COMPUTATIONAL TOXICOLOGY SUBCOMMITTEE MEETING SUMMARY

U.S. Environmental Protection Agency (EPA)
Research Triangle Park, North Carolina
December 17–18, 2007

Monday, December 17, 2007

Welcome and Introductions

George P. Daston, Subcommittee Chair, Miami Valley Laboratories, The Procter & Gamble Company

The meeting was called to order at 1:06 p.m. by Dr. George Daston, Chair of the Computational Toxicology Subcommittee of the Board of Scientific Counselors (BOSC). He introduced himself and welcomed participants to the third face-to-face meeting of the Computational Toxicology Subcommittee. The Subcommittee was formed by the BOSC to provide advice to the Office of Research and Development's (ORD) Computational Toxicology Research Program and the newly created National Center for Computational Toxicology (NCCT). Presiding over the meeting was Ms. Lorelei Kowalski, Designated Federal Officer (DFO) for the BOSC Executive Committee and the Computational Toxicology Subcommittee. Dr. Robert Kavlock, Director of the NCCT, was in attendance for the entire meeting.

Dr. Daston explained that this Subcommittee was created as a standing subcommittee so that the members could become knowledgeable about the Center and its programs to provide ongoing input and useful advice rather than just a one-time judgment about the progress of the NCCT. He reminded the Subcommittee that it has a privileged role in commenting on the strategic direction of the NCCT and praised the Center for its tremendous progress over the past year. He noted that Dr. Kavlock structured the meeting to focus on five specific programs. Following the presentation on each program, there was a panel discussion between the Subcommittee members and the Center staff involved in that program. The panel discussions provided opportunities for the Subcommittee members to pose questions to Center staff about the programs under review. Dr. Daston explained that, following this meeting, the Subcommittee would develop a letter report that contains recommendations for the Center. He welcomed Dr. Cynthia Stokes to the Subcommittee and then asked the Subcommittee members to introduce themselves. A list of the members and other attendees is attached.

DFO Remarks

Ms. Lorelei Kowalski, DFO for the Subcommittee, Office of Research and Development (ORD), U.S. Environmental Protection Agency (EPA)

Ms. Kowalski thanked the Chair, the Subcommittee members, and others present for their attendance at the meeting, and welcomed Dr. Stokes to the Subcommittee (Dr. Stokes' biosketch

was distributed to meeting attendees). Ms. Kowalski then explained her role as the liaison between the Subcommittee and EPA and her responsibility for ensuring that the Subcommittee and its meetings comply with the requirements of the Federal Advisory Committee Act (FACA). By law, the meeting must be open to the public. Notes of the meeting presentations and discussions are being taken by a contractor, Beverly Campbell, and a summary of the meeting will be made available to the public on the BOSC Web Site following final approval by the Subcommittee Chair. Ms. Kowalski explained that the purpose of the meeting is to provide advice and to account for the NCCT's progress in the past year. The meeting also is designed for the Subcommittee to answer the charge questions and submit a letter report. The federal docket for this meeting can be found on the Web at http://www.regulations.gov, docket number EPA-HQ-ORD-2007-1148. Members of the audience were asked to be recognized by the Subcommittee Chair before speaking and to identify themselves for the record.

Ms. Kowalski confirmed that the Subcommittee members had received the required ethics training and filed the necessary financial disclosure forms. She asked that members notify her during the meeting should the discussions pose any potential conflict of interest. The previous Subcommittee meeting was held in June 2006 in Research Triangle Park (RTP). The letter report that resulted from last year's meeting was approved by the BOSC Executive Committee in October 2006 and submitted to ORD. A response to the report was prepared by ORD and submitted to the BOSC in January 2007. Both the BOSC letter report and ORD's response were provided in the meeting materials distributed several weeks ago. Ms. Kowalski also sent the public comments she received prior to the meeting to the Subcommittee members by e-mail on Friday, December 14. For members who did not receive those e-mails, a copy of the comments are available at the registration table. She noted that one of the individuals who asked to make a public comment will be present in person.

Ms. Kowalski reminded the members to complete their homework forms and submit them to her as soon as possible, preferably before the end of the meeting. She then asked those present to sign in at the registration desk if they had not already done so.

Dr. Daston reviewed the agenda for the afternoon, noting that it will include on overview presentation by Dr. Kavlock, followed by presentations on ToxCast (by Dr. David Dix), Information Management/Information Technology–Informatics (by Dr. Richard Judson), and Subcommittee working time.

National Center for Computational Toxicology (NCCT) Overview
Robert Kavlock, Director, NCCT, ORD, EPA

Dr. Kavlock thanked the Subcommittee members for their willingness to attend a meeting at this time of the year. He opened his presentation with a definition of computational toxicology and its purpose within EPA: "to integrate modern computing and information technology with molecular biology to improve Agency prioritization of data requirements and risk assessment of chemicals."

The meeting agenda includes this general overview of NCCT's activities and status, and five presentations on ToxCast, Information Management, Virtual Liver, Virtual Embryo, and Arsenic Biologically Based Dose Response (BBDR) Model. Each of the five 45-minute presentations

will be followed by 45 minutes of round table discussion that will allow interaction between Center staff and the Subcommittee members.

The Subcommittee has been tasked with addressing the following charge questions:

1. Does the scope and involvement of expertise in the project reflect activities consistent with the function of a Center?

2. Are the goals and milestones suitably described, ambitious, and innovative?

3. Are there significant gaps in the approach that can be pointed out at this point in the evolution of the project?

4. Does the work offer to significantly improve environmental health impacts and is the path toward regulatory acceptance and utilization apparent?

5. Have appropriate data management and analysis tools been incorporated into the project?

6. How would you assess the outreach to other groups?

Dr. Kavlock explained that NCCT is one of the four centers within EPA's ORD, which also includes three laboratories. He noted that 10 percent of all EPA staff members are employed by ORD. NCCT is the smallest of the seven ORD entities, with about 20 full-time equivalents (FTEs), and the National Health and Environmental Effects Research Laboratory (NHEERL) is the largest, with 680 FTEs.

NCCT is trying to integrate modern computing and information technology with molecular biology and bring them into the Agency's risk assessment framework. Dr. Kavlock noted that computational toxicology is about the digitization of data. The Agency has legacy data that are dispersed and not readily available. The Center is seeking new techniques to assess a large number of chemicals, a wide variety of targets, and at different levels of biological organization. Computational toxicology requires quantification of physiology, biochemical pathways and networks, and biology and use of data mining and management. NCCT is interested in using high throughput screening (HTS) assays and high content assays and using a systems biology approach to bring the data together to prioritize chemicals. This was the focus of the Center's first 2 years of effort. Now, NCCT is looking at applying the tools to the risk assessment process.

Dr. Kavlock presented a timeline of program development. The Initializing phase began in Fiscal Year (FY) 2002 with congressional redirection and the endocrine disrupting chemicals (EDCs) proof of concept efforts. The Building Foundation phase began in FY 2003 and included formation of the design team, development of the framework document, Science Advisory Board (SAB) and BOSC reviews, the RTP workshop, and release of the Science To Achieve Results (STAR) HTS Request for Applications (RFA). In FY 2004, the Implementing phase began. This included the formation of the Computational Toxicology Implementation Steering Committee (CTISC), expansion of the proof of concept work, and release of the STAR Systems Biology RFA. During the Institutionalizing phase, which began in FY 2005, the NCCT was formed, the prioritization initiative was implemented, the first face-to-face meeting of the

standing BOSC Computational Toxicology Subcommittee was conducted, and the ToxCast concept was initiated. In FY 2006, the Operating phase began, which included grants to the STAR Computational Toxicology Centers, development of the Implementation Plan (which was included in the meeting notebook), completion of the ToxCast design, and the second face-to-face meeting of the BOSC Computational Toxicology Subcommittee. The Growing phase, in FY 2007, included the launch of ToxCast, the hiring of Title 42 employees, the initiation of the Chemoinformatics, Virtual Liver, and Arsenic Model efforts, and the hosting of the EPA Science Forum. In FY 2008, the Impacting phase began. In this phase, the Center completed its staffing, the National Academy of Sciences (NAS) published its vision for toxicity testing in the 21st century, ToxCast entered the next phase, a second STAR Center will be awarded shortly, and the third face-to-face meeting of the BOSC Computational Toxicology Subcommittee was held. The Implementation Plan for 2009-2011 will be completed later this year and it will be designed to show EPA and the scientific community that the Center was a wise investment.

