B·O·S·C

BOARD OF SCIENTIFIC COUNSELORS

May 27, 2008

Dr. George Gray
Assistant Administrator
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

Dear Dr. Gray:

The Board of Scientific Counselors (BOSC) has completed a requested review of ORD's Human Health Risk Assessment (HHRA) Research Program. A seven-member BOSC Subcommittee, including one consultant to the Subcommittee, was charged to conduct the Program review. The Chair of the Subcommittee also is a BOSC Executive Committee member.

The BOSC Subcommittee conducted four teleconferences from October 2, 2007, through January 28, 2008, two of which were after the face-to-face meeting to review and revise the draft report. The 3-day face-to-face review meeting of the HHRA Program was held November 14-16, 2007. This review was a detailed retrospective and prospective analysis of the Program and included extensive material describing the programmatic long-term goals, individual research programs, and budgetary and bibliographic information. The BOSC Executive Committee reviewed the Subcommittee's report and requested appropriate clarification and revision prior to approving transmittal of the final report to ORD in May 2008.

The principal charge of the review was to evaluate the HHRA Program from a program assessment framework relative to program relevance, structure, performance, quality, leadership, communication, and outcomes. A second priority was to provide a summary assessment and performance ranking for each of the three long-term goals identified with the HHRA Research Program.

The BOSC has duly discharged its responsibilities and, as an outcome of the review, has developed recommendations for ORD to consider based on the materials provided and the discussions organized as part of the review. The text of the report provides the full context and details for these comments as well as other specific recommendations. This report is anticipated to further assist ORD in longer term program enhancement,

comparative analysis with other programs, and intermediate research investment decision-making.

On behalf of the BOSC Executive Committee and the HHRA Review Subcommittee, it is my pleasure to transmit this program review report to you. We expect the report will assist ORD in continuing to improve its science, and assist and inform clients within and outside EPA of the significance of its research and its utilization.

Please feel free to contact me if you have any questions concerning this report. We look forward to your response.

Sincerely,

Gary S. Sayler, Ph.D.
Chair, BOSC

B·O·S·C

BOARD OF SCIENTIFIC COUNSELORS

REVIEW OF THE OFFICE OF RESEARCH AND DEVELOPMENT'S HUMAN HEALTH RISK ASSESSMENT PROGRAM AT THE U.S. ENVIRONMENTAL PROTECTION AGENCY

Final Report

Office of Research and Development
U.S. Environmental Protection Agency

BOSC HUMAN HEALTH RISK ASSESSMENT SUBCOMMITTEE

George Daston (Chair), Miami Valley Laboratories
Bruce Allen, Consultant
Henry Anderson, Wisconsin Division of Public Health
Richard Corley, Battelle Pacific Northwest National Laboratory
John Stephen Evans, Harvard School of Public Health
Mark Utell, University of Rochester School of Medicine
Lauren Zeise, California Office of Environmental Health Hazard Assessment

EPA CONTACT

Joanna Foellmer, U.S. EPA, Designated Federal Officer

April 1, 2008

This report was written by the Human Health Risk Assessment (HHRA) Subcommittee of the Board of Scientific Counselors, a public advisory committee chartered under the Federal Advisory Committee Act (FACA) that provides external advice, information, and recommendations to the Office of Research and Development (ORD). This report has not been reviewed for approval by the U.S. Environmental Protection Agency (EPA), and therefore, the report's contents and recommendations do not necessarily represent the views and policies of the EPA, or other agencies of the federal government. Further, the content of this report does not represent information approved or disseminated by EPA, and, consequently, it is not subject to EPA's Data Quality Guidelines. Mention of trade names or commercial products does not constitute a recommendation for use. Reports of the Board of Scientific Counselors are posted on the Internet at http://www.epa.gov/osp/bosc.

TABLE OF CONTENTS

I. EXECUTIVE SUMMARY

Background and Charge to the Subcommittee

The Human Health Risk Assessment (HHRA) Program provides EPA programs and offices with risk assessments on chemicals of critical importance, and advanced risk assessment tools that improve the quality and certainty of those assessments. The principal products of the HHRA Program are comprehensive reviews of the toxicity of chemicals, which are compiled in the Integrated Risk Information System (IRIS); Integrated Science Assessments (ISAs) on criteria air pollutants; and Provisional Peer-Reviewed Toxicity Values (PPRTVs) for environmental contaminants found at Superfund sites. The HHRA Program has an established goal for each of these: 16 new IRIS and 50 new PPRTV assessments each year, and updates of the ISAs for key atmospheric pollutants on an ongoing 5-year cycle. The Program also has made significant efforts in developing and applying improved risk assessment methods that make better use of experimental data and reduce the uncertainty in risk assessment.

Although entities like IRIS and ISAs are not new, the National Center for Environmental Assessment (NCEA) consolidated them into a single program, HHRA, in 2004 as a way to emphasize human health risk assessment as one of the strongest pillars of support for EPA's core mission to protect public health. The focus on risk assessment has allowed the HHRA Program to serve as a center for risk assessment methodology development. It also serves as a key risk assessment resource for the Agency, a resource that is tapped into on a regular basis to handle emergencies. HHRA expertise has played a crucial role in national emergencies such as the aftermath of the 9/11 terrorist attacks on the World Trade Center and Hurricane Katrina.

The HHRA Program has three long-term goals (LTGs) that encompass the development and application of risk assessment methods. The LTGs are:

- LTG 1—IRIS and Other Priority Health Hazard Assessments
  The goals for this aspect of the Program are to provide a set number of health hazard assessments on priority substances. These assessments are used to set standards for exposure and cleanup. Currently, the success criteria for this LTG are 16 new IRIS assessments and 50 new PPRTVs per year. There also is an expectation that existing IRIS assessments will be reviewed and updated on a regular basis.

- LTG 2—State-of-the-Science Risk Assessment Models, Methods, and Guidance
  The purpose of this goal is to develop, assess and adopt new risk assessment methods that reflect advances in the life sciences and statistics. These methods will improve the quality of risk assessments and provide greater support for decision-making.

- LTG 3—ISAs, Formerly Air Quality Criteria Documents
  ISAs are updated on a recurring 5-year cycle to include the best and most up-to-date science on the effects of exposure to the criteria air pollutants. The ISAs are critical to

EPA's Office of Air and Radiation (OAR) as it reviews the National Ambient Air Quality Standards (NAAQS), intended to protect public and environmental health with an adequate margin of safety.

A subcommittee of the BOSC was formed to review the progress of the HHRA Program toward its LTGs, and to make recommendations for improvements in the Program. The members of the Subcommittee are listed in Appendix A. The Subcommittee was tasked with responding to a number of charge questions that address aspects of the Program's relevance, structure, performance, quality, scientific leadership, coordination and communication, and outcomes; and to make summary evaluations for each LTG. The full charge is provided in Appendix B.

The Subcommittee met by conference call twice in October 2007, and for a face-to-face meeting in November 2007, in Bethesda, Maryland. The face-to-face meeting consisted of an in-depth review of all aspects of the Program. NCEA Director Peter Preuss presented an overview of the HHRA Program. Each LTG was introduced by a leader of the respective programs, followed by poster sessions that highlighted some of the work being performed within each LTG. The Subcommittee also heard from the key customers in the Agency's offices and regions who rely on the information and scientific expertise provided by the HHRA Program, as well as external users of HHRA products. The Subcommittee began drafting its report at the face-to-face meeting. A draft report was reviewed by the Subcommittee in December 2007 and again in January 2008.

Relevance

The Subcommittee concluded that the Program's goals are fully consistent with the Agency's strategic mission and with the Program's multi-year plan (MYP). The products from LTG 1 and LTG 3 are critical to EPA's regulatory mission and form the foundation for regulatory decisions and policies in a variety of program offices and regions. IRIS assessments are critical to a number of goals and objectives listed in EPA's 2006-2011 Strategic Plan, including:

Goal 1: Clean Air and Global Climate Change
- Objective 1.1: Healthier Outdoor Air
- Objective 1.2: Healthier Indoor Air

Goal 2: Clean and Safe Drinking Water
- Objective 2.1: Protect Human Health

Goal 3: Land Preservation and Restoration
- Objective 3.1: Preserve Land
- Objective 3.2: Restore Land

Goal 4: Healthy Communities and Ecosystems
- Objective 4.1: Chemical, Organism, and Pesticide Risks
- Objective 4.2: Communities

IRIS also serves as the internationally recognized standard in chemical risk assessment for other federal, state, local, and international regulatory bodies, and the private sector. The comprehensiveness, transparency, and consistency of the IRIS approach have made it into the internationally recognized standard in hazard characterization. The IRIS Web site receives about 8 million visits annually, a testament to the value of IRIS as a resource.

The Annual Performance Goal (APG) for IRIS is 16 new assessments per year. The HHRA is meeting this goal on a year-by-year basis. Given the complexity of IRIS reviews and the need for them to undergo rigorous internal and external review, 16 assessments per year should be considered to be an ambitious goal given current staffing and resources. The Subcommittee, however, was concerned that 16 assessments may not be enough to meet the Agency's (or indeed, the world's) needs for reliable assessments on chemicals of concern. The Subcommittee heard from high-level managers from EPA regions and a contractor commissioned by NCEA to survey the user community that a higher number of IRIS assessments would be of great value to them. The Subcommittee realizes that the HHRA staff is working at capacity (and indeed is leveraging the expertise of others within the Agency) to produce 16 assessments annually. The Subcommittee suggests that EPA consider allocating more resources to the HHRA Program if it feels that more IRIS assessments would facilitate EPA's mission. The Subcommittee also noted that there is a commitment to update IRIS assessments on a 10-year cycle or to remove them from the database. The Subcommittee understands the philosophy behind the policy, but believes that, given current resources, it may be impractical to update all of the assessments but inappropriate to remove any unless made irrelevant or incorrect because of data that have become available since the review. The Subcommittee recommends that NCEA develop a mechanism for prioritizing assessments for updating, and a mechanism for retaining assessments in the database that have not been updated within 10 years, particularly when new data on that chemical are limited.

LTG 3 is aligned with the requirements for assessment of criteria air pollutants as mandated by the Clean Air Act (CAA). The importance of the HHRA Program in meeting CAA requirements cannot be overstated. The Program has instituted internal and external feedback mechanisms that ensure that its work remains relevant to the Agency's needs, particularly via the ongoing scientific review provided by the Clean Air Scientific Advisory Committee (CASAC).

LTG 2 is to provide models, methods, and guidelines that improve the quality and reduce the uncertainty of the work done under the other LTGs, and for other programs in the Agency that conduct risk assessments. The research conducted under LTG 2 focuses on critical needs. Good strategic choices have been made to concentrate research in areas that are likely to bear fruit in ways that will markedly improve risk assessment. A good example of this type of relevant research is the development of benchmark dose guidance and software, which is now used in every IRIS assessment and is the standard benchmark dose methodology used worldwide. The Subcommittee saw numerous examples of other projects that have been responsive to the Agency's needs, such as the development of age-specific pharmacokinetic modeling.

The HHRA Program has been highly responsive to the needs of the program offices and regions. The Subcommittee heard testimonials from a number of highly placed officials in regions and program offices who strongly value the work and expertise of the HHRA, both in providing risk

assessment products (IRIS, PPRTVs, ISAs) and in supporting emergency responses to crises like the 9/11 terrorist attacks on the World Trade Center and Hurricane Katrina. Expertise in risk assessment within the HHRA Program has been critical in dealing with these and other disasters, yet neither the maintenance nor the application of this expertise has been captured in Annual Performance Measures (APMs). In one way this is understandable; disasters are, by nature, unpredictable. The Subcommittee, however, encourages the HHRA Program to find a way to catalog this important application of its expertise in its APMs.

The Subcommittee heard numerous examples illustrating how the HHRA Program has been responsive to recommendations from outside advisory boards and stakeholders. The extensive process by which CASAC reviews the plans and work products under LTG 3 demonstrates a commitment and responsiveness to outside advice. The research undertaken within LTG 2 is responsive to recommendations that have been made to the Agency by the Science Advisory Board (SAB) and the National Academies' National Research Council (NRC), and stakeholders, with regard to pharmacokinetic modeling, the limitations of biologically-based dose-response models, additivity to background, and others. These examples indicate the responsiveness of the Program to external advice; the Subcommittee, however, did not undertake a comprehensive review of external reviews and responses.

The Subcommittee saw strong evidence that the HHRA Program's work is integral to decisions that have public benefits. The ISAs produced under LTG 3 are the foundation for setting acceptable exposure levels for criteria air pollutants and have directly improved public health. Likewise, the IRIS and PPRTV assessments produced under LTG 1 form the basis for regulatory decisions made both inside and outside of the Agency on chemicals of concern. The tools and methods provided by the research undertaken within the HHRA Program have led to improved risk assessments.

Program Structure

The Subcommittee believes that the HHRA Program has a comprehensive and logical framework for producing high-quality risk assessments and for managing internal and external review processes. The consolidation of staff from multiple groups into a single core program under the HHRA rubric has facilitated communication and the adoption of standard practices in the groups that are responsible for EPA's risk assessments. Interaction and cooperation between the HHRA Program and other ORD Programs appears to be occurring at a high level.

The Program's structure appears to have facilitated meeting the goals of LTG 1 and LTG 3 in an efficient manner. The APGs have been met consistently. Having a core competence in human health risk assessment has been critical to the Agency in responding to emergencies such as Hurricane Katrina, the Libby, Montana asbestos cleanup, and others. These emergency activities are not captured in the long-term plans for the HHRA Program, but are of critical importance to the Agency. Every effort should be made to communicate the significance of these efforts to ensure that the program receives credit for, and continues to get support for these vital activities

LTG 1 calls for the HHRA Program to produce 50 new PPRTV assessments on an annual basis. The Subcommittee heard from the regional offices that these values are crucial to support the cleanup of contaminated sites. In the absence of IRIS values for a chemical, PPRTVs can have a significant impact on regulatory decisions. The absence of either IRIS or PPRTV values for a chemical may result in that chemical being inappropriately prioritized at cleanup sites. PPRTVs, however, are not publicly accessible. Therefore, it is worth considering making PPRTVs available to the general public.

The Subcommittee found the science that the HHRA Program is using in its risk assessments to be state-of-the-art, and that the research conducted under LTG 2 is forward looking and cutting-edge. The Subcommittee notes that the scientists within the HHRA Program are developing ties with the EPA's National Center for Computational Toxicology (NCCT). The Subcommittee believes that this is appropriate, and that many of the upstream projects within NCCT will eventually be transferred to HHRA for development into practical tools for risk assessment.

The MYP for HHRA logically lays out the workflow to meet each LTG. The targets for LTG 1 and LTG 3 are explicit and are being met in a consistent way. The work performed under LTG 2 is on target and is providing a steady pipeline of new methodologies. The Subcommittee heard from many stakeholders that they highly value HHRA's work products and would very much like to have additional IRIS and PPRTV assessments beyond the 16 and 50, respectively, which are added annually. There also is a strong desire to see PPRTV assessments upgraded to IRIS assessments on a more routine basis. The Subcommittee realizes that the HHRA Program is working at capacity and that it may not be possible to meet these requests but expressed an interest in knowing whether any analysis had been done to determine what resources would need to be allocated to increase the number of assessments done.

