 appears to be inverted. “Protective strategies” are not a design objective but a means to an end, which is satisfying some suitably defined safety criteria. It would seem more logical to put Section 6 up front and put qualitative discussion of acceptable strategies in guidance documents.

P2.12 “Within each protective strategy an approach can be taken that specifies certain deterministic requirements to help account for completeness uncertainties”. These “deterministic requirements” are a matter of judgment and may involve policy decisions, such as requiring that every design must have a containment. Since all PRA sequences are already analyzed deterministically, this feature alone is not the answer to incompleteness.

*Summary.*

This section should be a more “design-neutral” collection and analysis of what performance-based criteria can be derived in support of Objectives 1 and 2. As presented it mentions the QHOs as the only agency objective to be satisfied, which can be done much more economically than is described in the report. There is no development of subsidiary objectives, such as ALARA, which appear in Section 6, nor an assessment of how they compare with and complement the QHOs. The rest of this section is devoted to discussion of strategies, which are more the province of the designer and could fit better in guidance documents.

Elements of my Objectives 3 and 4, involving considerations such as efficiency, effectiveness, complexity, and suitability for convincing an informed public are mentioned in passing at times in the report. Such statements appear empty, as there seems to be no place where they are used to make choices about what to include in the framework and how to implement it.

### SECTION 3 SAFETY, SECURITY AND PREPAREDNESS EXPECTATIONS

This section is concerned with further development of ideas and specifications for the framework based on the “overall level of safety demanded by the NRC” (p3.1).

P3.2 A key new feature of the framework seems to be clearly stated as the treatment of safety goals as actual regulatory requirements. This excludes the three region approach. Why is it presented in Figure 3-2 as if it were a feature of the framework? It is then dismissed on the next page but invoked further down the page as demonstrating “margins”. There seems to be significant ambivalence. Is the figure ever used later in the document for some purpose?

P3.3 “The current PRA technology is relatively mature”. Is it mature if the output is a continuous variable such as “dose”?

P3.4 The replacement of CDF and LERF by Figure 3-3 is a huge development which cannot be justified in a few paragraphs that supply little rationale. What is the justification for “imposing additional constraints in addition to satisfying the QHOs” when these are the only basic safety objectives that have been quantitatively articulated up to this point? How is imposing criteria on the outputs of individual PRA sequences justified? The explanation on p3.6 that “it has been established to support achievement

of the overall safety objective” provides no justification for this microscale regulation at the level of individual sequences. Is it really “anchored in the safety goal QHOs” when these can be satisfied by evaluating overall risk without the need for any such process?

Figure 3-3 needs to follow logically in some way from some aspects of the articulated objectives. It appears to be pulled out of the air.

The discussion in the middle of p3.6 seems to indicate that additional calculations, besides what can be developed to satisfy Figure 3-3, are needed “to ensure the QHOs are met”. Then how does this key feature of the framework, Figure 3-3, help to satisfy the single pair of safety goals that has so far been treated as fundamental?

“Surrogate risk objectives” are mentioned as being useful. Are they developed further in any way in this document? How is their established “usefulness” over many years satisfied by some other part of the framework? What metrics will be used in Regulatory Guide 1.174, or some equivalent, for new reactors?

P3.7 A “risk-informed approach is to be taken to security”. Is this developed in the document? Does the framework provide performance-based criteria for security besides the discussion of desired qualitative features?

As in Section 2, preparedness is given only a discursive treatment.

#### *Summary.*

As in Section 2, there are insufficient linkages between the agency objectives and the structure that is described. Figure 3-3 comes out of the air and appears unrelated to any of the expressed goals.

## SECTIONS 4, DEFENSE IN DEPTH and 5, SAFETY FUNDAMENTALS

Both chapters provide informative discussion of design features and strategies that might play important roles in achieving adequate public safety.

They provide little substance by way of risk-informed or performance-based criteria and methods.

They are too qualitative for inclusion as part of a “framework” where the emphasis is more appropriately on essential functional requirements, leaving the engineering designer latitude for creativity in meeting them. The NRC is not in the design business.

These sections appear to be more appropriate as part of guidance documents.

A more useful Section here, helping to introduce Section 6, would be derived from Section 8 and Appendix F. It would explain the central role to be played by the PRA,

how its technical analysis is sufficiently upgraded to remove the need for further analysis and so on. It might also derive generic functional requirements for new reactors that are more clearly related to the features sketched out in Section 6 and would provide a rationale for them.

## SECTION 6 DESIGN CRITERIA AND GUIDELINES

This is where the framework is made quantitative. Up to here the document is mostly descriptive (too long and detailed?) and includes little in terms of definite measures of performance, so it is hard to identify performance-based criteria to which this section responds.

It is perhaps the most important chapter in the NUREG.

It starts by repeating the objectives, somewhat differently phrased. This is unnecessary if an appropriate set has already been articulated.

### *Figure 6-2: The f-C curve*

Figure 6-2 is the basis for the framework and is the centerpiece of the report. It is called an F-C curve, which is misleading as it is not the classic F-C curve, which involves cumulative probabilities. It is used to define the upper limit of allowable outcomes of individual PRA sequences. Since these have a frequency  $f_i$ , one could call the allowable upper bound  $f$  and plot it versus consequence,  $C$ , as the staff have done. I will therefore call their curve the  $f$ -C curve. End-states of the PRA will appear as a swarm of individual points on an “f-C map” in the region below the curve.

Each range of dose “is assigned” a frequency on page 6.3 and in Table 6-1. The first one is said to be derived from the ALARA cumulative dose limit of 5mrem/ry. It is not explained how this leads to the decision that “doses in the range of 1mrem-5mrem are assigned a frequency of 1 per year”. It appears from the figure that one dose of 5mrem in a year is allowed and is sufficient to reach the allowable cumulative limit. Since the criterion is applied to all the (tens of thousands?) of PRA end-states, there may well be many that correspond to doses in the 1 to 5mrem range. Therefore this is not a way to meet the ALARA criterion, which corresponds to a cumulative yearly dose.

The next range of doses, from 5 to 100mrem is assigned a frequency of 1E-2/ry. The public dose limit of 100mrem/ry is a cumulative dose. Using the criterion in the figure, one hundred individual doses of 100mrem each with a frequency of 1E-2/ry, would be needed to reach the 100mrem/ry annual limit. This is a different treatment than was accorded the ALARA range, where one end-state, rather than one hundred, could cause the cumulative limit to be reached. A different logic appears to be being used to derive the two limits.

It makes sense to have the allowable frequency decrease with dose, but clearly some other reasoning is at work than is explained in the text.

There appears to be a basic anomaly about specifying a cumulative ALARA dose of 5mrem/ry over a range of frequency of five and a cumulative dose of 100mrem/ry over the next range of twenty in frequency. If the regulations were “risk neutral” the cumulative risk over each order of magnitude of frequency would be constant. Over a range of frequency it would be proportional to the logarithm of the ratio of the lowest to the highest frequency in the range.  $\log(20)/\log(5)$  is less than 2. In this case the higher dose range is acceptable with 20 times the value allowed for the lower dose range. This is counter to the risk neutral basis by over a factor of ten and even more at odds with a “risk averse” approach, in which risk decreases with dose, which has sometimes been suggested to reflect public preference.

The next downward steps in the curve are related to increasing severity of an event, which triggers different responses from the licensee and the regulator. Though these are indications of the NRC’s view of the importance of such events, there is no basis in the arguments supplied in the text for the assignment of specific frequencies to them until the range 300-500rem, where the NRC’s early fatality safety goal is invoked.

Another place where similar frequencies are mentioned in the document is in reference to CDF and LERF, which are cumulative measures. Perhaps the values of 1E-4 and 1E-5 are being transferred to this figure because of some unspoken association of different levels of dose with core damage or significant release? In any case, such criteria should refer to cumulative probability of all events rather than being a criterion for each end-state on the PRA trees. If there are ten sequences, each producing 10rem, from the tree for a single accident type, each with probability 1E-4/ry, then the net risk from this accident would be 1E-3rem/ry. If there are ten accident types like this one, the net plant risk would be 1E-2rem/ry. Public risk is better defined in terms of the total risk from all accidents and should not be allowed to accumulate depending on how many kinds of accident are defined or how the PRA trees are subdivided.

At the right hand end of the curve the dose causing early fatality is reached. The region between 100 and 300rem is given the frequency 1E-6/ry, which might be related to the corresponding QHO criterion, though this is not offered as a rationale in the text. On page 3-5 the Safety Goal Policy Statement is said to correspond to 2E-6/ry individual cumulative risk of latent fatalities, which would only allow two sequences to approach the limit of 1E-6/ry. Again, this needs to be expressed in terms of the total plant risk since more than one end-state can give doses in this range.

