Conclusion: Challenges for the Future
by D. Warner North

The title “Challenges for the Future” implies the challenge to summarize a very complex meeting. Of necessity, I will present a personal impression. My interest is in risk assessment, which I define as a process for summarizing science in support of decision making. Risk assessment is sometimes regarded as arcane numerology, a rigid process of computing risk numbers in which much available science is unused. I am a strong advocate for the broader definition of risk assessment. It is encouraging to learn how much science is becoming available for use in risk assessment for gasoline, its components, and alternative fuels.

The Decision Context

At the beginning of the symposium, John Graham noted that we are making multibillion dollar decisions on the basis of the studies presented. How do we summarize the science for a multibillion dollar decision? Carefully, with the participation and critical review of the best people available. This is what we have been doing. I want to extend my compliments and thanks to the sponsors and the symposium planning committee. Much new scientific information was presented, and the discussants, panelists, and audience have reacted to this information with perceptive questions, challenges, and suggestions for further analysis and research.

I do not have a large part of my professional life committed to the topic of this symposium. That puts me at a disadvantage in terms of the depth of my knowledge on many of the details. The symposium planning committee asked me to give this wrap-up talk in the hope that I can provide a perspective to help integrate all that we have heard over the past 4 days and to assess the implications of this material. I hope I can stimulate your thinking about the big issues: what science can tell us now about the health effects of gasoline; what science might be able to tell us in the future, given the expenditures that are made for further research, data acquisition, and analysis; how science can be summarized and communicated to those responsible for decision making; and how our institutions, procedures, and attitudes might evolve to facilitate better decision making.

Gasoline is perhaps the most familiar “chemical” for citizens of industrialized societies. Some of our largest industrial firms refine gasoline from petroleum and market it to hundreds of millions of consumers whose mobile lifestyle depends on it. So as manufacturers, vendors, and consumers, we have many decades of experience with gasoline. We have familiarity, but also fear and mistrust. Many of you know far better than I how the public feels about oil companies and government regulatory agencies. As several of you have pointed out, risk communication is a difficult business. The National Academy of Sciences report “Improving Risk Communication” indicates that the best way to make progress is assiduous attention to detail in the process of building trust.

Likewise, many of you may understand better than I do how strongly the public feels about increases in the price of gasoline. And how the public feels about reducing smog: the ozone and other pollutants formed from reactive hydrocarbon emissions and oxides of nitrogen. There is also the international competition issue: oil production and refining are moving offshore because of the regulatory environment. A lot of money and jobs are at stake. And then there is the greenhouse issue, which is linked to energy efficiency. How does it all fit together?

The challenge to all of us is to increase the extent to which decisions reflect an understanding of science. I think this implies we must do a better job of educating the public and our leaders in both the public and private sector. It is an awesome challenge. I submit that meeting the challenge will require that a large number of us become risk communicators as well as risk assessors and scientists. This meeting has been an important milestone. But to impact the decisions, the messages from this meeting need to reach a wide audience.

Let me now turn from the decision context to my summary of the meeting, of what was said, and the topics I think deserve more exploration in future meetings of this type. First, let me state my conclusion on the health effects of gasoline. We have decades of experience with gasoline in its present form. My impression is that the health effects are modest, but there are some. High doses of gasoline are clearly dangerous. At the same time, the data and analyses presented at this conference suggest that the health risk to the public from gasoline properly used in motor vehicles is small. For alternative (or reformulated) fuel, the health...
risks could be larger. What are the potential serious risks? And how can we facilitate careful assessment of these risks as input to the decision process?

**Exposure**

The symposium began with a session on exposure. Paul Lioy stated in summarizing the conclusions from last year’s gasoline exposure workshop: “There is a need for realistic exposure scenarios, and for understanding the uncertainties.” This theme has been echoed in various forms throughout the meeting. Gerry Aklad noted the need to look at microenvironments and human activity patterns. Jack Moore introduced the afternoon session by telling us that “Exposure data can be the criteria by which you decide whether you need toxicity data.”

