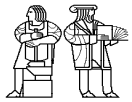

Massachusetts Institute of Technology
Engineering Systems Division

Working Paper Series

ESD-WP-2003-05

HOW USEFUL IS QUANTITATIVE RISK ASSESSMENT?

Professor George E. Apostolakis

July 2003

Perspectives

How Useful is Quantitative Risk Assessment?

Professor George E. Apostolakis

apostola@mit.edu

Abstract

This article discusses the use of Quantitative Risk Assessment (QRA) in decision-making regarding the safety of complex technological systems. The insights gained by QRA are compared with those from traditional safety methods and it is argued that the two approaches complement each other. It is argued that peer review is an essential part of the QRA process. The importance of risk-*informed* rather than risk-*based* decision-making is emphasized. Engineering insights derived from QRAs are always used in combination with traditional safety requirements and it is in this context that they should be reviewed and critiqued. Examples from applications in nuclear power, space systems, and an incinerator of chemical agents are given to demonstrate the practical benefits of QRA. Finally, several common criticisms raised against QRA are addressed.

1. INTRODUCTION

It has been about thirty years since Quantitative Risk Assessment (QRA) was first applied to large technological systems[1]. Since then, we have seen many methodological advances and applications to nuclear power reactors (where it is called Probabilistic Risk Assessment - PRA), space systems, waste repositories (where it is called Performance Assessment), and incinerators of chemical munitions.

In every application, I have observed a familiar pattern of progress. At first (Phase 1), the safety community of that industry is very skeptical about the usefulness of this new technology. Then (Phase 2), as engineers and decision makers become more familiar with the technology, they begin to pay attention to the insights produced by QRA. Typically, the decision makers first pay attention to the "negative" insights, i.e., those that reveal failure modes of the system that had not been identified previously. Actions are taken to make these failure modes and their consequences less likely. With time (Phase 3), confidence in QRA increases as more safety analysts use it and they begin to pay attention to the "positive" insights, i.e., that some of the previously imposed safety requirements can be relaxed because either they do not contribute to safety or they contribute a very small amount that can not be justified when it is compared with the corresponding cost. Entering Phase 3 usually requires a cultural change regarding safety management. This change is not always easy for engineers who have been using traditional "deterministic" methods for years. In all three phases, risk insights alone are never the sole basis for decision-making.

Of course, there are no sharp lines dividing the three phases. The phase in which a particular industry is depends on the extent to which it is regulated. Thus, the heavily regulated US nuclear power industry entered Phase 1 in 1975 and Phase 2 a few years later. The first regulatory guidance on how risk information could be used in regulatory affairs was issued in 1998[2], which marks the beginning of Phase 3. It took about a quarter century for Phase 3 to begin. This is evidence of the caution with which QRA is accepted by real decision-makers. Most foreign nuclear regulatory agencies are still in

Phase 2 and are watching carefully the US experiment with Phase 3. In my view, NASA is in Phase 1 at this time and is cautiously experimenting with more substantive use of QRA insights.

Citizen groups that oppose a particular decision or industry frequently raise questions about the validity of QRA, especially the "Q" part (see, for example, reference 3). Although there is occasionally a reply by responsible groups[4], these criticisms, especially articles in the popular press, a recent example being reference 5, remain largely unanswered.

My purpose in this article is to address some of these criticisms and to place the use of QRA in perspective. It is my thesis that QRAs should be viewed as an additional tool in safety analysis that improves safety-related decision-making. QRA is not a wholesale replacement of traditional safety methods or philosophies. QRA analysts will be the first to admit that this tool is not perfect, yet it represents tremendous progress toward rational decision-making.