The HHRA Program is logically laid out and consistent with Agency priorities. While the MYP is clear with regard to processes and timelines for assessments, it was not clear in the presentations provided to the Subcommittee how chemicals are prioritized for IRIS and PPRTV assessments. The Subcommittee notes that representatives from two EPA regions who spoke at the meeting both felt that the current prioritization process worked for them. It appears that the program and regional offices are providing the HHRA program with information about the public health importance of the compounds that are prioritized, but it is not clear whether all public health concerns are being considered in the prioritization process. Therefore, it would be of value for the program to make its prioritization decisions more transparent.

Program Performance

The HHRA Program has defined 23 APGs across the three LTGs for 2006-2012. Within each of these are APMs. The three LTGs are making great progress toward their goals. For all three LTGs, the APMs were met in nearly every instance (in 2006, 2/2 for LTG 1, and 7/8 for LTG 2, and 2/2 for LTG 3). This success rate does not include acute assessments. Although these were included in the MYP, EPA has been instructed not to conduct these assessments. Therefore, the Subcommittee does not feel that it is reasonable to count the APMs for acute assessments in the Program's performance. At the time of review, performance toward 2007 goals appeared to be satisfactory. Most goals were completed or nearing completion. The Subcommittee also notes

that this level of performance does not include the emergency support provided by the HHRA staff, as it is not captured in the APMs. As in other parts of the report, the Subcommittee has heard from HHRA's customers that the work produced is of high value and that they would like even more IRIS and PPRTV assessments, if possible.

Program Quality

The quality of the HHRA Program was judged primarily on the basis of the global acceptance and use of the risk assessments, and the presentation and publication of research undertaken by HHRA staff. The very high quality of these products is evident. Internationally, IRIS assessments are considered to be of the highest quality and reliability. The research efforts presented to the Subcommittee had a high degree of scientific relevance and merit. The review of criteria air pollutants has an excellent record of past performance. The ISA procedure is relatively new, and given the fact that only one such assessment has been fully completed under the new process, it was difficult for the Subcommittee to assess the quality of the products produced under the new procedure in comparison to the old. However, in light of the fact that the revised procedure was developed and refined with the input of CASAC, stakeholders, and Congressional staff, the Subcommittee believes that the ISA work products will continue to be of high quality. That said, the process of reviewing criteria air pollutants has been substantially changed and the HHRA Program should be particularly vigilant in assuring that the new process provides assessments of comparable quality to those developed under the previous process. Although the Subcommittee believed that the input from CASAC, stakeholders and Congressional staff provided confidence that ISA work products will continue to be of high quality, it also thought there were several important issues that needed attention. These included: (1) the need for specific criteria that clearly articulate a process for inclusion or exclusion of scientific studies; (2) a strategy for identifying gaps in the science and a plan for filling in the gaps; and (3) the potential challenges associated with creating a single document, the ISA, that both summarizes the recent research and prioritizes the "key" studies across the older and more recent research.

The Subcommittee found that the HHRA Program makes extensive use of internal and external peer review to ensure that its outputs are of high quality. The extent of peer review for IRIS and ISA assessments exceeds most other examples with which the Subcommittee members were familiar. The review process appears to be effective.

Scientific Leadership

The HHRA Program is internationally recognized as a leader in risk assessment methods development and implementation. The Subcommittee was provided with numerous examples of scientific leadership, including: (1) the preeminent role in the translation of research findings into environmental health policy; (2) the role of NCEA in creating academic Particulate Matter Research Centers to help strengthen high quality science for use in the ISAs; (3) the development of mode of action data to inform the quantitative approach for low-dose extrapolation of risk; and (4) the development of physiologically-based pharmacokinetic (PBPK) model applications for IRIS and other priority health hazard assessments. HHRA scientists have broad-based expertise including environmental engineering, environmental health science, risk assessment,

epidemiology, geology, microbiology, physiology, statistics, toxicology, and management with senior and junior scientists and 34 postdoctoral trainees. HHRA scientists have a strong record of scientific service on journal editorial boards, in professional societies, and as adjunct faculty at universities, and have won numerous awards from EPA and external organizations. The HHRA Program produced almost 300 published journal articles between 2001 and 2007. The HHRA program has been viewed as a major source of core expertise for EPA. To build on that strength NCEA should consider recruiting one or two additional senior scientists, especially into the LTG 2 programs where efforts are underway to integrate emerging technologies into the risk assessment processes. Experienced investigators with proven track records are likely to more rapidly move these fields forward and serve as catalysts to junior scientists (LTG2-2).

Coordination and Communication

The HHRA Program has done an excellent job of engaging scientists and managers from other ORD Programs as well as program and regional offices in its planning. Procedures have been established for seeking input on IRIS and PPRTV priorities and are working well. The ISA process has a highly structured work and review process that includes continuing input from the program office, CASAC, and others. The ORD organizational structure of National Program Directors (NPDs) assists in facilitating coordination across labs, centers and programs, and assures a free flow of ideas to HHRA.

The HHRA Program also has formal relationships with external groups that facilitate coordination and avoid duplication of effort. This includes a memorandum of understanding (MOU) with the Agency for Toxic Substances and Disease Registry (ATSDR) that emphasizes the sharing of information on substances being evaluated by both organizations. HHRA also has close working relationships with the World Health Organization's International Program on Chemical Safety (WHO/IPCS), the International Agency for Research on Cancer (IARC), and the United Nations Environment Program (UNEP). HHRA regularly seeks input on its programs from outside stakeholders.

HHRA expends considerable effort to assure widespread awareness of and access to its work products. IRIS alone has almost 8 million visits to its Web sites, and other work products like the Exposure Factors Handbooks are available to the public and contain information that is not available from other sources. HHRA offers tools such as free benchmark dose software through its Web site. Of course, in addition to its Web site, HHRA scientists share their research through more traditional publications.

One set of work products that are not widely shared are the PPRTV assessments. PPRTVs far outnumber IRIS assessments and are being developed at four to five times the rate of IRIS assessments. They may be useful to other programs at the state or territorial level. Currently, PPRTVs and their supporting documentation are only available on a Web site restricted to use by EPA staff or to those who obtain special permission from EPA. The Subcommittee encourages EPA to make the PPRTVs publicly available for use in hazardous waste site risk assessment and promote their use where appropriate

Outcomes

The outcome measures are extremely well-defined for each LTG. For LTG 1 and LTG 3, these consist of annual targets for an explicit number of assessments. LTG 2 covers the development of state-of-the-art risk assessment methods, models, and guidance. The outcome measures for the projects in LTG 2 are more varied, but are well described. NCEA has made good, logical choices about areas in which to focus their research efforts. The Program staff for LTG 3 has developed a well-defined system for updating ISAs on an ongoing, recurring basis. The outcome measures for LTG 1 state a target value for the number of assessments to performed in any given year, but do not provide the identity of the chemicals to be assessed. The Subcommittee agrees that this is the only practical way to set performance measures that cover a 5-year period, as priorities are likely to change over time. The staff members who prepare IRIS and PPRTV values have established procedures for prioritizing chemicals. The IRIS prioritization process seeks input from multiple stakeholders both inside and outside of the EPA. The procedure appears to be well considered and to work well, but it is not completely transparent as to how decisions are made.

It was clear from the Subcommittee's review and the multiple testimonials from program and regional offices that the IRIS staff is viewed by the rest of the Agency as the source of risk assessment expertise to contact when crises arise. While it is clear that none of these crises can be anticipated, it seems to the Subcommittee that it is possible to anticipate that environmental crises will arise in a country the size of the United States on a regular basis. It is also possible, on a retrospective basis, to determine the amount of staff time spent annually in dealing with crises, and therefore, to estimate the level of effort that will be needed in the future.

The HHRA Program's products are routinely used by environmental decision makers to support and inform decisions and achieve results. Environmental decision makers inside and outside of the EPA are highly dependent on the work products of the HHRA Program. The Subcommittee received input during the course of the review from program offices, regional offices, and academia, as well as a survey that included information on state and international regulators, industry, and scientific consults to attorneys. All of these varied groups found the work products to be of immense value.

The HHRA Program does not have immediate access to the information it needs to evaluate the impact of its products on public health. Other programs have the responsibility of evaluating the effect of their decisions on public health risks and benefits. That said, the Subcommittee felt that it would be useful for the HHRA Program to obtain synopses of this information from the program and regional offices that are applying its values, if only to facilitate communicating the value of the Program.

Summary Assessments

LTG 1: Meets Expectations

LTG 1 has highly prescribed APGs requiring the production of a set number of IRIS reviews and PPRTVs each year. The HHRA Program is meeting those goals by producing the mandated number of reviews. These reviews appear to be of a consistent, high quality and are used as the

basis for decision-making in EPA program offices, regions, and outside of the Agency by other federal and state agencies as well as by private companies and organizations. The scientists who work in the HHRA Program are nationally and internationally recognized experts in risk assessment. They are well-connected with the rest of the Agency and do a good job of leveraging their limited resources by selectively enlisting experts from other parts of the Agency to assist in specific reviews. The scientists in this group are often leaders in their fields and participate in numerous external activities that advance the science.

In addition to the measured activities under LTG 1, the scientists who carry out the assessments are relied upon to provide their expertise on an emergency basis as crises arise for the various offices and regions. The NCEA staff performed admirably in responding to these crises. Their contributions are important and valuable in supporting rapid decision-making. The NCEA management should consider whether there is a way to account for the time spent in dealing with unforeseen events in its APMs.

The Subcommittee heard clearly from the Agency customers for LTG 1 products that they value the projects, but wish they could have more. The progress under LTG 1 could exceed expectations if the review process was more efficient (while maintaining its high level of quality) or if the HHRA Program had the staff to increase the number of reviews. Regardless, the level of effort will need to increase as newer, but potentially more time-consuming, methods of improving risk assessment become available. Furthermore, IRIS values are to be updated on an ongoing basis, which will put an additional burden on the staff that is not necessarily accounted for in the APMs.

LTG 2: Exceeds Expectations

NCEA is providing under LTG 2 a well-conceived portfolio of methods, models, and guidance to enable the performance of quality hazard and dose-response assessments under LTG 1 and to address critical elements of exposure assessment not addressed by other agency research efforts or the program offices. In establishing the goals for research and Program output, NCEA collaborates with other Agency national laboratories, particularly the NCCT and the National Health and Environmental Effects Research Laboratory (NHEERL), and receives input from the NRC. Over the past 3 years, NCEA has released a number of critical work products for guiding and conducting Agency assessments and is conducting research with the intention of output for 2008-2012 on several key areas of hazard and dose-response assessment. The Program is asking the right questions and engaging in the right activities to answer them.

Over the 2005-2007 timeframe NCEA has:

- Finalized the Guidelines for Carcinogen Risk Assessment and the supplemental guidance to address early life cancer susceptibility in risk assessment
- Released guidance on the conduct of PBPK modeling,
- Developed a framework for assigning childhood exposures,
- Released the Child-Specific Exposure Factors handbook for external review,

- Issued a final report on aging and toxic response,
- Updated the benchmark dose models for analysis of continuous non-quantal endpoints,
- Further developed its categorical regression models for analyzing endpoints across multiple toxicity domains, and
- Resurrected to the point of release for review of its All-Ages Lead Model.

In all of these efforts, the quality of the products has been very good.

Additional near term output is expected in the areas of mode of action data and their use in risk assessment decision-making, further guidance in PBPK modeling, further development of methods for categorical regression and meta-analysis, and the assessment of age susceptibility and less-than-lifetime exposures.

The characterization of risk information and uncertainty will need continuous refinement and careful consideration. The Program is aware of those needs and has initiated efforts to support fuller variability and uncertainty evaluations. A very impressive scientific staff has been assembled at NCEA to pursue those questions. While it is too early to fully judge these efforts, significant progress has been made and the presentations to the Subcommittee members indicated an approach that is sound and likely to have a positive effect on the quantitative analysis of uncertainty in decision-making on environmental health risks.

To further support its efforts and to provide the basis for its scientific judgment, the Program conducts workshops and also sponsors the NRC to conduct workshops, organizes symposia, and commissions white papers by expert scientists and more in-depth work by the NRC.

All of the evidence presented to the Subcommittee suggests that the end results of all of these efforts are exceptional work products of high scientific quality. Examples include the development of, and expansions to, approaches for incorporating more scientific information through models (PBPK, fate and transport, and dose-response) and methods for the quantitative representation of risk assessment findings (through analysis of variability and uncertainty and attention to the implications of mode of action). High quality work such as this indicated to the Subcommittee that the Program has exceeded expectations.

The Program's guidance documents, reports, and models are widely used in the conduct of risk assessment, by NCEA staff in the conduct of LTG 1 work, by the program offices and regions, and by other assessors at the local, state, national, and international levels. The work products of NCEA developed under LTG 2 are used extensively in the development of risk assessments used to inform environmental decision-making. The reliance of clients and users world-wide on the methods, models, and guidance produced by HHRA attests to their appropriateness and quality.

LTG 3: Meets Expectations

The assessment process under LTG 3 is scientifically advanced, both at NCEA and by an experienced peer review process. LTG 3 is designed to meet the mandates of the CCA for

conducting air quality health assessments and establishing the NAAQS for six criteria pollutants. The mandate also requires EPA to review the scientific basis for the air quality standards every 5 years to ensure that EPA is using the latest scientific knowledge. The HHRA Program has recently instituted a new process for obtaining input from stakeholders and the public as well as for conducting extensive internal and external peer reviews at numerous steps in the process before a standard is promulgated. The HHRA's product has recently changed from the long-standing comprehensive Air Quality Criteria Documents to the Integrated Science Assessment. As this is a relatively new process where external review drafts have only recently been completed, it is too soon to evaluate its performance other than to note that the HHRA Program has met its expectations as exemplified by completing the assessment for ozone in 2006 and lead in 2007 using the Criteria Document process while developing drafts for external review of the new ISAs for Nitrogen Oxides and Sulfur Dioxides on time in 2007. It is expected that much will be learned as these new ISAs undergo external peer review, which will assist the HHRA Program in adjusting their process where necessary. Nevertheless, the process for summarizing and integrating the health science data on criteria pollutants is generally quite mature.

The Subcommittee is encouraged that this new plan and the established timelines for completing external review drafts of ISAs that are under their direct control has the potential for improving the Agency's overall compliance with CAA mandates for incorporating the latest scientific advances on a 5-year cycle. Of particular note is the assistance from LTG 1 and LTG 2 in developing tools that facilitate the integration of science along the exposure-dose-response continuum and in the development of an annotated "living" database specific to the needs of LTG 3 to keep up with the latest advances in research associated with the six criteria pollutants. These processes are commendable and should lead to improvements in the HHRA Program's ability to maintain state-of-the-art assessments. It is also noteworthy that the HHRA Program has fostered the development of an interdisciplinary team that leverages expertise across the Agency as they work together towards the common LTG.

The Subcommittee saw evidence of exceptional work related to: (1) the appropriateness of the effort to integrate the science transmitted to the Office of Air and Radiation for their review in the rulemaking for criteria pollutants; (2) the proactive approach involving the scientific community and interested stakeholders in the development of scientific questions prior to the initiation of the specific review process; and (3) the comprehensive peer review process at many stages throughout the process. Because the LTG 3 process is so new for the Agency, however, it was not possible to assess the milestones or the quality of the product.