The NCCT annual budget is $7.9 million, based on the President's budget for FY 2008. Although the Center has not conducted a traditional bibliometric analysis of its publications, the publication rate averages 4 publications per senior investigator per year. A list of NCCT's publications was included in the meeting notebook.

Dr. Kavlock provided the Merriam-Webster dictionary definition of a center. He emphasized the second definition—a point, person, or thing that is most important or pivotal in relation to an indicated activity, interest, or condition; a source from which something originates; a faculty providing a place for a particular activity or service.

Dr. Kavlock noted that NCCT's budget has been protected from the reductions that have limited other programs. The Center staff has been innovative and tried things that probably would not have made it through the traditional planning process.

The efforts of the Center translate into three kinds of activities: service, education, and research. Service includes the participation of research originating largely outside of the Center (e.g., Arsenic BBDR) and coordinating the Communities of Practice (CoPs). With respect to education, the Center wants to create a new generation of environmental scientists and this must include postdoctoral fellows, predoctoral fellows, undergraduates, and interns. The Center wants to offer both traditional and Web-based specialty courses. The research activities largely originating as a result of NCCT include ToxCast, Aggregated Computational Toxicology Resource (ACToR), Virtual Liver, and Virtual Embryo.

Dr. Kavlock then summarized the major comments from the second letter report of the BOSC Computational Toxicology Subcommittee.

BOSC Recommendation: NCCT should establish a Cumulative Risk CoP and consider the formation of other CoPs (Cross-Species, Multi-Media Fate).
NCCT Response: The Center considered the formation of additional CoPs but has found that it is difficult to keep the individuals involved. NCCT does not have the personnel to lead all of these CoPs at the present time.

BOSC Recommendation: The STAR Bioinformatics Centers should be integrated with the Center's program.

NCCT Response: NCCT's interactions with the STAR Centers will be described in the meeting presentations.

BOSC Recommendation: 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.
NCCT Response: This will be addressed in the meeting presentations.

BOSC Recommendation: The Center should consider additional training mechanisms beyond postdoctoral fellowships, such as instituting a career development award similar to the National Institutes of Health (NIH) "K" awards.
NCCT Response: The Center Director proposed the institution of an award similar to the NIH "K" awards to the National Center for Environmental Research (NCER) and NCER replied that it is not in its statutory authority. NCCT plans to revisit this issue with NCER in the future.

BOSC Recommendation: NCCT should develop a more detailed work plan for the virtual liver model, and submit this plan to the Computational Toxicology Subcommittee for review during its next face-to-face meeting.
NCCT Response: This will be addressed in the meeting presentations.

BOSC Recommendation: The Center should consider hiring staff members with expertise in bioinformatics, and one of them should have strong skills in data management systems.
NCCT Response: The Center has hired individuals with the required expertise.

BOSC Recommendation: NCCT should establish a regularly scheduled plan for communication and updates.
NCCT Response: The Center has developed a plan that includes a Web site and seminars for communication. NCCT also is bringing in a communication specialist to help the Center improve communications.

BOSC Recommendation: The Center should invite the program offices to BOSC reviews and ask them to share how they are using NCCT's methods, tools, and information. NCCT also should communicate with and engage the regional risk assessors.
NCCT Response: The Center Director has met with the program offices and invited them to this BOSC Subcommittee meeting, but no one was available to attend.

Dr. Kavlock reported that there had been a number of new hires since the 2006 Subcommittee meeting as well as a few departures. The new hires include three Title 42 positions: Richard Judson, Bioinformatician; Imran Shah, Systems Biologist; and Thomas Knudsen, Systems Biologist. Drs. Judson and Shah joined the Center in August-September 2006, and Dr. Knudsen was hired in September 2007. The postdoctoral fellows employed by the Center include: David Reif, Vanderbilt Department of Genetics; John Wambaugh, Duke University Physics Department; Rocky Goldsmith, Duke University Chemistry Department; Fathi Elloumi, University of Tunisia (via NIH); Jason Pirone, North Carolina State University Department of Mathematics; and Anne Marie Petrozelli, Duke University. The Center team also includes Amar Singh, a Scientific Systems Analyst from Lockheed Martin. Two postdoctoral fellows and one graduate student have left the Center since the last BOSC review—Melissa Pasquinelli, Michael Breen, and Amber Goetz.

Dr. Kavlock stated that NCCT draws support from the recent NAS report on toxicity testing in the 21st century. This report envisions a new toxicity-testing system that relies mainly on understanding "toxicity pathways"—the cellular response pathways that can result in adverse health effects when sufficiently perturbed. Such a system would evaluate biologically significant alterations without relying on studies of whole animals. The key elements of the vision include some targeted testing in animals (for the foreseeable future), dose-response and extrapolation modeling, population-based and human exposure data, and collection of biomonitoring data. The report also emphasizes the importance of evaluating risk contexts—common decision-making scenarios—for which toxicity testing is being conducted. Some risk contexts require rapid screening of thousands of environmental agents, while others require highly refined dose-response modeling for an individual agent.

Dr. Kavlock was pleased that the NAS report endorsed much of what the Center is doing. EPA has signed a Memorandum of Understanding (MOU) with NIH's National Institute of Environmental Health Sciences (NIEHS)/National Toxicology Program (NTP) and NIH's National Human Genome Research Institute/NIH Chemical Genomics Center (NCGC) to determine how to move forward on HTS, toxicity pathway profiling, and biological interpretation of findings. This MOU brings together three very strong organizations to work on the future of toxicity testing.

EPA's annual Science Forum was held in RTP in 2007 and the focus of the forum was computational toxicology. The forum's 400 attendees were from 16 different countries and it was the largest meeting ever held on EPA's RTP campus. A mini-review of the state-of-the-science based on the presentations from the forum has been accepted for publication in the journal *Toxicological Sciences*.

NCCT helped organize the International Workshop on Uncertainty and Variability in Physiologically Based Pharmacokinetic (PBPK) Models, which was held October 31-November 2, 2006 in RTP. The workshop focused on model specification, calibration, and uncertainty. It was attended by 120 individuals from 9 countries and the outcomes of the workshop were reported in *Toxicological Sciences* (*Toxicological Sciences* 2007;99(2):395-402). The workshop identified best practices and the many challenges of using PBPK models.

Dr. Kavlock presented a slide that depicted international coordination via the Organisation for Economic Co-operation and Development (OECD). At the May 24, 2007 OECD/International Programme on Chemical Safety (IPCS) meeting, there was initial interest expressed in participating in data generation, sharing, and analysis among companies, academic institutions, member countries, and the Business and Industry Advisory Committee (BIAC). The Center is working on an agreement for collaboration work (chemicals, assays, endpoints, and data sharing) for Molecular Screening/ToxCast. Prior to the second OECD/IPCS meeting (late 2008 or early 2009), which will focus on data analysis, Phase II of ToxCast will have been launched and a white paper describing proof of concept will have been prepared.