Between 300 and 500rem the “assigned frequency” is 5E-7/ry “to meet the NRC early fatality goal”. Shouldn’t this be a cumulative frequency for doses above the fatality threshold, not the frequency of individual outcomes?

The curve is capped at 500rem. It appears from Figure 6-2 that the curve ends there, but it should be continued along the axis to show that doses larger than 500rem are possible

and must have a frequency below  $1E-7/ry$ . This may be a reasonable way of capping individual risk at the site boundary, or somewhere else such as ten miles away, but it does not make much sense from the point of view of real public risk from the largest possible accidents, such as at Chernobyl in 1986. The accident affected much of Europe. The fact that people would have died at the site boundary is no measure of the severity of that event.

Though the QHOs are given prominence at the outset on page 6-1, they are not used to create any part of Figure 6-2, which is surprising.

The early fatality QHO goal is  $5E-7/yr$ . It is a cumulative frequency goal. It is not clear how this relates to the  $1E-6$  frequency assigned to 100-300rem, where “the threshold for early fatality is exceeded” as well as the QHO frequency. 300-500rem is assigned a frequency of  $5E-7$ , allowing only one event in this range is the goal is to be met. Above 500rem the allowed frequency is  $1E-7$  for all events, no matter how severe, though the number of fatalities would be expected to play a role in public acceptance.

On page 6-7 it is claimed that accident sequences that lie below the  $f$ - $C$  curve will satisfy the QHOs of the safety goal policy individually. There is no clear demonstration of how this statement is justified in relation to the latent cancer fatality goal. It would appear simple to do so. Using the latent cancer risk of  $2E-6/yr$  from Appendix C and dividing by the latent cancer fatality risk coefficient of  $5E-4/rem$  gives a cumulative dose limit of  $4E-3rem/yr$ . This appears incompatible with the ALARA range at the low end of the frequency spectrum, where a single dose of 5mrem will use up all of the allowable cumulative dose and leave nothing available for more severe events.

The cancer fatality QHO is equivalent to a total expected dose, or expected risk, per year,  $R_q = (\sum f_i C_i)_q$  if the zero threshold assumption is adopted. If some other assumption is used, the summation starts at the threshold. For a single outcome this is simply a certain value of frequency times consequence, the cancer risk,  $(fC)_c rem/ry$ . For multiple outcomes it is the sum of their individual risks. The curve in Figure 6.2 lies almost entirely between lines of constant risk,  $fC=1E-4rem/ry$  and  $fC=1E-3rem/ry$ . These values are easily compared with the value,  $R_q$ , corresponding to the QHO of 4mrem/yr. This gives an idea of how many individual outcomes can be allowed to approach the curve, since the sum of their risk values must be less than  $R_q$ .

Specifying limits to be satisfied by each PRA sequence allows a great deal of flexibility to not meet the QHOs on a total plant basis. Figure 6-2 is given prominence as a tool for evaluating individual sequences while satisfying the QHOs is a criterion mentioned in passing on the middle of page 6-7. I would expect satisfying the QHOs to be the primary acceptance criterion. How it is achieved is up to the designer, in a performance-based regulatory system.

*If the QHOs can be met by requiring that a limit of cumulative expected dose be met below 500rem and that a cumulative frequency be met for doses above 500rem (or some conservative lower value), what is the rationale for devising more elaborate criteria based on the f-C map?*

Figure 6-2 seems to be partly justified as a tool to create licensing basis events (LBEs), but the need for such a category and its use in regulation needs to be explained.

#### *Use of a classical F-C curve*

If it is truly necessary to specify dose versus frequency, rather than overall measures of compliance with the QHOs, the authors might have been better advised to introduce a classical *F-C* curve, describing the allowable total plant risk profile. In this case “F” is the cumulative frequency of events with consequences greater than C. This could form the primary basis of regulations, instead of what is presented in Figure 6-2, or be used to supplement it. I have used the italic *F* to distinguish the regulatory safety limit (not “assured” by DID as stated on page 1-3) from the actual value of F for the plant.

The cumulative ALARA dose of 5mrem/yr, if assumed to be resulting from end-states with consequences in the range 1 to 5mrem, would be represented as the area to the left of the *F-C* curve between those dose levels.

The cumulative public dose of 100mrem/yr, if appropriate and not incompatible with the latent cancer QHO, would be represented as the area to the left of the curve in the range 5 to 100mrem.

The individual cancer risk, according to the zero-threshold hypothesis, is proportional to the total dose and the constant of proportionality can be determined. If there is believed to be a threshold, it could be used to select where to start the “dose” on that axis. In terms of all events at the plant, the expected value of the yearly dose, or the net plant risk, gives the individual cancer risk. This dose is presently derived at 10 miles from the plant, which may or may not be the “dose” used by the staff on their figure. If it is, then this QHO is satisfied by having the area to the left of the entire *F-C* curve (or the area starting at the cumulative frequency of the threshold dose) equal to this acceptable cumulative dose, per year. Therefore this QHO is directly related to the curve.

If the dose axis refers to dose at some other location, it would have to be transformed appropriately to give the 10 mile dose.

The prompt fatality risk can similarly be represented on an *F-C* curve. If a dose greater than 300rem (or is it 500rem?) is fatal, then the cumulative frequency of outcomes of accidents causing larger doses should be limited to agree with this QHO. This is simply  $F_D$ , the value of *F* at the dose causing death.

If some other criterion is considered, to limit the frequency of any single “accident” or class of accidents yielding doses greater than some chosen value, this is represented by the corresponding value of  $F$  at that dose.

Use of F-C curves allows adding up the effect of subgroups of sequences, events, accidents, and accident classes, since the  $F$  parameter has an additive property. This is not easily performed using  $f$  as the measure of individual sequences because it is paired with a value of  $C$  in two-dimensional space. Use of an additive property is probably a necessary specification for ensuring an internally consistent set of regulations.

### *Other key decisions*

Deciding on a suitable representation of allowable plant risk profile is just the beginning of a long haul. All the items italicized earlier when discussing the objectives (or some similar list) need to be addressed. Details of implementation need to be established at some stage, though this may be too great a level of detail for a “framework” document.

Regarding effectiveness and efficiency, some decisions have to be made about how detailed the regulations need to be. At the simplest level, one could simply say that any design must meet the QHOs. At the next level of detail, it could be required that the plant have a risk profile lying below some defined  $F$ - $C$  curve. This defines the level of expected overall performance. It is very efficient in terms of brevity, but may need more detail to reach a sufficient level of effectiveness. There may also be some safety objectives that do not fit description only in terms of cumulative frequency of dose at the site boundary.

Does anything further need to be defined about acceptable performance regarding specific accidents or accident classes, or about the key SSCs that are involved in the most risk-significant sequences? Should it be required that the PRA be more detailed and accurate in its modeling of the most significant sequences? At what level of detail should the NRC specify how licensees should meet the basic risk acceptance criteria?

Decisions about the appropriate answers to these, and many other primary questions, need a thorough justification in terms of the benefits and costs of imposing regulations based on them. The staff makes little effort to justify the details in its design except for describing them. Some of these may be justifiable in the light of experience, but some may not be. If future regulations are to be clearer, more effective and less burdensome to the NRC and the licensees, new alternatives to achieve these objectives need to be imagined and evaluated.

In line with my list of italicized items earlier, I would recommend working through them and trying to conceive of ways to perform those functions, both using present methods and possible improvements. The more important items might need at least the level of attention that I have given to risk above.

A hierarchical structure to the framework might help. First show how the top level criteria, such as the QHOs are to be met. Then explain how a traditional F-C curve is to be met at the level of accident sequences, classes and the overall plant response. Then get down to the level of regulating individual sequences, if this is necessary.

*LBEs: Section 6.4*

Two purposes for LBEs are described. The first (major) one is “to provide assurance that the design meets the design criteria for various accident challenges with adequate defense-in-depth (including safety margin) to account for uncertainties”. I do not understand how this objective is met.

What are the “design criteria for various accident challenges”? Figure 6-2 applies to individual end-states in a PRA. What are “accident challenges? I thought they would be a cluster of PRA end-states corresponding to some common feature, such as the initiating event? The properties of such a cluster have not been established in the report; the appropriate criteria will depend on what definitions are made.

I am unsure how DID and safety margin (p6-9) can be assessed in “performance-based” regulation without very clear definitions of what these are and how to quantify them.

I do not understand what distinguishes the LBEs, what purpose they serve, how they are evaluated, and how they are used. It appears that all of their functions can be performed directly by using the PRA sequences themselves.

On page xi it is stated that “the consequences for each event sequence from the PRA and each event sequence selected as an LBE must meet the  $f$ -C curve”. Since an LBE is a PRA sequence, what is being added by defining this subgroup that meets the same criteria?