There are three issues that the HHRA Program should consider under the spirit of "continuous improvement." These include the need for: (1) increasing the transparency in the selection of studies utilized in the integrated summary. Specific criteria need to be developed that clearly articulate a process for inclusion or exclusion of scientific studies; (2) developing a strategy for identifying gaps in the science and then, wherever possible, creating a plan to fill those gaps; and (3) developing a process that deals with the potential challenges associated with creating a single document, the ISA, that both summarizes the recent research and prioritizes the "key" studies across the older and more recent research. For these issues, the Subcommittee recommends that the HHRA Program revisit this issue with CASAC and other potential customers and be prepared to accommodate their communication needs.

Summary of Recommendations (by LTG)

Overall Recommendation

- The Subcommittee considers the responsiveness of the staff members to national emergencies and the HHRA Program's contributions to particularly difficult cleanup sites as being of such high value that this should somehow be captured in the APGs.

Recommendations for LTG 1

- The HHRA Program needs to consider information on the potential public health concern of various chemicals as it prioritizes them for IRIS or PPRTV review. It appears that some of this information is being provided by the program and regional offices, but it would be of value for the Program to make transparent the basis for its prioritization decisions for IRIS and PPRTVs.

- The Subcommittee recommends that, in addition to the goals of 16 new IRIS and 50 new or revised PPRTV assessments per year, aspirational goals be established for increasing the number of IRIS assessments. The Subcommittee recognizes that it may not be possible to do more, given current staffing and budgetary limitations, but there is clearly a significant demand for these products.

- NCEA should assess what needs to be done to increase the Program's ability to produce more IRIS and PPRTV assessments per year, not only to meet their own stated objectives but also to satisfy the needs of their clients. This could either be in the form of a recommendation to the Agency for more resources, or the development of a more streamlined process.

- Mechanisms should be considered for retaining IRIS assessments older than 10 years that have not been updated, rather than allowing these assessments to expire and be removed from the IRIS database and Web site. One option is to simply annotate them as such.

- The Subcommittee recommends that well-developed PPRTVs be considered as a source of prioritization in the development of full IRIS documents. This should assist the HHRA Program in meeting its goal of producing 16 IRIS assessments per year, but also should facilitate the accomplishment of stretch goals for completing additional assessments

- In order to maintain the high level of quality that is evident in the HHRA work products, the Subcommittee strongly recommends that steps be taken to ensure the transparency of decisions made in the process of performing IRIS and PPRTV assessments

- PPRTVs far outnumber IRIS assessments and are being developed at four to five times the rate of IRIS assessments. They have been developed specifically to address the site-specific needs of EPA's Superfund Program. They have not undergone the Agency and interagency review required for toxicity values to be placed in IRIS, but have been peer

reviewed. However, they are developed using the same type of data sources, analyzed with the same level of scrutiny, and developed specifically for use at hazardous waste sites. They may be useful to other programs. Currently, PPRTVs and their supporting documentation are only available on a Web site restricted to use by EPA staff or to those who obtain special permission from EPA. The Subcommittee encourages EPA to make the PPRTVs publicly available for use in hazardous waste site risk assessment and promote their use where appropriate.

Recommendations for LTG 2

- The HHRA Program should continue to develop ties with NCCT, and should provide formal input to that Program on the aspects of its research that will be of value to HHRA.

- The HRRA Program should consider using available resources to recruit one or two additional senior scientists, especially into the LTG 2 Program where efforts are underway to integrate emerging technologies into the risk assessment processes. Experienced investigators with proven track records are likely to more rapidly move these fields forward and serve as catalysts to junior scientists.

Recommendation for LTG 3

- In order to maintain the high level of quality that is evident in the HHRA work products, the Subcommittee strongly recommends that steps be taken to ensure the transparency of decisions made in the process of performing ISA assessments

II. INTRODUCTION

The Human Health Risk Assessment (HHRA) Program provides EPA programs and offices with risk assessments on chemicals of critical importance, and advanced risk assessment tools that improve the quality and certainty of those assessments. The principal products of the HHRA Program are comprehensive reviews of the toxicity of chemicals, which are compiled in the Integrated Risk Information System (IRIS); Integrated Science Assessments (ISAs) on criteria air pollutants; and Provisional Peer-Reviewed Toxicity Values (PPRTVs) for environmental contaminants found at Superfund sites. The HHRA Program has an established goal for each of these: 16 new IRIS and 50 new PPRTV assessments each year, and updates of the ISAs for key atmospheric pollutants on an ongoing 5-year cycle. The Program also has made significant efforts in developing and applying improved risk assessment methods that make better use of experimental data and reduce the uncertainty in risk assessment.

Although entities like IRIS and ISAs are not new, the National Center for Environmental Assessment (NCEA) consolidated them into a single program, HHRA, in 2004 as a way to emphasize human health risk assessment as one of the strongest pillars of support for EPA's core mission to protect public health. The focus on risk assessment has allowed the HHRA Program to serve as a center for risk assessment methodology development. It also serves as a key risk assessment resource for the Agency, a resource that is tapped into on a regular basis to handle emergencies. HHRA expertise has played a crucial role in national emergencies such as the aftermath of the 9/11 terrorist attacks on the World Trade Center and Hurricane Katrina.

The HHRA Program has three long-term goals (LTGs) that encompass the development and application of risk assessment methods. The LTGs are:

- LTG 1 – IRIS and Other Priority Health Hazard Assessments
  The goals for this aspect of the Program are to provide a set number of health hazard assessments on priority substances. These assessments are used to set standards for exposure and cleanup. Currently, the success criteria for this LTG are 16 new IRIS assessments and 50 new PPRTVs per year. There also is an expectation that existing IRIS assessments will be reviewed and updated on a regular basis.

- LTG 2 – State-of-the-Science Risk Assessment Models, Methods, and Guidance
  The purpose of this goal is to develop, assess and adopt new risk assessment methods that reflect advances in the life sciences and statistics. These methods will improve the quality of risk assessments and provide greater support for decision-making.

- LTG 3 – ISAs, Formerly Air Quality Criteria Documents
  ISAs are updated on a recurring 5-year cycle to include the best and most up-to-date science on the effects of exposure to the criteria air pollutants. The ISAs are critical to EPA's Office of Air and Radiation (OAR) as it reviews the National Ambient Air Quality

Standards (NAAQS), intended to protect public and environmental health with an adequate margin of safety.

A subcommittee of the BOSC was formed to review the progress of the HHRA Program toward its LTGs, and to make recommendations for improvements in the Program. The members of the Subcommittee are listed in Appendix A. The Subcommittee was tasked with responding to a number of charge questions that address aspects of the Program's relevance, structure, performance, quality, scientific leadership, coordination and communication, and outcomes; and to make summary evaluations for each LTG. The full charge is provided in Appendix B.

The Subcommittee met by conference call twice in October 2007, and for a face-to-face meeting in November 2007, in Bethesda, Maryland. The face-to-face meeting consisted of an in-depth review of all aspects of the Program. NCEA Director Peter Preuss presented an overview of the HHRA Program. Each LTG was introduced by a leader of the respective programs, followed by poster sessions that highlighted some of the work being performed within each LTG. The Subcommittee also heard from the key customers in the Agency's offices and regions who rely on the information and scientific expertise provided by the HHRA Program, as well as external users of HHRA products. The Subcommittee began drafting its report at the face-to-face meeting. A draft of the report was reviewed by the Subcommittee in December 2007 and again in January 2008.

III. RELEVANCE

Question 1: How consistent are the LTGs of the Program with achieving the Agency's strategic plan and HHRA's MYP?

The LTGs are at the core of the HHRA's MYP, and as such are in every way consistent with it. The LTGs also are fully aligned with the 2006-2011 EPA Strategic Plan released on September 30, 2006.

Per LTG 1, IRIS hazard and dose response characterizations are produced for priority substances in response to the needs of the EPA programs and regions. Each IRIS characterization represents the Agency consensus position on the toxicological activity of the agent. IRIS is the main source of hazard identification and guidance values used by the EPA program offices to characterize risks from environmental exposures. Thus, IRIS is critical to meeting several of the goals and objectives in the 2006-2011 EPA Strategic Plan. These include:

Goal 1: Clean Air and Global Climate Change
- Objective 1.1: Healthier Outdoor Air
- Objective 1.2: Healthier Indoor Air

Goal 2: Clean and Safe Drinking Water
- Objective 2.1: Protect Human Health

Goal 3: Land Preservation and Restoration
- Objective 3.1: Preserve Land
- Objective 3.2: Restore Land

Goal 4: Healthy Communities and Ecosystems
- Objective 4.1: Chemical, Organism, and Pesticide Risks
- Objective 4.2: Communities

IRIS values are used by EPA program offices and regions to establish regulatory standards or guidance levels for non-criteria air pollutants, and for chemicals found in drinking water and at contaminated sites. IRIS values also are used to understand the risks from food residues not addressed by the Office of Prevention, Pesticides, and Toxic Substances (OPPTS), consumer product exposures, and cumulative risks experienced by communities.

The Strategic Plan goals and objectives are also furthered by the use of IRIS internationally and by other federal, state, and local agencies, regulated industries, environmental organizations, and academia. The Subcommittee agrees with the Agency observation that the roughly 8 million Web site visits per year to the IRIS site is an indicator of the relevance of the IRIS documents.

Under LTG 1, PPRTVs—and for data sparse chemicals, screening values—are produced in an expedited process to support the Superfund Program. This further enables meeting Objective 3.2– Restore Land.

An area of concern is the rather slow rate of production of assessments under IRIS. It is recognized that this rate is a function of NCEA resources, the generally slow process of peer review, and additional delay that may surround politically charged evaluations. The slow rate negatively impacts the overall relevance of the Program. Chemicals without toxicological evaluations are typically treated in quantitative characterizations as posing no risk or having little impact on a risk characterization. Residual risk calculations under the Air Program also are severely hampered by the lack of toxicity values for hazardous air pollutants. Risk estimates for environmental release scenarios can be strongly biased by the lack of information. This works against the ability of the Agency to meet the objectives in the Strategic Plan enumerated above. The Subcommittee heard from high-level managers in two EPA Regions about the critical need for greater output of IRIS values. The same message was heard from an NCEA contractor who evaluated the user community's impressions of IRIS. The Subcommittee understands that the IRIS Program is working at capacity and functioning well with existing resources; because of the shortfall between program and regional office needs and the capacity of the IRIS Program to produce new and revised assessments, the Agency should consider adding resources to the IRIS Program.

Under LTG 2, models, methods, guidelines, and parameter characterizations have been recently released or are under development to ensure that the science underlying Agency risk assessments is adequate and reflects the state-of-the-science. The Subcommittee was impressed by NCEA's well-reasoned and deliberate approach to determining areas of focus and levels of effort to address critical needs in this area, and to introducing new approaches into risk assessment practice within the Agency. The Strategic Plan areas met by this effort include: Objectives 1.6, 2.3, 3.3, and 4.4 – to enhance science and research to meet the clean air, clean water, land preservation and restoration, and healthy community goals. The NCEA's efforts will lead to more reliable risk characterizations aimed at meeting these goals.

The LTG 3 work is aligned with the assessment requirements for the CAA criteria pollutants and is designed to meet Objective 1.1 – Healthier Outdoor Air and Objective 1.2 – Healthier Indoor Air, under Goal 1 – Clean Air and Global Climate Change. It is difficult to overstate the importance of this Program for meeting Agency mandates under the CAA and ensuring public health and welfare. The Program has in place internal and external feedback mechanisms that ensure that Agency work remains relevant, especially the in-depth scientific peer review process by CASAC and the extensive opportunity for public and other expert input on LTG 3 activities.

Recommendation:

- NCEA should assess what needs to be done to increase the Program's ability to produce more IRIS and PPRTV assessments per year, not only to meet their own stated objectives but also to satisfy the needs of their clients. This could either be in the form of a recommendation to the Agency for more resources, or the development of a more streamlined process.

Question 2: How responsive is the Program focus to EPA's program office and regional health assessment needs?

The HHRA Program actively engages program offices and regions to understand and respond to health assessment needs. While the HHRA Program is responsive, the degree of response is limited by a lack of resources.

LTG 1 activities are designed to answer program and regional offices' health assessment needs: chemicals in need of assessment are identified; IRIS values and PPRTVs are developed on high priority chemicals; and health risk assessment advice and other support is provided in response to crises like the Libby, Montana asbestos cleanup and Hurricane Katrina. The Subcommittee heard from high-level managers in two regions that while NCEA is very responsive to their concerns, the extent of the response is necessarily limited by NCEA's ability to provide technical support due to limited resources. The emergency response and scientific advice to the regions is not spelled out in the LTGs, yet the work is critical for the protection of human health. One concern is the 10 year life span on IRIS assessments, and the 10 year re-review cycle. In a perfect world with unlimited resources, this would be wise to ensure that up-to-date science is used in assessments. Given current resources, however, this is impractical and has the potential for severe negative effects on the program offices and regional programs. This has the potential to substantially limit the number of new chemicals that can receive IRIS assessment, and removing assessments from IRIS simply by an arbitrary expiration date will result in a shrinking number of chemicals covered by IRIS. The resources committed to LTG 1 do not permit the HHRA Program to meet its goals for new assessments and for re-review of expiring assessments. Under questioning, Program staff members admitted that this was the case and that regarding these commitments, the numbers simply do not add up.

The LTG 2 effort is highly responsive to EPA program office and regional assessment needs. In the area of exposure assessment this is evidenced by the release of revisions to the Exposure Factors Handbook and the Child-Specific Exposure Factors Handbook, the scheduling of workshops, the research conducted to better understand exposures received by the elderly, the guidance provided on the monitoring of children for studying their exposure, and renewed efforts to describe the relationship between lead exposure and blood levels in adults with the All-Ages Lead Model. There is an impressive effort being made through the development of more nuanced pharmacokinetic modeling approaches that better articulate the linkage of external exposure to target tissue dose for the range of individuals and the uncertainty in these approaches. Under LTG 2, NCEA is beginning to explore approaches for characterizing toxicity on the basis of structure activity and non-apical endpoint data for data sparse chemicals; several scientists at NCEA and the NCCT are collaborating on projects that may eventually lead to new tools for toxicity assessment. Efforts are underway to develop better approaches to addressing sensitive groups and inter-individual variability. These various efforts will enable program and regional offices to understand more fully the level of protection afforded by actions contemplated or taken. Through careful work on mode of action and on biologically-based models, the Agency is developing an appreciation of these approaches and the usefulness and limits of different applications. This will improve the confidence in the assessments produced under LTG 1, an

important consideration as program offices and regions make decisions on the basis of these assessments.
It is too early to say whether the new process and the new form of the LTG 3 work products, the ISAs, will be more or less responsive to program office and regional health assessment needs in terms of transparency and communication of the scientific basis behind the criteria pollutant standards. Previously. the Criteria Documents, Staff Paper, and rulemaking packages together provided the documented scientific and policy rationale for the standard. This is now supplanted by the ISA and the Air Office's Advance Notice of Proposed Rulemaking. Clearly, the ultimate outcomes of NCEA's LTG 3 efforts address the needs of EPA's Air Program for scientific analysis to support the establishment of the standard, and of the health assessors in the regions and states for the articulation of the relationships between criteria pollutants and health effects.