Dr. Kavlock's next slide described the framework for the Mechanistic Indicators of Childhood Asthma (MICA) study being conducted by the Center. This project is exploring the use of genomics in conjunction with a variety of biomonitoring tools to address major research questions (e.g., Are ambient pollutant exposures reflected in clinical/biological markers of

exposure?) in a population of third and fourth grade children in Detroit.

Other activities within the NCCT include: (1) cumulative risk of pesticides (organophosphates, carbamates, pyrethroids, conazoles), (2) molecular docking of perfluorinates and pyrethroids, (3) Steroidogenesis Model for the "Small Fish Project," and (4) the Administrative Efficiencies Project and information management/information technology activities in ORD. Dr. Kavlock noted that seven staff members of the Center received nine ORD honor awards (medals) since the last BOSC review. He concluded his presentation by stating that the Center has continued to build out the "plane" that Paul Gilman proposed when he was the Assistant Administrator for Research and Development. He also noted that the program will undergo a more traditional BOSC program review in February 2009.

Dr. Daston asked Dr. Kavlock if the Center had been too successful in getting involved with other projects. Does NCCT have the staff to maintain this level of involvement? Dr. Kavlock responded that if NCCT does not reach out and create partnerships, it cannot accomplish its intended purpose as a center. Collaboration is encouraged; he acknowledged that it would be great if the Center had more staff to pursue collaborative projects.

Dr. John Quackenbush asked if the STAR Centers are coordinating with each other. Dr. Kavlock replied that NCCT planned to meet with the STAR Center investigators on a monthly basis. Dr. Quackenbush asked if these meetings involved all of the STAR Centers together or just one at a time. Dr. Kavlock answered that they had planned to meet with each Center individually.

ToxCast: Developing Predictive Signatures for Chemical Toxicity
Dr. David Dix, NCCT, ORD, EPA

Dr. Dix opened his presentation with some background on the ToxCast project. ToxCast is intended to address the Agency's chemical screening and prioritization needs. The goal of ToxCast is to use computational methods to build models to predict potential for toxicity of environmental chemicals. ToxCast is being developed to provide a means to efficiently and quantitatively prioritize the thousands of chemicals of concern to EPA for additional toxicological evaluation based on computational models using chemical descriptors and biological activity profiling. The screening approach is based on the experience of the pharmaceutical industry. It involves a comprehensive use of a broad range of current HTS technologies. The Center is using a phased approach to evaluate the utility of ToxCast and is committed to stakeholder involvement and publication of the data. Dr. Dix mentioned that the Chemical Prioritization CoP continues to function and it includes representatives from EPA, NTP, and NCGC.

The problem to be addressed by ToxCast is that there are too many chemicals to test, the cost of testing these chemicals in terms of dollars and animals is much too high, and there are not enough toxicity data on the chemicals. ToxCast will derive classifiers or signatures from hundreds of HTS, high content screening (HCS), and genomics assays to predict toxicity, and these toxicity predictions will be used for prioritizing further testing of environmental chemicals. These collections of assays result in predictions of cancer and non-cancer endpoints. The project is being implemented in phases and now is in the latter half of Phase I. This first phase involves signature development for about 300 pesticides based on known toxicity and using data from more than 400 assays at a cost of $20K per chemical. This phase will be completed in FY 2008.

Phase II, will involve the expansion and validation of ToxCast signatures with more than 1,000 additional chemicals using 300 assays at a cost of $15-20K per chemical. Phase III involves applying the toxicity prediction for thousands of chemicals using less than 300 assays at a cost of $10-12K per chemical. ToxCast is designed to be an affordable science-based system for categorizing chemicals. Confidence in the system will increase as the database grows, and potential mechanisms of action (MoAs) are identified. ToxCast will refine and reduce animal use for hazard identification and risk assessment.

In Phase I, the responses of more than 300 chemicals (mainly pesticides and other select chemicals) are being profiled. More than 400 HTS and HCS assays are being provided by nine extramural contracts capable of handling up to 10,000 chemicals over a 5-year period. An Interagency Agreement has been established with the NCGC to provide access to the novel technologies developed by the NIH Molecular Libraries Initiative. The second phase of ToxCast is projected to involve up to 1,000 additional chemicals representing broader chemical structure and use classes, in order to evaluate the predictive bioactivity signatures developed in Phase I.

To use the millions of data points stemming from ToxCast, it is necessary to have electronic information available in formats that will allow sophisticated data mining and analysis. The Toxicological Reference Database (ToxRefDB) provides a relational database of standard toxicity test results for pesticides, making it possible to link toxicity information with the HTS and genomic data of ToxCast. The creation and population of ToxRefDB has been a collaborative effort between NCCT and EPA's Office of Pesticide Programs (OPP), and initially ToxRefDB is being populated with 20 years of toxicology data from pesticide registration studies (OPP's Data Evaluation Records [DERs]). These records contain high quality, comparable data with significant quality control.

ToxRefDB has a complete data package for 308 unique chemicals, which includes subchronic, chronic, cancer, reproductive, and developmental toxicity data. The data are curated into a relational model with expert-developed controlled vocabularies. Comparable toxicity data from other Agency databases (e.g., High Production Volume Information System [HPVIS]), the NTP, European agencies, and other sources also are being captured in ToxRefDB. The Aggregated Computational Toxicology Resource (ACToR) will be used to manage the large-scale sets of ToxCast assay data. ACToR is comprised of several independent data repositories tied together to a common database of chemical structures and properties. ACToR includes information from multiple sources both within and external to EPA and its first use will be to provide a repository and context for data from the ToxCast program's *in vitro* biochemical and genomics assays. ToxMiner also will be used for predictive modeling. The toxicity predictions will be specific to test, target, and type.

There are 308 unique structures in the ToxCast_320 (Phase I), including 291 pesticide actives, 9 industrial chemicals, and 8 metabolites. Dr. Dix noted that 55 of the 73 chemicals proposed for Tier 1 testing in the Endocrine Disruptor Screening Program (EDSP) are included in Phase I.

Using ToxRefDB investigators can observe and calculate novel relationships between effects. The database formats toxicity data in a manner that is conducive to linking HTS and genomic data and it provides an unlimited means for looking at toxicity data by chemical(s), study type(s), effect(s), species, and dose.

Two-hundred sixty-eight (268) of the 308 chemicals in ToxCast_320 have been assessed in rat (liver, kidney, and lung) and 256 have been assessed in mouse (liver, kidney, and lung). Dr. Dix presented a figure that depicted rodent non-neoplastic and neoplastic pathology for ToxCast_320 chemicals. To explain the figure, he stated that for 135 chemicals, there were no non-proliferative effects in rat lung, but there were non-proliferative effects for 133 chemicals; 64 of the 133 caused proliferative effects in rat lung; and 28 of the 64 caused neoplasms in rat lung.

A table of chronic/subchronic endpoints for prediction derived from ToxRef was presented. The toxic effects were selected for relevance, potency, and power. Cancer and non-cancer effects and groups of effects were included. The target organs include liver, kidney, lung, mammary gland, testes, and thyroid. Expansion to include reproductive and developmental tests is underway.

Nine contracts, one Interagency Agreement, and one Memorandum of Understanding (MOU) provide for chemical procurement and hundreds of biochemical, cellular, tissue, and genomic assays. These mechanisms give the Center the capacity to screen up to 10,000 chemicals. Dr. Dix presented a table with some detail on the types and number of assays being run by these organizations as well as a table on the status of their efforts.