2. TRADITIONAL SAFETY ANALYSIS

It is important to recognize that even traditional safety analyses must deal with probabilities, but, unlike in QRA, these probabilities are not quantified[1]. The evaluation of safety is typically bottom-up, i.e., it starts with postulated failures and proceeds to identify their consequences. If an event, such as the failure of a component, is judged to lead to unacceptable consequences, measures are taken either to make it less likely (without knowing quantitatively by how much) or to mitigate its potential consequences. Typically, these actions include the introduction of redundant elements and additional safety margins, i.e., the difference between the failure point and the anticipated operating point is made larger. These actions are based on engineering judgment informed by

[1] Traditional safety analyses are often called "deterministic." This is a misnomer. Uncertainties are always present.

analyses, tests, and operating experience. The result is frequently a complex set of requirements for the design and operation of the system.

A facility that meets these requirements is judged "acceptable" in the sense that there is no "undue risk" to the public or the crew. What "undue risk" is remains unquantified. The presumption is that meeting the requirements guarantees adequate protection of public and crew health and safety, i.e., the (unquantified) risk is acceptably low.

3. QUANTITATIVE RISK ASSESSMENT

In its bare essence, QRA answers the three questions posed in Reference 6, namely: i. What can go wrong? ii. How likely is it? and iii. What are the consequences? It is a top-down approach that proceeds as follows:

1. A set of undesirable *end states* (adverse consequences) is defined, e.g., in terms of risk to the public, loss of crew, and loss of the system. These answer the 3rd question of Reference 6.

2. For each end state, a set of disturbances to normal operation is developed which, if uncontained or unmitigated, can lead to the end state. These are called *initiating events (IEs).*

3. *Event* and *fault trees* or other logic diagrams are employed to identify sequences of events that start with an IE and end at an end state. Thus, *accident scenarios* are generated. These scenarios include hardware failures, human errors, fires, and natural phenomena. The dependencies among failures of systems and redundant components (common-cause failures) receive particular attention. These scenarios answer the first question.

4. The probabilities of these scenarios are evaluated using all available evidence, primarily past experience and expert judgment. These probabilities are the answer to the second question.

5. The accident scenarios are ranked according to their expected frequency of occurrence.

A peer review by independent experts is an essential part of the process. Depending on the significance of the issue, this review may involve national and international experts occasionally supported by staff. The first PRAs for nuclear power reactors and severe accidents were subjected to such detailed reviews[7-8]. NASA's QRA for the International Space Station was also subjected to peer review[9]. The QRA for the Tooele incinerator of chemical munitions, sponsored by the US Army, was reviewed by independent experts[10] and the whole operation was under the oversight of an independent committee of the National Research Council[11]. These reviews had significant impact on the respective QRAs. The PRA Standard issued by the American Society of Mechanical Engineers contains a peer review as an integral part[12]. Guidance on peer reviews can be found in reference 13. Insights from QRAs should not be used in decision-making unless they have been subjected to a peer review by independent experts.

4. QRA BENEFITS

QRA has been found useful because it

1. Considers thousands of scenarios that involve multiple failures, thus providing an in-depth understanding of system failure modes. Such an enormous number of possible accident scenarios is not investigated by traditional methods. The completeness of the analysis is enhanced by the QRA investigation significantly.

2. Increases the probability that complex interactions between events/systems/operators will be identified.

3. Provides a common understanding of the problem, thus facilitating communication among various stakeholder groups.

4. Is an *integrated* approach, thus identifying the needs for contributions from diverse disciplines such as the engineering and the social and behavioral sciences.

5. Focuses on uncertainty quantification and creates a better picture of what the community of experts knows or does not know about a particular issue, thus providing valuable input to decisions regarding needed research in diverse disciplines, e.g., physical phenomena and human errors.

6. Facilitates risk management by identifying the dominant accident scenarios, so that resources are not wasted on items that are insignificant contributors to risk.