Recommendations:

- Mechanisms should be considered for retaining IRIS assessments older than 10 years that have not been updated, rather than allowing these assessments to expire and be removed from the IRIS database and Web site. One option is to simply annotate them as such.

- The HHRA Program should continue to develop ties with NCCT, and should provide formal input to that Program on the aspects of its research that will be of value to HHRA.

Question 3: How responsive is the HHRA Program to recommendations from outside advisory boards and stakeholders?

The outputs of the HHRA Program are heavily scrutinized by stakeholders and subject to external peer review by advisory boards or individual experts. The Subcommittee saw various indications of responsiveness, and individual Subcommittee members have had a generally positive experience with this issue as peer reviewers of Agency products. The importance afforded to peer review and the commitment to high-quality science were obvious in discussions with Program staff. Clearly, the extensive process put in place by CASAC for the review of LTG 3 plans and work products demonstrates a commitment and responsiveness to outside advice. Process changes reflect Agency resource needs as well as the input received from CASAC, stakeholders, and the general public, although the reservations expressed by some advisors about these changes are acknowledged. The HHRA Program under LTG 2 is engaging in a number of critical areas identified in reports by the SAB and the NRC and stakeholders including: in pharmacokinetic modeling work to address model quality, physiologic parameter standardization, inter-individual human variability, uncertainty characterization in dose-response modeling, model validation, ascertainment of the uses and limits of biologically-based dose-response models, consideration of issues of background additivity, and in the PPRTV Program, an effort to explore approaches for developing toxicity measures for chemicals with sparse datasets. While these examples all indicate responsiveness, a thorough review by this Subcommittee of products pre- and post-review and the pursuit of recommendations from reviewing boards and individual experts was not undertaken. Finally, it is important to recognize the importance of receiving input from stakeholders while being circumspect and cautious in response.

Question 4: How clearly evident are the public benefits of the HHRA Program?

With regard to LTG 1, the Subcommittee heard testimony about the usefulness of IRIS assessments for decision-making directly from high-level management in two regions as well as in a report by a consultant who interviewed stakeholders about the IRIS Program. With roughly 8 million hits per year, it clear that the products of this Program are widely used and seen as important. IRIS and PPRTVs are heavily utilized to guide regulatory decisions; the Subcommittee also heard testimony regarding how critical some of the IRIS values have been for the clean-up of Libby, Montana, and the strong need for more as guides. The inefficiencies in this case, when no clean-up target values are available, were made pretty clear. Again, the removal of IRIS values that have not been updated in over 10 years from the IRIS Web site could result in major public costs, inefficiency, and could have health consequences.

LTG 2 efforts will lead to higher quality IRIS and PPRTV documents. While it is difficult to measure these impacts, it is nonetheless clear that overestimates of risk can result in undue economic cost and underestimates may result in misguided actions which also can be costly to the public's health.

Under LTG 3 the standards for the criteria air pollutants in large part are driven by the scientific analyses. Because certain current ambient levels of criteria pollutants have significant health impacts, the HHRA Program's LTG 3 output has directly impacted public health. The substantial reduction of criteria pollutant exposures over time reflects the actions by EPA, states, and regions to reduce levels in response to air standards, again based on an understanding of effect. With greater study of the criteria contaminants comes a greater understanding of their health impacts, including mortality, along with more subtle adverse effects. The public benefits of this Program are quite clear and reinforced by the Congressionally-mandated benefits assessment conducted by the Agency.

IV. PROGRAM STRUCTURE

Question 1: How clearly do the LTGs provide a logical framework for organizing and planning the health assessment activities and demonstrating outcomes of the Program?

Overall, the HHRA Program has developed a comprehensive and logical framework for not only producing the assessments that they are charged with but also in responding to the increasing demands for internal and external peer review. It also is clear that the consolidation of staff members that were located in multiple groups into a single core program is now bearing fruit and greatly facilitates communication, development, and adoption of standard practices, and a focus on continuously improving processes to respond to new challenges. Cooperation between the HHRA Program and the other research groups in ORD, the program offices, and regions appears to be occurring at a higher level than before and, based upon evidence provided to the Subcommittee, the products of the HHRA Program appear to be satisfying customer needs, at least in terms of quality, if not in the quantity of the assessments; this is a chronic issue that still needs to be addressed.

The IRIS Program under LTG 1 has a very clear framework and guidelines for organizing and presenting information leading to the development of cancer and non-cancer risk assessments. IRIS documents are posted on EPA's Web site and are thus available to any interested party. The sheer number of site visits and downloads are a testament to the importance of these documents to risk assessments. Given the reality that developing an IRIS assessment takes significant effort and time, the HHRA Program prepares PPRTVs using a streamlined version of the IRIS process so that their customers in the program offices and regions have at least some basis for carrying out their mandates. It does not appear, however, that the PPRTVs are available outside of the Agency nor is the annual prioritization list for the selection of chemicals publicly available, although it was evident from the regional offices dealing with Superfund sites that they felt the prioritization process was working for them.

The MYP plan has set forth that 50 new PPRTVs and 16 new IRIS assessments will be developed each year under LTG 1. The plans for reaching this goal are clear. It is hoped that as efficiencies are gained and additional staff members are hired, these numbers can increase over time. This is especially important for the IRIS documents, which are considered to be the gold standard. What is not clear in the organization plan is how the HHRA Program is going to revisit PPRTVs every 5 years and IRIS documents every 10 years. Furthermore, HHRA staff members have provided invaluable service to program offices, regions, states, etc. in responding to emergencies (e.g.. the 9/11 terrorist bombings, Hurricane Katrina) or assisting in difficult cleanup activities (e.g., asbestos cleanup in Libby, Montana). These high value activities are not captured in the overall framework and plans; the Subcommittee feels that the HHRA Program should be credited for these functions as well.

LTG 2 appears to have appropriately defined the critical needs for maintaining state-of-the-art risk assessments, particularly in the methods utilized in the toxicological assessments and the

establishment of handbooks and guidance documents. The Subcommittee believes also that validating new methods that are closest to being ready for use, such as PBPK modeling, is a rational emphasis. There are numerous assessments in progress that could benefit from the appropriate application of these models.

LTG 3 has embarked upon a new process for development of NAAQS. While the process is clear and logical, the Program has just entered the stage where the new style ISAs have been drafted for external peer review for two criteria pollutants, Nitrogen Oxides and Sulfur Dioxides. Much will be learned about how well this new process is going to function in the MYP as these two documents are externally reviewed and finalized. Thus, it is too early to determine the effectiveness of the new plan and the ISAs. One issue that is not clear is the process by which HHRA Program staff will evaluate and choose new science as it occurs to maintain transparency in the decisions that are reached within the ISA documents. Given the extensive experience of the staff dealing with the six criteria pollutants, they are fully aware of each pollutant's critical issues and data gaps and can, therefore, with modest effort, develop a more prospective process to document the decision-making process. During this review, it was apparent that HHRA Program staff members are clearly willing and able to be flexible if changes to the assessment and review process are deemed necessary.

Recommendation:

- the Subcommittee considers the responsiveness of the HHRA staff members to national emergencies and their provision of assistance at particularly difficult cleanup sites as being of such high value that it must somehow be captured in the APGs.

Question 2: How appropriate are the assessments and science used to achieve each LTG (e.g., Is the Program asking the right questions, or has it been eclipsed by advancements in the field?)?

Overall, the science that the HHRA Program uses and has currently developed to achieve the LTGs is considered state-of-the-art. IRIS documents prepared under LTG 1, in particular, are considered the gold standard. Although the Subcommittee found that the HHRA Program is addressing the right questions and using the right science, the prioritization process was not adequately covered in the presentations. However, representatives from the two EPA regions indicated that the prioritization process currently in place was working for them. Thus, while there are clearly not enough IRIS assessments to satisfy the demand, there appears to be an effort to work on higher priority chemicals, with PPRTVs being developed to at least provide clients with information until a full IRIS assessment can be developed.

For LTG 3, the criteria pollutants are mandated and reviewed on a 5-year cycle to ensure the latest scientific knowledge is used. It remains to be seen whether the new ISAs developed under LTG 3 will be met with the same level of enthusiasm or be as clear in process as the previous criteria documents, although it is clear that they too will have an enormous impact.

Under LTG 2, new methods are being explored to assist both the IRIS and PPRTVs in LTG 1 and ISAs in LTG 3 in keeping up with the latest advances that have achieved a reasonable level

of scientific consensus. Of particular note is the recent creation of the PBPK working group which now has critical mass in the field and consists of motivated staff members who are working on the major issues that have prevented the application of PBPK models in past assessments. This group is likely to move the field forward in a positive way that will create more consistency and transparency in PBPK model development and provide criteria for model evaluation and application in the risk assessment process. It also was apparent to the Subcommittee that the assistance from LTG 1 and LTG 2 in developing tools that facilitate the integration of science along the exposure-dose-response continuum and in the development of an annotated "living" database specific to the needs of LTG 3 to keep up with the latest advances in research associated with the six criteria pollutants are commendable and should lead to improvements in the HHRA Program's ability to maintain state-of-the-art assessments for criteria pollutants.

It also is appropriate that the Agency has developed a computational toxicology program (NCCT) that will eventually assist the HHRA Program in assessing and incorporating the latest advances in genomics, proteomics, and metabonomics that are rapidly transforming modern biology. While many of these approaches are not yet ready for use in a risk assessment, it is anticipated that they soon will be used to inform mode of action, dose-response, and sensitive population analyses. Thus, under LTG 2, close communication with NCCT will become increasingly important.

Recommendation:

- The HHRA Program should continue to develop ties with NCCT, and should provide formal input to that Program on the aspects of its research that will be of value to HHRA.

Question 3: Does the MYP describe an appropriate flow of work (e.g., the sequencing of related activities) that reasonably reflects the anticipated pace of scientific progress and the timing of client needs?

For LTG 3, a new process in place with established timelines for completing external review drafts of ISAs has the potential for improving the Agency's overall compliance with CAA mandates for incorporating the latest scientific advances on a 5-year cycle. The HHRA Program has extensive experience in complying with CAA mandates and it is expected that the new process can maintain, if not improve, the efficiency in incorporating the latest important scientific knowledge in reassessing each criteria pollutant every 5 years. As stated elsewhere, the Subcommittee feels that efforts should be put forward under LTG 3 to increase the transparency in the evaluation of the scientific merits of the latest research and the rationale for choosing which studies are most critical to their assessments and which ones are supportive, complementary, or even contradictory.

LTG 1 has a potentially greater challenge in meeting the needs of its customers given the sheer number of chemical assessments that are needed by both internal and external customers. For IRIS, no other risk assessment process receives more scrutiny through the numerous internal and external reviews. That is both the strength and liability of the IRIS Program; the documents are the gold standard in risk assessments performed by program offices, other federal agencies,

states, and even international organizations, but each IRIS document requires extensive effort to develop and finalize and there simply are not enough IRIS documents to satisfy the current demand. The PPRTVs developed under LTG 1 are a short-term effort to provide clients with initial assessments so they can respond to their own mandates. For some important chemicals, however, these PPRTVs can create a life of their own if they are not followed by a more extensive IRIS review within a reasonable period of time.

While the processes for developing IRIS documents and PPRTVs under LTG 1 are clear, it is noteworthy that the LTGs call for a static number of assessments to occur each year rather than calling for stretch goals to increase the number of assessments each year. The HHRA Program already has demonstrated that it can achieve its goal of preparing 16 IRIS assessments for external peer review and 50 PPRTVs even while responding to unplanned emergencies such as Hurricanes Katrina and Rita. Thus, it seems plausible that these numbers could be increased in the coming years as additional staff members are added and as the Program adapts to the more stringent review processes. Even with the increased demands for meeting the more stringent review requirements, it is notable (and laudable) that the Program met its APGs for producing the number of IRIS assessments, PPRTVs, and ISAs stated in the APMs.

Recommendations:

- The Subcommittee recommends that, in addition to the goals of 16 new IRIS and 50 new or revised PPRTV assessments per year, aspirational goals be established for increasing the number of IRIS assessments. The Subcommittee recognizes that it may not be possible to do more, given current staffing and budgetary limitations, but there is clearly a significant demand for these products.

- The Subcommittee further recommends that well-developed PPRTVs be considered as a source of prioritization in the development of full IRIS documents. This should assist the HHRA Program in meeting its goal of producing 16 IRIS assessments per year, but also should facilitate the accomplishment of stretch goals for completing additional assessments.

Question 4: Does the HHRA Program use the MYP to help guide and manage its health assessment activities?

The MYP was developed appropriately and in a logical manner and it was clear that the HHRA Program has developed its processes according to this plan. Although the MYP represents a long-term plan, it must inevitably be adapted over time to remain relevant to Program needs, scientific advancements, and resource constraints. What is not accounted for in the plan are the resources needed to respond to emergencies (e.g., the 9/11 terrorist bombings, Hurricane Katrina) and to provide expert scientific assistance to program offices, regional offices, and states. Although it is difficult to articulate specific APGs for these activities, the value of these contributions to the Agency needs to be acknowledged.

Question 5: How logical is the Program design, with clearly identified priorities?

The MYP has two major priorities: development of IRIS and PPRTV assessments for prioritized chemicals (LTG 1) and ISAs for criteria pollutants (LTG 2). In support of LTG 1, an additional long-term goal (LTG 2) was developed specifically to advance the scientific state-of-the-art in the risk assessment process. Within LTG 3, advances are being made to integrate information across the exposure-internal dose-response continuum and to develop a living, annotated database that, if successful, will facilitate the integration of new science on an ongoing basis to more efficiently comply with the 5 year reassessment mandates for the criteria pollutants. It also was evident that the MYP as it is being implemented is making gains in the development of IRIS documents and PPRTVs, although the demand, especially for IRIS documents, is still much greater than the supply.

The MYP is clear with regard to the processes and timelines that will be followed to develop and review IRIS documents, PPRTVs, and ISAs. What was not clear in the presentations provided to the Subcommittee was the prioritization process for the selection of chemicals for IRIS documents and PPRTVs. However, as stated elsewhere in this report, representatives from at least two of EPA's regions felt the prioritization process worked for them.

It also was obvious in this review that HHRA Program staff members are highly capable, dedicated, and motivated to achieving excellence in their work practices and to keeping up with the latest scientific advances in the development of their products even while "living in fishbowl." It also is obvious that HHRA Program staff have chosen to validate new risk assessment procedures that are close to being ready for actual use, while keeping informed of new developments that could eventually impact risk assessments.

V. PROGRAM PERFORMANCE

Question 1: How much progress is the Program making on each LTG based on clearly stated and appropriate milestones?

NCEA has defined 23 APGs across the three LTGs for 2006-2012. Within each of the APGs are APMs that have been used as milestones to quantify or measure progress on each LTG.