Up to page 6-5 there is no mention of uncertainties, so presumably what has to lie in the “acceptable region” is the entire cloud of the myriad results from the *point estimate* PRA.

The discussion on pages 6-8 and 6-9 addresses “probabilistically selected LBEs”. Nothing on these pages indicates that LBEs are selected based on the statistics of uncertainties in the PRA, so the reader is likely to assume, as I initially did, that LBEs are selected based on the statistics of the cloud of point estimates, which also have means and 95<sup>th</sup> percentiles.

After a very long time I came to the conclusion that what must be going on is that the PRA sequences end, not with a single frequency, but with a range of frequencies and consequences (determined how? By Monte Carlo analysis? This represents an extra computational burden) from which the means and 95<sup>th</sup> percentiles can be determined at some confidence level. This is not explained in the text, but it appears to be implemented in Appendix E.

Using statistics to evaluate uncertainty gives a lot of opportunity for creative manipulation of results. For example, if one puts only the uncertainty of pipe break frequency, based on the expert elicitation, into the DEGB analysis, the 95<sup>th</sup> percentile value may be an order of magnitude or more different from the mean. If this is carried through the PRA, a similar range of frequency outputs is obtained. Is it intended that similar ranges will distinguish some LBEs and move a point by orders of magnitude on the f-C map? One can make the huge range of pipe break frequency less important by combining a range of pipe sizes into a LBLOCA, bringing in other uncertainties, and doing a 95/95 analysis in which the extreme pipe break diameter might not show up.

The LBEs seem to be introduced to take account of the uncertainty in frequency. The “more stringent criteria” on page 6-8 are presumably *statistical* criteria, though they could be requirements for better thermal/hydraulic analysis or something else that would improve confidence in the modeling of the event. This is not explained. It is also not explained how one can evaluate a suitable mean frequency without taking account of consequences, since the frequency of higher consequences is presumably more important than the frequency associated with lower consequences. How does one define 95<sup>th</sup> percentile in a two-dimensional space of frequency and consequence?

A series of Steps in the LBE selection process is presented in Figure 6-3. It would be useful to have more discussion and rationale for why these are needed, what function they serve, and what alternative, possibly more efficient and straightforward, approaches to satisfy this function, were rejected.

Step 1 is to credit only safety-significant SSCs. What is the point of this? Is it to make the PRA simpler by discarding pieces that have no significant effect on f and C, thereby reducing the computational effort? Without some measure of significance, it is unclear how to decide what to discard. By including more detail the PRA is more complete and effective. By excluding some detail it becomes more efficient. How is the tradeoff made?

On page 6-18 it is stated that “the term ‘safety significant’ is assigned to those SSCs whose functionality plays a role in meeting the acceptance criteria imposed on the LBEs”. Something circular seems to be going on, whereby Step 1 selects based on the properties of LBEs that are not constructed until Step 6. How does one measure “plays a role”?

The process described on page 6-19 for determining risk importance appears based on the entire PRA and does not seem to make use of the LBEs.

Step 2 is to discard sequences with point estimates  $<1E-8$ . This might lead to discarding something like the DEGB, which could leap into significance if one evaluated its 95<sup>th</sup> percentile, perhaps several orders of magnitude higher.

In Step 3, the mean and 95<sup>th</sup> percentile frequency are determined. Of what? I at first thought this meant the mean of several frequencies of a set of end-states corresponding to the same “event”. Perhaps it does? More likely, it refers to the statistical spread in a single end-state, as discussed above.

Though it is perhaps imperfectly stated, I assume that the 95<sup>th</sup> percentile of the predicted frequency of a single unique sequence is to be computed, based on the uncertainties in the PRA. If these uncertainties lie solely in the probabilities of taking various branches in an event tree of a physically deterministic scenario, then the uncertainty is only in frequency. If the uncertainties include physical properties, such as uncertainty in the size of a pipe break in the LBLOCA category, then the consequences are also uncertain unless the choice of size is built into the logical tree structure to define separate unique sequences.

Model uncertainty, as in the thermal/hydraulic phenomena, further influences consequences as well as frequency. It could also cause a switching from one “sequence” to another, ending up with an entirely different end-state and significantly different consequences (e.g. switching the particle size of CalSil arriving at a screen could switch the probability of achieving sufficient NPSH). The question “95<sup>th</sup> percentile of what?” might require a careful answer.

It is possible that when model uncertainty is included, 95% of the results lead to no consequences while 5% with a suitable combination of circumstances lead to disaster. How does one compute a meaningful 95<sup>th</sup> percentile? In two-dimensional space there are various possible measures of 95<sup>th</sup> percentile. One measure is 95% confidence that there is 95% probability of not crossing the  $f$ -C curve by random probabilistic displacement in any direction from the point value.

In Step 5, page 6-11, the LBEs are chosen by grouping similar accident sequences into an event class”. These “still satisfy the  $f$ -C curve” and there are more of them for “higher dose sequences”.

So, what has to satisfy the curve? Is it the 95<sup>th</sup> percentile of frequency and consequence for a single PRA sequence in this “class”? (There appears to be nothing in the text about using the 95<sup>th</sup> percentile to satisfy the curve until page 6-16). Is this selected as the one that comes closest to the curve with its mean value(s) or with its 95<sup>th</sup> percentile value(s)? How many LBEs are there per class? There seems to be an unstated implication that the 95<sup>th</sup> percentile of *all sequences* must satisfy the curve (they are all computed in Step 3). In that case, there is no need to require, as earlier in the chapter, that the point values do so.

Perhaps what is being done by defining event classes is to take the large number of end-states and group them into clouds on the  $f$ -C map that each contain fewer points, but still involve a large number. The emphasis seems then to be on outliers from these clouds, not on any cumulative properties, which might be more safety-significant measures. As

mentioned earlier, these outliers can be manipulated by suitably designing the structure of the PRA.

Those events about which there is large uncertainty will move around significantly on an f-C map, both in frequency and consequence, when inputs are varied randomly. It may be that some sequences with point estimates that approach the limits will still be OK, whereas those that were orders of magnitude away will now exceed the limits in extreme cases. The method for choosing LBEs, though not clear to me from Step 5, seems to be based on the latter. This is perhaps a good way of showing up sequences that may require consideration for improved modeling or DID add-ons, but it appears to neglect those sequences which are more “risk significant” on the average. How does this fit with the introductory statement “LBEs represent all potentially risk significant accident challenges” on page 6-8?

Since the NRC regulations refer mostly to cumulative effects of many accidents and of a myriad of PRA sequences ( $\sim 1E-5$  for the extended scope PRA envisaged?) it is possible that what is going on here is an attempt to accumulate frequency within a class and a range of consequences. If so, how do groupings of outcomes satisfy an f-C curve that is a limit on an f-C map describing individual PRA sequence outcomes? These individual values cannot be simply added up to give an effective f and C for an event or class of events. Besides, when many are added up, the resulting f and C, if they can be defined, are likely to be significantly greater than the individual values, so it might make sense to revert to a simpler “limiting curve” defined in terms of cumulative variables, such as risk.

What are “higher dose sequences” in the context of events and classes. Individual end states have consequences and some of these involve higher doses. Many “events” may have one or more end-state with a high dose. How many higher individual doses does it take for the “class” to be selected for more LBEs? Should these “higher doses” be cumulative in some way?

Step 6 includes terms such as “bounding consequence”, “selected event sequence”, “bounding event”, and “frequency of the event class”. None of these is defined and as a result it is impossible to determine how the authors propose to aggregate or select from the large number of individual values of  $f_i$  and  $C_i$  in order to evaluate them. Presumably each event and each event class include many individual PRA end-states. How are events compared to find the “bounding” one? Is it the one with the greatest cumulative risk, the greatest maximum value of C, irrespective of its frequency, or what? Is it the largest mean value or the largest 95<sup>th</sup> percentile value?

In Step 7 “the LBEs mean frequency is the highest mean frequency of the event sequences in the event class”. Mean frequency of a set of event sequences has not been defined. Is it independent of consequence? If there are N end-states in the corresponding PRA, is the mean frequency  $\Sigma f_i/N$ ? How can this mean anything when it is independent of consequences? Besides, the PRA event tree can always be subdivided to get N as large as desired while not changing  $\Sigma f_i$  much, so the result can be manipulated.

Perhaps “mean frequency” means something else?

Step 8. “The LBEs have to meet the  $f$ - $C$  curve”. What does this mean? Is it a cloud of points, the maximum points, an envelope, some average, some aggregated values, or what that must lie in the acceptable region? If the LBEs meet the curve, are there some sequences whose 95<sup>th</sup> percentile, if it can be defined, does not meet the curve?

The simplest interpretation is that LBEs are simply “bounding events” from the whole set of PRA sequences that are already required to meet the curve at the 95% level. If so, then they already meet the  $f$ - $C$  curve.

