B·O·S·C
BOARD OF SCIENTIFIC COUNSELORS

HUMAN HEALTH SUBCOMMITTEE

Conference Call Summary
Tuesday, April 21, 2009
1:00 – 2:00 p.m. Eastern Time

Welcome

Dr. James Klaunig, Indiana University School of Medicine, Subcommittee Chair

Dr. James Klaunig, Chair of the Board of Scientific Counselors (BOSC) Human Health Subcommittee, welcomed the Subcommittee members to the teleconference and provided an overview of the agenda, a copy of which is attached to this summary.

Subcommittee Designated Federal Officer (DFO) Remarks

Ms. Virginia Houk, U.S Environmental Protection Agency (EPA)/Office of Research and Development (ORD), Subcommittee DFO

Ms. Virginia Houk, Subcommittee DFO, thanked the Subcommittee members for their attendance, took roll—noting that Drs. Paul Blanc and Christopher Portier had prior commitments and could not attend—and reviewed the Federal Advisory Committee Act (FACA) procedures that are required for all BOSC Subcommittee meetings. Ms. Houk explained that as the DFO, she ensures that all FACA requirements are met and that records of board deliberations are made public. All meetings and teleconferences involving substantive issues, whether in person, by phone, or by e-mail, that include one-half or more of the Subcommittee members must be open to the public. All public meetings and teleconferences of the Subcommittee must be announced in the *Federal Register* at least 15 days prior to the meeting or teleconference; the notice for this conference call was published on March 25, 2009, and an electronic docket was established. The docket is available at http://www.regulations.gov; the docket number is EPA-HQ-ORD-2008-0649. The minutes of the meeting will be prepared by Kristen LeBaron of The Scientific Consulting Group, Inc. and, following review by the Subcommittee members and certification by the Chair, will be available on the BOSC Web Site.

This conference call was convened to finalize the draft report in preparation for submitting it to the BOSC Executive Committee. Ms. Houk worked with EPA officials to ensure that appropriate ethics regulations were satisfied. Subcommittee members must inform the DFO if they discover a potential conflict of interest with regard to any of the topics under discussion during this call. Although there were no advance requests for comment from the public, time for public comment is scheduled for 1:55 p.m. She asked that comments be limited to 3 minutes each.

Executive Summary Discussion

Dr. James Klaunig, Subcommittee Chair

Dr. Klaunig explained that the executive summary was organized by the topics that the Subcommittee was charged with evaluating (e.g., program relevance, structure, and quality), and he is amenable to any changes the Subcommittee members suggest. The salient comments and strengths and weaknesses from the Long-Term Goal (LTG) sections were included in the executive summary.

Dr. George Daston thought that the recommendations should be included in the executive summary. The advantage to this approach is that all of the recommendations would be up front and in the section of the report that most people read; the disadvantage is the resulting redundancy. He asked for clarification of the term "ecological health studies" on line 37 of page 3. Dr. Joel Schwartz explained that this concern originated from the LTG 4 discussions and the concern that the Office of Management and Budget would require proof of public health impacts following EPA action that depend on a multitude of additional variables that fluctuate over time. This was mentioned to provide perspective and ensure that this fact is not forgotten. Ecological health studies examine the health of an entire population (rather than an individual) and correlate this health to another factor.

Dr. Henry Falk commented that the executive summary reads well and he thought that adding the recommendations would increase its specificity and be helpful to ORD. Dr. Klaunig agreed that it is important to include the recommendations in the executive summary because many people will read only that section.

Dr. Schwartz noted that, in terms of LTG 1's summary assessment, additional details should be added to explain the issue of integrating mode of action with risk assessments generated from epidemiology studies. The specific concern was that individuals who understand epidemiology-based quantitative risk assessments are not on staff, and this absence impedes integration. The Program does not need to conduct research in this area, but this expertise is needed to ensure integration.

The Subcommittee discussed where the recommendations should be placed within the executive summary (i.e., within each summary assessment or grouped together in their own section). The Subcommittee decided to place the recommendations in their own section following the summary assessments; this is necessary to ensure that the recommendations are not taken out of context following this move.
In summary, Dr. Klaunig stated that the executive summary will contain the following components in the following order: (1) introductory paragraph, (2) sections on charge topics, (3) summary assessments of each LTG, and (4) final section with all recommendations.

LTG 1 Discussion
Dr. David Hoel, Medical University of South Carolina, Workgroup Lead

Dr. David Hoel noted that many of the comments regarding the section on LTG 1 indicated that the section was too brief. He noted that more specific examples could be added regarding better interaction between groups within ORD. Dr. Edo Pellizzari said that examples of how these groups could benefit from better interaction could be included. Dr. Schwartz stated that specific cases were discussed at the face-to-face meeting, including a discussion regarding the sophisticated modeling being performed within LTG 1 that was not being shared with researchers within other LTGs, who were subsequently trying to "reinvent the wheel." Dr. Klaunig added that there were specific comments captured in the face-to-face meeting summary that indicate that the Subcommittee members thought that certain LTGs would be more successful if the researchers from each LTG worked together. Dr. Pellizzari noted that the draft report identifies potential beneficial interactions within other LTGs; these identified interactions also could apply to this LTG. Perhaps under LTG 1, the following should be stated: "Better interactions among the groups within ORD would be beneficial as indicated by the discussion found below under other LTGs." Referring to the other sections will decrease redundancy.

In terms of the section on coordination and communication, Dr. Schwartz noted that ORD researchers are communicating externally, but the level of effort was unclear. Dr. Klaunig commented that strengths and weaknesses of outside collaborations could be included. Dr. Pellizzari mentioned one specific communication example that had an appropriate level of effort, but the researchers clearly did not have the proper knowledge or guidance to successfully carry it out. Dr. Schwartz added that it was unclear how much real interaction was occurring with the Children's Environmental Health Research Centers and other partners. Dr. Hoel noted that it was unclear what type of collaboration mechanisms are in place

(e.g., cooperative agreements), and there was no information regarding the level of effort for extramural work. The Subcommittee discussed a general recommendation regarding information that would be useful for future reviews.

Dr. Schwartz agreed with the additions that had been made to the summary assessment and suggested encouraging people to start using basic mechanistic studies in risk assessments. Dr. Hoel noted that these are generated by the Integrated Risk Information System (IRIS), and it is unclear whether the components work together to have the modes of action modify the quantitative risk estimates; this sort of example would be beneficial.

Dr. Klaunig stated that recommendations are needed for LTG 1. Dr. Schwartz agreed to write the recommendation regarding mode of action/epidemiology, and Dr. Hoel will write the mode of action/IRIS recommendation. Dr. Klaunig stated that he would place the recommendation regarding which materials should be provided to reviewers by the Program in the overall recommendations.

LTG 2 Discussion

Dr. Edo Pellizzari, RTI International, Workgroup Lead

Dr. Pellizzari explained that in response to the discussion from the previous conference call, this section was pared down, and the specific recommendations were brought out from the text. There were no subsequent comments to address. Dr. Klaunig thought that the recommendations should remain at the end of the section, but the recommendations listed within the text should be removed to remain consistent with the other sections. Dr. Pellizzari agreed to make that change.

Dr. Schwartz commented that, in terms of the uncertainty analyses for the source-to-dose model, the desirable outcome is source-dose-health; therefore, the approach is reasonable but must be extended to include the health component. The epidemiological studies include different exposure characterizations and, therefore, different uncertainties. More work needs to be performed to characterize the uncertainties. Dr. Pellizzari noted that some of the emphasis was lost in the general issue of describing the uncertainty in each stage of the source-dose-health continuum; the example may be more specific than was originally intended. Dr. Schwartz noted that the discussion must not convey the impression that the Subcommittee is limiting the directions in which the work can progress (i.e., to the health portion of the continuum or other directions).

LTG 3 Discussion

Dr. George Daston, Procter and Gamble, Workgroup Secondary

Dr. Daston noted that there were few changes from the previous draft; citations had been added to broaden the horizon of different age groups and diseases in terms of the susceptibility and vulnerability. The most significant change was to group all of the recommendations at the end of the section. It was determined that no further changes are necessary.

LTG 4 Discussion

Dr. Joel Schwartz, National Institutes of Health, Workgroup Secondary

Dr. Schwartz noted that LTG 4 exceeded expectations and summarized some of the recommendations, including moving responsibility for the *Report on the Environment* to a better location within the Agency where it would receive more emphasis. Dr. Falk thought that this recommendation needed to be expanded because the justification for moving the report is not present in the current text. The reasoning for why it would be better placed elsewhere in the Agency must be explained. The recommendation is valid, but more justification is needed. Dr. Schwartz explained that the *Report on the Environment* is globally relevant to EPA, as it describes the state of the environment, including various details and analyses. The report contains key information that should be included in the Agency's central decision-

making process. If the report is obscured too deep within ORD, it is more difficult to get the information to the rest of the Agency. Dr. Falk agreed that this was a good explanation, but he was unclear whether the report itself was being moved or whether the personnel responsible for the drafting the report would be moved as well. These details need to be included in the recommendation. Dr. Schwartz agreed to expand and clarify the recommendation.

Dr. Falk noted that it would be helpful to add details within the summary assessment to explain why the Subcommittee thought that this LTG exceeded expectations; it is not obvious within the current text.

Final Draft Preparations
Dr. James Klaunig, Subcommittee Chair

Dr. Klaunig instructed the Subcommittee members to send their editorial comments to Ms. Houk via e-mail. Ms. Houk explained that the goal was to have the report ready for the next BOSC Executive Committee face-to-face meeting, which will be held June 4–5, 2009. If Subcommittee members can provide her with their comments by May 5, 2009, she will send the revised draft to the members by May 6, 2009. If final approval is received by May 20, 2009, then the report should be ready for the BOSC Executive Committee meeting. The Subcommittee members agreed to this timeline. Ms. Houk noted that Subcommittee members could copy other members of their workgroups on the e-mails if they so desired; she will e-mail the schedule for finalizing the report to the Subcommittee members. Dr. Falk noted that the current draft was a significant improvement from the previous draft, and the Subcommittee's impression of the Program is clear.

Public Comment
Ms. Virginia Houk, Subcommittee DFO

Ms. Houk called for public comments at 1:49 p.m. No comments were offered.

Dr. Klaunig thanked the Subcommittee members for their efforts and participation throughout the review process and adjourned the teleconference at 1:51 p.m.