Last year’s workshop and this year’s exposure session illustrated that better exposure data can be obtained. We can get measurements of the levels and the time at self-service pumps for consumers and at full service pumps for occupational exposure. I heard a lot of reinforcement for the value of old-fashioned industrial hygiene: break exposure situations down into tasks; then use modeling, measurements, extrapolation, and data in old files—whatever works. The data can always be made better for a price. What level of accuracy do we need? Roger McClellan and Peter Rombout stressed the interaction of using exposure and toxicity information together in an iterative risk assessment process that helps us to understand where additional data are most needed.

**Noncancer End Points of Gasoline and Key Components**

The literature reviews presented at this symposium on acute toxicity, neurotoxicity, and developmental and reproductive toxicity are clearly important contributions toward understanding the potential problems and planning needed research. But what do these data tell us about impacts on public health? We need to look at a broad range of exposure scenarios.

The potential for aspiration pneumonia for inhaling a small amount of gasoline liquid impressed me. The potential for acute toxicity as the result of human exposure is clearly high. Ingestion of a teaspoon of methanol can kill a child or cause blindness in an adult. It seems clear to me that we should look carefully at end points and exposures involving misuse, spills, accidents, and contamination rather than limiting risk assessments to inhalation exposure under routine conditions.

Lead was mentioned in passing; it is still used in gasoline in Europe and in most of the world, but yet no one used lead as an example of the identification and characterization of the health impacts of a gasoline constituent, past or present. Are there more compounds like lead out there? Not too many years ago, EPA made a decision to remove lead from gasoline. We have since observed a major reduction in the blood lead levels of children. We have reason to believe that elevated blood lead levels translate into subtle effects on the nervous system. Would EPA have required the lead phase-out if it had not been for the impacts of lead in poisoning catalytic converters? Mexico is now taking decisions to remove lead from gasoline. Other countries may be doing so as well.

We have not yet found a new “bad actor” comparable to lead—some effect indicating that a key gasoline component needs to be reduced from present or projected future levels. But we have identified benzene, and butadiene is also under suspicion.

We need more focus on mixtures and synergisms. For example, $n$-hexane is potentially harmful as a pure substance at high exposure levels; but as a component of gasoline, the risk may be reduced, because the enzymes that form the damaging metabolites from $n$-hexane get distributed among many other components in the mixture. We need to spend more effort on these mixture issues and to share our understanding of these issues with the regulatory agencies and the public.

**Carcinogenicity of Gasoline**

As a member of the Environmental Health Committee of EPA’s Science Advisory Board, I participated in a review of the bioassay results from inhalation of fully vaporized unleaded gasoline shortly after these bioassay results became available to EPA. Several of us on the Committee calculated upper bound cancer risk from the exposure information and cancer potency as computed from the bioassay results. I recall that my calculation gave $6 \times 10^{-5}$, and I was awed by the magnitude of the risk management issue that this result posed for EPA. But at that time, information on the mechanism of action for the male kidney tumors in the male rat was beginning to emerge from the research by James Swenberg and his colleagues at the Chemical Industry Institute of Toxicology (CIIT). This research and the interpretation of the research results by EPA have set an important precedent for the use of mechanistic information in risk assessment. While some uncertainties remain, many in the scientific community now think that enough is known about the male rat kidney response from fully vaporized unleaded gasoline to conclude that this animal tumor response is not predictive of a human cancer risk.

What about the other result from the bioassay, the liver tumors in female mice? Little has been done on this issue, and further investigation is needed. An interesting poster by Judy MacGregor and colleagues indicates the possibility of a hormonal mechanism for these tumors.

**Epidemiology Studies**

It is gratifying to see the epidemiology on refinery and distribution workers. We have run a massive human experiment, at the parts per million level, in the occupational group. I offer as a simplified conclusion that there is no large effect in causing human cancer. There is strong evidence for an elevated risk for acute myeloid leukemia in groups occupationally exposed many decades ago to benzene, and there is substantial but perhaps not conclusive evidence that benzene causes elevated incidence of other leukemias and myeloma as well.
CONCLUSION

If there is an elevated incidence of kidney or other cancers from the occupational exposure, it is small enough to be hard to find. As a nonepidemiologist, I am impressed with the difficulty of interpretation, given the population of healthy workers, lack of smoking data, and confounding effects of lifestyle (smoking, drinking, and disease misclassification problems).