For all three LTGs, the Program appears to be making solid progress on the achievement of the LTGs. For all three LTGs, the APMs were met at or near 100 percent of the time for 2006 (2/2 for LTG 1; 7/8 for LTG 2; 2/2 for LTG 3); the one exception appears to be APG 7, APM 357 which calls for two final reports, one on dose-response models and population exposure methods for assessing microbial risks, and the other on defining microbial research needs for the Office of Water (OW). These above rates of APM success do not include acute assessments, which the Agency was directed not to do, after they had been planned and set up as APMs. The Subcommittee was informed that if those APMs were included in the determination of the rates of completion, the success rate for 2006 would be 63 percent.

The goals for 2007 include three APMs for each of the LTGs. Progress towards completion of those APMs appears to be satisfactory (including, at the time of this review, completion of all three goals for LTG 3).

With respect to LTG 1, the APGs for every year include the completion of 16 high priority health hazard assessments, and 50 new or renewed PPRTVs. It is noted that this rate of completion will not satisfy the stated goal to have no IRIS entries over 10 years old; because there are now over 540 IRIS chemicals, a renewal rate alone of 54 per year would be needed to achieve that goal.

The above appraisal specifically excludes consideration of the special requests for response that are not part of the planning process and are not defined in terms of APMs. They are not part of an appraisal of the performance of HHRA-related activities, but they do consume a sizable portion of the resources available, both in terms of time and staffing. The examples presented to the Subcommittee included support provided in relation to Hurricane Katrina, the 9/11 terrorist bombings, and lead levels in Washington, DC drinking water, as well as ongoing support in Libby, Montana. It is the belief of the Subcommittee that there should be recognition of these efforts, which often address the most pressing and urgent needs of the clients, the program and regional offices. In the view of the Subcommittee, the quantitative performance assessments underestimate the contributions of NCEA staff members by not factoring in these important efforts.

In summary, the Program appears to be making substantial and satisfactory progress on each LTG based both on the clearly defined milestones (APGs and APMs) and on providing the support requested in response to unscheduled emergency needs.

Recommendation:

- NCEA should assess what needs to be done to increase the Program's ability to produce more IRIS and PPRTV assessments per year, not only to meet their own stated objectives but also to satisfy the needs of their clients. This could either be in the form of a recommendation to the Agency for more resources, or the development of a more streamlined process.

VI. PROGRAM QUALITY

Question 1: How good is the scientific quality of the Program's health assessment and methods, models, and guidance products?

The quality of the products of the HHRA Program were judged primarily on the basis of the global acceptance and use of the risk assessments and the presentation of the research efforts completed and currently being pursued by staff scientists. On both counts, the very high quality of those products was evident.

Related to LTG 1, the IRIS assessments are the top-tier choice for use in most, if not all, site assessments. They also appear to be used extensively in other countries to support risk-related decision-making. These "customer satisfaction" indicators speak directly to the perceived quality of those documents.

For LTG 2, the research efforts presented to the Subcommittee also indicate a high degree of scientific merit and relevance to advancing the science of risk assessment. The staff pursuing advances in PBPK modeling, in particular, was noted by the Subcommittee to be excellent and to be producing extremely valuable models and modeling approaches. The Program efforts related to understanding, quantifying, and representing uncertainty and variability were noted to be of particular relevance and utility for enhancing the value of risk assessment products related to LTG 1 and LTG 3. The continuing methods-development and implementation (e.g., software) efforts related to benchmark dose modeling also were viewed as providing significant benefits both within and outside of the Agency.

Recent revisions in the format of the products of the ISA procedure, and the fact that only one such assessment had been completed fully under the new process, make it difficult to assess the quality of those products related to LTG 3. Based on the excellent past performance with respect to the criteria pollutant assessments, and the fact that the revised process was devised and refined with the input of CASAC, stakeholders, and congressional staff, the Subcommittee has confidence in the continuing quality of the ISA work products. Nevertheless, there were some factors that had the potential to affect Program quality and needed to be thoughtfully addressed. These included: 1) the need for specific criteria that clearly articulate a process for inclusion or exclusion of scientific studies; 2) a strategy for identifying gaps in the science and a plan for filling in those gaps; and 3) the potential challenges associated with creating a single document, the ISA, that both summarizes the recent research and prioritizes the "key" studies across the older and more recent research.

The high quality of all of the efforts also is supported by strong evidence of relatively high publication and citation rates in high-visibility journals of significant scientific reputation.

Recommendation:

- In order to maintain the high level of quality that is evident in the HHRA work products, the Subcommittee strongly recommends that steps be taken to ensure the transparency of decisions made in the process of performing IRIS, PPRTV, and ISA assessments.

Question 2: What means does the Program employ to ensure quality in its health assessment activities and products (including peer review, competitive funding, etc.)?

The Program utilizes an extensive system of both internal and external peer review to ensure the quality of its risk assessment products. For IRIS and other priority health hazard assessments (representing the products for LTG 1), fully five peer review steps are included in the process leading from the scoping meeting to the final assessment. Those include NCEA-wide consultations, an Agency review, an OMB/interagency review, external peer review (with public comment), and a final Agency and interagency review. The other primary LTG 1 product, the PPRTV, is intended to be a quicker and more expedited process than a full IRIS assessment. Still, it includes both an internal NCEA review and an independent external review.

For the ISAs (related to LTG 3), CASAC provides independent input and review at three distinct points in the process of their development. Public comments are solicited at those same points and subsequently following EPA's proposed decision on standards. Moreover, OMB and interagency reviews appear to be integral parts of the ISA process.

The Subcommittee noted that the level of peer review for the both the IRIS and ISA documents equals or exceeds any other examples with which the members were familiar.

Recommendation:

- NCEA should assess what needs to be done to increase the Program's ability to produce more IRIS and PPRTV assessments per year, not only to meet their own stated objectives but also to satisfy the needs of their clients. This could either be in the form of a recommendation to the Agency for more resources, or the development of a more streamlined process.

Question 3: How effective are these processes?

The effectiveness of the processes to assure quality can best be judged by the end product of the process. That is, assessing the quality of the science and personnel involved in the assessments and methods and models development products is the measure of the effectiveness of the process. Based on the above-cited perception of usefulness and relevance that users world-wide have expressed, and the relatively high publication and citation rates, these processes appear to be working very effectively.

VII. SCIENTIFIC LEADERSHIP

Question 1: Please comment on the leadership role the HHRA Program and its staff have taken in contributing to advancing the current state of the risk assessment science and solving important risk assessment problems.

There are important areas in which HHRA Program scientists have played leadership roles at both the national and international levels. The HHRA Program is clearly recognized as an international leader in risk assessment in both methods development and implementation. Areas of impressive leadership are related to IRIS and Air Quality Health and Environmental Assessments, which provide EPA's scientific positions on potential adverse health effects that may result from exposure to environmental chemicals and criteria air pollutants, respectively. For IRIS assessments, the toxicity values combined with specific exposure information provide the basis for risk assessments that are utilized not only in the United States but also throughout the world. The recent emphasis on uncertainty analysis and data-derived uncertainty factors has strengthened the risk assessment process. Another area of impressive leadership is the strategic planning regarding risk assessment methods development to incorporate state-of-the-science advances. For example, novel computational methods are being developed by NCCT, using toxicogenomics, structure-function, and systems biology approaches. These methods, which are at a very early stage of development, may eventually provide the basis for better identifying susceptible populations and the biological mechanisms responsible for susceptibility. Specifically, they provide a novel and state-of-the-art opportunity to explore the interface between genes and environment in disease causation. The HHRA collaborates with NCCT in this long-term research effort to promote the relevance and usefulness to risk assessment for this effort. In addition, the HHRA Program is providing high-quality leadership in the development of a new process for the ISAs for the criteria pollutants. These health evaluations and recommendations provide the underpinning for the NAAQS and have significant international impacts. In the recently revised NAAQS process, HHRA leadership has taken a proactive approach by involving the scientific community and interested stakeholders in the development of scientific questions prior to the initiation of the specific review process.

Additional examples of leadership include: (1) the preeminent role in the translation of research findings into environmental health policy; (2) the role of NCEA in fostering academic Particulate Matter Research Centers to help strengthen high-quality science for use in the ISAs; (3) the development of mode of action data to inform the quantitative approach for low-dose extrapolation of risk; and (4) the development of PBPK model application for IRIS and other priority health hazard assessments.

The HHRA Program scientists also demonstrated leadership in their respective disciplines. They are a multi-disciplinary group with broad-based expertise including environmental engineering, environmental health science, risk assessment, epidemiology, geology, microbiology, physiology, statistics, toxicology, and management with senior and junior scientists and 34 postdoctoral trainees. The Subcommittee was provided with tables of accomplishments, including leadership roles taken on by scientists, awards received, university appointments,

editorial work, invited presentations, and a bibliographic assessment. The documents demonstrate that the HHRA Program scientists provide extensive and highly regarded leadership in their fields. The following is a list of some of these accomplishments:

- Scientists serve as members of editorial boards for 14 journals and reviewers for 51 journals.

- Twenty seven of the Program scientists are or have served as adjunct professors or lecturers at universities across the United States.

- Scientists have received a number of EPA awards, including five gold medals for exceptional service, nine silver medals for significant service, 64 medals for commendable service, three President's awards for meritorious service, five Administrator's Awards for Excellence, and numerous scientific and technological achievement awards at all levels.

- Scientists have received a number of awards and honors from organizations and agencies outside of the EPA, including the United States Department of Agriculture (USDA) (for outstanding professional service), the U.S. National Parks Service, the U.S. Public Health Service (PHS), the National Society of Professional Engineers, the Society of Toxicology, the Society of Pathology, and the Society for Environmental Toxicology and Chemistry as well as a letter of appreciation from the Israeli Ministry of the Environment.

There is ample evidence documenting the leadership of HHRA scientists in providing advice and assistance to the Agency. This includes activities across the Agency, across programs and regions, and across ORD. Examples of this include serving on national and international advisory committees, organizing workshops, authoring strategic and research plans, and serving on review panels. Scientists provide advice and assistance in the areas of exposure and risk assessment to the Agency and program offices, notably to OAR, the Office of Solid Waste and Emergency Response (OSWER), OPPTS, and virtually all of the regional offices. This may be, for example, in the form of technical guidance, training, workshops, or the development of an action plan for asbestos in Libby, Montana. Across ORD, HHRA scientists are engaged in assisting in risk assessment applications, authoring MYPs, drafting research strategies in new areas such as nanomaterials, serving on expert panels, and organizing international symposia or interagency workshops.

There is an equally impressive listing of HHRA scientists' leadership activities external to EPA. They are invited presenters at professional conferences, program chairs for professional meetings (international institutes, Environmental Mutagen Society, and the Society of Toxicology (SOT)), serve as elected officers in professional societies, and organize and participate in national and international workshops. They provide assistance to the international scientific community such as WHO, Health Canada, the government of Brazil, and the Japanese Asbestos Research Delegation.

The Subcommittee also considered the quality and impact of peer-reviewed publications as an index of providing leadership to the scientific community. The bibliometric analysis covered 292

journal articles published from 2001 to 2007 on HRRA research. The analysis was completed using Thomson's Essential Science Indicators (ESI) and Journal Citation Reports (JCR) as benchmarks for the influence and impact measures. Journals can be ranked by their impact factor, which reflects the frequency of citation of published papers in a given year, and helps evaluate a journal's relative importance when compared to other journals in the same field. Forty-one percent (120) of the HHRA journal papers have been published in journals ranked in the top 10 percent of journals ranked by JCR, which is 4.1 times higher than expected. The number of times a paper has been cited can be compared to the expected number of citations, which is the average frequency of citation of papers in a particular journal. In 9 of the 14 fields in which the 292 HHRA journal papers were published, the ratio of observed to expected citations is greater than 1, indicating that most Program papers are cited more widely than the average paper. Forty-eight (16.4%) of the 292 HHRA journal publications qualified as highly cited when using the ESI criteria for the top 10 percent of highly cited publications. Further, the self-citation cite for Program papers is 5.8 percent, which is well below the accepted range of 10 to 30 percent author self-citation rate, indicating that the excellent statistics reflect the use of the HHRA research by other scientists.

Taken as a whole, the evidence speaks to a community of highly trained and productive scientists, many of whom are leaders in their field, who are providing leadership to the United States and international governments as well as scientific communities and are engaged in risk assessment science and in solving important risk assessment problems.

Question 2: How should the HHRA Program implement recruitment methods, incentives, or training procedures to maintain and increase leadership in the field and transition staff to emerging science fields and assessments?

The HHRA Program should continue to reward scientific excellence and keep administrative responsibilities to a minimum for junior scientists. Opportunities should be sought to provide young scientists with adjunct university appointments. This not only fosters collaboration but provides the opportunity for junior scientists to interface with the cutting-edge science that most frequently occurs in the academic environment. Recruitment and retention of the "best and brightest" is fundamental to success and is enhanced by a culture that allows access to more senior mentoring as well as bright and motivated students.

Recommendation:

- The HRRA Program should consider using available resources to recruit one or two additional senior scientists, especially into the LTG 2 Program, where efforts are underway to integrate emerging technologies into the risk assessment processes. Experienced investigators with proven research track records are likely to more rapidly move these fields forward and serve as catalysts to junior scientists.

VIII. COORDINATION AND COMMUNICATION

The Subcommittee was charged with a prospective and retrospective program review. As part of that review, the Subcommittee was to specifically evaluate the coordination and communication components of the HHRA Program. The factors considered in the evaluation included the linkages of the HHRA Program to other ORD programs and the MYPs for both the transfer of new science and data into assessments and to provide insight and direction for prioritizing research in support of improved risk assessments; the dissemination of assessments and peer-reviewed documents and risk assessment tools over the Internet; seminars and training for ORD, program and regional offices, at national and international meetings and workshops; and the published literature. The Subcommittee's discussion was organized to address coordination and communication issues concerning each of the LTGs and focused on the following three charge questions.

Question 1: How effectively does the Program engage scientists and managers from within ORD and other relevant program and regional offices in its planning?

The HHRA Program plays a unique role in serving the needs of the EPA programs and regions through incorporating, integrating, and coordinating the use of scientific information as a foundation for regulatory decision-making. To achieve success, communication and coordination have been effectively institutionalized within HHRA. Given the Program mission, communication and coordination are critical to success. These activities are well established and occur vertically and horizontally within NCEA and other relevant EPA programs and regional offices. Well-documented systems are in place and have operated for many years to provide a systematic, structured prioritization and communication strategy to assure EPA program and regional office scientists and managers are effectively involved in setting priorities for development and HHRA activities such as IRIS and PPRTV assessments that reflect the client's needs. Priorities are set annually and planning processes provide active rather than passive priority identification so that needs that surface during emergent incident response assessment activities can be efficiently and systematically addressed in a timely fashion. The ORD organizational structure of NPDs who meet regularly assists in facilitating coordination and assures a free flow of ideas and needs to, and responses from, HHRA. This high-level communication and coordination strategy is fundamental to enabling the identification of data gaps, the highest priority research to be undertaken, and risk assessment tools developed that support on-the-ground risk assessment and problem solution needs. Seminars, training, and the development of guidance documents and Web-based information are all included in the internal EPA coordination and communication strategies adopted.