Action Items

- ✧ Dr. Klaunig will add the recommendations as the last section of the executive summary.
- ✧ Dr. Schwartz will write the LTG 1 mode of action/epidemiology recommendation.
- ✧ Dr. Hoel will write the LTG 1 mode of action/IRIS recommendation.
- ✧ Dr. Klaunig will include the recommendation regarding which information should be provided by the Program in future reviews in the overall recommendations.
- ✧ Dr. Pellizzari will remove the recommendations within the text of the LTG 2 section to be more consistent with the other LTG sections.
- ✧ Dr. Pellizzari will ensure that the specific example regarding the source-dose-health outcome conveys the correct impression.
- ✧ Dr. Schwartz will expand the LTG 3 recommendation regarding the *Report on the Environment*.
- ✧ Ms. Houk will e-mail the schedule for finalizing the report to the Subcommittee members.

- ✧ The Subcommittee members will send any editorial comments to Ms. Houk via e-mail by May 5, 2009.

- ✧ Ms. Houk will make the Subcommittee members' revisions and send the revised draft to the members by May 6, 2009.

PARTICIPANTS LIST

Subcommittee Members

James E. Klaunig, Ph.D., Chair
Robert B. Forney Professor
Department of Toxicology
School of Medicine
Indiana University

Henry Falk, M.D., Vice Chair
Director
Coordinating Center for Environmental Health and Injury Prevention
Centers for Disease Control and Prevention

George P. Daston, Ph.D.
Research Fellow
The Proctor & Gamble Company

David G. Hoel, Ph.D.
Distinguished University Professor
Department of Biostatistics, Bioinformatics, and Epidemiology
Medical University of South Carolina

Edo Pellizzari, Ph.D.
Senior Fellow
RTI International

Joel Schwartz, Ph.D.
Professor
Department of Environmental Health
Harvard University School of Public Health

Designated Federal Officer

Virginia Houk
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory (B305-02)

Other Participants

Steve Via
Regulatory Affairs Manager
American Water Works Association

Contractor Support

Kristen LeBaron, M.S.
The Scientific Consulting Group, Inc.

B·O·S·C
BOARD OF SCIENTIFIC COUNSELORS

HUMAN HEALTH SUBCOMMITTEE

AGENDA
April 21, 2009
1:00 – 2:00 p.m. Eastern Time

CONFERENCE CALL
Participation by Teleconference Only
866-299-3188
code: 919-541-7698#

1:00–1:05 p.m.	Welcome - Overview of Agenda	Dr. James Klaunig Subcommittee Chair
	BOSC DFO Remarks	Ms. Virginia Houk, Office of Research and Development
1:05–1:10 p.m.	Executive Summary Discussion	Dr. James Klaunig Subcommittee Chair
1:10–1:20 p.m.	LTG 1 Discussion	Dr. David Hoel LTG1 Workgroup Lead HH Subcommittee
1:20–1:30 p.m.	LTG 2 Discussion	Dr. Edo Pellizzari LTG2 Workgroup Lead HH Subcommittee
1:30–1:40 p.m.	LTG 3 Discussion	Dr. George Daston LTG3 Workgroup Secondary HH Subcommittee
1:40–1:50 p.m.	LTG 4 Discussion	Dr. Joel Schwartz LTG4 Workgroup Secondary HH Subcommittee
1:50–1:55 p.m.	Final Draft Preparations	Dr. James Klaunig Subcommittee Chair
1:55–2:00 p.m.	Public Comment	
2:00 p.m.	Adjournment	

B·O·S·C

BOARD OF SCIENTIFIC COUNSELORS

Human Health Subcommittee

Draft Report

A Federal Advisory Committee for the U.S. Environmental Protection Agency's Office of Research and Development

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

DRAFT VERSION
APRIL 20, 2009

BOSC SUBCOMMITTEE ON HUMAN HEALTH RESEARCH

James E. Klaunig (Chair) – Indiana University School of Medicine
Henry Falk (Vice-Chair) – Centers for Disease Control and Prevention
Paul D. Blanc – University of California San Francisco
George P. Daston – The Procter & Gamble Company
David G. Hoel – Medical University of South Carolina
Donald Mattison – National Institutes of Health, NICHD
Edo Pellizzari – RTI International
Christopher J. Portier – National Institute of Environmental Health Sciences
Joel Schwartz – Harvard University School of Public Health

EPA CONTACT

Virginia Houk, Designated Federal Officer

TABLE OF CONTENTS

I. EXECUTIVE SUMMARY

Independent expert reviews are used extensively by industry, federal agencies, Congressional committees, and academia, and have been recommended by the National Academy of Sciences as an approach for evaluating federal research programs. Accordingly, the Executive Committee of the Board of Scientific Counselors (BOSC) of the Office of Research and Development (ORD) within the U.S. Environmental Protection Agency (EPA) has agreed to undertake a series of reviews of major EPA research programs. It accomplishes this by forming subcommittees having appropriate expertise for the specific program. This report is a BOSC review of ORD's Human Health Research Program (HHRP). The members of the Human Health Subcommittee are **James E. Klaunig, Ph.D** (Chair), Henry Falk, MD, MPH (Vice Chair), Paul D. Blanc, MD, MSPH , George P. Daston, Ph.D. , David G. Hoel, Ph.D., Donald Mattison, M.D., Edo Pellizzari, Ph.D., Christopher J. Portier, Ph.D., and Joel Schwartz, Ph.D.. The affiliations of the members are listed in Appendix A.

The program review was structured to address charge questions provided by the BOSC (Appendix B) that relate to program relevance, structure, performance, quality, scientific leadership, coordination/ communication, and outcomes. To facilitate this review, the Subcommittee was provided with written materials and heard presentations on the goals, management, and research of the program. The Subcommittee members also reviewed posters and reports prepared and assembled by program staff related to research activities, accomplishments, and user applications. Presentations (testimonials) also were provided by major clients of the program. The Subcommittee has chosen to organize its response according to the Long Term Goals (LTGs) outlined in the *Human Health Multi-Year Plan*. Charge questions were addressed in the context of each LTG, and salient points within each LTG have been captured and aggregated across the Human Health Program as a whole in preparing this Executive Summary.

Executive Summary

Overall there was a consensus view of the subcommittee that there has been a maturing of the program (based on previous reviews). The program is much more integrated, and the level and quality of science has improved. There is considerably more emphasis on human health and human health-related issues, and there is an overall movement to more of a public health-themed program. The program, as a whole, appears to be robust and responsive to emerging issues. The scientific content is excellent and, compared to previous reviews, is more integrated within each Long Term Goal and between the Long Term Goals as well. The presentations of the poster session overviews by the Long Term Goal leaders and the posters session presenters were outstanding and were well received by the subcommittee, as were the presentations by senior EPA leadership during the conference calls. We also appreciated the general EPA staff attendance at the poster sessions. There was notable enthusiasm in the presentation and answering of questions. The subcommittee recognizes and appreciates the extensive work that was required to organize and present this review, including the poster sessions, and applauds these efforts by all those involved. There appears to be good evidence for strong scientific productivity and a formidable impact of the work produced by the program overall. In general, we felt the leadership is excellent to outstanding from the senior level to the laboratory.

<u>Program Relevance</u>

The subcommittee felt that the current HHRP objectives for achieving the USEPA's strategic plan were appropriate and that each of the LTGs used suitable science to address the objectives of the HHRP and of it stakeholders. There was appropriate responsiveness of the research plan of the LTG to outside advisory groups and stakeholders. The program scientists were participating in and contributing to Agency workgroups engaged in identifying and addressing research needs. The subcommittee concluded that the current HHRP objectives are appropriate in achieving the Agency's strategic plan and providing a clear public benefit. These objectives are essential for the Agency to improve its current risk assessment methodologies and to incorporate chemical mixtures and other exposures that influence risk. The goal to reduce the uncertainty in exposure assessments used for assessing risk is an appropriate one. A challenge for the program is to successfully integrate mode of action (MOA) results into the quantitative risk assessments. In future reviews, the subcommittee recommends that more evidence be provided on the use of completed research products in cumulative risk assessments.

In general, the science selected by the program was deemed appropriate. The computational toxicology and reproductive toxicology efforts were particularly well received by the subcommittee. The studies of MOA and biologically based dose response (BBDR), while needing more integration to provide biological plausibility to the epidemiology studies, are nonetheless important and timely. The subcommittee noted that the HHRP was fully aware of the changing nature of risk assessment and the need to incorporate new tools to evaluate chemical mixtures and complex exposure patterns. There is clear evidence that the HHRP recognizes the need to include susceptibility and the role of other types of stressors (e.g., social class, economic factors, age, sex, disease presence in the individual) into its research

base. As such, the subcommittee suggests the further incorporation of epigenetics and genetic polymorphisms in human susceptibility into the program. The program in general appears very responsive to the issues raised by the stakeholders. The Subcommittee notes that with increasing use of dose-response estimates in epidemiology, there comes a greater need for interface with MOA. It was not clear, however, to the Subcommittee that the program has shown to its stakeholders specific examples of how the incorporation of MOA considerations can benefit quantitative risk estimates. The incorporation and use of the pharmacological literature of human data should also be explored to support human risk evaluations. Finally, suggestions that the program needs to establish the validation of its models through the use of human data (e.g., NHANES and pharmacological data) was made.

Program Structure

In general, the structure of the program is well organized and clearly defines the priorities and outcomes of the overall program. The multi-year plans showed appropriate work flow and as outlined reflected a reasonable pace of scientific progress. The program overall showed the potential to respond to changing priorities and areas of need in order to help fulfill the client's needs. Overall, the leadership of the HHRP has done an excellent job in bringing the programs together. Research planning has been guided by the MYP and thus has tended to be very vertical-centric.

The framework of the proposed research and the planning of the specific research activities from basic research to modeling to human health effects to risk assessment appear to be properly structured. The planning and organization of the overall program is logical and linked to the MYP. However, the subcommittee felt that the program could improve on it comprehensiveness. There is evidence of cross-LTG planning; however, this is an area in need of further improvement. Although many successes were noted, the further interaction and linkage with other federal agencies could be enhanced. It was also a concern that a small number of investigators with experience in epidemiology and biostatistics reside within the program. These skills are important for integration and some modeling issues, and the subcommittee suggests this is an area of recruitment need.