NovaScreen is running 240 HTS assays for ToxCast, including 30 cytochrome P450, 81 G-protein coupled receptor (GPCR), 22 ion channels, 28 kinases, 24 nuclear receptors, 19 phosphatases, 9 transporters, and 27 other enzymes assays. These are binding or inhibition assays and the sources are human, rodent, or other species. NovaScreen initially screened the ToxCast_320 at 25 μm and they are working on concentration-response follow-up.

Through the EPA-NCGC agreement, nuclear receptors for chemical agonism and antagonism are being screened. A titration-based approach that efficiently identifies biological activities in large chemical libraries is being used. Dr. Dix summarized the NCGC efforts as follows: Invitrogen assays, human targets, 1536 well format, and 1408 chemicals at 11 concentrations. He noted that this effort is a precursor to Phase II; all of the other organizations are running only the 320 chemicals.

Dr. Dix presented a diagram that illustrated the *in silico* affinity of ToxCast chemicals to *in vitro* assay targets. The ligands are the ToxCast_320, including sterioisomeric forms and the targets include approximately 300 unique structures (180 unique sequences; nuclear receptors, kinases, oxidoreductases, phosphatases, esterases, and mixed function oxidases [MFOs]; >90% human; co-crystallized [holo] with ligands in agonist, antagonist, substrate, and inhibitor forms).

There are two collections of HepG2 assays being run by Attagene—Cis-Factorial Biosensors and Trans-Factorial Biosensors. Attagene is looking at 67 different molecular targets (43 transcription factors and 27 nuclear receptors) and how they are modulating the reporter gene. The chemicals are initially screened at LC_{50}/LC_{10} (maximum of 50 μM). Attagene plans to rerun this assay in concentration response, which will be useful for EC_{50}s.

Dr. Stokes asked if data were taken at a single time point and Dr. Dix replied in the affirmative. There is an option to look for variation over time but it is not certain if there is enough funding at this point to do that. He thought that concentration response, however, would be more useful.

IVAL is conducting toxicogenomic profiling of hepatocyte-Kupffer cell co-cultures. The company is using a collagen-matrigel sandwich culture and testing human, mouse, and rat with nuclear receptor modulators. IVAL is looking at time points, replicates, and concentration response and is screening the ToxCast_320 at five concentrations (mRNA by microarray or miRNA by polymerase chain reaction [PCR] or microarray). Dr. Dix noted that there is collaboration with the Center's Virtual Liver Project—Illumina is helping with the microarrays.

Dr. Dix presented a graph that depicted the TNF alpha response at diluted lipopolysaccharide and Kupffer cell concentrations. Dr. Quackenbush asked if they had looked at non-hepatotoxic compounds in the pilot to see if useful signals are obtained. Dr. Dix replied that they had looked only to see if the Kupffer cells were responding. They do not expect to see measurable toxicity; they will continue to measure cytotoxicity.

Expression Analysis is doing rat arrays for 22,000 transcripts using customized chips (up to 1400 genes in 96 well format). These are individual or multiplexed PCR arrays ($\leq$ 48 transcripts in parallel). ACEA is running assays with A549 human lung cells at eight concentrations with 48-hour monitoring. IVAL's human primary cell co-culture (integrated multiple organ culture) assays are looking at cytotoxicity using a 96 well format with 6 wells forming a chamber in which media are shared across all 6 wells, allowing potential for biotransformation. IVAL plans to do concentration response format. Cellumen is conducting high content cytotoxicity assessments with 11 endpoints, multiple modes of toxicity, 3 time points, and 10 concentrations. The company is considering running the assay in the rat hepatocyte model that has been developed since Cellumen was awarded the contract. BioSeek is identifying biomarkers of cell function in complex human *in vitro* systems. BioSeek is using 8 different assays, 87 endpoints, and 4 concentrations. Two hundred nineteen of the ToxCast_320 chemicals were active in at least one of the BioSeek assays.

The Hamner Institutes for Health Sciences is conducting a two-part project for ToxCast—Part I is the assessment of rat target organ toxicity and Part II is the evaluation of dosimetry and exposure. Hamner Institutes is trying to create a PBPK model that will connect *in vitro* rat results to the exposures in humans that would be required to generate that tissue level in humans. The project involves organ slice cultures for rat liver, lung, and kidney and chemicals at five concentrations. Hamner Institutes is measuring cytotoxicity at 24 hours, conducting a standard dose-response analysis, and estimating EC_{20}s. High-throughput PBPK modeling of non-volatile organics and physiological parameters from literature or measurements are being used. Chemical-specific rate constants and partition coefficients are derived from animal studies or estimated using quantitative structure activity relationship (QSAR). Hamner Institutes is trying to develop a PBPK model to predict exposure conditions required for target tissue doses equal to EC_{20} values in tissue cytotoxicity experiments.

Dr. Dix stated that the deluge of data has started with more than 100,000 data points from ACEA, 22,720 from Attagene, 111,360 from BioSeek, 105,600 from Cellumen, 30,976 from NCGC, 115,200 from NovaScreen, and so on. He presented a ToxCast data matrix that identified the types of assay (chemical, HTS, HCS, genomics, toxicity) data generated for each of the ToxCast_320 chemicals.

Dr. Dix identified a number of limiting issues, including: solubility and volatility, metabolism and biotransformation, selecting the right assays and being able to afford them, ability to assess

and infer mechanisms of toxicity, and *in vitro* concentration-response relative to *in vivo* dose response. He noted that the last item is the most difficult point and it may not be as important for chemical prioritization.

ToxCast is moving into areas of predictive modeling, looking at direct molecular interactions, biological pathways, cellular processes, and tissue/organ endpoints (M-P-C-E). The ToxCast assay data are being organized by ontology. The ToxCast biological ontology data are used in ToxMiner predictive modeling. Each chemical will have a spectrum of activities from M-P-C-E nodes. Predictive classifiers will include features from multiple data levels. The goal is to find classifiers that accurately predict toxicity endpoints. To reach this goal, the project will use all available data—ToxRefDB; ToxCast HTS, HCS, and genomics; physicochemical properties and calculated properties; and data from ToxCast partners. The properties of an ideal classifier include: (1) low false negative and false positive rates, (2) inexpensive and easy to measure for new chemicals, and (3) biologically plausible and provide mechanistic insight.

The next steps for ToxCast are to: (1) complete the ToxRefDB data population for ToxCast Phase I, (2) capture and quality control all ToxCast Phase I assay data, (3) complete toxicity signature discovery for ToxCast Phase I, (4) prioritize chemicals for ToxCast Phase II, (5) add the human exposure component to ToxCast, (6) capture toxicology data for ToxCast Phase II, and (7) generate HTS/HCS/genomics data for ToxCast Phase II.

Dr. Dix identified a number of internal EPA partnerships associated with ToxCast. OPP is partnering on ToxRefDB, antimicrobials, and inerts; the Office of Pollution Prevention and Toxics (OPPT) is involved with high production volume (HPV), moderate production volume (MPV), and perfluorocarbons (PFCs); the Office of Water (OW) is partnering on chemical contaminant candidates; and ORD's NHEERL, National Exposure Research Laboratory (NERL), and National Center for Environmental Assessment (NCEA) are involved in the program. Dr. Dix hopes that the ToxCast partnerships will be amplified by the March 2008 EPA ToxCast workshop.