The same problem with undefined properties of event classes occurs on page 6-16 with the statement “The PRA has to meet the  $f$ - $C$  curve in terms of the mean with respect to frequency and dose of the various event classes in which the accident sequences are grouped while the LBEs have to meet the  $f$ - $C$  curve with the 95% probability value with regard to both frequency and consequences”. Not only is it unclear how the frequency and dose of event classes are defined, but it is not explained what is meant by the mean value and 95% probability value in the two-dimensional  $f$ - $C$  space. It is meaningless to average the  $f$  and  $C$  values independently, since a value of  $f$  has much more risk significance if it is associated with a high value of  $C$ , and vice versa. In any case, cumulative frequency may be more meaningful than any “mean” at a certain value of  $C$ .

Step 9 adds to my confusion.

“at what level are the selected sequences defined: cut-set, systemic, or functional?” (Page 6-12). How are these three alternatives related to the PRA structure? Is the term “cut-set” defined when the consequences are expressed as a continuous parameter such as dose?

“The LBEs are selected at the ‘systemic’ level in terms of front-line systems that provide the needed safety functions” (page 6-12). What does this mean unless it is related quantitatively to the PRA results? Functions are “needed” if they help to reduce some suitable aggregation of frequency and consequence at some specified level. This implies the existence of a measure of “need”.

On page 6-13 “other sequences besides the ones shown in Figure 6-4, which belong in the same event class, will contribute in terms of frequency to the LBE frequency in that class”. This seems to imply that the LBE frequency might be the sum of a set of frequencies. Does this mean that there is some consideration of cumulative frequency involved in this process?

On this same topic, I am puzzled by the “Additional Dose Criteria” in Section 6.4.2.3. “an additional requirement for the LBEs with frequencies greater than  $1E-3$  per year (why just in this range?) is that they meet the cumulative dose requirements”. “This means a frequency weighted summing of the doses of all the LBEs in the range”.

This process seems very odd. Cumulative doses to the public must include all sequences, not the sample represented by LBEs, which apparently are a small set of all single PRA sequences.

At the bottom of page 6-15: “Specifically, the use of sequence specific source terms requires the applicant to do sufficient testing to confirm (a list of things)”. This implies a significant burden (Level 2 or 3 PRA?), since it must presumably be done for all sequences with any consequences, in order to plot the results on Figure 6-2.

P6-16. “The deterministic LBE event is to be analyzed mechanistically to determine the timing, magnitude and form of radionuclide released from the reactor building. ....established such that the worst two-hour dose at the EAB and the dose at the outer edge of the LPZ for the duration of the event do not exceed 25rem TEDE”.

Apart from the TEDE criterion, isn't such an analysis required for all PRA sequences in order to plot them on Figure 6-2?

P6-17. Table 6-3 states that all sequences must meet the  $f$ -C curve with the mean value and the LBEs must meet it with the 95% probability value (determined with what confidence?). This indicates that a statistical uncertainty analysis is needed for every PRA sequence, a significant computational burden.

If LBEs are truly bounding, it would seem that all PRA sequences will meet the same criteria as the LBEs do. In that case one could work directly with the PRA sequences without the need to define a separate set of LBEs?

Section 6.4.2 states (p6-27) that “A risk-informed and performance-based approach has been taken in the development of security performance standards”. Does this mean that there has to be a “security PRA”? Or are these just qualitative statements?

Many terms and manipulations described in the text in Section 6.4 appear ambiguous or undefined and need to be cleared up. The rationale may be consistent, but it needs to be made much clearer.

### *Licensing and design bases*

Since “licensing basis” and “design basis” are terms used to qualify “accidents”, I tried to determine what these terms mean in a basic regulatory sense, apart from helping to develop a routine for enforcement or describing the content of an SAR. 10 CFR 50 is extraordinarily silent on the matter. In 50.2 “*Design bases* means that information which identifies the specific functions to be performed by a structure, system, or component of a facility, and the specific values or ranges of values chosen for controlling parameters as reference bounds for design”. This very general definition would seem to apply to any engineering device, such as a car, which was rationally designed. For example, seat belts perform the function of decelerating a passenger without serious injury over a range

of collision speeds, impact impedance, passenger weight and so on and are tested to specified failure criteria.

The term “design basis” appears almost nowhere else in 10 CFR 50. There is some mention or inference in the GDCs, Appendix A, where required or desirable features of various systems and their functions are described (many of them appear to apply to any design and to be useful for defining a “framework”), but there is no definition of how they fit into an operationally defined “design”, or “licensing”, “basis”.

In order to understand the basis for any framework that uses such terms, I need to see an exposition of what function LBEs perform in relation to a clear definition of the “design basis” or “licensing basis”, and how this cannot be performed by using the entire PRA.

### *Defense in depth*

DID is discussed in Sections 4.3 and 4.4 and also in Section 6.4.2.2 on page 6-14.

The long discussions in 4.3 and 4.4 describe various strategies and approaches for incorporating design features that may contribute to DID. They provide no “performance-based” evaluation criterion. What if a designer supplies good arguments for pursuing some other novel strategy and approach?

On page 2-5 it is stated "The ability to quantify risk and estimate uncertainty using PRA techniques and taking credit for DID measures in risk analysis allows a better answer to the question of how much DID is enough". The same (promising?) idea is repeated in other words on pp 4-1, 4-3 and 4-4. Yet this does not seem to be followed up later in the text. In the analytical Section 6, where more explicit development of the concept might be expected in the context of the proposed framework, DID is associated with *deterministic criteria* and a rather vague description of *safety margin*.

### *Section 6.4.2.2: Additional Deterministic criteria*

It seems reasonable to use judgment to impose some additional acceptance criteria in order to account for limitations in the completeness of the modeling of technical and human responses to events. The proposal is to do so by imposing additional “deterministic” criteria on the LBEs.

It is not clear why these criteria are imposed on LBEs and not on all PRA sequences, unless there is some unspoken expectation that some greater level of attention to technical detail will be given to LBEs than is present in the PRA.

Criteria such as “no barrier failure occurs” and “a coolable geometry is maintained” may be technology-specific and require careful definition.

No additional deterministic criteria are imposed on the rare range, frequency less than  $1E-5/\text{yr}$ , though this might be thought to encompass accidents where DID is most appropriate because the experiential basis for assessing these events is sparse and there is more opportunity for inadequate modeling. This is perhaps where design requirements, such as requiring a containment or multiple barriers no matter what the PRA predicts, are probably appropriate, rather than additional criteria imposed on a PRA sequence that may not include such features.

*Section 6.4.3: Deterministically selected LBE.*

Each design needs to have a “controlled leakage barrier”, “based on a process that defines an event representing a serious challenge to fission product retention in the fuel and coolant system”.

It could be more straightforward to explicitly insist on an *additional* leakage barrier, besides the fuel and coolant system, since it may be argued that the PRA predicts that the probability of leakage from the fuel and/or coolant system is so low as to present no “challenge” (as appears to be the case with some already approved designs, such as the AP1000, which still have a containment). Then the “serious challenge” is presumably created by changing some of the low probabilities in the PRA sequence to 1, which begins to look like the creation of a design basis accident.

Requirement for containment/confinement would appear to be a significant feature of future regulation, with significant implications for perceptions of safety. *It should be given prominence early on in the report* and explained in much greater detail in terms of the kinds of “challenges” to be considered and the means and criteria for their evaluation.

“The deterministic LBE is to be analyzed mechanistically”. The PRA already contains a mechanistic analysis. Is there some implication that the analysis of this “deterministic LBE” should be more thorough, as with the present DBAs, and include some conservative assumptions? In that case, perhaps it is advisable to move in the direction of more thorough analysis of all significant “serious challenges to fission product retention” rather than just one. If this route is taken, there need to be detailed definitions of what new assumptions, methods of analysis and criteria are to be used.

It would be useful if this process were clarified by an example, perhaps extending the LWR example in Appendix E to include this “deterministically selected LBE” in detail and comparing it with the present treatment of design basis accidents which challenge the containment.

*Safety margin: Section 6.6*

The “safety margin” defined here seems simply to be a way of accounting for uncertainty in the inputs to the PRA without improving its related physical analyses.

How this can be done, how to select the best approach, and what to do with the results would probably require a whole NUREG.

The discussion in the text does little to help define a useful performance-based safety margin, how to calculate it and how to apply criteria to it.

*A high level view of Figure 6.2 and its relationship to the latent cancer QHO*

A glaring omission from NUREG-1860 is the absence of any expression of the latent cancer QHO in terms of the measures presented in Figure 6.2.

Plotting  $fC = 0.004$  rem/yr on Figure 6.2 gives the condition for an individual PRA sequence to just meet the QHO. The first step crosses the curve, showing that only one low dose of about 5mrem and a frequency of 1 would be sufficient to violate the QHO. Does the staff really intend to apply this QHO to such low doses?