Program Quality

The overall quality of the programs within the HHRP is excellent to outstanding. The subcommittee, however, did recognize a need for improvement on some of the projects within the program. A concern was raised that in several of the projects, outdated statistical methods were being used. Similarly, modeling and statistical tools were being created that had already been developed. There is also extensive expertise on the uses and limitations of ecologic health studies that does not appear to have been fully tapped. The subcommittee noted some excellent success stories within the program. For examples, the cumulative risk assessments that have been completed for OPP (with the OPs, carbamates and pyrethrin insecticides) appear to be well-conducted and at an appropriate level of complexity. The use of the SHEDS model to drive the

exposure scenarios is a good example of cooperation across multiple labs and of using regulatory-validated models as part of the assessment. The ROE was also an excellent effort to track trends in exposure and health, and to begin to provide a framework for closing the loop.

The subcommittee suggests that the program consider incorporating additional sets of data sources for looking at health trends (i.e., Centers for Medicare and Medicaid Services (CMS) data on MEDICARE and MEDICAID, Homeland Security monitoring networks for ER visits, state databases on all hospital admissions, NIH databases, etc.). The subcommittee was also impressed in the early results of community-based studies. These studies combined science with appropriate interactions with community leaders, resulting in policy-relevant information being transferred. One concern that was noted was the limitation of susceptibility/vulnerability studies in humans to the childhood life-stage. Further consideration of additional life stages should be made.

Coordination and Communication

Overall, the research within the long term goals (intra-LTG) appears to be very well coordinated. The scientific leadership is to be commended for their attempts to enhance the coordination and communication efforts with program offices and through interagency collaborations. There is good evidence of interactions between ORD and program staff. The subcommittee had concerns on the extent of the interactions across long term goals, specifically with regard to LTG2. Similarly, the subcommittee noted that the tools being developed that allow for the evaluation were being developed in a way that will allow them to be shared with other groups who would want to use them to expand these activities. The Subcommittee believes that better translation of the potential impact of MOA on quantitative risk estimation and management be made to the program offices. In the case of LTG3, the utilization of the combined strengths on both the intramural and extramural fronts was noted. This coordination serves as a model within the program for intramural – extramural coordination and interaction. It is evident that program scientists are very much engaged in communicating knowledge developed from their research endeavors to the national and international scientific community. This occurs through publications, presentations at national and international conferences, briefings and seminars, and preparation of reference compendiums, e.g., pesticide exposure factors for children.

Program Performance

Overall, there was evidence of significant progress by each of the LTGs in addressing their milestones. The programs outcome measurement is, for the most part, well defined and appropriate. Excellent progress has been made to demonstrate the performance of the formulated concepts and approaches. The program has efficiently managed resources for its long term goals.
The subcommittee strongly encourages the HHRP to continue thinking outside of the box regarding how to evaluate the impact of policy and regulations on human health and thus bring

accountability to the decisions made about environmental health.

The indicators selected in the LTGs for monitoring exposures or adverse effects have been selected on solid scientific principles. As an example, the subcommittee noted the success of using observing decreasing levels of cotinine, a nicotine metabolite, in urine as a monitor of the impact of public education on smoking and environmental tobacco smoke (ETS) exposure. The subcommittee found it difficult to fully evaluate the productivity of the overall program based on the bibliography provided. While there appears to be good evidence for strong scientific productivity and a formidable impact of the work produced by the program overall, there were, however, elements of the bibliographic analysis presented in the review material that were difficult to interpret and understand. Moreover, the co-mingling of intramural and extramural publications made it difficult to evaluate the overall contribution of the Long Term Goals to the scientific program, and the relative contributions of intramural and extramural research to each of the goals.

Scientific Leadership

In general, the subcommittee felt that the program leadership is excellent to outstanding from the senior level to the laboratory. Substantial evidence exists that researchers within the HHRP are providing thought leadership by participating in a variety of boards, panels, workshops, and in presentations at conferences and through publications. The HHRP is providing scientific leadership to the entire field of toxicology in areas such as reproductive and developmental toxicology, computational toxicology, and respiratory health effects. The subcommittee also noted the important role of the HHRP specifically and USEPA in general as a scientific leader in regard to the National Children's Study. As is evident within the HHRP, the coordination of extramural and intramural efforts has produced significant results in childhood vulnerability to environmental stressors, in particular in terms of childhood asthma. In addition, the community-based tools and websites being developed within the program are an example of leadership providing tools to local communities that can be used to make local decisions on environmental health. Similarly, the role of HHRP in epidemiological research represents a substantive maturation of the program. There is always a need to cultivate new leaders within the program, and this is acknowledged by the senior leaders of the program. Continued efforts should be made to ensure that new leaders are developed or recruited to the program and that the institutional memory of the retiring leaders is captured. The subcommittee also strongly recommends that added resources into developing the science in cumulative risk assessments be made.

Summary Assessments of Long Term Goals

Summary Assessment of LTG 1: (Meets Expectations)

The scientific quality of the program and its outstanding leadership makes it an essential component of the Agency's human health research program. The program is at the forefront in

computational biology as well as the traditional areas of developmental and inhalation toxicology. What is needed is better integration of MOA with the quantitative risk assessment generated by the epidemiology studies. In particular it is important to demonstrate the value and impact that the basic mechanistic studies of MOA have on the Agency's quantitative risk assessments

Summary Assessment of LTG 2: (Meets Expectations)

The leadership and scientists of this Long Term Goal are commended for their accomplishments. They recognized the need and demonstrated the ability to move from single chemical with multiple routes of exposures to multiple chemicals with similar mode/mechanisms of action. They have successfully incorporated sophisticated modeling concepts into N-methyl carbamate risk assessments and have done so in partnership with the Program Offices. The effort in this LTG has remained true to the two major research goal on cumulative risk and susceptible populations as described in the MYP. However, the Subcommittee believes that even though the planning and organization has been logical, this LTG could achieve greater benefits from more cross-LTG planning. The coordination and communication effort with program offices is laudable; however, the Subcommittee believes that more attention should be given to the needs of Regional Offices. Overall there is substantial evidence that LTG2 scientists are providing thought leadership through participation in a variety of boards, panels, workshops, and in presentations at conferences. The bibliometric data indicates that they are creating new knowledge, transferring this knowledge to the public domain and adroitly applying it to environmental health issues.

Summary Assessment of LTG 3: (Meets Expectations)

LTG3 was assessed as meeting program expectations based on the inarguable population health and public policy relevance of this area of research, the LTG's commendable coordination and communication efforts with program offices and the scientific leadership role manifest, the excellent to outstanding scientific quality of the specific endeavors, the high level of productivity within the areas in which it has LTG has focused. The programmatic structure was assessed as over-weighting childhood health within its life-stage construct of vulnerability additionally treating asthma, one if its major foci, as little more than a surrogate of childhood risk. Absent this serious limitation this LTG would have been assessed as "exceeding expectations."

Summary Assessment of LTG 4: (Exceeds Expectations)

LTG4 was assessed as being an integral part of closing the loop created when we identify a hazard and develop/implement decisions related to that hazard by working to develop the tools necessary to determine if the management decisions were warranted, effective and should be continued. Many times this critical aspect of environmental health decision making is overlooked and programs are put into place that are unnecessary or no longer effective. Having the tools to evaluate risk management decisions must be a priority and the Subcommittee is

pleased this is being undertaken with regard to the long-terms impacts on human health. Even though the program is rather new, we see enthusiasm in the staff involved, early successes in the approaches chosen and the beginnings of a very successful activity for the Agency.

II. INTRODUCTION

The overall goal of the HHRP, as defined in the current Multi-Year Plan (June 2006), is to characterize and ultimately reduce uncertainties in extrapolations inherent in the risk assessment process by providing a greater understanding of the fundamental determinants of exposure and dose and the basic biological changes that result from exposures to environmental toxicants. An overarching theme is to improve understanding of the linkages in the exposure-to-dose-to-effect continuum. It is, of necessity, an inter-disciplinary research program that develops the methods, models and data needed to characterize uncertainties in each of these linkages and apply the information to the real world to elucidate exposures and risks in communities. Research projects are integrated across the intramural and extramural grants programs and are currently organized around four LTGs. The four goals are inter-related by design.

Long Term Goal 1 (LTG1): Risk assessors and risk managers use ORD's methods, models or data to reduce uncertainty in risk assessment using mechanistic (or mode of action) information. *Fundamental research in this goal elucidates mechanisms of action of priority environmental contaminants and related families of contaminants, explores toxicity pathways that are perturbed by these contaminants, and uses this information to develop and link pharmacokinetic and pharmacodynamic models for use in risk assessment. These models are applied to reducing uncertainties associated with extrapolating from high to low dose, from test species to humans, from in vitro data to in vivo exposures, and between cancer and non-cancer effects. Progress is measured by the extent to which this information is being used in Agency risk assessments and rulings. A new direction in this goal is to develop a systems biology approach and apply novel models such as a virtual liver to predict toxicity and estimate risk.*

Long Term Goal 2 (LTG2): Risk assessors and risk managers use ORD's methods, models, and data to characterize aggregate exposure and cumulative risk in order to inform risk management for humans exposed to multiple environmental stressors. *Research in this goal develops and applies biomarkers to assess cumulative exposure and risk; develops and applies source-to-dose models for cumulative risk assessment and dose reconstruction; and creates tools for community-based exposure and risk assessments of complex mixtures. The long term objective is to produce a research framework outlining tools and approaches to characterize and assess aggregate exposures and cumulative risks, especially for vulnerable populations, based on a full range of both chemical and non-chemical stressors.*

Long Term Goal 3 (LTG3): Risk assessors and risk managers will use ORD's methods, models and data to characterize and provide adequate protection for susceptible populations. *This goal focuses on susceptibility as a function of life stage with a strong*

emphasis on children and older Americans as potentially vulnerable populations. Fundamental research characterizes real-world exposures and the key exposure factors for these populations. Research is designed to examine how developmental exposures during pregnancy and early childhood may impact health later in life, and how life stage affects responsiveness to environmental contaminants, particularly in children and older adults. Tools and methods for longitudinal epidemiology studies developed in this research are applied in STAR-funded Children's Environmental Health Centers and translated to other national longitudinal studies on children's health. A specific strategy is being applied to understand the predisposing factors for asthma as a function of life stage, considering interactions with contaminants in both outdoor (e.g., diesel particles) and indoor air (e.g., mold) environments.