There are a number of external partnerships associated with ToxCast, including:

- ✧ NIH Chemical Genomics Center (MOU pending)
- ✧ NTP HTS Initiative (MOU pending)
- ✧ OECD Molecular Screening Program (second workshop in June 2008)
- ✧ Security and Prosperity Partnership: North American Coordination on Industrial Chemicals
- ✧ MOU with the Hamner Institutes for Health Sciences
- ✧ Cooperative Research and Development Agreement (CRADA) with L'Oreal (pending)
- ✧ CRADA with Illumina
- ✧ Material Transfer Agreement (MTA) with Invitrogen (pending)
- ✧ MOU with North Carolina Central University (NCCU) (pending)
- ✧ Society of Toxicology 2008 workshop (proposal for 2009)
- ✧ Public ToxCast meeting presenting data and ToxMiner predictions
- ✧ Public access to ToxRefDB and ToxCast Phase I data.

Dr. Dix concluded his presentation with the timeline for ToxCast Program development. The NCCT was formed in FY 2005 and the ToxCast concept was developed. In FY 2006, the ToxCast implementation phase began with hiring staff, project planning, procuring contracts, and

filling Title 42 positions. The operating phase began in FY 2007 when the contracts were awarded, the quality management plan was developed, and Phase I was initiated. The impacting phase began in FY 2008 and it will continue into the future. This phase includes ToxRefDB, making Phase I data public, identifying predictive classifiers for Phase I, EPA and OECD workshops, and expansion to Phase II. In closing, Dr. Dix provided the URL for the ToxCast Web site (http://www.epa.gov/ncct/toxcast).

Panel Discussion

Referring to the three-dimensional plot Dr. Dix used to depict the universe of chemicals of interest, Dr. Daston asked how multidimensional this needs to be and if those are the right three axes. Dr. Keith Houck replied that it is as good as any method for representing the diversity of chemicals, but there are many other methods that could be used. He added that LogP was one of the factors used in selecting the 320 Phase I chemicals, because the Center did not want the chemicals to be insoluble in water assays. Dr. Daston stated that, at some point, the program will have more data than it needs. When will you reach that point? He mentioned that the data for the 320 Phase I chemicals cost the program about $6 million. Dr. Dix responded that they hope to reduce the number of assays over time as they determine those assays that are providing the most useful information. By decreasing the number of assays, the cost per chemical should decline.

Dr. Daston mentioned REACH (Registration, Evaluation, Authorisation and Restriction of Chemical substances), a new European community regulation on chemicals and their safe use, which will require manufacturers and importers to gather information on the properties of their chemical substances and to register the information in a central database. He noted that there soon should be a robust set of toxicity data on HPV chemicals. This information could save EPA a lot of time. Is there a way to work with the European Chemicals Agency (ECA) to share this information? Dr. Dix replied that Dr. Kavlock went to Helsinki to meet with the ECA and there are plans to work with ECA and OECD in the future. He mentioned that the data captured in the REACH database may not be the quantitative data that EPA would like to obtain for ToxCast. He was not sure that it would give the program the dose response data that are needed. Dr. Dix also was not certain that the program would have the needed confidence in the data. Dr. Daston commented that, for REACH, data for one compound are being applied to a group of compounds (categories); perhaps ToxCast can assess the validity of that approach. At some point, it will be very important to be engaged with the program offices as this starts to become part of regulatory decisions.

Dr. Richard Di Giulio asked if all of the Phase I chemicals are pesticides and Dr. Dix replied that 290 of the 320 are pesticides. Dr. Di Giulio then asked if the Center was comfortable with that choice. He noted that pesticides are designed to have a biological effect but many chemicals in the environment are not; therefore, EPA may not want to base the model on bioactive chemical data. Dr. Dix agreed that was a valid point that will need to be addressed in Phase II. He added that there are some industrial chemicals included in Phase I, but in Phase II, the program should expand into other chemicals. Dr. Jim Clark stated that as the program moves toward chemicals that are environmentally dispersed, it will be more difficult because the impacts being examined will broaden. Dr. Dix acknowledged that it will be challenging, but if they can get the testing cost down to about $10K per chemical, then *in vitro* testing will be possible. Dr. Kavlock mentioned that there is a project at the Athens laboratory involving a metabolic simulator. Dr.

Dix said that the Center will think about microbial metabolism as well as metabolism within the mammalian system when bringing in the human exposure component to ToxCast.

Dr. Stokes asked if all of the 320 chemicals in Phase I are toxic. Are any inactive, nontoxic? Dr. Dix responded that the toxicity depends on the dose. The way the studies are being run, all 320 chemicals cause some effect (e.g., reduced body weight), but the majority do not cause serious effects. They are looking at dose response. Dr. Stokes commented that if the program is trying to create a signature, chemicals that do not cause an effect are needed too. Dr. Dix replied that the substantial detail in ToxRefDB allows the program to go the level needed (a level where 10-30% of the chemicals cause an effect) with good specificity. Dr. Stokes noted that, with some assays, time will not have as much effect but in cellular and tissue assays, time of measurement can be important. Understanding that there is a limited budget, she asked if the Center had considered taking measurements at several time points. Dr. Houck replied that there is some concern about limited time points for the genomic assays, but most of the other assays are using different time points. He mentioned that the biochemical assays were in equilibrium conditions. Dr. Dix commented that one assay uses one time point; the company already had determined the best time to take the measurement.

Dr. Moiz Mumtaz said that he is a proponent of modeling tools and as long as inherent qualifiers of the chemical can be defined, they are useful to risk assessment. He was pleased to hear that the program is looking at biological systems. Dr. Mumtaz asked if the Center had plans to increase the relevance of the work to risk assessment. He did not think NCCT was working on any projects to accomplish this. Dr. Dix responded that the Virtual Liver and Virtual Embryo projects relate to risk assessment. The Virtual Liver Project is expanding to a more detailed analysis of a small number of chemicals and will generate information that will be useful for risk assessment. There will be presentations on those two projects tomorrow.

Dr. Dix stated that the ToxCast Project has proceeded better and faster than the Center staff expected so the staff has begun to think about relevance to human exposures. Dr. Kavlock added that the results being generated by ToxCast will "feed" toxicologists for quite some time. He thinks they will find targets that previously were not thought to be targets. Dr. Jim Rabinowitz pointed out that there is an experience issue involved; an assay will become more than a screen with increased experience. Dr. Daston agreed that the ToxCast information will feed many projects, but he stressed the importance that each project remains focused on its own goal (i.e., prioritization). Therefore, assays that have high correlation with each other should not be done all the time, but could be useful some times. Dr. Daston asked when the Center will start to narrow down the list of chemicals for Phase II. Dr. Dix responded that Dr. Richard Judson has done a comparison for one type of data. This is not ready for presentation but most of the data will be in hand by April 2008 and the first round of predictions should be ready by the June 2008 OECD meeting. Dr. Dix wants to make the data and predictions available to the public by September 2008. Center staff members are working to close out the first round of predictions and plan to kick-off the new enterprise in FY 2009. A workshop is planned for March 2009 that will allow the Center to hear what other groups are doing with the data.

Dr. Quackenbush said he was impressed with the scope and depth of the data to be generated but he cautioned that it is important to capture the data systematically. Has the Center started to build a relational database to analyze the data? Dr. Dix replied that Dr. Judson has some slides in his presentation that will address this question. Dr. Quackenbush asked about the integration

of the ontologies being developed. Dr. Dix answered that the Center does not plan to create new ontologies. Dr. Quackenbush asked if the work will be useful to others. He was concerned that the gene expression profiling work was all in rats and hepatocyte cultures when not all of the chemicals may be hepatotoxic in rats. Dr. Dix commented that if he had to choose a rodent species, he would select rats. Dr. Quackenbush asked if they could use human cell lines. Dr. Dix replied that human cell lines offer their own challenges. Dr. Quackenbush asked about using non-hepatic cell lines. He cautioned against holding that as a gold standard, adding that much better secondary effects that are more dramatic than the primary effects seen in the assay may be observed.