A fault of the “step” or “staircase” form is apparent. For the  $1E-4$  frequency, the lowest dose of 1rem meets the QHO by a factor of 40, but the high dose of 25rem only meets it by a factor of 1.6, allowing only one sequence of this type. It seems better to have a continuous curve or line rather than a staircase.

The lowest points on the staircase are close to  $fC=1E-4$  and the average is close to  $fC=4E-4$ .

One could imagine a simple guidance (even part of a rule) which went something like this (a possible alternative among many which could be imagined and evaluated on the basis of decision criteria):

- The sum of the risk,  $R=fC$  rem/yr, from all sequences must be less than  $4E-3$  in order to satisfy the latent cancer QHO. (Perhaps some measure of confidence needs to be attached to this).
- Any sequence having a risk  $>1E-4$  requires extra attention to the mechanistic modeling (e.g. thermal/hydraulics) including treatment of model uncertainty.
- Any individual initiating event leading to the sum of the risk from all its sequences exceeding  $4E-4$  requires comprehensive mechanistic treatment of the modeling of its significant sequences.
- No class of events (e.g. LOCA) may exceed a total risk of  $1E-3$  without special justification (for example, showing that there is only one significant class of accident for the particular design).

This structure is so simple that no figure is needed to describe it, though an f-C map might be useful in order to make the most significant sequences apparent. Working in

terms of “risk = frequency times consequence” provides an additive property which enables summation of the safety metric “risk” (measured by units of dose per year, not CDF) over any number of events, classes of events, similar sequences, or whatever, and is much preferable to use of f and C independently. The total risk of the plant can be expressed on the same scale and might even be a replacement for CDF and/or LERF in the sorts of uses to which Regulatory Guide 1.174 is put today.

The present requirement for “deterministic” analyses of selected sequences called DBAs is no different in principle from the above. It makes up for crudity in the physical modeling in the PRA. For a particular design, the requirements may well be technology specific, since the uncertainties, lack of knowledge, and “what if” scenarios, are particular features and not easily expressed in some general formula.

## SECTION 7. PRA TECHNICAL ACCEPTABILITY

This section describes high level requirements for PRA scope and technical acceptability.

It also describes the key role played by the PRA as the basis for the entire framework. This is very useful to the reader. It would be well to put this section, perhaps in condensed form, much earlier in the document, perhaps as Section 4, replacing the present Sections 4 and 5, which do not lead up to Section 6 in the way that Section 7 does. Some parts of Appendix F should also be included earlier in the document.

Until reading this section and Appendix F, it was unclear to me that the intent is to base essentially all safety evaluation on the results of the PRA. Deterministic approaches only come in as a few additional evaluation criteria. The technical analysis is all in the PRA. Since this is a revolutionary development, which may have significant implementation difficulties, it needs to be explained clearly right at the beginning of the report.

P7-6. “As discussed in Chapter 6, LBEs are bounding event sequences that are subjected to additional analysis.”.

*I am unable to find any description of such “additional analysis” anywhere in the report. Perhaps the authors are uncertain about this issue.*

Apparently, LBEs do not resemble the present DBAs, which are given extra attention in the form of more complete technical analysis in the present regulations. Appendix F appears to clarify this, requiring that thermal/hydraulic codes be used in the PRA and not in some separate technical evaluation. This may be difficult to implement unless the use of codes is restricted to a few of the most important PRA sequences. Integrating codes and the PRA has not yet been attempted. There are significant structural and computational hurdles to be overcome in order to do so.

P7-7 Why are LBEs needed in order to characterize safety significant SSCs? Cannot the entire set of PRA sequences be used, particularly the most significant ones?

P7-10 “Appendix F identifies the high level requirements necessary to ensure the technical adequacy of a PRA”. These requirements, which seem to imply a significant amplification of the technical analyses in the PRA, should not be relegated to an appendix. They are a key feature of the framework and should be given prominence in the opening Sections of the report.

## SECTION 8. REQUIREMENTS DEVELOPMENT PROCESS

This section resembles a regulatory guide about numerous items and considerations to be considered in the SAR for the plant, which establishes a “licensing basis”. It would help if it were more structured and less discursive. In particular, the roles played by significant features of the proposed framework need to be made more specific and thereby justified.

There is mention of a “licensing analysis”. There is no explanation of why LBEs are called “licensing basis events” and what role they play in a “licensing basis”.

The justification for LBEs is weak throughout the entire document. Perhaps it can be established in this section by summarizing how they are used and what functions they perform. It appears that most roles of LBEs, such as showing conformance with the QHOS, meeting an F-C curve, defining safety significant SSCs, and meeting additional deterministic criteria, can be performed by the PRA sequences themselves. If LBEs are not subject to additional analysis, as DBAs are now, it is not clear why they should be defined at all.

Table 8.1 includes “provision to establish a containment functional capability”. This important requirement should be given prominence earlier in the report and not hidden here.

## APPENDICES TO NUREG-1860

The appendices are very useful for clarifying parts of the text and reassessing some of the comments made at the start of this review.

### APPENDIX A. *Safety characteristics of the new advanced reactors.*

This appendix describes safety characteristics as well as important design considerations and strategies. There are useful descriptions of important features to consider in design and how these are implemented in advanced reactor concepts.

These discussions do not lead to identification of any Top Level Regulatory Criteria that can be shown to have generic applicability. Nor do they indicate how the proposed framework is particularly appropriate for regulating these designs. Therefore they do not help the reader to understand the justification for features in the framework, particularly what criteria are to be applied to decide that each particular design is “safe enough”.

### APPENDIX B. *Relationship to 10 CFR*

This expands on the links to parts of 10 CFR that were sketched out in Section 1.

A long list is presented, with a description of the content of the link to very large number of existing regulations.

No assessment is made of how any of these many items influence the design of the framework. If they had no influence, then this list adds nothing to the rationale for the particular structure that was chosen. It would be helpful if high level criteria could be extracted from this wealth of regulations. If this is not possible, this is perhaps also a useful conclusion.

### APPENDIX C. *Protection of the environment.*

This topic appears sufficiently important to be included in Sections 1 or 2 of the main text.

Actually, this appendix says little about the environment, as far as land contamination, loss of property and opportunity, social disruption and so on are concerned. It is based on dose to individuals in the environment, in particular the relationship to ALARA and the QHOs. The very short treatment of cleanup costs appears inadequate as a comprehensive evaluation of environmental effects.

The most interesting part of this appendix is in the middle of page C-3 where the latent cancer QHO is compared to dose.

The calculation is based on the 20rem dose which has a frequency of 1E-5/yr on Figure 6-1 if only one dose is considered, which seems invalid as this figure applies to individual PRA sequences and several may produce doses in this range. The resulting expected dose of  $20 \times 1E-5 = 2E-4$ rem/yr is multiplied by the latent cancer fatality risk coefficient of 5E-4/rem to obtain 1E-7/yr, which is cited as being “much less than the latent fatality QHO individual risk of 2E-6/yr.

This appears to be a strange and misleading conclusion, as it only considers a single 20rem dose. All ranges of dose (perhaps above a threshold) contribute to latent cancer, so what matters is the cumulative dose from all events. We can work back to this cumulative dose by dividing the QHO risk value of 2E-6/yr by the latent cancer fatality risk coefficient of 5E-4/rem to get:

$$2E-6/5E-4 = 4E-3\text{rem/yr}$$

as *the cumulative dose consistent with the QHO*. This value is compatible with the ALARA 3mrem/yr total body dose quoted as item (1) on page C-1 (how can the total body dose be less than the individual organ dose of 10mrem/yr?) or with the competing 5mrem/yr attributed to ALARA in Figure 6-2.

*This is exactly the calculation that one would have expected the staff to make at the very beginning of NUREG-1860*, since the link between this QHO and the design criteria must be made. It has a profound effect on the key features of the framework in Section 6. Indeed, if one accepts the cumulative ALARA dose of 5mrem that is cited there, this uses up all the available QHO dose, leaving nothing at all available as allowable dose from more serious events, apparently invalidating the entire rest of the curve.

#### APPENDIX D. *Derivation of risk surrogates for LWRs*

It is argued that CDF of 1E-4/yr and LERF of 1E-5/yr are acceptable surrogates for the QHOs for LWRs.

Early fatalities are most likely with serious accidents involving major releases. The staff argues that a LERF of 1E-5/yr is equivalent to an early fatality risk of 3E-7/yr which is “less than the early fatality QHO of 5E-7/yr by a factor of about two”. Since this estimate is based on a single accident sequence (Figure E-1 shows several in this range) the “factor of two” may be unjustified. One may also doubt if the conditional probability of early fatality given a LERF is always as low as 3E-2.