Long Term Goal 4 (LTG4): Evaluation of the Impact on Human Health of Risk Management Decisions. *Research in this goal develops and tests indicators for gauging the effectiveness of risk management decisions and pollution mitigation efforts. This research makes use of fundamental information generated by the other three goals. Current efforts focus on real world scenarios and include projects developed in collaboration with EPA regional offices and by NCER grantees. These projects test the hypothesis that measured changes in community and personal exposures result in improvements in human health that can be measured and confirmed by using appropriate environmental health indicators. This research both contributes to and draws from issues raised in EPA's Report on the Environment.*

Research products are typically not program office or media-specific. Rather, HHRP research is designed to produce knowledge and tools that are generalizable to the needs of multiple Program Offices, Regions, other parts of ORD including the National Center for Environmental Assessment (NCEA) and The National Center for Computation Toxicology (NCCT), and other Federal Agencies (e.g. NIH/NICHD) and international groups (e.g. OECD) to further their goals.

For the present review, a nine-member Subcommittee was formed, the members of which are listed in Appendix A. The charge to the Subcommittee is provided in Appendix B and includes questions that originate with and relate to the Office of Management and Budget (OMB) Program Assessment Rating Tool (PART). The Subcommittee was provided with a number of documents related to the HHRP as well as several presentations made during public teleconferences and during the face-to-face meeting (see Table 1 for the dates of these events).

Table 1. Summary of HH Subcommittee Meetings

[illegible]	[illegible] [illegible] [illegible]
October 10, 2008	Administrative Call
October 10, 2008	Conference Call
December 1, 2008	Conference Call
January 7, 2009	Administrative Call
January 13-15, 2009	Face-to-Face Meeting
February 27, 2009	Conference Call

The following responses of the Subcommittee, by LTG, are organized according to the major topics of program relevance, structure, quality, coordination and communication, program performance, and scientific leadership.

III. LONG-TERM GOAL 1: USE OF MECHANISTIC DATA IN RISK ASSESSMENT

Program Relevance

The current HHRP objectives are essential for the Agency to improve their current risk assessment methodologies. The challenge for the program is to successfully integrate mode of action (MOA) results into the quantitative risk assessments. The choice of science used is appropriate. The computational toxicology and reproductive effects were particularly well received by the Subcommittee. Further, it was recognized that it has been an important shift to a pathways orientation instead of traditional single enzyme analyses. Additional scientific areas should be incorporated into the program, namely, epigenetics and genetic polymorphism in human susceptibility population studies. Regarding "asking the right questions," the subcommittee has some specific comments:

a. Epidemiology is being used to generate more dose-response curves for risk assessment, a trend that is likely to continue. The studies of MOA and biologically based dose response (BBDR) need to integrate more with this trend. Specifically, there is a need to provide biological plausibility to the epidemiology studies (e.g., potential mechanisms), but more importantly, use differences in the top-down and bottom-up approaches to generating dose-response curves to generate hypotheses that further our understanding. Moreover, epidemiologic studies will often have poor power at low dose, suggesting that an emphasis on low-dose modeling would be most useful for risk assessment.

b. The algorithm behind the choice of chemicals to study is not clear, and should be made explicit. What approaches will be used to adjust this over time? How can we review this process?

The program understands the importance of the MOA approach to risk assessment. It is not clear, however, to the Subcommittee that the program has shown to its stakeholders specific examples of how the incorporation of MOA considerations can benefit quantitative risk estimates, and we would like to see such examples. Also, the program needs to establish the validation of its models through the use of human data (e.g., NHANES and Pharmacological data). How to evaluate uncertainty is an additional important issue, and attention needs to be paid to modeling the variance as well as the mean.

The program is responsive to both the stakeholders and their issues, but leads in establishing emerging issues in risk assessment. The Subcommittee notes that with increasing use of dose-response estimates in epidemiology there becomes a greater need for interface with MOA. The use of the pharmacological literature of human data should also be explored.

We would like to see a specific example of a risk assessment whose results have changed as a result of this work. We are concerned that program offices may not be utilizing these tools optimally and that training programs may be needed to allow this technology transfer.

Program Structure

The structure of the program is good. It has a large number of parts. The level of effort on the individual parts is not clear to the Subcommittee. It was not possible to evaluate whether the distribution of skills among the personnel was appropriate. However, there was some concern with the small number of investigators with experience in epidemiology and biostatistics. These skills are important for integration and some modeling issues. Although milestones were given, there was not a clear way to evaluate their success. Again, it was not possible for the Subcommittee to evaluate the use of MYP in guiding its research.

Program Quality

The scientific quality of the program is high, as expected by the Subcommittee. The publication quality is good, but the Subcommittee cannot answer questions about competitive funding. What is very important is that the program is leading the entire field of toxicology in the areas of reproductive and developmental toxicology, computational toxicology, and also respiratory toxicology.

Coordination and Communication

There is good evidence of interactions between ORD and program staff. We assume this also applies to issues of planning. Better interactions among groups within ORD (e.g., the people in LTG 1 with LTG 2 and LTG 3) would be beneficial. There is critical help through the recently sponsored National Research Council (NRC) studies. Specific projects often involve collaboration with outside scientists. The level of their effort in the individual studies is not known. Generally, with a few exceptions, the quality of the outside collaborations has been very good. The Subcommittee believes that better translation of the potential impact of MOA on quantitative risk estimation and management be made to the program offices. Basically, an advocacy approach should be considered.

Program Performance

Considerable scientific progress has been made since the previous review. Measures of outcome were not evaluated by the Subcommittee. The use of program results by decision makers is very important, but we do not have specific examples of this at this time. Further, without a clear understanding of the resources involved, resource efficiency is difficult to evaluate.

Scientific Leadership

The leadership is outstanding, and the quality of the staff is generally first rate. The program is providing scientific leadership to the entire field of toxicology in areas such as reproductive and developmental toxicology, computational toxicology, and, to some extent, respiratory toxicology.

Summary Assessment

The scientific quality of the program and its outstanding leadership makes it an essential component of the Agency's human health research program. The program is at the forefront in computational biology as well as the traditional areas of developmental and inhalation toxicology. What is needed is better integration of MOA with the quantitative risk assessment generated by the epidemiology studies. In particular it is important to demonstrate the value and impact that the basic mechanistic studies of MOA have on the Agency's quantitative risk assessments.

Meets expectations.

Recommendations

IV. LONG-TERM GOAL 2: CUMULATIVE RISK

Program Relevance

In general, the goals of this LTG address research on cumulative risk assessments and susceptible populations as set forth in the MYP. These goals are grounded in and are responsive to several key legislative mandates (SDWA, FQPA, etc.).

The research program is clearly aware of the evolving nature of risk assessment, and the enlarging responsibility of the agency for these tools – from single chemical to complex exposure patterns and complex exposures. The research team has demonstrated the ability to move from single chemical with multiple routes of exposures to multiple chemicals with similar mode/mechanism of action. There is clear evidence of attempts to enhance the risk assessment methods to include susceptibility and the role of other types of stressors (e.g. social class, economic factors, age, sex, disease presence in the individual).

The HHRP objectives are appropriate. They address improving risk assessment methods by incorporating multiple interacting chemicals and non-chemical exposures that influence risk and are totally in line with objectives of the agency. However, except for one example [public health tools provided by CDC],the extent that approaches are translated to the regional offices and states was not apparent to the subcommittee.

RECOMMENDATION: The subcommittee recommends that the MYP include a concerted educational outreach effort to the Program Offices, Regional Offices and states regarding the use of sophisticated models and new knowledge developed through its research.

ORD's long-range research objective on forward and reverse prediction of source-to-dose-to health effects paradigm through development and performance evaluation of models and methods is a laudable one. This vision has great potential for impacting on the basis for Agency decision making, and elements from it will be very useful to LTG4's goal where indicators are identified within this paradigm for its Report on the Environment. Selecting stochastic and physiologically based pharmacokinetic (PBPK) models to integrate the source to health paradigm at this time appears to be relevant, while there are some concerns regarding the specific models and approaches selected. The goal to reduce the uncertainty in exposure assessments used for assessing risk is an appropriate one. Estimating distributions of human exposure is difficult with models, particularly when estimating the tails of the distribution.

RECOMMENDATION: The subcommittee recommends that goals or guidelines be

> defined that describe the threshold of acceptable accuracy for source-to-dose models and methods used in making assessments. Further characterization of the uncertainty of models as, for example, described in the paper by Ozkaynak, et al.[1] is highly endorsed.

The models and methods that are being developed are sophisticated, and perhaps even esoteric. The utility of the models and methods for use by the various partners throughout EPA remains to be fully appreciated. Future products will be potentially very useful and align with the needs of the ORD and the Program Offices; however, these products are of less interest and benefit to the Regional Offices or Homeland Security, where the latter often respond to emergencies or acute issues.

> **RECOMMENDATION:** As part of future BOSC reviews and as an accountability goal, the subcommittee recommends that evidence (in summary narrative form) is provided on the use of completed research products in cumulative risk assessments.

In summary, the research conducted under this LTG is staying true to the two major research goals on cumulative risk and susceptible populations as described in the 2006-2013 Multi-year Plan. The MYP is heavily influenced by the needs of the Program Offices.

Program Structure

It is evident that the planning and organization of research to accomplish the goals of LTG2 is guided by the MYP. The planning and organization is logical from this perspective, but it is not as comprehensive as it could be because it appears that the cross-LTG planning is minimal, if non-existent in some cases. As such, optimal value of their accomplishments is less than it could be.

The framework of the proposed research and the sequencing of related activities, e.g., the modeling from source to dose to health effect and the development of new models and tools for cumulative exposure assessments, appear to be properly structured so that new knowledge will inform corollary research activities occurring within this LTG in a timely fashion. Research planning has been guided by the MYP and thus, has been highly vertical-centric. The subcommittee believes that the science developed in LTG1 and LTG3 could even benefit more the research conducted in LTG2 if there were greater planning efforts and knowledge sharing between them. For example, knowledge about modes of action and corresponding modulators of action would benefit LTG3 in the development of exposure-to-dose models with more accurate predictability. Such knowledge would also facilitate interpretation of biomonitoring data and

[1] Analysis of Coupled Model Uncertainties in Source-to-Dose Modeling of Human Exposures to Ambient Air Pollution: A $PM_{2.5}$ Case Study. Ozkaynak, Haluk., Frey, H. Christopher, Burke, Janet, and Robert W. Pinder. Atmospheric Environment, 43(9), 1641-1649, 2009.

designing the biomonitoring strategies in epidemiological studies.

ORD's research in support of the N-Methyl Carbamate cumulative risk assessment has been timely for the Office of Pesticide Programs (OPP). Using similarly linked human exposure and dose models developed and applied for carbamates, it has been nicely dovetailed and logically extended to informing the cumulative risk assessment of Pyrethroid insecticides.