I return to the theme of exposure scenarios. As an example, truck drivers who wore gasoline-soaked clothing for years inside a confined space should be good candidates for acute myeloid leukemia. This hypothesis seems well worth continued investigation with case-control studies. My suspicion is that if you find an elevated leukemia incidence, it will be associated with very high exposure levels well above today's occupational limits. Then we will need to ask what the implications are for today's occupational and ambient exposure levels.

Epidemiological data may not resolve the extent of the risk at low doses, and uncertainties will remain about the extent to which gasoline components cause various health endpoints. It is useful when informed expert judgment can be summarized in the form of a quantitative statement about the uncertainty. Bernard Goldstein gave an example when he told us that in his judgment the probability that benzene causes myeloma is greater than 50%. That type of statement conveys important information to those with risk management responsibility. Similarly, it is good to hear the epidemiologists discussing the need to include judgments on biological plausibility and not restricting themselves to a discussion of the p-values. Philip Enterline shared with us his classroom illustration, a medical decision situation involving a close relative, where an experimental treatment shows an expected improvement that is not statistically significant. In the scientific context, we do not yet meet the usual standard for concluding that the experimental treatment is effective. But in the decision context, the experimental treatment may be the best alternative available for the patient. Epidemiological data are important in the public health decision context even if statistical significance has not been obtained.

The presentations and discussion on formaldehyde, butadiene, and benzene shows the progress being made in risk assessment. A major theme is the importance of research to identify and characterize the underlying biological mechanisms.

Hazard Identification versus Exposure–Response Relationships

Roger McClellan stressed that the goal of protecting public health motivates increased emphasis on the steps beyond hazard identification in the risk assessment paradigm from the 1983 National Academy Redbook. These steps are dose– or exposure–response assessment, exposure assessment, and risk characterization. As Jack Moore noted, the information from occupational exposures is in the parts per million range while public exposures are in the parts per billion range. What is the shape of the dose–response curve, and what are the implications of a better understanding of mechanism for the extrapolation from parts per million to parts per billion? An increased understanding of what is biologically plausible may be critical for public health risk management.

Formaldehyde and butadiene provide examples of progress on developing mechanistic understanding that can be used in exposure–response assessment. The bioassay data for formaldehyde are highly nonlinear. A mechanistic understanding for these data is being sought in the research work at CIIT. Linda Birnbaum's presentation on butadiene summarized the progress being made in understanding metabolic differences between species. Both case studies illustrate how pharmacokinetic and pharmacodynamic approaches are being applied to model the relationship between exposure and response.

The third case study, on benzene, also stressed the importance of mechanistic information. Dennis Paustenbach concluded his presentation with the opinion that "Only better mechanistic data incorporated into a physiologically based pharmacokinetic (PBPK) model plus more robust epidemiological data will change [regulatory] decisions in 1992." Paustenbach noted the potential importance of information on short-term (peak) exposure to benzene. Peak rather than cumulative exposure might be the appropriate measure of dose for benzene.

The default procedure for dose–response assessment used by EPA and many other regulatory agencies is the application of the linearized multistage model to administered dose cumulatively over a lifetime. This model was intended for screening, based on one plausible mechanism for the cancer process. All three case studies involve biological information that may motivate replacement of the default procedure with an alternative based on more detailed understanding of mechanism for that case-study chemical. Key questions for the risk assessment are: Are we measuring the dose in the right way for the biological effect? (e.g., should we be using cumulative dose or peak dose?) Do we have the correct shape for the dose–response relationship? (e.g., should the relationship be linear through zero or nonlinear, or no response at all until a threshold has been exceeded?)

The implications of such departures for risk management can be profound. The poster by Cox and Ricci provides an example. They propose a cubic relationship for benzene dose–response. With a cubic relationship, risk drops much more quickly than with a linear model as one goes from occupational levels to the levels of benzene in ambient air. Would we then conclude that 50–ppb levels of benzene are acceptable, as opposed to the levels of 10–37 ppt that Goldstein and Paustenbach cited, assuming a linear model? What would be the implications of a cubic dose–response relationship for decisions on the reformulation of gasoline? While benzene levels of less than 1% are being contemplated for California, Europe continues to use gasoline with even higher benzene levels than in the U.S. How are decisions on the appropriate benzene content of gasoline going to be made? Will U.S. regulatory agencies continue to use cumulative lifetime-dose and low-dose linearity for benzene? Departures from these defaults will need to be supported by mechanistic information. Given the potential for altering the multibillion-dollar decisions
on gasoline reformulation, is an appropriate amount of mechanistic research now underway on benzene? The magnitude of the risk management decision suggests that the research strategy should be more aggressive.