Comparing with Figure 6-2 it is not clear how this relates to the final three steps in the range above 100rem. The staff needs to clarify what is the acceptable way to compute early fatalities, perhaps with some conservatism, and relate it to an allowable cumulative frequency of such events, which is not directly derivable from Figure 6-2 as it applies to individual PRA sequences and not their cumulative effect.

The argument that the early fatality QHO is roughly equivalent to the present LERF is probably valid. It indicates that an output of a future PRA that measures the overall risk status of the plant, replacing LERF, could well be “cumulative early fatality frequency”. It can be computed directly from the entire PRA. Any additional requirements governing individual sequences would be secondary.

Regarding the latent cancer QHO, the comparison with CDF is more tenuous, since there are many kinds of core damage, some of which lead to little effect on public health (e.g. TMI2). However, if dose is to be used as the technical measure of public impact of events, it appears reasonable to replace the present CDF metric with the “cumulative expected yearly dose” from all events and require that it be less than the 4mrem/yr in order to meet this QHO.

These two metrics, representing the QHOs, would then be candidates for use in evaluating the overall risk status of the plant, enforcement actions, and allowable changes, as is now done using CDF and LERF with RG1.174, for example.

#### APPENDIX E. *Example of LBE and safety classification selection process.*

The proposed framework is applied to an existing LWR plant, based on a Level2 PRA model. The process described in the text is followed, helping to clarify what is stated there.

A long list of accident sequences with a point estimate of frequency exceeding 1E-8/yr is developed in Table E.6. Since the point estimate, mean and 95<sup>th</sup> percentile are quoted for all sequences, this indicates that there is a considerable computational burden, especially if there is a comprehensive treatment of uncertainty, including model uncertainty. It is not explained how these values are derived. They should presumably reflect the statistical uncertainties in the prediction of the scenario, which one would expect to be highly dependent on the constituent events, some of which (e.g. DEGB) are much more uncertain than others. Yet the 95<sup>th</sup> percentile doses in the large majority of cases appear to be 4 to 5 times the mean, which itself is close to the point estimate, suggesting that some simpler process has been used.

From this table it is possible to add up the yearly probabilities of fatalities and latent cancers, a task which surprisingly is not performed, though the QHOs are ostensibly the (only?) quantitative high level measures of adequate safety.

The events that are darkened in this table have 95<sup>th</sup> percentile frequency less than 1E-7 and are discarded in order to create Table E.7. This leads to removal from consideration of several events leading to fatal doses.

Since the doses are not very precise, I considered a mean dose greater than 200rem as representing a slightly conservative estimate of a fatal dose (the corresponding 95<sup>th</sup>

percentile doses are definitely fatal). Adding up the point estimate of frequency for those in this category from the shaded events leads to a total fatality expectation of  $2.7E-6/\text{yr}$ , significantly exceeding the early fatality QHO of  $5E-7/\text{yr}$ .

Since this group of rare events alone appears capable of violating the early fatality QHO by about a factor of five, it may well be judged *unacceptable to cut out PRA sequences with a 95th percentile frequency  $<1E-7/\text{yr}$  (step 5)*.

The events in Table E.7 are further reduced by selecting “bounding events” in order to create Table E.8, licensing basis events.

It appears from Figure E.1 that six such events violate the frequency-consequence curve at the 95<sup>th</sup> percentile value and four of them violate it at the mean value. *Why is the design then acceptable? What has been gained by plotting LBEs rather than all the PRA sequences, to see which ones violate the criterion?*

If one draws the line  $fC=4E-3$ , corresponding to the latent cancer QHO, on Figure E.1, it appears that the single events 14 and 34 are each sufficient to violate this cumulative dose requirement. The three events 14, 16 and 34 individually violate the early fatality requirement. *Why is this design acceptable?*

In addition to frequency-consequence criteria the framework requires that some “deterministic” criteria be applied in the interest of defense in depth. These lead to two further sequences, LBE-21 and LBE-07, failing to meet acceptance criteria. *What conclusion is drawn from this? Is the design acceptable?*

This example illustrates how the staff has redefined the “deterministic-probabilistic” duo. The traditional deterministic analysis of DBAs has been discarded. Everything is based on the physical analyses in the PRA, with its somewhat approximate modeling and success criteria. “Deterministic requirements” are used at the end of the process to impose additional acceptance criteria on a few sequences. For example, LBE-07 is found not to have a coolable geometry. Yet this conclusion is not based on the type of thorough thermal/hydraulic analysis typical of DBAs. The proposed process is utterly different from the present regulatory system in which “deterministic” means analyses of the type presented in Chapter 15 of the SAR, such as those responding to the LOCA rule, 50.46. These traditional “deterministic” analyses involve the use of thermal/hydraulic codes and uncertainty assessments that are thoroughly reviewed by the NRC, and by the ACRS. They are far more complete technically than has been possible to date in PRA models and should represent a superior representation of “what happens” during the more important events. *Does the staff really intend to abandon the traditional “deterministic” approach?*

It would appear that the classical deterministic analysis representing a “best estimate”, using appropriately sophisticated technical tools, of what happens physically during significant events cannot justifiably be thrown away unless an equivalent thorough analysis, with adequate level of detail and consideration of uncertainty, is somehow

incorporated into the PRA. This seems to be required in Appendix F, but with no assessment of the feasibility of doing so, which is presently beyond the capability of any computer code(s).

The list of Chapter 15 events in Table E.14 is not used to make a comparison between the present regulatory system and the proposed one, helping to justify the latter. If this were done, it would be clear that Chapter 15 has its own requirements for quality of analysis and for evaluation criteria which create a “separate regulatory space” from the PRA. As an example, the peak clad temperature of 2200F required in DBA LOCA analysis appears not to be directly connected with the core damage criteria in the PRA.

*It appears that if the proposed process were to be used to license existing reactors Chapter 15 analyses would not be required. How is the regulatory function presently performed by Chapter 15 events and DBAs to be performed under the proposed framework?*

This Appendix would benefit from explanation of what is to be concluded from the results that are displayed and how these would lead to regulatory decisions. It would also benefit from detailed comparisons with the present system for regulating this same reactor, explaining why the proposed framework is an improvement.

#### APPENDIX F *Scope of the PRA.*

Table F-24 lists requirements for accident progression analysis. These include “Use verified and validated accident analysis codes to evaluate the progression of the accident”, “Use verified and validated codes to evaluate the vessel, confinement/containment, and other barrier capacity to withstand the challenges introduced by accident phenomena”. Tables F-25 and F-26 contain similar requirements for the use of codes to compute the source term and consequences.

These requirements, within a description of the scope of the PRA, seem to suggest that detailed analyses, of the DBA type, are to be incorporated into the PRA. This would require a technical revolution. Up to now, it has been far beyond the range of computational feasibility to combine the probabilistic event tree structure with comprehensive technical analysis using codes. Practical ways to do so are not yet evident even at the research stage.

*How is it going to be possible to incorporate the level of technical analysis represented by accident analysis computer codes into the PRA of future reactors?*

## APPENDIX

### PRA, DBAs, LBEs AND ALL THAT

#### EVENTS

An event, occurrence, or accident at a nuclear power plant consists of a scenario in which various happenings, such as loss-of-feedwater, pumps starting or not, valves opening or not, operator actions and so on, take place. Throughout this drama numerous thermal/hydraulic, mechanical, electrical, chemical, human, and information processing phenomena occur and influence the course of the scenario.

The progression of this event may be represented on a tree diagram in which a sequence of paths, selected by probabilities of physical alternatives, human actions, or of one of several choices in the other parameters defining what happens, leads to a final condition of the system. This end state may involve hazards to the public, activated either on the way to achieving it or thereafter. The hazards of particular concern with nuclear power are radioisotope releases and resulting damage to health, environmental, and other measures of value to society.

Each PRA sequence is a unique path through this tree diagram. So is a design basis accident (DBA), which has traditionally been elevated to a greater level of importance than the other sequences.

#### PRA, f-C MAPS, F-C CURVES ETC.

PRA is intended to provide quantitative measures of the probability of following each of the various paths during an event and of the resulting consequences.

For each distinct scenario, or PRA sequence, it is useful to have a single measure of consequences,  $C$ , such as dose at some place, curies released beyond some region, or perhaps dollars in order to combine all ill effects on a common scale. (The simplest measure of consequence is binary, such as core damage, which is defined so that it either occurs or does not. As this is not used in NUREG-1860, it will not be discussed further here).  $C$  may be plotted versus the frequency of the sequence,  $f$ , (should be "probability", as a measure of our state of knowledge?), on an "f-C map". Since there are many individual scenarios for a certain initiating event, and some of these lead to damage, the f-C map may be covered by many points derived from a single event tree. If uncertainties are included, the points become regions in which there are additional probabilities of achieving certain values of  $f$  and  $C$  for a given sequence. These probabilities can lead to measures of confidence that one has assessed the range of possible values of  $f$  and  $C$ .