While the MYP has provided the overarching long-term goal of addressing cumulative risk and susceptibility assessments, the APG has provided the specific work product to the Program Offices in a highly successful manner.

> **RECOMMENDATION:** The subcommittee recommends the continuation of the general framework for planning with the inclusion of greater planning efforts and knowledge sharing between LTG1, LTG3, and LTG2, and with other agencies.

Program Quality

The research in LTG2 was divided into two basic areas: cumulative risk assessment and community-based exposure and risk screening. The cumulative risk assessment was also sub-divided into methods and statistical models for dose-additivity, full cumulative risk assessments for high priority environmental exposures, and methods development addressing a variety of questions.

Overall, the subcommittee finds the research that has already been conducted has mixed quality. There were a number of projects for which the methods being used were either using outdated statistical methods or were trying to create tools that have already been developed elsewhere. This is especially true for the statistical models for dose-additivity where there is a rich statistical literature that appears to have been ignored in favor of a linear models approach. This issue is also apparent for the community-based research where there is a tremendous literature on "non-chemical stressors" and the inclusion of researchers with experience in these areas should be included in the overall plan. There is also extensive expertise on the uses and limitations of ecologic health studies that does not appear to have been fully tapped.

> **RECOMMENDATION:** The subcommittee recommends that researchers who have extensive experience in "non-chemical stressors" be included in the overall plan for community-based research.

The examples in which cumulative risk assessments have been completed or are being conducted for OPP (OPs, carbamates and pyrethrins) appear to be well-conducted and at an appropriate level of complexity for the question at hand. The use of the SHEDS model to drive the exposure scenarios is a good example of cooperation across multiple labs and of using regulatory-validated models as part of the assessment. There is a tendency to think that every issue requires

a PBPK model, but these are highly specialized modeling forms that require an expertise that may not be available in other programs for an extended period of time. The quality of the research could be improved if a broader array of models and in some cases simpler ones, are included in the arsenal of tools being used for the analyses.

Coordination and Communication

The scientific leadership is commended for their coordination and communication efforts with program offices. There is evidence that scientists in this LTG are responsive to the Office of Water (OW) and the Office of Pesticide Programs (OPP). The needs of the offices are clearly part of the fabric of the research thrust that is undertaken in this LTG. OPP has expressed great satisfaction with the research products provided by ORD. These observations are to some extent reinforced in the results of the Partner Survey.

Based on testimonials, planning occurs between OPP and ORD. For example, ORD is taking a systems biology approach in defining underlying biological mechanisms of chemical mixtures and understanding the of dose and mixture composition on chemical interactions and joint toxic action of mixtures.

Even though ORD is providing some tools for examining community exposures that may be useful to the Regional Offices, there is less apparent interaction and thus coordination of research efforts that serve these offices. The very low satisfaction exhibited in the Partner Survey seems to support this observation.

> **RECOMMENDATION:** As a future goal, the subcommittee recommends more engagement of the Regional Offices in planning and identifying areas where they need tools, methods, and data from ORD.

Based on discussions with the poster presenters, training and communication through outreach efforts with the stakeholders of models, methods, and data are evident but somewhat rudimentary.. Training and outreach efforts will become even more important as ORD scientists develop comprehensive, complicated and even esoteric models and methods for use by the stakeholders (see Recommendation a).

ORD scientists are very much engaged in communicating knowledge developed from their research endeavors to the national and international scientific community. This occurs through publications, presentations at national and international conferences, briefings and seminars, and preparation of reference compendiums, e.g., pesticide exposure factors for children.

Program Performance

Overall, program performance was coherent, within a somewhat bifurcated programmatic context. This division was mandated by the focus on cumulative pesticide exposure-effect studies driven by OPP needs that accounted for one major component of the output. In contrast, other consideration of cumulative exposure-response appeared to be either theoretical/conceptual (or to some extent, hypothesis generating) or, on the other hand, service oriented. The latter, a programmatic approach to risk mapping at the local level, appears to be popular with field offices. For this program, in performance, to succeed in the "response" component of the exposure-response dyad, may require considerably more evolution, given the difficulties in quantifying such outcomes. Program performance, in terms of intramural-extramural balance, was felt to be meritorious, as was peer-review publication productivity.

This LTG has become more focused and organized during the past couple of years. It is gaining traction where it seemed to be lacking before.

Scientific Leadership

In examining scientific leadership, the subcommittee addressed the question of what is the overall role of HHRP in promoting the improved use of science in cumulative risk assessment. And in looking at this issue, the main question is what should have been done 5 years ago, what should be the state of the use of the science today, and what will be used in 5 years? The subcommittee is pleased that the leadership of the ORD in moving this issue to the forefront and has made great strides during the past 4 years by completing some very high profile and clear cumulative risk assessments for pesticides. However, the subcommittee believes this issue is still behind where it could be, and it will continue to lag if the EPA does not invest in the necessary resources in both dollars and skills that are necessary to move this issue forward. The work, both intramural and extramurally, on the interpretation and use of biomarkers of exposure is needed and also shows good scientific leadership by the HHRP. Finally, the community-based tools and websites being developed are a good step forward in providing tools that local communities can use to make local decisions. This activity can also move more rapidly than it has and begin to incorporate a broader community of scientists to aid in the program.

> **RECOMMENDATION:** The subcommittee suggests that an added influx of resources into developing the science in cumulative risk assessments if such assessments are to be effective in a reasonable timeframe.

Substantial evidence exists that researchers are providing thought leadership by participating in a variety of boards, panels, workshops, and in presentations at conferences and through publications.

Summary Assessment

The leadership and scientists of this Long Term Goal are commended for their accomplishments. They recognized the need and demonstrated the ability to move from single chemical with multiple routes of exposures to multiple chemicals with similar mode/mechanisms of action. They have successfully incorporated sophisticated modeling concepts into N-methyl carbamate risk assessments and have done so in partnership with the Program Offices. The effort in this LTG has remained true to the two major research goal on cumulative risk and susceptible populations as described in the MYP. However, the Subcommittee believes that even though the planning and organization has been logical, this LTG could achieve greater benefits from more cross-LTG planning. The coordination and communication effort with program offices is laudable; however, the Subcommittee believes that more attention should be given to the needs of Regional Offices. Overall there is substantial evidence that LTG2 scientists are providing thought leadership through participation in a variety of boards, panels, workshops, and in presentations at conferences. The bibliometric data indicates that they are creating new knowledge, transferring this knowledge to the public domain and adroitly applying it to environmental health issues.

Meets expectations

Recommendations

1. The subcommittee recommends that the MYP include a concerted educational outreach effort to the Program Offices, Regional Offices and states regarding the use of sophisticated models and new knowledge developed through its research. [Ref.: Program Relevance; Coordination and Communication]

2. The subcommittee recommends that goals or guidelines be defined that describe the threshold of acceptable accuracy for source-to-dose models and methods used in making assessments. Further characterization of the uncertainty of models as, for example, described in the paper by Ozkaynak, et al. (1) is highly endorsed. [Ref.: Program Relevance]

3. As part of future BOSC reviews and as an accountability goal, the subcommittee recommends that evidence (in summary narrative form) is provided on the use of completed research products in cumulative risk assessments. [Ref.: Program Relevance]

4. The subcommittee recommends the continuation of the general framework for planning with the inclusion of greater planning efforts and knowledge sharing between LTG1, LTG3, and LTG2, and with other agencies. [Ref.: Program Structure]

5. The subcommittee recommends that researchers who have extensive experience in “non-chemical stressors” be included in the overall plan for community-based research.

6. As a future goal, the subcommittee recommends more engagement of the Regional Offices in planning and identifying areas where they need tools, methods, and data from ORD. [Ref.: Coordination and Communication]

7. The subcommittee suggests that an added influx of resources into developing the science in cumulative risk assessments if such assessments are to be effective in a reasonable timeframe. [Ref.: Scientific Leadership]

V. LONG-TERM GOAL 3: SUSCEPTIBLE AND VULNERABLE POPULATIONS

Program Relevance

There was consensus on the Subcommittee that the subject area of susceptible and vulnerable subpopulations is highly relevant to the Human Health Research Program. The previous Program Review (2005) specifically commented positively on the relevance of assessing the effects of low-level environmental exposures among susceptible subpopulations insofar as such effects might be only manifest within those subpopulations, and that for that reason such groups should be a priority for study. Although this is one argument in favor of this research focus, its rationale is more broadly based. For example, a higher frequency of responses in susceptible subpopulations provides more statistical power to epidemiological investigations and controlled human exposure studies, allowing detection of effects with enrollment of smaller numbers of subjects. Therefore, such studies are more feasible to carry out and more likely to provide data relevant to risk assessors seeking to protect both more- and less-susceptible groups. Moreover, the responses of one susceptible subpopulation may provide insights into mechanisms of action and cumulative exposure effects relevant to other subgroups and to the population as a whole. One example of insights into mechanism of action derived from studying vulnerable subpopulations may be found in genetic polymorphisms associated with increased risk of disease onset or worsening.

Program Structure

The program structure in relation to the overarching Long Term Goal raised a series of questions within the Subcommittee. These questions related principally to the relative weighting of the research priorities, in particular the programmatic dominance of children as a vulnerable subpopulation within the "life-stage" research track, and, beyond that, the life-stage track in relation to the two other structural components of program: track two, methods for longitudinal research (which is, in effect, an extension of childhood life-stage activities) and track three, asthma (which, as will be discussed below, is approached by the Human Health Program as a surrogate measure of childhood vulnerability).

Confusion as to the underpinning for the programmatic structure may stem in part from a lack of clarity in the Agency's scientific justifications, as elucidated. In that regard, the comments of the 2005 Review on this aspect of the LTG 3 are noteworthy: "Although the Agency's focus on children as a susceptible population subgroup appears well justified, the justification was based on a consensus of recommendations across external advisory bodies (e.g., Office of Pollution

Prevention and Toxics, National Research Council). This justification can be strengthened by the Agency's own scientific assessment of the public health benefit to be achieved through a research focus on children as a particular subpopulation. *Such justification is likely to become more important in considering potential subpopulation research foci that may be less obvious than children.*" (page 32 of Report, emphasis added) At the face-to-face review session, Dr. Devon Payne-Sturges, EPA/ORD/National Center for Environmental Research (NCER) commented that, "The definitions of vulnerability and susceptibility do not come from the Multiyear Plan, but from clients, including EPA's National Environmental Justice Advisory Council and Risk Assessment Forum." These comments further underscore the need for the EPA to re-examine internally (with the Agency itself) how its sees the definition and scope of "susceptible and vulnerable" populations.