**Risk Characterization**

It is clear that we need to do more to treat uncertainty explicitly and to improve upon single-number body counts as the input to risk management. I will defend EPA's defaults as appropriate for screening, but they can become highly inappropriate for chemicals for which we have developed an understanding of pharmacokinetics and mechanism of action. We have the example of the male rat kidney tumors involving the $\alpha_{2u}$-globulin mechanism, and formaldehyde and butadiene are emerging as examples where mechanistic understanding may support departure from default regulatory assumptions. Perhaps in a few years, we may have achieved a mechanistic understanding of benzene that will support departure from the default assumptions. Even if we can make a lot of progress on mechanistic understanding, uncertainties will remain. It will be important to explain to the public the scientific basis for departing from the usual assumptions in cancer risk assessment.

**Regulatory Decisions as a Context for Valuing Research**

While commending progress on the case studies, I also note that they represent substantial amounts of money and time. It took more than 10 years and millions of research dollars from initiation of research on the $\alpha_{2u}$-globulin mechanism to the EPA Risk Assessment Forum document that Imogene Rogers described. But the improvement in the decision seems to me clearly worth the time and money. Consider what it might have cost if a risk estimate based on the bioassay data had been the basis for decisions to impose tight controls on all U.S. gas pumps and on-board vapor control systems for fuel systems on cars.

The advances in biology have opened new opportunities, and I believe there is a strong case for vigorously pursuing these opportunities, even if they are expensive. For example, the stem cells for the human blood-forming system have been transplanted into mice, providing an excellent vehicle for leukemia research. Such mice could be used to study the mechanism by which benzene causes human leukemia. Perhaps by this or other new approaches we might get a firm biological basis for understanding whether the dose–response relationship for benzene-induced leukemia is linear, cubic as Cox and Ricci propose, or threshold in character. If we could conclude and persuade others that there is no leukemia risk for benzene exposures below 50 ppb, we might save billions in refining costs for reformulated fuels.

I do not want to conclude on the optimistic note that our problems will be solved by scientific breakthroughs. Much of the time the scientific research raises as many questions as it answers, and anticipated breakthroughs prove elusive. Research may tell us that some subportion of the population really is at high risk from exposure to gasoline, its components, or reformulated fuels. Then the savings from research will be valued in improvements to public health.

**Conclusion**

We should conduct risk assessment iteratively and share the data openly. This meeting and last year's exposure meeting seem like an excellent way to proceed. Industry has money and access to technical skills. Regulatory agencies have a mandate to develop understanding and to implement measures needed to protect health. The public and its leaders need to have trust that the science is reliable and that they can trust the regulatory agencies to do their job properly. Rather than have the regulatory agencies and the regulated industry in an antagonistic relationship, we should promote cooperation to obtain better science and to overcome public mistrust and fear. This last objective will require a continuing effort over a long period.

We are making progress. The example of the $\alpha_{2u}$-globulin mechanism for the male rat kidney tumor shows that industry and regulators can cooperate on the research and implementation of risk assessment advances based on improved mechanistic understanding. Furthermore, the mechanistic research has been complemented by large-scale epidemiological investigations to determine if an excess of kidney tumors and other forms of cancer are appearing in groups with high exposure to gasoline. Both types of research have significantly enhanced our ability to assess the risk to human health posed by exposure to gasoline.

This symposium has brought together a large and distinguished group of people from industry, academic and research institutions, and state and Federal regulatory agencies. Reflecting on the discussions we have had over the past 4 days, I believe this meeting reflects important accomplishments. We are making progress in improving the science, and we are making progress in translating these improvements into the risk assessments and the decisions on risk management for gasoline, its components, and alternative fuels. This progress should continue, and we must continue our efforts to communicate an understanding of the science to the wider audience—all those concerned with the risk management decisions regarding gasoline, its components, and its substitutes.