Dr. Daston commented that the Center's goal was to initiate some projects that could yield results in the shorter term and one of those projects was ToxCast. He noted that the Center should be pleased with the program's progress.

Information Management/Data Management
Dr. Richard Judson, NCCT, ORD, EPA

Dr. Judson said that the computational domains within NCCT include: (1) chemical toxicity data mining (ToxCast, ToxRefDB, and DSSTox), (2) biological pathway modeling and organ modeling (Virtual Liver, development toxicology modeling), (3) computational chemistry (structure prediction, docking, QSAR), and (4) BBDR and PBPK modeling (arsenic modeling). He identified six chemical toxicity data mining projects:

- ACToR—Aggregated Computational Toxicology Resource
- DSSTox—Distributed Structure Searchable Database
- ORD Genomics Data Management
- ToxRefDB—Reference *In Vivo* Toxicology Database
- ToxCast Data Management Workflow
- ToxMiner—ToxCast Data Analysis Pipeline.

The Center staff members working on chemical toxicity data mining are Richard Judson, Ann Richard, Keith Houck, David Dix, Matt Martin, Imran Shah, and Tom Knudsen; contractor staff members include Fathi Elloumi, Tommy Cathey, Tom Transue, Marti Wolf, and Amar Singh.

The NCCT information technology (IT) structure is scalable and easy to upgrade. It includes resources at the ORD Computer Center, the EPA National Computer Center, and the University of Louisville.

With respect to software, NCCT chooses open source where appropriate. The Center has access to a wide range of database, modeling, and analysis software and the system environments include Linux and Windows, Java, Perl, R, MPI, Lisp, and Fortran.

The complete toxicity package includes acute, subchronic, chronic/cancer, developmental, reproductive, immunotox, neurotox, genotox, dermal, respiratory, nephrotox, endocrine, cardiotox, hepatotox, and ecotox. All of these are required for pesticide food use actives at a cost of millions of dollars per chemical. Five of these (acute, developmental, reproductive, genotox, and ecotox) are required for Toxic Substances Control Act (TSCA) chemicals. Only a small percentage of environmental chemicals are well characterized, primarily chemicals in the

Integrated Risk Information System (IRIS) and Toxics Release Inventory (TRI), and pesticide actives. There are many data gaps for widely-used environmental chemicals; less than 10 percent have data for most tests. ACToR is the one-stop shop for these data.

ACToR is aggregating the world's chemical structure, bioassay, and toxicology data for environmental chemicals. It includes data on more than 15 million chemicals from more than 150 sources, all of which are publicly available. ACToR will be used to manage all of the ToxCast data. The Center plans to release it to the public by June 2008 and there currently is a prototype on the EPA Intranet at http://134.67.216.45:22722/servlet/ActorPrototype13.

ACToR includes the chemical name and Chemical Abstracts Service (CAS) number, chemical structure, *in vitro* and cell-based assay data, and *in vivo* chemical toxicity data. It can be browsed by data source and searched by name, CAS, and structure. The assay types included in ACToR are:

- ✧ PhysicoChemical (logP, MW)
- ✧ Biochemical (Ki–ToxCast, PubChem)
- ✧ Cellular (cytotoxicity–ToxCast, PubChem)
- ✧ Tissue (tissue slice assay)
- ✧ *In vivo* toxicology
    - Tabular – primary (NTP, OPP Information Network)
    - Tabular – secondary (IRIS)
    - Summary calls (Scorecard, CalEPA)
    - Summary report via URL (INCHEM)
- ✧ Category (from OPPT, Health Canada)
- ✧ Regulation (TSCA, Federal Insecticide, Fungicide, and Rodenticide Act [FIFRA])
- ✧ Description (Inventory Usage Report [IUR] usage levels).

Many data collections have been loaded into ACToR and it aggregates data from global sources. To illustrate the types of information available in ACToR, Dr. Judson presented the chemical summary profile of propiconazole.

DSSTox provides access to standardized structure-annotated toxicity data files for use in modeling studies. The data are subjected to extensive quality assurance review procedures and file documentation. The published inventory exceeds 7,000 unique compounds in 11 data files, spans multiple toxicity endpoints and study areas, and includes multiple collaborative sources and link outs to source data. The DSSTox Structure-Browser allows text (CAS, name) and structure (exact, substructure, similar) searching. The identifiers ensure PubChem and ACToR compatibility. The DSSTox chemical quality assurance procedures include: chemical identification, structure annotation, internal consistency, and quality review.

Dr. Judson presented a slide of the search page on the DSSTox Structure Browser, noting that a user can search all files or selected files and conduct text or structure searching. The help feature links to supporting information and help resources. His next slide illustrated the integration of numerous data sources in DSSTox, such as IRIS, EPA HPVIS, NTP, Carcinogen Potency Database, PubChem, and ACToR.

NCCT is working with NHEERL to consolidate R&D genomics data. Many experiments have been conducted and more are being planned; there is a need to archive, annotate, and disseminate valuable datasets. Data are being moved into ArrayTrack, a public software tool with a wide array of capabilities for data analysis and pathway mapping of microarray data. The Genomics Catalog will track database content and all EPA staff members will have access to the data. ArrayTrack allows first order analysis of the data and offers tools that are simple to use. Users can merge datasets so that they can look across the data and do meta-analysis.

ToxRefDB is being populated with OPP's DERs, which contain 20 years of key primary toxicology data for industry submitted toxicology studies. It will include the complete data package for more than 300 chemicals. The data are high quality with significant quality control. It incorporates a relational model with expert-developed controlled vocabularies and can be integrated with other primary toxicology sources (NTP, OPPT, European agencies). This project is the first effort to capture, tabulate, and mine this unique resource.

Dr. Judson presented a sample DER and illustrated how the data are extracted from the document and captured in the database. The data captured include the chemical name, the purity of the chemical, the type of study, description of the study, and the results of the study. All of this information maps into specific fields in the relational database. The database captures data for the type of study, the target (e.g., liver, kidney, heart, lung), and the description of the study. Controlled vocabularies are used.

There are a number of customers that need the ACToR data, including:

- ToxCast—for "signature training."
- EPA's OPP—for data mining trends and patterns.
- Outside Regulatory Agencies—because combined data sets provide more power for retrospective analyses and modeling.
- Other Toxicology Database Developers—ACToR offers a successful model for capturing large-scale toxicology data sets.
- Outside Collaborators—extend ToxCast predictive toxicology efforts.

Dr. Judson then described the ToxCast data management workflow. The Center receives raw HTS/HCS data from eight vendors that are using approximately 400 different assays in many different formats and multiple plate layouts per vendor. The data need to be quality controlled, normalized, and filtered and then put into standard form for input to ACToR. The infrastructure needs for the project are three people for a period of 3 months.

Toxicology data, *in silico* assays (chemical descriptors, QSAR, bioaccumulation), HTS assays from vendors, and genomics all are entered into ACToR. A lot of the HTS data from the vendors are running through GeneData Screener, an open API commercial application widely used by pharmaceutical companies. All of the data flows into ACToR and it can be viewed as a large table of data. The Center is setting up a separate database schema to analyze endpoints. There are many different ways to analyze data, many different endpoints, and a lot of different

algorithms and assays that can be used. The schema allows a user to queue up the analysis he wants to do and then access the data to run the analysis. This runs 24-hours a day. This system uses ToxMiner as the analysis engine. The goal of the data analysis is to identify "classifiers" that accurately predict ToxRefDB endpoints. For example, if the chemical binds to receptor x and is given enzyme y and the assay turns blue, there is an 80 percent chance that the chemical will cause a carcinoma. The properties of an ideal classifier include: accurate (high sensitivity and specificity), easy to interpret (provides biological insight), and inexpensive and easy to measure for new chemicals. Dr. Judson's hope is that the easy to interpret ones will lead to hypotheses that will be helpful to the Virtual Liver and Virtual Embryo projects, allowing researchers to focus on the most interesting points in the process.

