December 12, 2006

Dr. George Gray
Assistant Administrator
Office of Research and Development
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Dr. Robert Kavlock
Director
National Center for Computational Toxicology
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dear Drs. Gray and Kavlock:

This is a letter report from the Board of Scientific Counselors (BOSC) reviewing the Computational Toxicology Research Program conducted by the National Center for Computational Toxicology (NCCT). The Computational Toxicology Subcommittee of the BOSC reviewed NCCT's progress and plans during a 2-day meeting on June 19-20, 2006, at the EPA facility in Research Triangle Park, North Carolina. The BOSC Subcommittee consists of George Daston (Chair), James Clark, Michael Clegg, Richard DiGiulio, Muiz Mumtaz, and John Quackenbush.

This is the second review of the NCCT. The first review of the Center was conducted in May 2005. The Subcommittee was very pleased with the progress that the NCCT has made towards its goals since that first review. The Center first became operational in February 2005; during the 16 months between its establishment and this review, the NCCT has made substantial progress in: (1) establishing priorities and goals; (2) making connections within and outside EPA to leverage the staff's considerable modeling expertise; (3) expanding its capabilities in informatics; and (4) significant contributions to research and decision-making throughout the Agency.

Many of the recommendations made by the BOSC during its first review have been acted on by NCCT. This includes expanding its capabilities in bioinformatics through the funding of two external centers and through staff hires, expansion of its technical approaches to even more programs within the

Agency, and the development of communities of practice (CoPs) throughout the EPA research community in chemoinformatics, biological modeling, and chemical prioritization. CoPs are cross-organizational groupings of experts who share an interest in a common technology.

The Subcommittee addressed a number of charge questions during its review, the responses to which provide a basis for comments on progress as well as specific recommendations.

Question 1: *What progress has been made in the last year in developing/maximizing connections and collaborations within ORD and the Agency, through communities of practice and other interactions? Are there notable examples of collaborations that have been established to increase the reach and effectiveness of NCCT? Are there additional collaboration opportunities NCCT should explore?*

During the review, the Subcommittee members heard reports on three active CoPs: (1) Chemoinformatics; (2) Biological Modeling; and (3) Chemical Prioritization. They also heard about one proposed CoP, Cumulative Risk.

All active CoPs have formal memberships and are chaired by NCCT staff. The Center also has observed active participation among numerous EPA laboratories and centers and several program offices. The Chemoinformatics and Chemical Prioritization CoPs already have demonstrated outreach to outside agencies, such as the National Institutes of Health (NIH), National Institute of Environmental Health Sciences (NIEHS), and National Toxicology Program (NTP). Some are working with or soliciting international and private sector collaboration. The CoPs have been effective in focusing on defining problems and suggesting solutions, agreeing on modeling approaches and database issues, and setting up forums and workshops for discussions. They will be responsible for leading a better coordinated effort within EPA and among agencies.

The Subcommittee believes that establishing a Cumulative Risk CoP is worthy of pursuit. Such a CoP would provide significant opportunities to define areas for improvement in risk assessment practices and could provide inventory tools and other benefits. NCCT should consider whether it would like to provide a facilitator role or leadership role in this area.

With regard to other opportunities for exploration, the Subcommittee suggests that NCCT seek broader program office input. Additionally, CoPs covering areas such as Mixtures, Cross-Species Extrapolation, Population/Systems Dynamic Models, and Multimedia Fate and Effects Modeling should be considered for either NCCT use or ORD's broader use.

Question 2: *How does the work of the new extramural bioinformatics centers complement the intramural program, and how should the outputs best be integrated into NCCT strategic direction?*

ORD funded two extramural Bioinformatics Centers, one at the University of North Carolina directed by Fred Wright and a second at the University of Medicine and Dentistry of New Jersey headed by William Welsh. The Centers are used to extend the

capabilities of the intramural program. Individually, the Bioinformatics Centers were viewed to be excellent choices, each providing expertise and resources largely complementary to each other and to the NCCT with little overlap. Although both Centers are just beginning their work with EPA, there is great opportunity for synergy in developing new approaches for the analysis of toxicogenomics data and the integration of diverse information necessary to place these data into an appropriate context. In addition, each Center has existing links to risk managers and risk management groups, providing additional potential avenues for outreach to link the research programs of the NCCT and the Centers to real problems.

Integration of the external Bioinformatics Centers and the programs within NCCT will occur following the hiring of one senior and one junior bioinformatics scientist. This may not represent sufficient personnel, however, to allow NCCT to fully support its overall mission. Much of NCCT's program is focused on development of predictive models using systems biology approaches. Although this is a laudable approach, it ultimately will be driven by the availability of high-quality, well-annotated data and their integration with a wide range of other information. This will require significant effort. Although there are efforts underway under the direction of various NCCT personnel to begin this process, a more integrated approach is needed.