This f-C map is useful for describing individual PRA sequences. It is not a measure of the overall response of the plant or of public safety. If there are enough events and uncertainties to form a continuum of consequences, their cumulative effect may be represented on a separate plot by the probability density distribution,  $p(C)$ , which adds up the probabilities of all sequences in the interval  $dC$  such that the probability of all external effects in this interval is  $p(C)dC$  and does provide a measure of public safety.  $p(C)$  has the units of  $1/C$ . If the probabilities are added up in finite ranges of  $C$ , or bins, rather than over a continuum of outcomes, the equivalent representation is a histogram.

These approaches are fundamental and are more basic than any regulatory criteria that may be imposed on the overall picture. They can be used to derive additional concepts and definitions which need to be selected carefully for their appropriateness and utility, and used unequivocally in a report such as NUREG-1860 if it is to be authoritative and persuasive.

The design of an event tree is an art form and requires choices that are considerably influenced by perception or experience of qualitative differences between the several hypothesized scenarios. For an LWR LOCA, for example, it would appear necessary to subdivide the tree into S, M, and L branches, because the events and phenomena that occur in each class are almost as different as Hamlet and Macbeth in the category of Shakespearian tragedy. Each of these classes of LOCA lead to a number of outcomes and therefore to a number of points on the f-C map.

Points resulting from a given initial event, within a class, or within several classes taken together, *cannot be combined into a single effective (f, C) point* without making some arbitrary definitions. Therefore one cannot plot such composites on an f-C map. This fundamental constraint restricts the utility of any process, such as that described in NUREG-1860, which attempts to represent all levels of detail on a single “frequency-consequence” figure.

As a measure of success at achieving public safety, one might try to decree that no single accident should have an effect outside some acceptable region on the f-C map. If this “accident” is defined by a unique PRA sequence, a single completely defined path through an event tree, and uncertainties are not considered, it can be represented by a single point on this map. Such a path describes a particular historical happening, but its representation by a sequence of finite branches on a PRA tree is not unique. For example, an SBLOCA could be defined as a break between 1 and 3 inches and the first branch point might involve an expected operator action. If more detailed description is desired, the first branch point might involve selection of the break size in one of the two ranges 1-2 and 2-3 inches (It is not an action in the usual sense, but could be thought of as the pipe’s decision to break into one size range or the other), thus subdividing the tree into two major branches and splitting the frequency of similar outcomes into two smaller parts. The frequency (probability) of a happening is not independent of the structure of the parameters used to describe it.

For example, the probability of Tennessee winning a basketball game differs from the probability of them winning by 10 points, or the probability of the actual score, or the probability of the particular sequence of scoring that led to this score. The probabilities in the previous sentence decrease as one progresses through it, though it is the same game that is being described. A question such as “what is the probability of what happened in that game?” is meaningless. Similarly, asking “what was the probability of the accident at TMI2?” is meaningless until one defines the parameters used to describe that event and the precision with which they are represented. There is a kind of uncertainty principle at work. The more complete and precise the definition, the lower the frequency.

A difficulty of using the f-C map as a basis to satisfy evaluation criteria is that one can always make a set of sequences “acceptable” simply by subdividing them into a finer set by including more branches, thereby reducing the frequency of the offending points, and to some lesser extent modifying their consequences.

If it is not a particular sequence that is to be acceptable, but the outcomes of a given initiating event, or the outcomes from a class of events such as a SBLOCA, then one has to figure out how to apply acceptance criteria to a group of points on the f-C map. *They cannot be added up without some format for aggregating them.* One way to do this is to define “risk”,  $R$ , as frequency times consequence. (Some other ways involve averaging or binning. If used, such methods need to be defined and evaluated). Risk then becomes an additive property of the scenarios which can be added up to give the net risk from the “accident” as  $R_A = \sum f_i C_i$ . It is relatively insensitive to how a comprehensive PRA tree is branched and would be expected to converge to a limiting value as the structure is made finer, as long as no new phenomena are introduced. If this measure is chosen, then the acceptance criterion needs to be congruent with it, perhaps by decreeing that exceeding a certain level of risk is unacceptable for each accident or class of accident, which is not representable on an f-C map and is inherently simpler. How to pick the definition of an “accident” or “class” involves a degree of judgment, with the general idea that the scenarios and consequences should not vary too much within a given class or that an accident should be suitably representative of that class.

The probability density  $p(C)$  is also relatively independent of the tree structure as it adds up all the events in an interval  $dC$ , no matter how they are described. The existence of a continuous probability density implies suitable continuity in the distribution of outcomes, either by having very many of them or by having the uncertainty included in each. A histogram is a “binning” representation of the same thing and loses information the coarser the subdivision into bins is.

The sum  $\sum f_i C_i$ , or the integral of  $f(C)C dC$  over *all* scenarios is the “expected value” to society of the consequences of all accidents,  $\langle C \rangle$ , and is a useful concept as part of the assessment of how safe is safe enough. Adopting the above definition of risk, this is the same as the total risk from all accidents at the plant,  $R_T = \sum f_i C_i$ . If  $C$  is dose, this measure can be directly related to the latent cancer QHO.

Another useful concept, *which allows adding up effects at any desired level of detail*, is the complementary cumulative distribution function,  $F$ .  $F$  is a function of  $C$  and is equal to  $\sum f_i$  (frequency alone can be added) for all consequences greater than  $C$ . It has units of frequency. (In the expert elicitation report on pipe break frequency, what is being plotted is presumably  $F$  versus  $D$ , though this was not clearly apparent in the draft version that we saw). If the representation on the  $f$ - $C$  map consists of a set of discrete points, the  $F$ - $C$  curve will have steps in it. A change  $dF$  for a change  $dC$  is equal to  $\sum f_i$  in that interval. If the probabilities are continuous enough to be representable by a probability density function, then  $F$  is the integral of  $f(C)dC$ .

If it is desired to specify that any accident, rather than a specific sequence, with consequence exceeding  $C$  should be limited to a certain probability, then the appropriate measure to satisfy that criterion is  $F$ , derived from the PRA tree(s) describing all the sequences associated with that particular accident.

The curve is also useful for determining the probability,  $F_1-F_2$ , of any event with consequences between  $C_1$  and  $C_2$ . The text describing the meaning of at least some of the steps plotted in Figure 6.2 of NUREG-1860 appears to indicate that this representation might be more appropriate than the  $f$ - $C$  map that is presented there.

Because  $dF$  is equal to  $\sum f_i$  in the interval  $dC$ , the *area to the left* of a piece of the  $F$ - $C$  curve between  $C_1$  and  $C_2$  is the expected value of the consequence of all events in that range: the range risk value,  $R_R = \sum f_i C_i$ . (This may appear counterintuitive. The outcomes in the interval between  $C$  and  $C + dC$  appear in the vertical slice with width  $dC$  on an  $f$ - $C$  map, but their frequencies add up to an increment  $dF$  on the  $F$ - $C$  curve and the risk contribution  $CdF$  is the area of a horizontal slice to the left of the  $F$ - $C$  curve). For example, if  $C$  is dose, then  $R_R$  is the expected total dose resulting from all events in the range. This interpretation is useful for describing some of the regulations, such as the ALARA limit, more consistently than in Figure 6.2 of NUREG-1860, where a cumulative dose requirement is reinterpreted as frequency of dose from a single PRA sequence.

The area to the left of the entire  $F$ - $C$  curve (which is the same as the area under it if it is bounded) is the expected consequence from all events, the same as  $\langle C \rangle$  or  $R_T$ .

If one has the  $F$ - $C$  curve for the plant, including all accidents, and consequences above some value,  $C_L$ , are called “large”, then  $F_L$  is the “large release frequency” or LRF. It is the corresponding area under the probability density function, plotted against  $C$ .

Similarly, if consequences above some value  $C_D$  are all associated with core damage, which is a prerequisite for most significant releases, then the corresponding value of  $F_D$  is the core damage frequency.

The  $F$ - $C$  curve is relatively insensitive to the fineness with which the PRA tree is defined, but not entirely so. Dividing an event with consequence  $C$  and probability  $f$  into separate events with consequences  $C_j$  and probability  $f_j$  will produce a different set of steps on the  $F$ - $C$  curve and one may not end up in quite the same place by climbing these two

staircases. If the steps are small enough and convergence is achieved, then the F-C curve is well-defined as the integral of the probability density distribution.

One could plot  $f$  and  $F$  on the same figure. There would then be a large cloud of individual outcomes towards the bottom of the picture, each with their own  $f$  and  $C$ , and an  $F$  curve (far) above these with each point on it representing the sum of all the  $f$ 's to the right of that point. This sum significantly exceeds the value of any  $f$ 's below it unless there are a few singularly dominant outcomes. If the PRA is more detailed and has a more refined branch structure, the cloud of  $f$ - $C$  points retreats downwards, though the  $F$ - $C$  curve is relatively unchanged.