This is particularly relevant to the need for a better thought-out scientific justification for the selection of asthma as a disease of primary study interest for the Agency. Operationally, this appears to be viewed as little more than a surrogate measure of childhood-associated life-stage vulnerability. Indeed, by way of background (Multiyear Plan, page 55), after describing the prevalence of asthma in children (as of 1995), it is stated "Although children appear to be the population most at risk, there is growing concern that new cases are also arising in adults." This conceptual weakness fails to recognize fully the importance of asthma across various life stages, does not appear to separately concern itself with asthma exacerbation or recrudescence as opposed to initiation, or take into account the potential vulnerability of the high risk subpopulations of adults of working age as well as the elderly (those no longer in the labor force). An outgrowth of such conceptual shortcomings is a missed opportunity to consider asthma and chronic obstructive pulmonary disease (COPD) within a spectrum of airway disease, particularly insofar as this may be relevant to susceptibility of exacerbation (not causation). For example, it is well established that persons with increased non-specific airway responsiveness are more susceptible to broncho-constriction following exposure to certain air pollutants (e.g., sulfur dioxide), thus constituting a subpopulation that should be of great interest across a range of ages. In addition, susceptibility to environmentally-related disease worsening or complications is not limited to pre-existing asthma. Diabetes has been demonstrated to convey susceptibility to the effects of air pollution, for example.[2,3] And there are now well established genetic risk factors for neurocognitive effects of lead.[4,5] The choice of ages and predisposing conditions examined does not appear to be consistent with Agency risk assessments, such as the "812" studies required by the Clean Air Act to examine the health benefits gained through legislation,

2 Liu L, Ruddy TD, Dalipaj M, et. al. Influence of personal exposure to particulate air population on cardiovascular physiology and biomarkers of inflammation and oxidative stress in subjects with diabetes. J Occup Environ Med 2007; 49:258-265

3 Zanobetti A, Schwartz J. Are diabetics more susceptible to the health effects of airborne particles? Am J Respir Crit Care Med. 2001; 164:831-3.

4 Wang FT, Hu H, Schwartz J, Weuve J, Spiro AS III, Sparrow D, Nie H, Silverman EK, Weiss ST, Wright RO. 2007. Modifying Effects of the HFE Polymorphisms on the Association between Lead Burden and Cognitive Decline *Environ Health Perspect*: 2007; 115:1210-5.

5 Rajan P, Kelsey KT, Schwartz JD, Bellinger DC, Sparrow D, Spiro III A, Smith TJ, Wright R, Nieh H and Hu H. Lead Burden and Psychiatric Symptoms and the Modifying Influence of the |*delta*|-Aminolevulinic Acid Dehydratase (ALAD) Polymorphism: the VA Normative Aging Study. Am J Epidemiol, 2007; 166:1400-8.

for example, insofar as the burden of air pollution health effects appears to fall disproportionately on the elderly rather than children.

It is interesting to note that the introduction to the posters subsumed under LTG3 took a far broader conceptual view of the criteria upon which subpopulation vulnerability might be defined and seemed to contextualize asthma as an example of a preexisting condition that could impart vulnerability. This was not reflected, however, in the content of the poster presentation itself nor by the portfolio of research upon which it was based.

Program Quality

The program quality within LTG3 on a project by project basis is excellent to outstanding. For example, among the extramurally funded projects there is a level of epidemiological sophistication that represents an impressive gathering of resources. Exemplars of this can be found in the Duke Center for Children's Environmental Health Research and the University of Southern California's Children Health Study. Innovation and creativity is not limited to the extramural projects, but is also evident within the intramural components. Indeed, it is likely that additional resources directed at areas currently under-studied (e.g., aging human populations) would also result in high quality outputs.

Although the quality of the individual parts is not at issue, the programmatic quality as a whole is compromised to the extent that susceptibility/vulnerability has been defined in an overly narrow dense to virtually exclude rigorous investigation beyond the childhood life-stage aspect of such risk. This issue has been addressed elsewhere in this review and will not be revisited in greater detail as a quality issue per se.

Coordination and Communication

As with other Long Term Goals, the scientific leadership is to be commended for their coordination and communication efforts with program offices and, beyond that, in terms of interagency collaborations with the National Institute for Environmental Health Sciences (NIEHS) and the Centers for Disease Control (CDC). Such coordination has been particularly relevant to the development and funding of RFA's for extramural research and to the complexities of the National Children's Study.

LTG3 was particularly noteworthy for combined strengths on both the intramural and extramural fronts. Coordination among the various and diverse research projects that fall within the scope of LTG3, particularly interactions between intramural and extramural undertakings, would present challenges to any organization. The general recommendation made previously in regard to strengthening internal EPA epidemiological resources is particularly relevant to this specific point.

Although the within-Long Term Goal coordination indicated programmatic coherence, questions arose in regard to interactions across goals, specifically in regard to the "cumulative exposure" focus of LTG2. Despite that fact that this component of LTG2, on its face, seems directly related to the core of LTG3, there did not appear to be much in the way of additive, let alone synergistic interactions, between the two working groups. This may point more to the aspects of the LTG structure that are inherently artificial, rather than reflecting on the content of the work involved.

Program Performance

The childhood and the asthma components of LTG3 have been highly productive, with particular note of the extramurally funded (jointly with the NIEHS) children's centers and the intramurally-based research program on the developmental (pre-natal and early childhood) origins of adult diseases. This successful performance can be measured in peer-reviewed and governmental publications, as well as in practicable applications, such as the "Relative Moldiness Index." Although the National Children's Study, by design, will not have endpoint data for analysis, methods-related research and documentation has also reflected excellent performance.

The aging component of the life-stage track, by contrast to the childhood-asthma components, has been a relatively weak performer. In scope, a similar program of work was described in the last review, five years ago. In particular, there appears to be very little epidemiological or controlled human exposure activity within this track, either currently or planned for the immediate future. Paralleling this, the commendable interagency communications alluded to above in regard to the NIEHS and CDC do not seem to have extended from childhood susceptibility (including asthma) to the aging, nor is their evidence of active coordination with other NIH agencies that could have a role (in particular, the National Institute on Aging). This may reflect the priorities of those outside bodies as much or more than those of the EPA.

It is difficult to evaluate how program performance in LTG3 more narrowly translates into risk assessment activities. This may be an artifact of the current structure of the four Long Term Goals as formulated, given that programmatic outcomes are more clearly addressed in LTG4. Also, the "partner testimony" as presented was not particularly germane in this regard. Nonetheless, the trajectory of the program performance of LTG3 indicates that risk assessment, as it touches on childhood as a susceptibility factor, will benefit from this work.

Scientific Leadership

The EPA is inarguably a scientific leader in regard to the National Children's Study. Its role is formidable in the extramural funding of childhood vulnerability in general and in particular in terms of childhood asthma. This includes both independent leadership and in coordination with

the NIEHS and CDC. The EPA's role in extramural epidemiological research in this arena represents an area of substantive maturation within the Human Health Research Program. At the same time, intramural laboratory-based scientific leadership for the same foci has not been compromised.

Summary Assessment

LTG3 was assessed as meeting program expectations based on the inarguable population health and public policy relevance of this area of research, the LTG's commendable coordination and communication efforts with program offices and the scientific leadership role manifest, the excellent to outstanding scientific quality of the specific endeavors, the high level of productivity within the areas in which it has LTG has focused. The programmatic structure was assessed as over-weighting childhood health within its life-stage construct of vulnerability additionally treating asthma, one if its major foci, as little more than a surrogate of childhood risk. Absent this serious limitation this LTG would have been assessed as "exceeding expectations."

Meets expectations.

Recommendations

1. Develop a more fully elucidated conceptual framework for vulnerability and susceptibility.

2. Redress program imbalance within the life-stage arm of LTG3 such that the strengths of the childhood susceptibility research thrust are matched with an expanded research program addressing the elderly as well as potential subgroups across the entire age range.

3. Rethink the approach to asthma a target condition so that it is not simply approached as a surrogate of childhood susceptibility to new disease onset, but rather considered across the entire age range and considered also in terms of vulnerability in pre-existing disease.

4. In addressing preexisting conditions, consider expansion beyond asthma to encompass other airway disease, in particular COPD, and, beyond lung diseases consider other classes of disease such as neurological and endocrine disorders.

5. Attempt better integration across long term goals, in particular with LTG2 in terms of cumulative exposure.

6. Parallel successful intra-agency collaborations with the NIEHS and the CDC in regard to childhood asthma to address other vulnerable subpopulations, for example the National Institute on Aging in terms of potential susceptibility of the elderly to certain environmental exposures.

VI. LONG-TERM GOAL 4: DEVELOPING TOOLS TO EVALUATE RISK MANAGEMENT DECISIONS

Program Relevance

The program is structured around three themes: the development of means and methods to measure impacts, research studies of impact measurements, and finally the Report on the Environment (ROE).

Program Structure

The leadership of the HHRP has done an excellent job in bringing these groups together, including the regional offices, and moving forward on the overall goals of LTG4. Hence, we find they have done a good job in linking these various activities into a single long-term project that holds great promise for addressing the utility of EPA's programs to improve health and the environment.

There are some activities that can be developed to strengthen the structure of this LTG. The BOSC Subcommittee feels that the HHRP could do a better job of interaction and linkage with other federal agencies. While this activity uses quite a bit of information generated by the CDC and the U.S. Geological Survey (USGS), they need to work closely with these agencies to improve their products to have greater utility for the needs of the HHRP. In addition, the HHRP should partner with other granting agencies such as the NIH and the National Science Foundation (NSF), and the Center for Medicare and Medicaid Services, monitoring their activities and taking advantage of opportunities to partner and leverage resources. In this specific area, where you are attempting to evaluate risk management decisions, testing new technologies or procedures in concert with epidemiology or laboratory studies just beginning could both improve the activities and extend the resources.

There was some concern that, in this specific area where long-term trends are being evaluated over extended periods of time, loss of institutional memory could lead to unexpected problems in completing research or errors in the interpretation. The HHRP is encouraged to look into this issue and find ways to limit its impact on the program.