5. QRA LIMITATIONS

Several items that are either not handled well or not at all by current QRAs are:

1. *Human errors during accident conditions*. There is general agreement among analysts that the Human Reliability Handbook[14] provides adequate guidance for human errors during routine operations, e.g., maintenance. For an accident in progress, we can distinguish between errors of omission (the crew fails to take prescribed actions) and errors of commission (the crew does something that worsens the situation). These errors, especially those of commission, are not handled well and research efforts are underway to improve the situation. I point out, however, that this is a good example of the 5th benefit that I mentioned in Section 4. Even in times of very limited resources, the Nuclear Regulatory Commission has expended significant resources on research on errors of commission[15]. This work has benefited tremendously from the insights developed

by human error theorists, e.g., reference 16. It is also important to point out that experience has shown that the crews often become innovative during accidents and use unusual means for mitigation. These human actions are not modeled in QRAs.

2. *Digital software failures*. This is an active area of research. The aim is not to quantify the failure probabilities but, rather, to understand the kinds of failure modes that may be introduced. It is to the credit of QRA analysts that they have not rushed to use any of the many models available in the literature that treat software as black boxes and assign failure rates to them based on dubious assumptions. Protecting against digital software errors is still within the realm of traditional safety methods, e.g., by requiring extensive testing and the use of diverse software systems.

3. *Safety culture*. When asked, managers of hazardous activities or facilities say that they put safety first. Unfortunately, experience shows that this is not always the case. While it is relatively easy to ascribe an accident that has occurred to a bad safety culture, the fact is that defining indicators of a good or bad safety culture in a predictive way remains elusive. QRAs certainly do not include the influence of culture on crew behavior and one can make a good argument that they will not do so for a very long time, if ever.

4. *Design and manufacturing errors*. These are especially important for equipment that would be required to operate under unusual conditions, such as accident environments. Traditional safety methods of testing and equipment qualification address these errors. It is encouraging to note that a study by the Nuclear Regulatory Commission[17] found that, of the design basis violations reported by nuclear power plants in 1998, only 1% had some safety significance.

6. EXAMPLES OF QRA INSIGHTS

The publication of the Reactor Safety Study (RSS)[1] in 1975 and subsequent industry-sponsored PRAs[18] had a tremendous impact on the thinking of nuclear safety experts. Several major new insights were as follows (see, for example, reference 19):

- Prior thinking was that the (unquantified) frequency of severe core damage was extremely low and that the consequences of such damage would be catastrophic. The RSS calculated a core damage frequency on the order of one event every 10,000 to 100,000 reactor-years, a much higher number than anticipated, and showed that the consequences would not always be catastrophic.

- A significant failure path for radioactivity release that bypasses the containment building was identified. Traditional safety analysis methods had failed to do so.

- The significance of common-cause failures of redundant components and the significance of human errors were demonstrated quantitatively.

An early QRA for the auxiliary power units (APUs) on the shuttle's orbiter produced some interesting results[20]. A traditional qualitative method for developing failure modes (Failure Modes and Effects Analysis) had been employed to identify hazards. Of the 313 scenarios identified, 106 were placed on the so-called critical items list using conservative screening criteria.

- The QRA showed that 99% of the probability of loss of crew/vehicle was contributed by 20 scenarios, two of which had not been on the critical items list.

- Clearly, risk management is much more effective when one has to deal with 20 scenarios as opposed to the 106 of the critical items list, which, in addition, missed two important scenarios.

- The QRA demonstrated, once again, the significance of common-cause and cascading failures. These failure modes are handled very poorly in traditional safety methods.

Unlike the QRAs for nuclear power reactors and the space shuttle, the QRA for the Tooele incinerator of chemical munitions was performed while the facility was still under construction. This allowed the use of QRA insights in design and operation changes[11, 21]. For example, the QRA identified a scenario that dominated worker risk, which involved the potential for buildup and ignition of agent vapors in a furnace feed airlock. As a result of this finding, a hardware change was implemented to vent the airlock and preclude agent buildup. In addition, the operations were changed to limit the time any item is held in the feed airlock.