Besides representing the total plant risk, an  $F$ - $C$  curve can also be used to represent the net consequences from all outcomes of a single "accident" or class of accidents, which cannot be done on a composite  $f$ - $C$  plot.  $F$  is then the sum of the  $f_i$  for all PRA outcomes with consequences within the PRA tree representing that accident, or the net sum over several trees representing a class of accidents. If there is a family of  $F$ - $C$  curves which each describe separate accidents,  $j$ , within a class, then they can be added together simply by adding the cumulative frequencies  $F_j$  of each accident at a particular level of consequence  $C$ , to give the  $F$ - $C$  curve for the class, where  $F = \Sigma F_j$ . Adding up the  $F$ s from all classes gives the  $F$ - $C$  curve for the plant. This additive property is useful when deriving compatible regulations at different levels of detail.

The  $F$ - $C$  curve is one of the most useful tools for describing the overall public safety impact of a plant, or of a selected class of accidents, clearly and compactly.

## DBAs

DBAs were developed before PRA techniques evolved, but have some features in common with them.

A DBA is essentially a unique sequence, or perhaps a group of sequences, on a PRA event tree. Both the DBA and the corresponding PRA sequence are analyzed using deterministic methods to describe the physics. In the DBA, the branches that are taken (e.g. a certain size or type of pipe breaks, or the worst single failure occurs) are specified by the regulator rather than having a probability of being selected. There may also be rules about what is to be assumed about certain phenomena, e.g. the correlation for critical flow, other selections of "branches" in the input parameters, such as decay heat, and operator actions. Success criteria are usually prescribed to define the regulatory condition for zero or suitably limited undesirable consequences.

Replacing the large number of PRA sequences by a dozen or so particular sequences which are called DBAs is a bold, perhaps risky, step since a great deal of possibly useful information is thereby lost. Considerable thought must be given to selecting the DBA(s) within some class that is considered to be important so that it will be typical or "bounding" in some aspects. DBAs cannot be bounding of all PRA sequences if it is

specified, as is sometimes the case, that there should be no adverse consequences. In that case there would be no predicted public consequences whatever, and no risk from the plant.

A DBA analysis provides a certain sense, perhaps illusion, of clarity and “determinism” because the conditions of the problem to be analyzed are prescribed to make it so. This apparent certitude is bought at the price of a leap of faith that the DBA analysis and its results are sufficiently representative for its conclusions to be extrapolated as a basis for “design” against all similar accidents in its class, and that all important classes of events have been covered. It provides no measure of consequence or risk, and therefore no measure of public safety. Indeed, by decreeing that DBAs should have no, or sufficiently low, adverse public consequences one removes the very measures, such as f and C, which may be of most interest to the public. A category of “beyond design basis” accidents may be formed to account for this.

The sequences of most concern for their effects on public safety are those leading to damage that is forbidden in at least some of the DBAs. There is a risk that they will be missed and the design will be inadequate for coping with them if the design basis is too restrictive.

With many years of experience with a certain design one may develop confidence that this extrapolation of “DBA space” to “risk space” is valid as a practical way to proceed. With new designs one must inherently be less sure and there is more motivation to evaluate all events. This appears to be one motivation for the staff to “risk-derive” the aspects of design which are most worthy of attention. Another purpose might be to avoid having two separate incommensurate frameworks for regulation, a feature that has complicated efforts to risk inform the current Part 50.

One way that a DBA has historically differed from the exact same PRA sequence is in the level of attention given to the analysis of phenomena. All PRA sequences must necessarily contain an assessment of “what happens” by way of analysis of thermal/hydraulics and so on. At the crudest level this is done by taking the vendor’s word that certain simplified rules apply. For example, 2 trains out of 3 are sufficient to cool the core, or that the opening of 5 out of 8 valves leads to successful depressurization. Now, this assurance from the vendor must be based on some suitable analysis, which might need to be every bit as elaborate as that which is appropriate for the DBA. One is taking a step of faith that the vendor has performed a sufficient and adequate analysis to assess the various PRA success criteria. This becomes more problematic if it is desired to include considerations of model uncertainty in the PRA.

Since a large effort, using codes such as RELAP, TRAC, TRACE etc., goes into analyzing DBA sequences that by definition contribute little or nothing to risk, it might make sense to apply a similar effort to those sequences that contribute the most to risk. If only one of these is selected in each accident class, the amount of computational effort involved is essentially the same as in the DBA case (somewhat more if consequences are investigated in more detail and given quantitative measures). Since some DBAs are now

investigated by making 59, 93 or 124 independent runs on computers to enable statistical investigation of uncertainties, it might no longer be considered “burdensome” to move in this direction. It might even be reasonable to apply “better” thermal/hydraulic analysis within the PRA, to a number of sequences, perhaps a few dozen, of the most risk significant ones out of the tens of thousands, most of which have insignificant consequences. This might require significant research and development

It appeared that the staff might be moving in this direction by defining LBEs based on risk and treating them much as DBAs are treated today. However, it appears clear from Appendix E that the LBEs are not reanalyzed with more complete physics. The PRA model is the only basis for the prediction of what happens physically during an event. It appears from Appendix F that thermal/hydraulic codes are to be used within the entire PRA, which may not be feasible.

## DETERMINISTIC and PROBABILISTIC ANALYSIS

The regulations in Part 50 were originally based on “deterministic” analyses of chosen DBAs with a prescribed format and conservative success criteria. These are exemplified by the Chapter 15 analyses in current SARs. Uncertainties are accounted for by making suitably conservative assumptions. The analyses use elaborate thermal/hydraulic codes in an effort to make the analyses as complete and realistic as possible.

With the advent of PRA, an independent structure of “probabilistic” analysis was developed and matured. Rather than prescribing a conservative path to an outcome, the probability of various outcomes is computed as the synthesis of the probabilities of various events or branch points in the scenario. In order to keep the computation manageable, the criteria for taking various branches in the event tree, or the success criteria, are simplified, based on a condensation of the results of more complete analysis of the type used in the analysis of DBAs. The advantage of a PRA is that it is able to give an estimate of the probability of undesirable outcomes, and hence measures of the “risk” associated with the plant, which is a direct metric of public safety.

While it is conventional to distinguish PRAs from “deterministic” approaches, the physical modeling in each case is deterministic. Each distinct PRA sequence represents a deterministic analysis. A probabilistic element is added by assessing the probability that the assumptions and inputs used in this deterministic analysis will be valid. The result is not only a deterministic prediction, but an assessment of how likely it is to represent reality. The latter feature adds significant useful information.

Both of these approaches are useful for regulatory purposes. Because of some inherent weaknesses of each (in the extrapolation of a few selected and idealized DBAs to all accidents, or in the adequacy of the PRA success criteria, the assessment of probabilities, and the completeness of the modeling) they have come to be regarded as complementing each other, forming two independent bases for decision making. This may sometimes

lead to conflicts, as the physical modeling and success criteria in the two systems are sufficiently different that they may lead to different conclusions.

It appears from Appendix E that the proposed framework excludes the traditional “deterministic” analysis entirely. Everything is based on the PRA models. The only place where a “deterministic” element is introduced is in the addition of a few prescribed elements of defense in depth to outlaw certain sequences if they fall into certain ranges of frequency and consequence and also contain some key physical feature, such as the failure of a barrier.

## DEFENSE IN DEPTH

Defense in depth (DID) provides additional assurance that really bad things won’t happen. It is there to help prevent losing the war, or a major battle, rather than to win a minor skirmish. It may be introduced to reduce the chance of frequent occurrences escalating into major accidents or to protect against uncertainties in the course of events that are initiated by rare but large challenges to safety.

The reason that frequent events generally are predicted to have low consequences is that several barriers, mitigating systems, and redundancies prevent them from escalating into major disasters and these all operate well within their design capabilities. Since there is experience with these events and a basis for estimating failure modes and probabilities, this sort of DID has a chance of being modeled probabilistically. Some “unknown”, such as unanticipated sequences or common cause failures or operator action (the traitor in the medieval castle who opens doors in a series of barriers) can still intervene. Some measure of DID may be achieved by exploring what happens if some probabilities are arbitrarily set to 1 or 0 and requiring that a certain level of safety be maintained. The analysis of the sequences remains “deterministic” and this is not the feature that by itself achieves DID.

Perhaps the best DID protects against a whole gamut of outcomes. An example of this is a containment, which it may be well to require for any system no matter what the PRA or other assessment of risk says.

A useful principle might be to provide greater DID against the low frequency events with large consequences, since these are the ones with which there is no experience, and therefore less ability to predict what will happen or its likelihood.