While the flagship HHRA products are IRIS documents, PPRTVs and guidance documents, it is the broad senior scientist staff expertise that is the critical glue to responding efficiently and effectively to Program needs on a day-to-day basis as well as over the long-term. The availability of internationally recognized expertise in HHRA to advise and assist in site-specific and substance-specific program solutions further assures coordination and communication. It is the

well-functioning HHRA team that provides the capability to respond effectively to unique situations that require creative solutions. This experience is then translated into the development of long-term solutions to resolve gaps identified.

Question 2: How effectively does the Program engage outside organizations, both within and outside of the government, to promote collaboration, obtain input on Program goals and assessment priorities, and avoid duplication of effort?

The HHRA Program has developed formal relationships with outside organizations to facilitate coordination and avoid duplication of effort. Specifically, EPA-IRIS has an MOU with ATSDR. ATSDR develops toxicological profiles including quantitative minimal risk levels (MRLs) for non-cancer effects. The MOU emphasizes the sharing of information on substances under evaluation by both organizations.

The HHRA Program maintains close relationships with international organizations addressing environmental health risks such as the WHO/IPCS, IARC, and UNEP. In addition to these formal agreements, HHRA scientists belong to scientific organizations, sponsor workshops and use other methods to obtain programmatic recommendations and facilitate collaboration with academic institutions. They also frequently seek direction and share information with state agency toxicologists and risk assessors. These collaborative relationships have resulted in HHRA scientists being recognized as global leaders in conducting state-of-the-science health risk assessments and assures that the field of health risk assessment remains dynamic and addresses identified gaps and methodological needs.

In addition to targeting specific allied programs and scientists, the HHRA Program regularly publishes notices in the Federal Register requesting input from anyone wishing to contribute. Such notices provide the broadest opportunity to reach scientists and interested parties who might not be reached via other communication routes. After they are developed, HHRA products undergo extensive peer review to assure that the risk assessments and methods reflect the scientific consensus. Thus, HHRA seeks external advice at both the front-end planning and priority-setting stage and the final product evaluation. NCEA has led efforts to obtain external consultations with the National Academy of Sciences and works closely with state agencies to develop approaches to fill data gaps and develop needed risk assessment tools and approaches. Such collaborations have advanced the scientific community's understanding of how to characterize risk to susceptible populations. One example is the use of mode of action (MOA) and innovative approaches for dosimetric adjustment for inhalation exposures for application in risk assessments.

Question 3: How effective are the mechanisms that the Program uses for communicating assessment results both internally and externally?

Effective external communication is critical to the success of the HHRA Program and considerable effort is expended to assure widespread awareness of the Program's products and efforts. HHRA external communication occurs primarily in the scientific arena through active

presentation of state-of-the-science research and innovative methods at scientific conferences and through placing a priority on scientific peer review publications. With the exception of the PPRTV support documents, assessments (such as IRIS), methods, guidelines, and reference documents such as the Exposure Factors Handbooks, are all available to the public on the Internet and provide information not available from any other source. IRIS alone has more than 7 million website downloads a year. Software, such as that developed to generate benchmark doses and points-of-departure is available via the Web site for use by risk assessment practitioners. This software is free with user registration. Providing free software and training workshops to promote the use of such methodological advances has been instrumental in promoting the use of benchmark methodology and its replacement of the more traditional no observed adverse effect level (NOAEL) and lowest observed effect level (LOAEL) methods. The efforts to facilitate access to such methods are a classic "research to practice" integrated activity. HHRA has been a leader in developing methodological advances. Without the communication and collaborative efforts to aggressively disseminate and promote what HHRA has developed for use by others, however, adoption of new techniques by the risk assessment community would be impeded and would be less well-accepted.

While the Web has been an effective tool, more traditional peer reviewed scientific publications are another means successfully used by the HHRA Program in increasing its reputation for innovation as well as communicating with other scientists. Systematic assessment of 420 HHRA/ORD scientific publications found that the HHRA publications were published in high impact journals and compared to other manuscripts in those journals, were more frequently cited. This evaluation is concrete evidence of the success of the communication strategy employed and the high value of the HHRA Program's research and products.

Recommendation:

- PPRTVs far outnumber IRIS assessments and are being developed at four to five times the rate of IRIS assessments. They have been developed specifically to address the site-specific needs of EPA's Superfund Program. While they have not undergone the Agency and interagency review required for toxicity values to be placed in IRIS, they have undergone critical external peer review. They are developed using the same type of data sources, analyzed with the same level of scrutiny, and developed specifically for use at hazardous waste sites. They may be useful to other programs. Currently PPRTVs and their support documentation are only available on a Web site restricted to use by EPA staff or to those who obtain special permission from EPA. The Subcommittee encourages EPA to make the PPRTVs publicly available for use in hazardous waste site risk assessments and encourage their use where appropriate.

IX. OUTCOMES

Question 1: How well defined are the Program's measures of outcomes?

The outcomes measures are extremely well defined for each LTG. LTG 1 has a goal of producing 16 IRIS and 50 PPRTV assessments each year for the foreseeable future. LTG 3 has the clear mandate of continuing updates of criteria air pollutants on a recurring basis. The annual measures are well described.

The Program staff members for LTG 3 have developed a well-defined system for updating the ISAs on an ongoing, recurring basis, allowing time for the multiple peer reviews and iterations that the ISA process requires. The Gantt charts presented to the Subcommittee made clear when each step of ISA updating is to be completed.

The outcome measures for LTG 1 state a target value for the number of assessments to be completed during the year, but do not provide the identity of the chemicals to be assessed for any given year. The Subcommittee agrees that this is the only practical way to set performance measures that cover a 5-year period, as priorities are likely to change over time. The staff members that prepare IRIS and PPRTV values have established procedures for prioritizing chemicals. The IRIS prioritization process seeks input from multiple stakeholders inside and outside EPA. The procedure appears to be well considered and to work well, but how decisions are made is not immediately transparent. The Subcommittee was particularly interested to know whether chemicals that had not reached a high enough priority level to be reviewed in a given year were carried over for consideration in ensuing years, and whether they were accorded a higher priority status by virtue of having been on the list for a period of time.

Recommendation:

- the HHRA Program needs to consider information on the potential public health concern of various chemicals as it prioritizes them for IRIS or PPRTV review. It appears that some of this information is being provided by the program and regional offices, but it would be of value for the Program to make transparent the basis for its prioritization decisions for IRIS and PPRTVs.

It was clear from the Subcommittee's review and the multiple testimonials from program and regional offices that the IRIS staff is viewed by the rest of the agency as the source of risk assessment expertise to contact when crises arise. High profile examples of these crises include the 9/11 terrorist bombings, Hurricane Katrina, and the asbestos contamination in Libby, Montana. These crises are unforeseen and, by their nature, unplanned for, but must be dealt with immediately and require a considerable amount of staff effort. This effort is not accounted for in the APMs but has the highest immediate impact of any of HHRA's work products.

While it is clear that none of these crises can be anticipated, it seems to the Subcommittee that it is possible to anticipate that environmental crises will arise on a regular basis in a country the size of the United States. It also is possible on a retrospective basis to determine the amount of staff time spent annually dealing with crises, and therefore, to estimate the amount of effort that will be needed in the future. This will be an inexact estimate, but NCEA management should consider whether there is a way to create performance measures to account for the effort. Otherwise, the work is an add-on to the staff and may overburden them as they strive to support the regional or program offices while also meeting their obligations to the stated LTG 1 goals.

Recommendation:

- The Subcommittee considers the responsiveness of the staff members to national emergencies and the HHRA Program's contributions to particularly difficult cleanup sites as being of such high value that this should somehow be captured in the APGs.

LTG 2 covers the development of state-of-the-art risk assessment methods, models, and guidance. The outcome measures for the projects in LTG 2 are more varied but are well described. NCEA has made good, logical choices about areas in which to focus their research efforts. The overarching principle in choosing research areas appears to have been to focus on areas that, with additional work, are scientifically mature enough to be incorporated on a regular basis to improve the risk assessments being prepared under the other LTGs. The APGs and APMs are well aligned with the principle of developing methods that will have practical application in NCEA's other efforts.

Question 2: How much are the Program's products being used by environmental decision makers to inform decisions and achieve results?

Environmental decision makers inside and outside of the EPA are highly dependent on the work products of the HHRA Program. The Subcommittee received input during the course of the review from program offices, regional offices, and academia, as well as a survey that included information on state and international regulators, industry, and scientific consults to attorneys. All of these varied groups found the products of the HHRA Program to be of immense value.

IRIS reviews are considered to be in the highest tier of information quality as the source for reference values. The ISAs for criteria air pollutants form the basis for decision-making by the program offices. Both IRIS documents and ISAs withstand and are improved by substantial peer review and have high credibility.

The PPRTVs have more restricted application but are viewed by OSWER, the Office of Superfund Remediation and Technology Innovation (OSRTI), and the regional offices as having great importance as the basis for making decisions regarding remediation of contaminated sites.

The products of LTG 2 are intended to be used to improve the processes that lead to the development of ISAs, IRIS values, and PPRTVs; therefore, they are one step removed from the immediate application of HHRA products to decision-making. It is clear that the products of LTG 2 projects are improving the quality of assessments, including better use of dose-response

information through the use of benchmark dose methodology, better extrapolation through PBPK modeling, and broader use of mode of action information in evaluating the utility of animal data for predicting human risk, and the choice of models to estimate that risk.

Question 3: How might the HHRA Program evaluate and compare public health risks and benefits to become more effective?

The HHRA Program's primary responsibility to the Agency is to provide expertise and high quality reviews on the potential toxicity of chemicals in the environment, and the dose at which those chemicals may pose a risk. This information is used by other programs and offices within the Agency to manage risks. Those programs also have the responsibility of evaluating the effect of their decisions on public health risks and benefits. The HHRA Program does not have immediate access to the information that would be needed to evaluate the impact of its products on public health. That said, the Subcommittee thought that it would be useful for the HHRA Program to obtain synopses of this information from the program and regional offices applying its values.

On a prospective basis, the HHRA Program needs to consider information on the potential public health concern of various chemicals as it prioritizes them for IRIS or PPRTV review. It appears that some of this information is being provided by the program and regional offices, but as noted in the response to question 2 in this section, it would be of value for the Program to make transparent the basis for its prioritization decisions for IRIS and PPRTVs.

X. SUMMARY ASSESSMENTS

LTG 1: Meets Expectations

LTG 1 has highly prescribed APGs requiring the production of a set number of IRIS reviews and PPRTVs each year. The HHRA Program is meeting those goals by producing the mandated number of reviews. These reviews appear to be of a consistent, high quality and are used as the basis for decision-making in EPA program offices, regions, and outside of the Agency by other federal and state agencies as well as by private companies and organizations. The scientists who work in the HHRA Program are nationally and internationally recognized experts in risk assessment. They are well-connected with the rest of the Agency and do a good job of leveraging their limited resources by selectively enlisting experts from other parts of the Agency to assist in specific reviews. The scientists in this group are often leaders in their fields and participate in numerous external activities that advance the science.

In addition to the measured activities under LTG 1, the scientists who carry out the assessments are relied upon to provide their expertise on an emergency basis as crises arise for the various offices and regions. The NCEA staff performed admirably in responding to these crises. Their contributions are important and valuable in supporting rapid decision-making. The NCEA management should consider whether there is a way to account for the time spent in dealing with unforeseen events in its APMs.

The Subcommittee heard clearly from the Agency customers for LTG 1 products that they value the projects, but wish they could have more. The progress under LTG 1 could exceed expectations if the review process was more efficient (while maintaining its high level of quality) or if the HHRA Program had the staff to increase the number of reviews. Regardless, the level of effort will need to increase as newer, but potentially more time-consuming, methods of improving risk assessment become available. Furthermore, IRIS values are to be updated on an ongoing basis, which will put an additional burden on the staff that is not necessarily accounted for in the APMs.

LTG 2: Exceeds Expectations

NCEA is providing under LTG 2 a well-conceived portfolio of methods, models, and guidance to enable the performance of quality hazard and dose-response assessments under LTG 1 and to address critical elements of exposure assessment not addressed by other agency research efforts or the program offices. In establishing the goals for research and Program output, NCEA collaborates with other Agency national laboratories, particularly the NCCT and the National Health and Environmental Effects Research Laboratory (NHEERL), and receives input from the NRC. Over the past 3 years, NCEA has released a number of critical work products for guiding and conducting Agency assessments and is conducting research with the intention of output for

2008-2012 on several key areas of hazard and dose-response assessment. The Program is asking the right questions and engaging in the right activities to answer them.

Over the 2005-2007 timeframe NCEA has:

- Finalized the Guidelines for Carcinogen Risk Assessment and the supplemental guidance to address early life cancer susceptibility in risk assessment
- Released guidance on the conduct of PBPK modeling,
- Developed a framework for assigning childhood exposures,
- Released the Child-Specific Exposure Factors handbook for external review,
- Issued a final report on aging and toxic response,
- Updated the benchmark dose models for analysis of continuous non-quantal endpoints,
- Further developed its categorical regression models for analyzing endpoints across multiple toxicity domains, and
- Resurrected to the point of release for review of its All-Ages Lead Model.

In all of these efforts, the quality of the products has been very good.

Additional near term output is expected in the areas of mode of action data and their use in risk assessment decision-making, further guidance in PBPK modeling, further development of methods for categorical regression and meta-analysis, and the assessment of age susceptibility and less-than-lifetime exposures.

The characterization of risk information and uncertainty will need continuous refinement and careful consideration. The Program is aware of those needs and has initiated efforts to support fuller variability and uncertainty evaluations. A very impressive scientific staff has been assembled at NCEA to pursue those questions. While it is too early to fully judge these efforts, significant progress has been made and the presentations to the Subcommittee members indicated an approach that is sound and likely to have a positive effect on the quantitative analysis of uncertainty in decision-making on environmental health risks.

To further support its efforts and to provide the basis for its scientific judgment, the Program conducts workshops and also sponsors the NRC to conduct workshops, organizes symposia, and commissions white papers by expert scientists and more in-depth work by the NRC.

All of the evidence presented to the Subcommittee suggests that the end results of all of these efforts are exceptional work products of high scientific quality.

The Program's guidance documents, reports, and models are widely used in the conduct of risk assessment, by NCEA staff in the conduct of LTG 1 work, by the program offices and regions, and by other assessors at the local, state, national, and international levels. The work products of NCEA developed under LTG 2 are used extensively in the development of risk assessments used

to inform environmental decision-making. The reliance of clients and users world-wide on the methods, models, and guidance produced by HHRA attests to their appropriateness and quality.

LTG 3: Meets Expectations

The assessment process under LTG 3 is scientifically advanced, both at NCEA and by an experienced peer review process. LTG 3 is designed to meet the mandates of the CCA for conducting air quality health assessments and establishing the NAAQS for six criteria pollutants. The mandate also requires EPA to review the scientific basis for the air quality standards every 5 years to ensure that EPA is using the latest scientific knowledge. The HHRA Program has recently instituted a new process for obtaining input from stakeholders and the public as well as for conducting extensive internal and external peer reviews at numerous steps in the process before a standard is promulgated. The HHRA's product has recently changed from the long-standing comprehensive Air Quality Criteria Documents to the Integrated Science Assessment. As this is a relatively new process where external review drafts have only recently been completed, it is too soon to evaluate its performance other than to note that the HHRA Program has met its expectations as exemplified by completing the assessment for ozone in 2006 and lead in 2007 using the Criteria Document process while developing drafts for external review of the new ISAs for Nitrogen Oxides and Sulfur Dioxides on time in 2007. It is expected that much will be learned as these new ISAs undergo external peer review, which will assist the HHRA Program in adjusting their process where necessary. Nevertheless, the process for summarizing and integrating the health science data on criteria pollutants is generally quite mature.