Finally, the BOSC Subcommittee sees the location of the ROE within the EPA bureaucracy as a critical component of LTG4 and, indeed, the management and evaluation of the entire EPA. This activity serves a critical role for summarizing the long-term trends in environmental quality

and as a communication tool to highlight the successes and areas for improvement of the EPA. As such, we feel the ROE should be as close to the Administrator as possible and we would suggest it be moved to the office of the Administrator's lead scientific advisor.

Program Quality

The program is in early stages, but the signs of quality are very encouraging. The ROE was an excellent effort to track trends in exposure and health, and begin to provide a framework for closing the loop. The program is cognizant of the potentials for confounding and methodological issues in ecological analyses. We recommend explicit incorporation of these understandings into documents to ensure that inappropriate pressure to look at the concordance or discordance of simple time trends are more easily resisted. The BOSC does think there is room to expand the set of data sources for looking at health trends, including Centers for Medicare and Medicaid Services (CMS) data on MEDICARE and MEDICAID, Homeland Security monitoring networks for ER visits, state databases on all hospital admissions, NIH databases, etc. It is not clear how other efforts, such as the Section 812 study, the OMB study of benefits of government regulations, etc. interact with this, and incorporation of some information from them, and generation of other cumulative burdens would be useful.

The BOSC was particularly impressed with the early results of community-based studies that close the loop (e.g., the exposure modeling in New Haven, the NYC study of pesticide exposure-- which resulted in a regulatory change in New York, etc.). These examples combine good science in the original study with appropriate interactions with community leaders, resulting in policy-relevant information being transferred, and demonstrations of the benefits of interactions. In part, these examples made use of studies that were funded under other rubrics (e.g. Children's Centers), which were subsequently integrated into the LTG. However the study of the impact of changes in drinking water disinfection on health seemed to be closer to a *de novo* exercise of this goal. The approach combined innovative science (developing a new approach for measuring antibodies to multiple infectious agents using a noninvasive approach), with an opportunistic intervention trial, which we think could be a model study for this LTG. As in comments on other products, however, we would like to see better evidence of a commitment by the program office to take the results of this research and use it more widely for regulatory purposes.

Coordination and Communication

This LTG appears to be very well coordinated and, certainly through the ROE, communicated to the public, other scientists and the regulatory community at the EPA. The one concern of the BOSC Subcommittee is whether the tools being developed that allow for the evaluation are being developed in a way that will allow them to be shared with other groups who would want to use them to expand these activities. It would be appropriate and useful to have detailed, publicly available databases providing the underlying facts supporting the ROE and the other evaluations. By making these widely available, the agency is allowing others to, not only review their

scientific interpretations of these data, but also to suggest alternative evaluations and interpretations that may better address the goals of LTG4.

Program Performance

This is a relatively new LTG that has gained considerable traction and has achieved several fine accomplishments in a short period of time. Overall, the program performance was considered exceptional.

As indicated above there are three themes of research. The review committee believes that the Report on the Environment (ROE) is asking the right set of science questions. Nineteen indicators have been identified to evaluate and answer questions related to exposure and effects of environmental contaminants as it is represented in the environmental public health paradigm consisting of source to transport/transformation to exposure, to dose, to altered structure/function to adverse health. Indicators are being used to help evaluate trends in human exposure to environmental contaminants, in the general health status of a population and in human disease and conditions, and thus, the impact of regulations.

In addition, excellent progress has been made to demonstrate the performance of the formulated concepts and approaches, one of which is how trends in human exposure to an environmental contaminant can be assessed by measuring an indicator through biomonitoring. For example, the impact of public education on smoking and environmental tobacco smoke (ETS) exposure in the human population has been determined by measuring and observing decreasing levels of cotinine, a nicotine metabolite, in urine. Another example is the decreasing levels of lead in blood subsequent to implementing lead regulations.

Trends in increased life expectancy, decreases in infant mortality, and increase in asthma prevalence in adults and children are indicators that are being used to assess health status. The ORD scientists are cognizant of the difficulties of relating trends in human exposure and body burdens to stressors with changes in a health status indicator. Burden of disease or some type of calculated national and global risk is useful and should be included with caveats. For example, correlating decreasing blood lead levels and changes in cognitive abilities in children is fraught with problems. Some are concerned that when doing these evaluations of risk management decisions, you could be forced into doing simple ecological evaluations of disease incidence or mortality rates that do not really describe that improvements have occurred. Never the less, there is a need to be direct in the estimate of the human health impact of environmental interventions and the agency is encourage dto partner with national and global agencies attemting to dot he same thing and use uniform measures such as DALY's. Overall, the indicators selected for monitoring exposures or adverse effects have been selected on solid scientific principles developed in LTG2 and LTG3.

The HHRP is strongly encouraged to continue thinking outside of the box regarding how to evaluate the impact of policy and regulations on human health and thus bring accountability to the decisions made about environmental issues. There are many different means available to

evlaute programmatic performance and these should be studied and applied where appropriate. Expanding beyond case studies to broader evaluations is also encouraged.

Scientific Leadership

The framework developed for assessing the public health impacts of risk management is a necessary means to move forward in this LTG, and it shows good leadership on the part of the HHRP. The nascent research studies of public health impact are timely and could be effective in advising the agency on how this can be done. The example studies are well positioned to address problems that will undoubtedly arise in these types of studies, such as stakeholder involvement, scientific quality, statistical power and linkage to future utility. Finally, the ROE is well done and is presented in such a way that it is likely to have the expected impact on the Agency that evaluations like this should. The leadership of the ROE has planned for changes to the ROE over time that mimic some of the recommendations of the Subcommittee such as a broader use of other databases, online tools for understanding environmental impacts at the local level and better indices of the health implications of risk management decisions displaying sound scientific leadership in this area.

Summary Assessment

LTG4 was assessed as being an integral part of closing the loop created when we identify a hazard and develop/implement decisions related to that hazard by working to develop the tools necessary to determine if the management decisions were warranted, effective and should be continued. Many times this critical aspect of environmental health decision making is overlooked and programs are put into place that are unnecessary or no longer effective. Having the tools to evaluate risk management decisions must be a priority and the Subcommittee is pleased this is being undertaken with regard to the long-terms impacts on human health. Even though the program is rather new, we see enthusiasm in the staff involved, early successes in the approaches chosen and the beginnings of a very successful activity for the Agency.

Exceeds expectations.

Recommendations

1. Improve interaction and linkage with other federal agencies and state agencies
2. Develop a means to capture and preserve institutional memory to improve long-term assessment of programs
3. Make the ROE more prominent and influential in the Agency by moving it into the Office of the Administrator or his/her chief scientific counsel.

4. Expand the use of health databases used to evaluate improvements in human health related to improvements in the environment, remaining cautious in interpreting these types of ecological analyses
5. Expand the use of direct estimates of the health implications of environmental interventions by calculating burden of disease or similar appropriate measures of risk.
6. Incorporate additional case studies into the LTG and attempt to extrapolate from existing case studies to other examples.

VII. APPENDIX A
HUMAN HEALTH SUBCOMMITTEE MEMBERS

Board of Scientific Counselors (BOSC)
2009 Human Health Subcommittee

Chair: **James E. Klaunig, Ph.D.**
School of Medicine
Indiana University

Vice Chair: **Henry Falk, M.D., M.P.H.**
Director, Coordinating Center for Environmental Health and Injury Prevention
Centers for Disease Control and Prevention

Members: **Paul D. Blanc, M.D., MSPH**
Chief, Division of Occupational and Environmental Medicine
Department of Medicine
University of California San Francisco

George P. Daston, Ph.D.
Miami Valley Laboratories
The Proctor & Gamble Company

David G. Hoel, Ph.D.
Medical University of South Carolina

Donald Mattison, M.D.
Senior Advisor to the Directors of the National Institute of Child Health and Human Development, and the Center for Research for Mothers and Children
National Institutes of Health, NICHD

Edo Pellizzari, Ph.D.
Senior Fellow
RTI International

Christopher J. Portier, Ph.D.
Associate Director, National Institute of Environmental Health Sciences
National Institutes of Health

Joel Schwartz, Ph.D.
Professor, Department of Environmental Health

1 Harvard University School of Public Health
2

VII. APPENDIX B
CHARGE

DRAFT
Program Review Charge
Human Health Research Program Subcommittee

1.0 Objective. The BOSC Human Health Research Program Subcommittee will conduct a retrospective and prospective review of ORD's Human Health Research Program, and evaluate the program's relevance, quality, performance, and scientific leadership. The BOSC's evaluation and recommendations will provide guidance to the Office of Research and Development to help:

- plan, implement, and strengthen the program;
- compare the program with programs designed to achieve similar outcomes in other parts of EPA and in other federal agencies;
- make research investment decisions over the next five years;
- prepare EPA's performance and accountability reports to Congress under the Government Performance and Results Act; and
- respond to assessments of federal research 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 Science 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 especially 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, especially in a research program like the Human Health Research Program that is specifically designed to address longer term, relatively intractable problems. Indeed, decades may be required for some research outcomes to be realized and measurable in terms of public health outcomes.

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 inter-disciplinary research programs such as these, progress toward outcomes can not 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 research 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 Agency partners, 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 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 coordination 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 four or five years is needed to characterize research progress, to assess how clients/partners are applying research to strengthen environmental decisions, and to evaluate their feedback about the usefulness of the research. Conducting program evaluations at this interval enables assessment of: research progress towards long term goals and the ability of the program to adjust its approaches and plans according to unanticipated results; the overall scientific quality and decision-making value of the research; and, to what extent the research progress has resulted in short-term outcomes for specific clients/partners.

As guidance for these periodic program evaluations and consistent with the recent NAS report "Evaluating Research Efficiency in the U.S. Environmental Protection Agency", ORD follows the STP/OMB *Research and Development Investment Criteria* appended to this document.

3.0 Background for ORD's Human Health Program and Draft Charge Questions

Background

The overall goal of the HHRP, as defined in the current MYP (June 2006), is to characterize and ultimately reduce uncertainties in extrapolations inherent in the risk

assessment process by providing a greater understanding of the fundamental determinants of exposure and dose and the basic biological changes that result from exposures to environmental toxicants. This research supports risk assessment activities conducted under the ORD Human Health Risk Assessment (HHRA) MYP and by Agency Program and Regional Offices. An overarching theme is to improve our understanding of the linkages in the exposure-to-dose-to-effect continuum. It is of necessity an inter-disciplinary research program that develops the methods, models and data needed to characterize uncertainties in each of these linkages and apply the information to the real world to elucidate exposures and risks in our communities. Research projects are integrated across the intramural and extramural grants programs and are currently organized around four Long Term Goals (LTGs). The relative effort and specific projects under each goal are adjusted on an annual basis in response to research findings as they become apparent and based on available resources.