Consequently, NCCT needs to develop a more comprehensive strategic plan for data collection, management, and integration through creation of databases that model the structure of the underlying information and its potential use. This will require a careful assessment of the capabilities extant in each center so that necessary components, as well as areas for future development, can be identified. Addressing these issues will provide the structured data needed by NCCT's Systems Modeling and Computational Chemistry groups.

It also was noted that there exists a need within the field for trained personnel in computational toxicology. In addition to the existing postdoctoral program, one feasible approach would be to institute a career development award similar to the NIH "K" awards that would provide mentored training and research to more senior personnel.

Question 3: *Although the intent is not to review individual research programs, do the research programs highlighted during this review offer the promise of increasing the use and effectiveness of computational methods in Agency research? Do the efforts fulfill the goal of leveraging the resources of NCCT to increase effectiveness?*

The long-term goals (LTGs) of the Computational Toxicology Research Program are to provide risk assessors with: (1) improved methods to understand the source-to-response continuum, (2) advanced hazard characterization tools for prioritization and screening, and (3) methods that enhance dose-response assessment and quantitative risk assessment. The research efforts that were highlighted as part of the review cover each of these LTGs, and have the potential to be broadly used within and outside the Agency. This included efforts in high-throughput screening (HTS), modeling of molecular interactions with biological targets, modeling of complex pharmacokinetic and pharmacodynamic behaviors of small molecules, and database development and management, among others. The portfolio provided a mix of short- and long-term deliverables. Many of the former stand a good chance for application within program offices or other parts of ORD within

months. The research programs included those from external institutions. The Subcommittee found that NCCT has effectively leveraged its limited resources.

One of the major aims of NCCT is to develop useful relational databases. This also presents a significant challenge in managing the information. The Center should develop a strategic plan for data integration and for constructing databases that should be considered as information models.

Question 4: *Because a large part of the mission of NCCT is to accelerate the use of computational tools in the mission of the Agency, please comment on:*

- ***Part A:*** *Do the proposed computational models have the potential to identify and reduce uncertainties associated with the risk assessment process?*

Yes, proposed computational models have the potential to identify and reduce uncertainties associated with risk assessment. Additional opportunities outside the mechanistic models (especially in biomarkers that indicate exposure but that are not immediately or directly linked to toxicological response) may exist to fulfill NCCT's mission.

- ***Part B:*** *Will these models be able to help identify susceptible populations and compare potential risks to those populations with risks to the general and less susceptible population?*

Ultimately, these and other models within NCCT and outside the Agency can help identify susceptible populations. Appropriately, models currently are being developed for use in computational toxicology. Within 3-5 years, some of these models likely will be sufficiently developed and validated to address susceptibility. "Susceptible populations" may be defined to include life stages, gender, race, socioeconomic group, species, and geographic distribution.

- ***Part C:*** *Is the coordination between model development and associated data collection sufficient to avoid problems with the models being either over- or under-determined?*

Overall, data collection appears appropriately coordinated with model development. It will be important to validate models based on genomic methodologies given the inherent constraints in sample sizes, and other challenges, with these approaches.

Question 5: *Please comment on the Computational Toxicology Implementation Plan, focusing on the NCCT and Science To Achieve Results (STAR) components. Does it set an achievable road map for accomplishing NCCT's major goals over the next 3 years, as described in "A Framework for Computational Toxicology Research Program"? Does it set realistic and relevant milestones, and clearly articulate projected program outputs that will result in environmental outcomes?*

The Implementation Plan consists of five research tracks that are intended to fulfill three long-term goals:

1. EPA risk assessors use improved methods and tools to better understand and describe linkages across the source-to-outcome paradigm;

2. EPA program offices use advanced hazard characterization tools to prioritize and screen chemicals for toxicological evaluation;

3. EPA risk assessors and regulators use new models based on the latest science to reduce uncertainties in dose-response assessment, cross-species extrapolation, and quantitative risk assessment.

The research tracks that will support these long term goals are: (1) development of data for advanced biological models; (2) information technologies development and application; (3) prioritization method development and application; (4) providing tools and system models for extrapolation across dose, life stage, and species; and (5) advanced computational toxicology approaches to improve cumulative risk predictions.

Each of the research areas is active. Tables 1 and 2 of the Center's Implementation Plan provide details of projects and the outputs/outcomes and expected impacts of the projects. NCCT has a core strength in modeling, and is expanding its expertise in informatics. The Center is leveraging its position by outreach to other EPA laboratories and programs via internal research funding and communities of practice, and externally via STAR grants, including the external bioinformatics centers. The addition of the informatics centers in particular strengthens NCCT's research in information technologies. This will be strengthened further through the hiring of NCCT staff with informatics expertise. The STAR grants greatly expand NCCT's capacities in the generation of high-information-content data sets that will be needed to support model development.