7. RISK-INFORMED DECISION-MAKING

I wish to make one thing very clear: QRA results are *never* the sole basis for decision-making by responsible groups. In other words, safety-related decision-making is risk-*informed*, not risk-*based*. The requirements of the traditional safety analysis that I described above are largely intact.

In all three examples that I have been using (Nuclear Regulatory Commission, NASA, and the Tooele incinerator), the safety requirements are overwhelmingly traditional. Only the Nuclear Regulatory Commission, which, as I have stated in Section 1, has entered Phase 3, has a formal process for modifying some traditional requirements using risk insights[2].

Even in the Nuclear Regulatory Commission's case, however, it is interesting to see how cautious the Commission was when it issued its policy statement in 1995[22]: "The use of PRA technology should be increased in all regulatory matters to the extent supported by the state of the art in PRA methods and data and in a manner that complements the NRC's

deterministic approach and supports the NRC's traditional defense-in-depth philosophy." The Commissioners made it very clear that PRA insights were to be scrutinized and were subordinate to the traditional safety philosophy of defense-in-depth.

At this time, formal methods for combining risk information with traditional safety methods do not exist. The process relies on the judgment of the decision-makers and is akin to the analytic-deliberative process that the National Research Council has recommended for environmental decision-making[23].

The continuous risk management process that NASA employs is similar in the sense that risk information is only one of the many inputs that the process considers.

8. COMMENTARY ON FREQUENTLY RAISED CRITICISMS

All QRA analysts care about is the bottom-line numbers.

I know of no QRA analysts who act this way (and I know a lot of them). The uncertainties about the results and the dominant scenarios are the results that experienced analysts look for. The lower the probabilities that are reported, the more suspicious these analysts become.

QRAs are performed to understand how the system can fail and to prioritize the failure modes, not to produce a set of numbers. The only QRA results that have any chance of influencing risk management are those that provide engineering insights.

NASA doesn't seem to be able to get the numbers for the Shuttle right.

This is an interesting comment that one encounters frequently. Old studies are cited[24] that apparently produced very low failure probabilities for the shuttle (as low as once

every 100,000 flights). Then, it is pointed out that there have been two failures in 113 shuttle flights; therefore, QRAs are no good.

An early QRA for the shuttle[25] gives a probability for the loss of vehicle that ranges from once every 230 flights to once every 76 flights. The analysts state that they are 90% confident that the failure probability is in this interval. The reported interval is on the right order of magnitude, especially for a first effort that has not been peer reviewed formally and in an industry that is still in Phase 1 of QRA development. I doubt very much that the old studies mentioned above had been subjected to the rigorous peer review process that modern QRAs receive.

But this criticism raises another important point. There is a presumption that QRA should "get the number right" right away. This does not allow time for the technology to mature in the NASA environment. As I mentioned in Section 1, the nuclear power industry has had plenty of time to improve its PRAs and one would indeed expect its numerical results to be consistent with operating experience. NASA, on the other hand, is still exploring what QRAs can do and how to apply them realistically to its systems, some of which are very different from nuclear power applications. For example, methodological improvements are needed to handle the dynamic nature of space flights. I believe that the agency should be given time to make this technology part of its standard analytical tools.

The fundamental question should not be whether old studies failed to accurately predict the number (although the results I cited above show that this is not quite true). The question should be: what did NASA do with these numbers? The answer is: very little, if anything. There are no guidelines as to which numbers are acceptable and, as I have said several times, the traditional safety requirements are intact. It is the impact on decision-making that matters, not that some study produced indefensible numbers.

Having seen first-hand how PRAs have improved safety in the nuclear power industry, I believe strongly that it would be a disservice to the nation, if NASA were discouraged from using this technology.

QRA results are not useful when they are highly uncertain. Why bother doing a QRA in those cases?