Long Term Goal 1 (LTG1): Risk assessors and risk managers use ORD's methods, models or data to reduce uncertainty in risk assessment using mechanistic (or mode of action) information. *Fundamental research in this goal elucidates mechanisms of action of priority environmental contaminants and related families of contaminants, explores toxicity pathways that are perturbed by these contaminants, and uses this information to develop and link pharmacokinetic and pharmacodynamic models for use in risk assessment. These models are applied to reducing uncertainties associated with extrapolating from high to low dose, from test species to humans, from in vitro data to in vivo exposures, and between cancer and non-cancer effects. Progress is measured by the extent to which this information is being used in Agency risk assessments and rulings. A new direction in this goal is to develop a systems biology approach and apply novel models such as a virtual liver to predict toxicity and estimate risk.*

Long Term Goal 2 (LTG2): Risk assessors and risk managers use ORD's methods, models, and data to characterize aggregate exposure and cumulative risk in order to inform risk management for humans exposed to multiple environmental stressors. *Research in this goal develops and applies biomarkers to assess cumulative exposure and risk; develops and applies source-to-dose models for cumulative risk assessment and dose reconstruction; and creates tools for community-based exposure and risk assessments of complex mixtures. The long term objective is to produce a research framework outlining tools and approaches to characterize and assess aggregate exposures and cumulative risks, especially for vulnerable populations, based on a full range of both chemical and non-chemical stressors.*

Long Term Goal 3 (LTG3): Risk assessors and risk managers will use ORD's methods, models and data to characterize and provide adequate protection for susceptible populations. *This goal focuses on susceptibility as a function of life stage with a strong emphasis on children and older Americans as potentially vulnerable populations. Fundamental research characterizes real-world exposures and the key exposure factors for these populations. Research is designed to examine how developmental exposures during pregnancy and early childhood may impact health later in life, and how life stage affects responsiveness to environmental contaminants,*

particularly in children and older adults.. Tools and methods for longitudinal epidemiology studies developed in this research are applied in STAR-funded Children's Environmental Health Centers and translated to other national longitudinal studies on children's health. A specific strategy is being applied to understand the predisposing factors for asthma as a function of life stage, considering interactions with contaminants in both outdoor (e.g., diesel particles) and indoor air (e.g., mold) environments.

Long Term Goal 4 (LTG4): Evaluation of the Impact on Human Health of Risk Management Decisions. *Research in this goal develops and tests indicators for gauging the effectiveness of risk management decisions and pollution mitigation efforts. This research makes use of fundamental information generated by the other three goals. Current efforts focus on real world scenarios and include projects developed in collaboration with EPA regional offices and by NCER grantees. These projects test the hypothesis that measured changes in community and personal exposures result in improvements in human health that can be measured and confirmed by using appropriate environmental health indicators. This research both contributes to and draws from issues raised in EPA's Report on the Environment.*

These four goals are inter-related by design. Findings in each goal continually enable progress in and adjustments to research in one or more of the others. For example, new biomarkers developed in LTG 2 may be used as indicators of children's exposures and health in LTG 3 and as measures of the impact of risk management decisions in LTG 4. Modes of action elucidated in LTG 1 are used to develop models for evaluating cumulative risk in LTG 2. Also, research products are typically not program office or media-specific. Rather, HHRP research is designed to produce knowledge and tools that are generalizable to the needs of multiple Program Offices, Regions, other parts of ORD including the National Center for Environmental Assessment (NCEA) and The National Center for Computation Toxicology (NCCT), and other Federal Agencies (e.g. NIH/NICHD) and international groups (e.g. OECD) to further their goals.

Draft Charge

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

Program Relevance

1. How appropriate are the current HHRP objectives for achieving the Agency's strategic plan (Safe Communities) and providing a clear public benefit?
2. How appropriate is the science used to achieve each LTG, i.e., is the program asking the right questions, and using the most appropriate methods?
3. How effectively does the program identify and respond to the needs of its stakeholders, i.e. EPA partners in the Program Offices, Regions, and ORD, and other partners outside EPA, and how effectively does it adjust to their changing needs.

4. How effectively does the program identify emerging issues relevant to its objectives and adjust its research strategy accordingly?

Factors to consider: the degree to which the research is driven by EPA priorities; the degree to which this research program has had (or is likely to have) an impact on Agency decision making; the appropriateness of the key science questions; the responsiveness of the research to the needs of EPA programs, regions, and other stakeholders within ORD (e.g. risk assessors); the responsiveness of the research plan to recommendations from outside advisory boards and stakeholders; the extent to which research program scientists participate on and contribute to Agency workgroups engaged in identifying and addressing research needs.

Program Structure

1. How clear a logical framework do the LTGs provide for organizing and planning the research, with clearly identified priorities and program outcomes?
2. 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?
3. Does the program use the MYP to help guide and manage its research? And is the program responsive to changing results and priorities as the science progresses?

Factors to consider: the scope of the LTGs in providing a logical framework for organizing the Human Health program to best meet its overall goals; the degree of clarity in the pathway to the performance goals specified for accomplishing the LTGs; the appropriateness of the LTG and associated Annual Performance Goals (APGs) identified in the MYP as the means to meet the overall objectives of the program.

Program Quality

1. How high is the scientific quality of the program's research products?
2. Are the means the program employs to ensure quality research (including peer review, competitive funding, etc.) sufficient?

Factors to consider: the scientific soundness of the research approaches used; the impact and use of research results by EPA program and regional offices and other organizations; the regularity with which papers on common themes are synthesized into documents more useful to decision-making; the degree to which peer reviewed publications from this program are cited in other peer reviewed publications, the immediacy with which they are cited, and their impact factor; the processes used to peer review intramural research designs and products (e.g., division-level or product-level reviews by independent panels); and the processes used in the competitive extramural grants program.

Coordination and Communication

1. How effectively does the program engage scientists and managers from ORD and relevant program offices in its planning?

2. How effectively does the program engage outside organizations, both within and outside government, to promote collaboration, obtain input on program goals and research, leverage the use of its resources with other organizations to achieve higher efficiency and avoid duplication of effort?
3. How effective are the mechanisms that the program uses for communicating research results both internally and externally?

Factors to consider: the extent to which program/regional office scientists/managers are involved in planning the research; the degree of collaboration and coordination with other federal agencies, academic institutions, industry partners, and/or other countries; the timeliness and means for making quality (peer-reviewed) information available to the Agency and scientific community (e.g., through peer reviewed publications, briefings, scientific meetings, seminars); the extent to which research reports are synthesized into review documents and/or guidance documents and made available to Agency partners?

Program Performance

1. How much progress is the program making on each LTG based on clearly stated and appropriate milestones?
2. How well defined are the program's measures of outcomes?
3. To what extent are the program results being used by environmental decision makers to inform decisions and achieve results?
4. How efficiently has the program invested and managed resources to achieve the LTGs?

Factors to consider: the degree to which scientific understanding of the problem has been advanced; the degree to which scientific uncertainty has been reduced; the extent to which the program demonstrates impact and its products are used by EPA program and regional offices, ORD partners, and other organizations; the effectiveness of the program in identifying and investing in the most promising lines of research to achieve the LTGs; the relative prioritization and allocation of resources and scientific staff among the LTGs; and the investment of resources in short-term vs. long-term research priorities

Scientific Leadership

1. Please comment on the leadership role the research program and its staff have in contributing to advancing the current state of the science and solving important environmental health research problems.

Factors to consider: the degree to which this program is identified as a leader in the field; the degree to which peer reviewed publications from this program are cited in other peer reviewed publications, the immediacy with which they are cited, and their impact factor; the degree to which Human Health scientists serve/are asked to serve on national/international workgroups and advisory groups, as officers in professional societies, on publication boards; the degree to which Human Health scientists lead national/international collaborative efforts, organize national/international

conferences/symposia, and are awarded for their contributions/leadership; and benchmarking of scientific leadership relative to other programs, agencies, and countries.

(B) Summary Assessment (rate program performance by LTG): A summary assessment and narrative should be provided for each LTG. The assessment should be based primarily on 3 of the questions included above, which are:

1. How appropriate is the science used to achieve each LTG, i.e., is the program asking the right questions, with the most appropriate methods?
2. How high is the scientific quality of the program's research products?
3. To what extent are the program results being used by environmental decision makers to inform decisions and achieve results?

Elements to include for Long-Term Goal 1:

The appropriateness, quality, and use of ORD science by Program and Regional Offices, ORD partners, and other organizations to characterize or reduce uncertainty in risk assessment by incorporating mode of action information and/or by taking a systems biology approach to model the dose to effect continuum and to enhance predictive toxicology.

Elements to include for Long-Term Goal 2:

The appropriateness, quality, and use of ORD science by Program and Regional Offices, ORD partners, and other organizations to accurately measure and assess the risks associated with complex exposures to individuals, populations and communities and relate these exposures to internal dose.

Elements to include for Long-Term Goal 3:

The appropriateness, quality, and use of ORD science by Program and Regional Offices, ORD partners, and other organizations to characterize susceptibility as a function of life stage and thereby contribute to protecting the health and well being of children and older Americans. The extent to which Human Health research informs activities of the Office of Children's Health Protection and the Office of Radiation and Indoor Air.

Elements to include for Long-Term Goal 4:

The extent to which "accountability" projects succeed in measuring exposures in communities at risk, before and after remediation, and relate those changes to indictors of public health impact. The degree to which research addresses gaps and needs identified in the EPA's Report on the Environment.

In developing the summary assessment for each LTG, the BOSC Human Health Subcommittee will assign a qualitative score that reflects the quality and significance of the research 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 scores should be in the form of the adjectives that are defined below 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 rating and the rationale for selecting a particular rating will be transparent. The rating may reflect considerations beyond the summary assessment questions, and will be explained in the narrative. 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 science being produced and the speed at which research result tools and methods are being produced. An exceptional rating also indicates that the program is addressing the right questions to achieve its 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 science or for the speed at which 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 scientific questions to meet their goals, and 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 questions being addressed are inappropriate or insufficient to meet the intended purpose. Questionable science is also a reason for rating 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.

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] Memorandum 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 Research Efficiency in the U.S. Environmental Protection Agency. National Research Council of the National Academies, The National Academies Press, Washington DC, 2008.

[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 II—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).

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 object of basic research, and some basic research investments may be justified 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 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 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.