

**Program Review Charge
Human Health Risk Assessment (HHRA) Subcommittee**

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

- Plan, implement, and strengthen the HHRA program;
- Compare the HHRA program with other programs designed to achieve similar outcomes both in other parts of EPA and in other federal agencies;
- Make ORD investment decisions over the next 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 and development programs such as those conducted by the Office of Management and Budget (OMB highlights the value of recommendations from independent expert panels in guidance to federal agencies^{1,2}).

2.0 Background Information

Independent expert review is used extensively in industry, federal agencies, Congressional committees, and academia. The National Academy of 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 very challenging⁴ when federal agencies conduct research to support regulatory decisions, and then rely on third parties⁵—such as state environmental agencies—to enforce the regulations and demonstrate environmental improvements. Typically, many years may be required for practical research applications to be developed and decades may be required for some research outcomes to be achieved in a measurable way.

Most of ORD's environmental research programs investigate complex environmental problems and processes – combining use-inspired basic research^{6,7} with applied research, and integrating several scientific disciplines across a conceptual framework⁸ that links research to environmental decisions or environmental outcomes. In multidisciplinary 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

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

In a comprehensive evaluation of science and research at EPA, the National Research Council⁹ 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 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 are applying research to strengthen environmental decisions, and to evaluate client feedback about the research. Conducting program evaluations at this interval enables assessment of: research progress, the scientific quality and decision-making value of the research, and whether research progress has resulted in short-term outcomes for specific clients.

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

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

Background

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

HHRA is a relatively new ORD program, commencing in fiscal year 2004 through consolidation and expansion of a number of health assessments and supporting efforts already underway in ORD's National Center for Environmental Assessment (NCEA). This consolidation was undertaken in order to foster a more integrated approach for resource allocation, prioritization and accountability of risk assessment within ORD. The HHRA program is distinct from, but linked to, other ORD programs such as the Human Health Research Program and the Drinking Water Research Program, etc., as its primary objective is the development of peer-reviewed health assessments and supporting methods, models and guidance rather than conducting research i.e. collection of laboratory and field data.

The principal clients for HHRA products are the EPA programs and regions, who request and receive qualitative and quantitative health assessment information on priority environmental

contaminants developed under the HHRA Program's three long-term goals:

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

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

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

Further details on components of the HHRA program are available in the HHRA MYP¹⁰.

Draft Charge

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

Program Relevance

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

Factors to consider: the degree assessments are driven by EPA priorities; the degree which this assessment program has had (or is likely to have) an impact on Agency decision-making; and the extent to which program scientists participate on or contribute to Agency workgroups and transfer products to program and regional customers.

Program Structure

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

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

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

Program Performance

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

Program Quality

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

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

Scientific Leadership

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

Factors to consider: the degree to which this program is identified as a leader in the field; the degree to which assessments and peer reviewed publications from this program are cited in Agency decisions and documents, other peer reviewed publications; the degree to which HHRA scientists serve/are asked to serve on national/international workgroups, officers in professional societies, publication boards; the degree to which HHRA scientists lead national/international collaborative efforts, organize national/international conferences/symposia, and are awarded for their contributions/leadership.

Coordination and Communication

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

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

Outcomes

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

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

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

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

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

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

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

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

Elements to include for Long-Term Goal 3: Criteria Air Pollutant Integrated Science Assessments [previously Air Quality Criteria Documents]:

The appropriateness, quality and use of HHRA's Integrated Science Assessments (formerly AQCDs) by EPA's Office of Air and Radiation in their review and developmental of national ambient air quality standards to protect human health and the environment.

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

- Exceptional: indicates that the program is meeting all and exceeding some of its goals, both in the quality of the health assessments* being produced, and the speed at which assessment tools and methods are being produced. . An exceptional rating also indicates that the program is addressing the right questions to achieve programmatic goals. The review should be specific as to which aspects of the program's performance have been exceptional.
- Exceeds Expectations: indicates that the program is meeting all of its goals. It addresses the appropriate scientific questions to meet its goals and the science is competent or better. . It exceeds expectations for either the high quality of the assessments or for the speed at which the work products are being produced and milestones met...
- Meets Expectations: indicates that the program is meeting most of its goals. Programs meet expectations in terms of addressing the appropriate health assessment questions to meet its goals, and that work products are being produced and milestones are being reached in a timely manner. The quality of the science being done is competent or better.
- Not Satisfactory: indicates that the program is failing to meet a substantial fraction of its goals, or if meeting them, that the achievement of milestones is significantly delayed, or that the health assessment questions being addressed are inappropriate or insufficient to meet the intended purpose. Questionable science is also a reason for evaluating a program as unsatisfactory for a particular long term goal. The review should be specific as to which aspects of a program's performance have been inadequate.

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

References

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

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

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

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

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

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

⁷ "Renewing the Compact between Science and Government," Stokes, D.E., in 1995 Forum Proceedings, Vannevar Bush 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).

¹⁰ Human Health Risk Assessment Multi-Year Plan September 2007

Appendix I

OSTP/OMB Research and Development Investment Criteria

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

I. Relevance

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

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

- A. Programs must have complete plans, with clear goals and priorities.** Programs must provide complete plans, which include explicit statements of: specific issues motivating the program; broad goals and more specific tasks meant to address the issues; priorities among goals and activities within the program; human and capital resources anticipated; and intended program outcomes, against which success may later be assessed.
- B. Programs must articulate the potential public benefits of the program.** Programs must identify potential benefits, including added benefits beyond those of any similar efforts that have been or are being funded by the government or others. R&D benefits may include technologies and methods that could provide new options in the future, if the landscape of today’s needs and capabilities changes dramatically. Some programs and sub-program units may be required to quantitatively estimate expected benefits, which would include metrics to permit meaningful comparisons among programs that promise similar benefits. While all programs should try to articulate potential benefits, OMB and OSTP recognize the difficulty in predicting the outcomes of basic research. Discovery is a legitimate 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.