These uncertainties exist independently of whether we do a QRA or not. The decisions that need to be made will be better if quantitative information that has been peer reviewed is available. Recall the misperceptions of the frequency and consequences of core damage that nuclear safety professionals had before the Reactor Safety Study was issued (Section 6).

By attempting to quantify the uncertainties and to identify the dominant contributors, QRA analysts contribute to the common understanding of the issues and, in addition, may provide useful input to the allocation of research resources.

Probabilities cannot be realistically calculated.

This criticism presumably means that one cannot use straightforward statistical methods and divide the number of failures by the number of trials to calculate "realistic" probabilities. This criticism appears to miss the point that QRAs are performed for systems that are highly reliable and well defended, so a plethora of failures does not exist (and is highly undesirable to begin with). QRA methods analyze rare events in a systematic way and use all the available information in evaluating probabilities. QRA analysts are fully aware of the extensive use of expert judgment and always look at the uncertainties associated with the results. As I have stated, peer reviews are essential. Ultimately, the decision-making process is risk-informed and not risk-based.

QRAs do not include "one technician's unbelievable stupidity"[5]

I don't know how one defines "unbelievable stupidity," so I have difficulty dealing with this criticism. Of course, these facilities are usually operated by crews of several people, so I am not sure how many trained persons would have to behave in an unbelievably stupid manner for something extraordinary to happen. As I mentioned above (Section 5), QRAs do not model acts of unbelievable intelligence either, which, as the operating experience shows, occur frequently.

ACKNOWLEDGMENT

My views on PRA and its proper use have been shaped by many stimulating debates with my colleagues on the Advisory Committee on Reactor Safeguards of the Nuclear Regulatory Commission. This article has also benefited from comments by H. Dezfuli, T. Ghosh, and L. Pagani. The tone of some of my comments and all potential errors are mine.

REFERENCES

1. U. S. Nuclear Regulatory Commission, "Reactor Safety Study, An Assessment of Accident Risks in U.S. Nuclear Power Plants, WASH-1400," Report NUREG-75/014, Washington, DC, October 1975.

2. U.S. Nuclear Regulatory Commission, *An Approach for Using Probabilistic Risk Assessment in Risk-Informed Decisions on Plant Specific Changes to the Licensing Basis*, Regulatory Guide 1.174, Washington, DC, 1998. Available at www.nrc.gov.

3. Union of Concerned Scientists, "Nuclear Plant Risk Studies: Failing the Grade," August 2000. Available at www.ucsusa.org.

4. Advisory Committee on Reactor Safeguards, "Union of Concerned Scientists, 'Nuclear Plant Risk Studies: Failing the Grade,'" October 11, 2000. Available at www.nrc.gov.

5. “NASA’s Nuclear-Fueled Oddsmaking. Assurances Can’t Calm Fears of a Chernobyl in the Sky,” *San Francisco Chronicle*, April 28, 2003.

6. S. Kaplan and B. J. Garrick, "On the Quantitative Definition of Risk," *Risk Analysis,* 1 (1981) 11-28.

7. H.W. Lewis, R.J. Budnitz, H.J.C. Kouts, W.B. Loewenstein, W.D. Rowe, F. von Hippel, and F. Zachariasen, *Risk Assessment Review Group Report to the U.S. Nuclear Regulatory Commission*, Report NUREG/CR-0400, U.S. Nuclear Regulatory Commission, Washington, DC, 1978.

8. W.E. Kastenberg, G. Apostolakis, J.H. Bickel, R.M. Blond, S.J. Board, M. Epstein, P. Hofmann, F.K. King, S. Ostrach, J.W. Reed, R.L. Ritzman, J.W. Stetkar, T.G. Theofanous, R. Viskanta, S. Guarro, *Findings of the Peer Review Panel on the Draft Reactor Risk Reference Document*, Report NUREG/CR-5113, U.S. Nuclear Regulatory Commission, Washington, DC, 1988.