Some challenges remain that will need to be overcome in the areas of database development and management. More details are provided in our response to Question 2. This will be especially important in the development and demonstration of biological models derived from complex data sets. The Center is encouraged to do whatever it can, within the boundaries of the grant process, to foster coordination of efforts between the two external bioinformatics centers and NCCT's internal program.

The research has milestones with nearer term and longer term time horizons, which is appropriate. It is clear that chemoinformatics tools and prioritization tools are well underway and are likely to be applied by risk assessors and regulators within the next few years. The timelines are realistic and the milestones will provide practical tools and methods to program offices. In the shorter term, information databases such as DSSTox and prioritization models such as ToxCast will be important tools for the pesticides and toxic substances programs, and will demonstrate the utility of computational toxicology in an applied setting. In the longer term, biological models such as the virtual liver, will improve mechanistic understanding of toxicological response and provide support for mechanism-based risk assessment.

The BOSC recommends that the NCCT develop a more detailed work plan for the virtual liver model, and that this plan be more extensively reviewed by the Computational Toxicology Subcommittee during its next annual review.

Question 6: *Please comment on the progress made in the five major research track thematic areas of the Computational Toxicology Research Program, and whether the current/planned research will address the major goals in the framework. The Center has made staffing additions and initiated new research over the past year. Based on these changes, what is the Subcommittee's view of the depth and breadth of the areas selected for emphasis?*

The Subcommittee believes that the research program covers the range of thematic areas. Some areas, however, have deeper coverage than others. The areas of cumulative risk assessment and cross-species extrapolation are still under-represented, but given the state-of-the-science, it is appropriate to place limited emphasis and continue to leverage research outside the Agency in these areas for the next 3-5 years. The staffing additions in HTS, toxicogenomics, and biological modeling are all strong and have improved the strength and breadth of NCCT. The planned staff additions in bioinformatics will be critical to the continuing success of the Center. One of these additions should have strong skills in data management systems.

Question 7: *What evidence exists that NCCT is responsive to program office and regional needs?*

Most of the presentations addressed program office input in planning priorities and approaches.

Some projects formed to support program office issues, such as carbamate cumulative risk, DSSTox, and RefTox DB. The Subcommittee noted program office and regional office staff as co-principal investigators on various projects. The Implementation Plan references a role for the Computational Toxicology Implementation and Steering Committee (CTISC), which could be useful, if sustained.

Question 8: *Please comment on how effectively NCCT is communicating its research program to EPA program offices, regional offices, and other stakeholders to inform their environmental decision making.*

NCCT has components of both a research and service center—it both initiates and receives new ideas. For a young organization, NCCT has done very well in establishing communication with its collaborators, contractors, and some stakeholders. The establishment of CoPs and participation of internal clients is a good start to communication within the Agency. Also of note is NCCT's establishment of monthly videoconference presentations. Within the past year, NCCT has commendably given 21 presentations to various offices within EPA to raise awareness. Most of the other communication activities seemed to be investigator-initiated. Given that the Center plans to develop tools and methods that will be used by ORD and other EPA staff, NCCT should establish a regularly scheduled plan for communication and updates. This process will convey the sense that new ideas are welcomed by NCCT and allow the Center to accept ideas and be aware of the needs of the program offices, regional offices, and

stakeholders. The establishment of such a process will enhance the marketing of tools and methods developed by NCCT. One way to give Agency clients part ownership in the Center is to invite them to BOSC reviews, such as this one, and ask them to share how they are using NCCT's methods, tools, and information. The Subcommittee recommends that NCCT communicate with the Regional Risk Assessor's Office and seek its representation.

Question 9: *Is the current research program designed to achieve environmental outcomes? Please provide recommendations on how the NCCT can best measure these outcomes.*

The current program is designed to achieve environmental outcomes that are appropriate to the Agency. Potential measures to determine these outcomes include:

- ✧ Use of screening models for chemical prioritization.
- ✧ Validation and use of genomics-associated biomarkers in field studies.
- ✧ Use of computational models in the risk assessment process in the long term.
- ✧ Success of databases (DSSTox, pesticides) in cleaning up and organizing disparate databases and making them widely useful to environmental science and regulatory communities.
- ✧ Use of specific models (such as virtual liver, pyrethroid metabolism, macromolecular modeling, physiologically based pharmacokinetic (PBPK) models, steroidogenesis models, cumulative risk models, and so forth, by broader environmental science and risk assessment communities.

In conclusion, the Computational Toxicology Subcommittee of the BOSC believes that NCCT is making exceptional progress towards its mission. We are pleased to provide advice on this important Center and look forward to future opportunities to offer suggestions for improving the NCCT.

Sincerely,

James R. Clark
Chair, Board of Scientific Counselors