The Subcommittee is encouraged that this new plan and the established timelines for completing external review drafts of ISAs that are under their direct control has the potential for improving the Agency's overall compliance with CAA mandates for incorporating the latest scientific advances on a 5-year cycle. Of particular note is the assistance from LTG 1 and LTG 2 in developing tools that facilitate the integration of science along the exposure-dose-response continuum and in the development of an annotated "living" database specific to the needs of LTG 3 to keep up with the latest advances in research associated with the six criteria pollutants. These processes are commendable and should lead to improvements in the HHRA Program's ability to maintain state-of-the-art assessments. It is also noteworthy that the HHRA Program has fostered the development of an interdisciplinary team that leverages expertise across the Agency as they work together towards the common LTG.

The Subcommittee saw evidence of exceptional work related to: (1) the appropriateness of the effort to integrate the science transmitted to the Office of Air and Radiation for their review in the rulemaking for criteria pollutants; (2) the proactive approach involving the scientific community and interested stakeholders in the development of scientific questions prior to the initiation of the specific review process; and (3) the comprehensive peer review process at many stages throughout the process. Because the LTG 3 process is so new for the Agency, however, it was not possible to assess the milestones or the quality of the product.

There are three issues that the HHRA Program should consider under the spirit of "continuous improvement." These include the need for: (1) increasing the transparency in the selection of studies utilized in the integrated summary. Specific criteria need to be developed that clearly

articulate a process for inclusion or exclusion of scientific studies; (2) developing a strategy for identifying gaps in the science and then, wherever possible, creating a plan to fill those gaps; and (3) developing a process that deals with the potential challenges associated with creating a single document, the ISA, that both summarizes the recent research and prioritizes the "key" studies across the older and more recent research. For these issues, the Subcommittee recommends that the HHRA Program revisit this issue with CASAC and other potential customers and be prepared to accommodate their communication needs.

XI. APPENDICES

Appendix A: BOSC HHRA Subcommittee Members

George Daston, Ph.D., Chair
Miami Valley Laboratories
The Procter & Gamble Company
11810 E. Miami River Road
Cincinnati, OH 45252
Phone: 513-627-2886
Fax: 513-627-0323
E-mail: daston.gp@pg.com

Bruce Allen
Consultant
101 Corbin Hill Circle
Chapel Hill, NC 27514
Phone: 919-401-5385
Fax: 919-401-5384
E-mail: bruce_allen@verizon.net

Henry Anderson, M.D.
Chief Medical Officer for Occupational and Environmental Health
Wisconsin Division of Public Health, Madison
1 West Wilson Street, Room 150
Madison, Wisconsin 53702
Phone: 608-266-1253
Fax: 608-267-4853
E-mail: anderha@dhfs.state.wi.us

Richard Corley, Ph.D.
Battelle Pacific Northwest National Laboratory
Biological Monitoring & Modeling
902 Battelle Boulevard
Richland, WA 99354
Phone: 509-376-8462
Fax: 509-376-9449
E-mail: rick.corley@pnl.gov

John Stephen Evans, Sc.D.
Harvard School of Public Health
Department of Environmental Health
One Lookout Lane
Portsmouth, NH 03801
Phone: 603-433-3956
Fax: 603-433-4174
E-mail: jevans@hsph.harvard.edu

Mark Utell, M.D.
University of Rochester
School of Medicine
Pulmonary and Critical Care Unit
Rochester, NY 14642-8692
Phone: 585-275-4861
Fax: 585-273-1058
E-mail: mark_utell@urmc.rochester.edu

Lauren Zeise, Ph.D.
Chief, Reproductive and Cancer Hazard Assessment
California Office of Environmental Health Hazard Assessment
515 Clay Street, 16th Floor
Oakland, California 94612
Phone: 510-622-3190
Fax: 510-622-3211
E-mail: lzeise@oehha.ca.gov

Appendix B: BOSC HHRA Subcommittee Draft Charge

Program Review Charge
Human Health Risk Assessment (HHRA) Subcommittee

1.0 Objective

The Board of Scientific Counselors (BOSC) Human Health Risk Assessment (HHRA) Subcommittee will conduct a prospective and retrospective review of the Office of Research and Development's (ORD's) HHRA Program, and evaluate the Program's relevance, quality, performance, and scientific leadership. The BOSC's evaluation and recommendations will provide guidance to ORD to help:

- Plan, implement, and strengthen the HHRA Program;
- Compare the HHRA Program with other programs designed to achieve similar outcomes both in other parts of EPA and in other federal agencies;
- Make ORD investment decisions over the next 5 years;
- Prepare EPA's performance and accountability reports to Congress under the Government Performance and Results Act; and respond to assessments of federal research and development programs such as those conducted by the Office of Management and Budget (OMB highlights the value of recommendations from independent expert panels in guidance to federal agencies[1,2]).

2.0 Background Information

Independent expert review is used extensively in industry, federal agencies, Congressional committees, and academia. The National Academy of Sciences has recommended this approach for evaluating federal research programs.[3]

Because of the nature of research, it is not possible to measure the creation of new knowledge as it develops—or the pace at which research progresses or scientific breakthroughs occur. Demonstrating research contributions to outcomes is very challenging[4] when federal agencies conduct research to support regulatory decisions, and then rely on third parties[5]—such as state environmental agencies—to enforce the regulations and demonstrate environmental improvements. Typically, many years may be required for practical research applications to be developed and decades may be required for some research outcomes to be achieved in a measurable way.

Most of ORD's environmental research programs investigate complex environmental problems and processes—combining use-inspired basic research[6, 7] with applied research, and integrating several scientific disciplines across a conceptual framework[8] that links research to environmental decisions or environmental outcomes. In multidisciplinary research programs such as these, progress toward outcomes cannot be measured by outputs created in a single year. Rather, research progress occurs over several years, as research teams explore hypotheses with individual studies, interpret research findings, and then develop hypotheses for future studies.

In designing and managing its research programs, ORD emphasizes the importance of identifying priority research questions or topics to guide its research. Similarly, ORD recommends that its programs develop a small number of performance goals that serve as indicators of progress to answer the priority questions and to accomplish outcomes. Short-term outcomes are accomplished when research is applied by specific clients (e.g., to strengthen environmental decisions). These decisions and resulting actions (e.g., the reduction of contaminant emissions or restoration of ecosystems) ultimately contribute to improved environmental quality and health.

In a comprehensive evaluation of science and research at EPA, the National Research Council[9] recommended that the Agency substantially increase its efforts to both explain the significance of its research products and to assist clients inside and outside of the Agency in applying them. In response to this recommendation, ORD has engaged science advisors from client organizations to serve as members of its research program teams. These teams help identify research contributions with significant decision-making value and help plan for their transfer and application.

For ORD's environmental research programs, periodic retrospective analysis at intervals of 4 or 5 years is needed to characterize research progress, to assess how clients are applying research to strengthen environmental decisions, and to evaluate client feedback about the research. Conducting program evaluations at this interval enables assessment of: research progress, the scientific quality and decision-making value of the research, and whether research progress has resulted in short-term outcomes for specific clients.

A description of the OSTP/OMB *Research and Development Investment Criteria* is included in Appendix I.

3.0 Background for ORD's HHRA Program and Draft Charge Questions

Background

Human health risk assessment is a process in which information is analyzed to determine if an environmental hazard might cause harm to exposed persons. It is the essential intermediary means by which primary data and published literature are compiled, analyzed, and summarized for application to decision-making in real world situations. Risk assessment is central to the implementation of EPA's statutory responsibilities and mission to protect human health and the environment. The HHRA Program and its Multi-Year Plan (MYP)[10] serve as a primary EPA mechanism to implement this process, linking laboratory and field science with the use of this information by EPA programs, regions, and the broader community.

HHRA is a relatively new ORD program, commencing in fiscal year 2004 through the consolidation and expansion of a number of health assessments and supporting efforts already underway in ORD's National Center for Environmental Assessment (NCEA). This consolidation was undertaken in order to foster a more integrated approach for resource allocation, prioritization, and accountability of risk assessment within ORD. The HHRA Program is distinct from, but linked to, other ORD programs such as the Human Health Research Program and the

Drinking Water Research Program, etc., as its primary objective is the development of peer-reviewed health assessments and supporting methods, models, and guidance rather than conducting research (i.e., collection of laboratory and field data).

The principal clients for HHRA products are the EPA programs and regions that request and receive qualitative and quantitative health assessment information on priority environmental contaminants developed under the HHRA Program's three long-term goals:

Long-Term Goal 1: **Integrated Risk Information System (IRIS) and other priority health hazard assessments:** EPA, state, and other risk assessors use the state-of-the-science health hazard assessment information provided on priority substances in their decisions and actions to protect human health from risks posed by environmental pollutants.

Long-Term Goal 2: **State-of-the-science risk assessment models, methods, and guidance:** IRIS and other EPA programs, states, and other risk assessors use the risk assessment models, methods, and guidance provided to enhance, through the incorporation of contemporary scientific advances, the quality and objectivity of their assessments and decision-making on environmental health risks.

Long-Term Goal 3: **Integrated Science Assessments**: As mandated in the Clean Air Act, the ambient air criteria pollutants are reviewed and Integrated Science Assessments [previously named Air Quality Criteria Documents (AQCDs)] are revised to reflect the best available scientific information on identifiable effects on public health and the environment from exposure to the pollutant, and this information is used by the EPA's Office of Air and Radiation in its review and promulgation of the National Ambient Air Quality Standards (NAAQS).

Further details on components of the HHRA Program are available in the HHRA MYP[10].

Draft Charge

(A) Program Assessment (evaluate entire program): The responses to the program assessment charge questions below should be in a narrative format, and should capture the performance for the entire program and all the activities in support of the program's Long-Term Goals (LTGs).

Program Relevance

1. How consistent are the Long-Term Goals (LTGs) of the Program with achieving the Agency's strategic plan and HHRA's Multi-Year Plan?
2. How responsive is the Program focus to EPA's program office and regional health assessment needs?
3. How responsive is the HHRA Program to recommendations from outside advisory boards and stakeholders?
4. How clearly evident are the public benefits of the HHRA Program?

Factors to consider: the degree assessments are driven by EPA priorities; the degree which this assessment program has had (or is likely to have) an impact on Agency

decision-making; and the extent to which program scientists participate on or contribute to Agency workgroups and transfer products to program and regional customers.

Program Structure

1. How clearly do the LTGs provide a logical framework for organizing and planning the health assessment activities and demonstrating outcomes of the Program?
2. How appropriate are the assessments and science used to achieve each LTG (i.e., is the program asking the right questions, or has it been eclipsed by advancements in the field?)?
3. Does the MYP describe an appropriate flow of work (i.e., the sequencing of related activities) that reasonably reflects the anticipated pace of scientific progress and timing of client needs?
4. Does the HHRA Program use the MYP to help guide and manage its health assessment activities?
5. How logical is the Program design, with clearly identified priorities?

Factors to consider: the linkage of annual products to accomplishing each LTG, the balance between assessments versus development of methods, models, guidance, and technical support to achieve each LTG and meet Agency needs.

Program Performance

1. How much progress is the Program making on each LTG based on clearly stated and appropriate milestones?

Program Quality

1. How good is the scientific quality of the Program's health assessment and methods, models, and guidance products?
2. What means does the Program employ to ensure quality in its health assessment activities and products (including peer review, competitive funding, etc.)?
3. How effective are these processes?

Factors to consider: the impact and use of assessment products by EPA program and regional offices, the processes used to develop and peer review assessments and science products (e.g., Agency, intra-agency, and independent panels reviews), and the extent of the bibliography of peer-reviewed publications.

Scientific Leadership

1. Please comment on the leadership role the HHRA Program and its staff have in contributing to advancing the current state of the risk assessment science and solving important risk assessment problems.
2. How should the HHRA Program implement recruitment methods, incentives, or training procedures to maintain and increase leadership in the field and transition staff to emerging science fields and assessments?

Factors to consider: the degree to which this program is identified as a leader in the field; the degree to which assessments and peer-reviewed publications from this program are

cited in Agency decisions and documents, other peer-reviewed publications; the degree to which HHRA scientists serve/are asked to serve on national/international workgroups, officers in professional societies, publication boards; the degree to which HHRA scientists lead national/international collaborative efforts, organize national/international conferences/symposia, and are awarded for their contributions/leadership.

Coordination and Communication

1. How effectively does the Program engage scientists and managers from within ORD and other relevant program and regional offices in its planning?
2. How effectively does the Program engage outside organizations, both within and outside of the government, to promote collaboration, obtain input on Program goals and assessment priorities, and avoid duplication of effort?
3. How effective are the mechanisms that the Program uses for communicating assessment results both internally and externally?

Factors to consider: the linkages of the HHRA Program to other ORD programs and MYPs for both the transfer of new science and data into assessments and to provide insight and direction for prioritizing research in support of improved risk assessments; the dissemination of assessments and peer-reviewed documents, and risk assessment tools over the Internet, seminars and training for ORD, program and regional offices, at national and international meetings and workshops as well as published literature.

Outcomes

1. How well-defined are the Program's measures of outcomes?
2. How much are the Program's products being used by environmental decision makers to inform decisions and achieve results?
3. How might the HHRA Program evaluate and compare public health risks and benefits to become more effective?

Factors to consider: the influence of HHRA products on key risk management decisions made by the Agency's program and regional offices; the outcome, output and efficiency measures developed by HHRA; alternative approaches for measuring progress in providing timely, high quality criteria especially for Integrated Science Assessments for the Office of Air and Radiation.

(B) Summary Assessment (rate program performance by LTG): A summary assessment and narrative should be provided for each LTG. The assessment should be based on three of the questions included above which are:

1. How appropriate are the assessments and science used to achieve each LTG (i.e., Is the Program asking the right questions and conducting the right assessments to inform client needs?)?
2. How good is the scientific quality of the Program's health assessment and method, model, and guidance products?
3. How much are the Program's results being used by environmental decision makers to inform decisions and achieve results?

Elements to Include for Long-Term Goal 1: IRIS and Other Priority Health Assessments:

The appropriateness, quality, and use of IRIS assessments, PPRTVs and other priority assessments by EPA's program offices and regions and other organizations to inform decisions and actions including: 1) Agency, state, and local risk assessors decisions and setting risk management goals, 2) Superfund Program actions regarding specific sites, 3) Agency needs by incorporation of scientific advancements into health assessments to protect human health from risks posed by environmental pollutants.

Elements to Include for Long-Term Goal 2: State-of-the-Science Risk Assessment Models, Methods, and Guidance:

The appropriateness, quality, and use of HHRA's methods, models, and guidance by IRIS and other EPA programs, states, and other risk assessors to enhance assessments including: 1) the science and objectivity of environmental health assessments, 2) characterization of risk information and uncertainty, and 3) quantitative analysis of uncertainty for decision-making on environmental health risks.