9. G. E. Apostolakis, H. Dezfuli, S. D. Gahring, M. B. Sattison, W. E. Vesely, *Report of the Independent Peer Review Panel on the Probabilistic Risk Assessment of the International Space Station. Phase II – Stage 7A Configuration*, Prepared for NASA Headquarters, Office of Safety and Mission Assurance, June 2002.

10. G.E. Apostolakis, R.J. Budnitz, P.O. Hedman, G.W. Parry, and R.W. Prugh, *Report of the Risk Assessment Expert Panel on the Tooele Chemical Agent Disposal Facility Quantitative Risk Assessment*, Prepared for Mitretek Systems, 1996.

11. R.S. Magee, E.M. Drake, D.C. Bley, G.H. Dyer, V.E. Falter, J.R. Gibson, M.R. Greenberg, C.E. Kolb, D.S. Kossen, W.G. May, A.H. Mushkatel, P.J. Niemiec, G.W. Parshall, W. Tumas, and J-S. Wu, *Risk Assessment and Management at Deseret Chemical Depot and the Tooele Chemical Agent Disposal Facility*, Committee on

Review and Evaluation of the Army Chemical Stockpile Disposal Program, National Research Council, National Academy Press, Washington, D.C., 1997.

12. American Society of Mechanical Engineers, *Standard for Probabilistic Risk Assessment for Nuclear Power Plant Applications*, ASME RA-S-2002.

13. R.J. Budnitz, G. Apostolakis, D.M. Boore, L.S. Cluff, K.J. Coppersmith, C.A. Cornell, and P.A. Morris, "Use of Technical Expert Panels: Applications to Probabilistic Seismic Hazard Analysis," *Risk Analysis*, 18 (1998) 463-469.

14. A. D. Swain and H. E. Guttmann, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*, Report NUREG/CR-1278, U.S. Nuclear Regulatory Commission, Washington, D.C., 1983.

15. M. Barriere et al, *Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA),* Report NUREG-1624, U.S. Nuclear Regulatory Commission, Washington D.C., 1998.

16. J. Reason, *Human Error*, Cambridge University Press, Cambridge, UK. 1990.

17. U. S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Draft Report, *Causes and Significance of Design Basis Issues at U. S. Nuclear Power Plants*, Washington D.C., May 2000.

18. Pickard, Lowe, and Garrick, Inc., Westinghouse Electric Corporation, and Fauske & Associates, Inc., *Zion Probabilistic Safety Study*, Report prepared for Commonwealth Edison Company, Chicago, 1981.

19. E. S. Beckjord, M. C. Cunningham, and J. A. Murphy, "Probabilistic Safety Assessment Development in the United States 1972-1990," *Reliability Engineering and System Safety*, 39 (1993) 159-170.

20. M.V. Frank, "Quantitative Risk Analysis of a Space Shuttle Subsystem," Proceedings of the International Topical Meeting on Probability, Reliability, and Safety Assessment, PSA '89, Pittsburgh, Pennsylvania, April 2-7, 1989.

21. Science Applications International Corporation, *Tooele Chemical Agent Disposal Facility Quantitative Risk Assessment*, Report SAIC-96/2600, Prepared for the US Army Program Manager for Chemical Demilitarization, 1996.

22. U.S. Nuclear Regulatory Commission, "Use of Probabilistic Risk Assessment Methods in Nuclear Activities: Final Policy Statement," *Federal Register*, Vol. 60, p. 42622, August 16, 1995.

23. National Research Council, *Understanding Risk: Informing Decisions in a Democratic Society,* National Academy Press, Washington, D.C., 1996.

24. "Columbia Disaster Underscores the Risky Nature of Risk Analysis," *Science*, 299 (2003) 1001-1002.

25. Science Applications International Corporation, *Probabilistic Risk Assessment of the Space Shuttle,* Report SAICNY95-02-25, Prepared for NASA Office of Space Flight, 1995.