Because it took some time to generate the data and it did not make sense to wait until they were available, the Center began running some simulations. They started by looking at the structure of the data that ToxCast would provide. This was used to develop methods to integrate the data and evaluate them to identify classifiers. The simulations were based on the direct molecular interactions between the chemical and the receptor or enzyme. These interactions trigger the cascade of events that lead to some endpoint. In the simulation model, a chemical is defined by a spectrum of interactions. Dr. Judson pointed out that there may be dozens of pathways to reach the same endpoint. In a situation where multiple pathways contribute to a given endpoint, the model must determine how well the different algorithms work. Dr. Judson presented a slide of simulation results that evaluated classification algorithms. He noted that feature selection is used to reduce the noise from the large amount of data.

Much of the recent work has focused on pulling data together and getting analysis tools set up. ACToR is pulling data from hundreds of data sources, whereas DSSTox, ToxRefDB, and ArrayTrack are pulling data from much fewer sources. ACToR and DSSTox are chemical structure based and include *in vitro* and *in silico* assays. ACToR, DSSTox, and ToxRefDB include *in vivo* toxicology data. ACToR and ArrayTrack include genomics data. Dr. Judson's point was that everything goes into ACToR because the quality control level is low; whereas it is high for DSSTox and ToxRefDB. ACToR, ToxRefDB, and ArrayTrack are relational and all eventually will be searchable on the Internet (ToxRefDB will not be on the Internet but the data will be available on the Internet through ACToR and DSSTox).

Dr. Judson presented a diagram that depicted the many databases and applications in the NCCT application framework. Over the next year, the Center will develop the CompTox API layer where most of these databases and applications will "talk" to the layer; making access and analysis of all the data much easier.

There are a number of NCCT information management collaborators. For ToxRefDB, the collaborators include OPP, OPPT, RIVM-Netherlands, BfR (Federal Institute for Risk Assessment), and Environment Canada. For ACToR, the collaborators are PubChem, Environment Canada, OPP, OPPT, OW, and NCGC/NTP HTS. For DSSTox, the collaborators include IRIS, NTP, and OPPT. For Genomics, the collaborators include MicroArray Quality Control (MAQC) project, NHEERL, and the ORD Genomics Task Force. There also is collaboration with the STAR Bioinformatics Centers at the University of Medicine and Dentistry of New Jersey/Princeton and the University of North Carolina (UNC). Dr. Judson stated that there are regular meetings with the UNC Bioinformatics Center staff and a meeting is planned with the New Jersey Center to discuss the ToxCast and Virtual Liver projects.

Dr. Judson identified the following FY 2008 major goals for chemical toxicity data mining: (1) complete ToxRefDB population for ToxCast Phase I, (2) capture and quality control all ToxCast Phase I assay data, (3) complete toxicity signature discovery for ToxCast Phase I, (4) prioritize chemicals for ToxCast Phase II, (5) capture toxicology data for ToxCast Phase II, (6) complete ACToR Version 1 and release it to the public, (7) apply the DSSTox chemical information approach across EPA, and (8) link DSSTox and ACToR with other public data sources.

There are a number of challenges to be faced by the program, including:

- ✧ Successful computational modeling requires lots of data and those data are in many different places, different formats, and of different levels of quality. The data must be located, organized, and quality controlled for ease of use.

- ✧ Real biology is complex. The Center's models must capture enough complexity to be useful and model-building must be modular.

- ✧ Computational toxicology is a rapidly evolving field; therefore, systems and models must be flexible and success requires collaborations inside and outside EPA.

- ✧ The NCCT approach is new to EPA. The Center is "pushing the envelope" on computational approaches and NCCT's open source/public access philosophy conflicts with Agency IT policies.

Panel Discussion

Dr. Quackenbush thought the Center had made a lot of progress on this front since the last BOSC review. He is convinced that NCCT is moving in the right direction. He expressed some curiosity about the simulation model and asked how the Center deduced the structure of the predictive network.

Dr. Judson responded that they would use Bayesian logic and draw a connectivity diagram in which each node is an assay. There will be some connectivity for some assays. Dr. Quackenbush commented that the problem with a Bayesian approach is that for 8 nodes there are 10^{12} possible topologies. Learning the structure is a challenge. Dr. Judson said that it will not be necessary to learn the structure. If it is possible to see the top and bottom layers, then he thinks the Center can produce an assay that shows what is going on even if the events in between are unknown. Dr. Quackenbush replied that it certainly gives the Center a place to start. Dr. Ann Richard commented that the Center is looking at this exercise relative to how the Agency currently prioritizes chemicals. She noted that QSAR is based solely on chemical structure. The Program is trying to predict endpoints and use a pyramid concept to aggregate data and define endpoints. For DSSTox, the North Carolina Bioinformatics Center plans to take slices of the pyramid. From purely an information standpoint, the Center is looking at aggregating biology information with the chemical data to make it easier to predict the endpoint. It is a more heuristic viewpoint.

Dr. Quackenbush stated that the Center has done some nice work on classifiers. There are diverse sources of input data and it may make sense to combine them in creative ways. Dr. Judson said that would be the next step.

Dr. Clark noted that the Center has an extensive list of collaborators and it appears that all of the key players are involved. With respect to making ACToR available to the public, Dr. Clark asked if there is a group outside of EPA that would use the tool to do something that the Agency would not do. Dr. Judson responded that by making ACToR publicly available to the public, the Agency meets its own needs but also makes the data available to others. What is the down side of allowing others to use the data? Dr. Clark answered that some groups could selectively use the data to further their causes. Dr. Judson pointed out that all of the data already are available; the Center is just bringing them all together.

Dr. Clark asked about the stakeholders targeted by the database. Dr. Kavlock responded that regulated industry and the public are the stakeholders. The more transparent the Center makes the data, the more comfortable stakeholders will become with these complicated concepts. Dr. Richard commented that the Center is looking across Europe and the United States and trying to develop weight of evidence approaches. That is what ACToR provides. Dr. Clark asked if there were criteria for source studies. Dr. Richard replied that ACToR includes public data; only data with high quality control goes into ToxRefDB. Dr. Clark posed the question: If a study shows up in ACToR, can a user assume it is a good study? Dr. Richard answered that, at some point, the Center depends on public literature that is peer reviewed. NCCT would like to make quality control judgments about the data but the Center does not have the staff members to make those judgments. Therefore, users will have to make those judgments. Dr. Judson commented that the Center is pulling in study reports, but is not endorsing the studies or judging their quality. Dr. Clark asked if there were any restrictions to what is being included. Dr. Mumtaz noted that only peer reviewed data are used for risk assessment. Dr. Judson replied that the Center is gathering data from trusted sources; there are no plans to go to the literature for data until Phase II. For the Phase II chemicals, it is likely that the Center will collect any data it can find and run them through signatures. He added that the Phase II chemical selection and data review process will be done in partnership with the EPA program office. Dr. Richard stated that the Center is adding value through the quality control of chemical annotation. Not all databases report precisely what chemical was tested so addition of this information adds value.