Elements to Include for Long-Term Goal 3: Criteria Air Pollutant Integrated Science Assessments (previously Air Quality Criteria Documents):

The appropriateness, quality, and use of HHRA's Integrated Science Assessments (formerly AQCDs) by EPA's Office of Air and Radiation in its review and development of National Ambient Air Quality Standards (NAAQS) to protect human health and the environment.

The BOSC HHRA Subcommittee will assign a qualitative score that reflects the quality and significance of the Program's health assessment activities*, as well as the extent to which the Program is meeting or making measurable progress toward the goal—relative to the evidence provided to the BOSC. The qualitative evaluation should be in the form of the following adjectives that are defined and intended to promote consistency among BOSC program reviews. The adjectives should be used as part of a narrative summary of the review so that the context of the evaluation and the rationale for selection will be transparent. The adjectives to describe progress are:

- Exceptional: Indicates that the Program is meeting all and exceeding some of its goals, both in the quality of the health assessments* being produced, and the speed at which assessment tools and methods are being produced. An exceptional rating also indicates that the Program is addressing the right questions to achieve programmatic goals. The review should be specific as to which aspects of the Program's performance have been exceptional.
- Exceeds Expectations: Indicates that the Program is meeting all of its goals. It addresses the appropriate scientific questions to meet its goals and the science is competent or better. It exceeds expectations for either the high quality of the assessments or for the speed at which the work products are being produced and milestones met.

- Meets Expectations: Indicates that the Program is meeting most of its goals. Programs meet expectations in terms of addressing the appropriate health assessment questions to meet its goals, and that work products are being produced and milestones are being reached in a timely manner. The quality of the science being done is competent or better.
- Not Satisfactory: Indicates that the Program is failing to meet a substantial fraction of its goals, or if meeting them, that the achievement of milestones is significantly delayed, or that the health assessment questions being addressed are inappropriate or insufficient to meet the intended purpose. Questionable science is also a reason for evaluating a program as unsatisfactory for a particular long-term goal. The review should be specific as to which aspects of a program's performance have been inadequate.

* Revision of wording to more accurately reflect the nature of the products and outcomes developed under this Program versus other ORD laboratories and centers and does not change the definition of the evaluation narrative.

References

[1] Budget Data Request 04-31. Executive Office of the President, Office of Management and Budget. March 22, 2004. "Completing the Program Assessment Rating Tool (PART) for the FY06 Review Process," pages 50-56.

[2] Memorandums for the Heads of Executive Departments and Agencies. Executive Office of the President, Office of Management and Budget. June 5, 2003. "FY 2005 Interagency Research and Development Priorities," pages 5-10.

[3] Evaluating Federal Research under the Government Performance and Results Act (National Research Council, 1999).

[4] The House Science Subcommittee. Letter to Dr. Bruce Alberts, President of the National Academy of Sciences, from F. James Sensenbrenner, Jr. and George E. Brown. October 23, 1997.

[5] The Government Performance and Results Act: 1997 Government wide Implementation Will Be Uneven. U.S. General Accounting Office. (GAO/GGD, 1997).

[6] Building a Foundation for Sound Environmental Decisions. (National Research Council, 1997).

[7] "Renewing the Compact between Science and Government," Stokes, D.E., in 1995 Forum Proceedings, Vannevar Bush IIC Science for the 21st Century. Pages 15-32. Sigma Xi, 1995.

[8] Risk Assessment in the Federal Government: Managing the Process. (National Research Council, 1983).

[9] Strengthening Science at the U.S. Environmental Protection Agency. (National Research Council, 2000, p 141).

[10] Human Health Risk Assessment Multi-Year Plan September 2007

Appendix I
OSTP/OMB Research and Development Investment Criteria

The relevance, quality, and performance criteria apply to all R&D programs. Industry-relevant applied R&D must meet additional criteria. Together, these criteria can be used to assess the need, relevance, appropriateness, quality, and performance of federal R&D programs.

I. Relevance

R&D investments must have clear plans, must be relevant to national priorities, agency missions, relevant fields, and "customer" needs, and must justify their claim on taxpayer resources. Review committees should assess program objectives and goals on their relevance to national needs, "customer" needs, agency missions, and the field(s) of study the program strives to address. For example, the Joint DOE/NSF Nuclear Sciences Advisory Committee's Long Range Plan and the Astronomy Decadal Surveys are the products of good planning processes because they articulate goals and priorities for research opportunities within and across their respective fields. Programs that directly address Presidential priorities may receive special consideration for support, with adequate documentation of their relevance to those priorities.

OMB will work with some programs to identify quantitative metrics to estimate and compare potential benefits across programs with similar goals. Such comparisons may be within an agency or among agencies.

A. Programs must have complete plans, with clear goals and priorities. Programs must provide complete plans, which include explicit statements of: specific issues motivating the program; broad goals and more specific tasks meant to address the issues; priorities among goals and activities within the program; human and capital resources anticipated; and intended program outcomes, against which success may later be assessed.

B. Programs must articulate the potential public benefits of the program. Programs must identify potential benefits, including added benefits beyond those of any similar efforts that have been or are being funded by the government or others. R&D benefits may include technologies and methods that could provide new options in the future, if the landscape of today's needs and capabilities changes dramatically. Some programs and sub-program units may be required to quantitatively estimate expected benefits, which would include metrics to permit meaningful comparisons among programs that promise similar benefits. While all programs should try to articulate potential benefits, OMB and OSTP recognize the difficulty in predicting the outcomes of basic research. Discovery is a legitimate objective of basic research, and some basic research investments may be justified based on external judgments of the opportunity for discovery.

C. Programs must document their relevance to specific Presidential priorities to receive special consideration. Many areas of research warrant some level of federal funding. Nonetheless, the President has identified a few specific areas of research that are particularly important. To the extent that a proposed project can document how it directly addresses one of these areas, it may be given preferential treatment.

D. Program relevance to the needs of the Nation, of fields of science and technology, and of program "customers" must be assessed through prospective external review. Programs must be assessed on their relevance to agency missions, fields of science or technology, or other "customer" needs. A customer may be another program at the same or another agency, an interagency initiative or partnership, or a firm or other organization from another sector or country. As appropriate, programs must define a plan for regular reviews by primary customers of the program's relevance to their needs. These programs must provide a plan for addressing the conclusions of external reviews.

E. Program relevance to the needs of the Nation, of fields of science and technology, and of program "customers" must be assessed periodically through retrospective external review. Programs must periodically assess the need for the program and its relevance to customers against the original justifications. Programs must provide a plan for addressing the conclusions of external reviews.

II. Quality

Programs should maximize the quality of the R&D they fund through the use of a clearly stated, defensible method for awarding a significant majority of their funding. A customary method for promoting R&D quality is the use of a competitive, merit-based process. NSF's process for the peer-reviewed, competitive award of its R&D grants is a good example. Justifications for processes other than competitive merit review may include "outside-the-box" thinking, a need for timeliness (e.g., R&D grants for rapid studies in response to an emergency), unique skills or facilities, or a proven record of outstanding performance (e.g., performance-based renewals). Programs must assess and report on the quality of current and past R&D. For example, NSF's use of Committees of Visitors, which review NSF directorates, is an example of a good quality assessment tool. OMB and OSTP encourage agencies to provide the means by which their programs may be benchmarked internationally or across agencies, which provides one indicator of program quality.

A. Programs allocating funds through means other than a competitive, merit-based process must justify funding methods and document how quality is maintained. Programs must clearly describe how much of the requested funding will be broadly competitive based on merit, providing compelling justifications for R&D funding allocated through other means. (See OMB Circular A-11 for definitions of competitive merit review and other means of allocating federal research funding.) All program funds allocated through means other than unlimited competition must document the processes they will use to distribute funds to each type of R&D performer (e.g., federal laboratories, federally funded R&D centers, universities). Programs are encouraged to use external assessment of the methods they use to allocate R&D and maintain program quality.

B. Program quality must be assessed periodically through retrospective expert review. Programs must institute a plan for regular, external reviews of the quality of the program's research and research performers, including a plan to use the results from these reviews to guide future program decisions. Rolling reviews performed every 3-5 years by advisory committees can satisfy this requirement. Benchmarking of scientific leadership and other factors provides an

effective means of assessing program quality relative to other programs, other agencies, and other countries.

III. Performance

R&D programs should maintain a set of high priority, multi-year R&D objectives with annual performance measures and milestones that show how one or more outcomes will be reached. Metrics should be defined not only to encourage individual program performance but also to promote, as appropriate, broader goals, such as innovation, cooperation, education, and dissemination of knowledge, applications, or tools.

OMB encourages agencies to make the processes they use to satisfy the Government Performance and Results Act (GRPA) consistent with the goals and metrics they use to satisfy these R&D criteria. Satisfying the R&D performance criteria for a given program should serve to set and evaluate R&D performance goals for the purposes of GPRA. OMB expects goals and performance measures that satisfy the R&D criteria to be reflected in agency performance plans.

Programs must demonstrate an ability to manage in a manner that produces identifiable results. At the same time, taking risks and working towards difficult-to-attain goals are important aspects of good research management, especially for basic research. The intent of the investment criteria is not to drive basic research programs to pursue less risky research that has a greater chance of success. Instead, the Administration will focus on improving the management of basic research programs.

OMB will work with some programs to identify quantitative metrics to compare performance across programs with similar goals. Such comparisons may be within an agency or among agencies.

Construction projects and facility operations will require additional performance metrics. Cost and schedule earned-value metrics for the construction of R&D facilities must be tracked and reported. Within DOE, the Office of Science's formalized independent reviews of technical cost, scope, and schedule baselines and project management of construction projects ("Lehman Reviews") are widely recognized as an effective practice for discovering and correcting problems involved with complex, one-of-a-kind construction projects.

A. Programs may be required to track and report relevant program inputs annually. Programs may be expected to report relevant program inputs, which could include statistics on overhead, intramural/extramural spending, infrastructure, and human capital. These inputs should be discussed with OMB.

B. Programs must define appropriate output and outcome measures, schedules, and decision points. Programs must provide single- and multi-year R&D objectives, with annual performance measures, to track how the program will improve scientific understanding and its application. Programs must provide schedules with annual milestones for future competitions, decisions, and termination points, highlighting changes from previous schedules. Program proposals must define what would be a minimally effective program and a successful program.

Agencies should define appropriate output and outcome measures for all R&D programs, but agencies should not expect fundamental basic research to be able to identify outcomes and measure performance in the same way that applied research or development are able to. Highlighting the results of basic research is important, but it should not come at the expense of risk-taking and innovation. For some basic research programs, OMB may accept the use of qualitative outcome measures and quantitative process metrics. Facilities programs must define metrics and methods (e.g., earned-value reporting) to track development costs and to assess the use and needs of operational facilities over time. If leadership in a particular field is a goal for a program or agency, OMB and OSTP encourage the use of benchmarks to assess the processes and outcomes of the program with respect to leadership. OMB encourages agencies to make the processes they use to satisfy GPRA consistent with the goals and metrics they use to satisfy these R&D criteria.

C. Program performance must be retrospectively documented annually. Programs must document performance against previously defined output and outcome metrics, including progress towards objectives, decisions, and termination points or other transitions. Programs with similar goals may be compared on the basis of their performance. OMB will work with agencies to identify such programs and appropriate metrics to enable such comparisons.

IV. Criteria for R&D Programs Developing Technologies That Address Industry Issues

The purpose of some R&D and technology demonstration programs and projects is to introduce some product or concept into the marketplace. However, some of these efforts engage in activities that industry is capable of doing and may discourage or even displace industry investment that would occur otherwise. Programs should avoid duplicating research in areas that are receiving funding from the private sector, especially for evolutionary advances and incremental improvements. For the purposes of assessing federal R&D investments, the following criteria should be used to assess industry-relevant R&D and demonstration projects, including, at OMB discretion, associated construction activities.

OMB will work with programs to identify appropriate measures to compare potential benefits and performance across programs with similar goals, as well as ways to assess market relevance.

A. Programs and projects must articulate the public benefits of the program using uniform benefit indicators across programs and projects with similar goals. In addition to the public benefits required in the general criteria, all industry-relevant programs and projects must identify and use uniform benefit indicators (including benefit-cost ratios) to enable comparisons of expected benefits across programs and projects. OMB will work with agencies to identify these indicators.

B. Programs and projects must justify the appropriateness of federal investment. Programs and projects must demonstrate that industry investment is sub-optimal to develop a technology or system and explain why the development or acceleration of that technology or system is necessary to meet a federal mission or goals.

C. Programs and projects must demonstrate that investment in R&D and demonstration activities is a more effective way to support the federal goals than other policy alternatives. When the federal government chooses to intervene to address market failures, there may be many policy alternatives to address those failures. Among other tools available to the government are legislation, tax policy, regulatory and enforcement efforts, and an integrated combination of these approaches. Agencies should consider that the legislation, tax policy or regulatory or enforcement mechanisms may already be in place to achieve a reasonable expectation of advancing the desired end.

D. Programs and projects must document industry or market relevance, including readiness of the market to adopt technologies or other outputs. Programs must assess the likelihood that the target industry will be able to adopt the technology or other program outputs. The level of industry cost sharing or enforceable recoupment commitments in contracts are indicators of industry relevance. Agencies must be able to justify any demonstration activities with an economic analysis of the public and private returns on the public investment.

E. Program performance plans and reports must include "off ramps" and transition points. In addition to the schedules and decision points defined in the general criteria, program plans should also identify whether, when, and how aspects of the program may be shifted to the private sector.

Appendix C: List of Acronyms

APG	Annual Performance Goal
APM	Annual Performance Measure
ATSDR	Agency for Toxic Substances and Disease Registry
BOSC	Board of Scientific Counselors
CAA	Clean Air Act
CASAC	Clean Air Scientific Advisory Committee
EPA	Environmental Protection Agency
ESI	Essential Science Indicators
HHRA	Human Health Risk Assessment
IARC	International Agency for Research on Cancer
IRIS	Integrated Risk Information System
ISA	Integrated Science Assessment
JCR	Journal Citation Report
LTG	Long-Term Goal
MOA	Mode of Action
MOU	Memorandum of Understanding
MRL	Minimal Risk Levels
MYP	Multi-Year Plan
NAAQS	National Ambient Air Quality Standards
NCCT	National Center for Computational Toxicology
NCEA	National Center for Environmental Assessment
NHEERL	National Health and Environmental Effects Research Laboratory
NPD	National Program Director
NRC	National Research Council
NOAEL	No Observed Adverse Effect Level
LOAEL	Lowest Observed Adverse Effect Level
OAR	Office of Air and Radiation
OMB	Office of Management and Budget
OPPTS	Office of Prevention, Pesticides, and Toxic Substances
ORD	Office of Research and Development
OSWER	Office of Solid Waste and Emergency Response
OSRTI	Office of Superfund Remediation and Technology Innovation
OW	Office of Water
PBPK	Physiologically-Based Pharmacokinetic
PPRTV	Provisional Peer-Reviewed Toxicity Value
SAB	Science Advisory Board
SOT	Society of Toxicology
WHO/IPCS	World Health Organization/International Program on Chemical Safety
UNEP	United Nations Environment Program
USDA	United States Department of Agriculture