Dr. Daston said that he sees a lot of value in making the information available and having multiple ways to search the database makes it much easier to find the information. As long as EPA makes clear the level of quality control on the data, he did not have any problem with including data from various sources. He noted that some users may find some beneficial discoveries using the data even though they are not subjected to rigorous quality control. Dr. Quackenbush agreed that having more data available could advance the field. Dr. Richard said that aggregating the data already has been helpful to the Center in assessing where the money has been spent. Dr. Quackenbush recognized the potential of collecting information that is not readily available in other places and pulling in signatures from papers that have not been published in a database.

Dr. Stokes asked about the customer service that the Center will provide after making the database available to the public. Can users enter additional data or new data before it is published? Dr. Judson replied that the Center does not have any plans for customer service or

for allowing users to enter data. Dr. Stokes suggested that the Center could publish guidelines so the researchers could dump data into the database. Dr. Richard commented that publishers are beginning to ask authors to fill out a form when they submit papers and the Food and Drug Administration (FDA) is starting to use forms for data entry. She added that the Center is trying to tap into some of these efforts. Dr. Kavlock noted that if users see the value of the data, it will build momentum to add to the database. He said that this is the first time OPP has had the data in DERs available online. The Center is suggesting to OPP that staff enter the data for new DERs directly into the database so that these new data are available online. The data could only be entered after the pesticide registration. Dr. Richard reported that OPP is finding the system very easy to use, and the staff thinks this information will be helpful for future risk assessments. Dr. Clark asked about metabolites and Dr. Richard responded that known pesticide metabolites have been incorporated; it may be possible to make inferences of families of compounds. She was not certain how this would play out in ToxCast. Dr. Kavlock commented that the Athens laboratory is extracting metabolites but that information is considered Confidential Business Information so it cannot be released to the public. Dr. Judson said that Center staff will try to predict metabolites as part of the modeling efforts. Dr. Clark noted that this will tie in nicely with the data on microbial metabolites.

Dr. Daston asked if ACToR can be searched by substructure. Dr. Judson replied in the affirmative but added that it is not ready for demonstration yet. Dr. Stokes said she was not clear on how the Center was getting the signatures. Dr. Judson responded that the Center is bringing together *in vitro* and *in vivo* data. There is a side schema with a program that builds the table out. Dr. Stokes asked if it is purely statistical and Dr. Judson confirmed that it was. The Center is looking at correlations; when things bubble to the top, the staff takes a closer look (e.g., upregulating of a particular nuclear receptor, which leads to upregulating of x, which leads to liver carcinoma). Dr. Stokes asked what happens if noise in the data makes it difficult to find anything. Dr. Judson said he thinks the staff will find many things even if they do not turn out to be valid. Dr. Richard commented that QSAR works to a point; this approach is a way to improve those decisions. Dr. Kavlock mentioned that researchers have asked the Center to run their chemicals through NCCT's system. The Center will do so as long as the researcher agrees that the data can be made public. Dr. Quackenbush stated that the researchers also could run the chemicals themselves and Dr. Kavlock responded that they could but the researchers find it beneficial to have EPA using their assay and showing that it works.

Dr. Daston commented that this may be an opportunity to recast what QSARs are all about. If the Center can increase our understanding of the relationships between molecular interactions, there may be value in developing QSARs to inform the guesses we make based on statistical calculations. The QSARs would provide weight of evidence and perhaps these efforts could help define the future focus of QSAR research. Dr. Judson agreed, stating that all of the data will be useful in helping the Center rule things out. Dr. Imran Shah confirmed that the Center staff will be better able to evaluate QSAR data. Dr. Judson noted that the Center will have to use structural information to inform how much extrapolation is possible from the 320 chemicals to the 11,000 chemicals. Dr. Richard said she is hoping that the biology will help the Center capture more islands of chemical space. The physical-chemical characteristics are extrinsic and cannot mirror the biology. The Center needs these biological metrics to augment the structural data.

Dr. Daston thanked the NCCT staff members for their presentations and the panel discussions. He stated that the remainder of the day would be devoted to Subcommittee working time.

Subcommittee Working Time

Computational Toxicology Subcommittee

Dr. Daston wanted to spend some time developing a plan to address the charge questions. Dr. Clark noted that there are six charge questions and six Subcommittee members. Each member could be assigned responsibility for drafting the response to a question. The following assignments were made:

- Dr. Jim Clark – Question #1
- Dr. Richard Di Giulio – Question #2
- Dr. Cynthia Stokes – Question #3
- Dr. Moiz Mumtaz – Question #4
- Dr. John Quackenbush – Question #5
- Dr. George Daston – Question #6

Dr. Daston stated that there will not be adequate time to draft the responses before the conclusion of the meeting; he suggested a working lunch on Tuesday to allow the Subcommittee some additional time to discuss the responses. Dr. Daston said he was counting on everyone to complete their assignments in a timely manner following the meeting. Dr. Di Giulio suggested going through the questions and sharing their initial thoughts. Dr. Daston agreed, and asked the members to share their initial responses based on the two projects that had been presented.

Question #1: Does the scope and involvement of expertise in the project reflect activities consistent with the function of a Center?

Dr. Mumtaz responded that he thought the Center had the expertise but commented that the people who have presented are researchers and not suppliers. He thought the Center would benefit from the involvement of other parts of the Agency (e.g., assessors). He was concerned that representatives from those groups were not present at the meeting to hear about the Center's planned research. Dr. Mumtaz stated that the support of other parts of the Agency will only come when they are involved with the Center. Dr. Daston responded that Dr. Kavlock invited them to attend the review but no one was available. Dr. Mumtaz suggested that the Subcommittee encourage the Center to pursue this involvement.

Dr. Clark thought the Center had the required expertise and is working with the right partners. Dr. Quackenbush agreed, stating that the Center had made substantial progress since the last Subcommittee meeting. Dr. Daston was impressed with ToxCast because it provides EPA with a logical scheme for prioritizing chemicals. ToxCast will serve an important function in the Agency.

Dr. Clark said that the Center indicated early on that it would focus on human health endpoints, but in today's presentations we are hearing about metabolites. He did not think the Center would be able to ignore environmental toxicology and fate. At some point the Center will have to integrate those data to stay on track.

Dr. Di Giulio commented that a researcher concerned about ecological endpoints could be informed by human health endpoints, but he agreed that the Center should include environmental toxicologists at some level as well as ecological data. Dr. Daston mentioned that the Center included ecological data in ACToR.

Question #2: Are the goals and milestones suitably described, ambitious, and innovative?

Dr. Di Giulio thought the goals and milestones were suitably described, ambitious, and innovative. Dr. Quackenbush agreed that the goals are well defined but he did not think the milestones were well defined. Dr. Clark mentioned that each presentation included some plans for the future and next steps. Dr. Quackenbush did not think the Center had thought enough about how to put all the data together. Dr. Di Giulio thought there was some disconnect between the two presentations—the projects did not seem well integrated.

Question #3: Are there significant gaps in the approach that can be pointed out at this point in the evolution of the project?

Dr. Stokes stated that the main goal of these projects is to pull together a huge amount of data; the secondary goal is to identify signatures. She did not know that running statistical analyses on the data would be sufficient to identify signatures. This did not seem to be a state-of-the-art approach.

Dr. Quackenbush commented that a database is fundamentally a model. The various contractors are generating data and the Center will pull it all together in a data collection. He cautioned that it will be difficult to make the links. There are serious design questions concerning how to link different phenotypic endpoints. Dr. Quackenbush did not think the Center staff had thought about all these questions. He mentioned that there is a difference between having a real database versus a Web-based database. If the Center wants to ask questions of the data, he thinks they will need a local copy. Dr. Stokes wondered what algorithms the Center will use to ask the questions. Even if the staff is not looking for mechanisms, this does not seem to be powerful enough to provide the answers being sought. Dr. Quackenbush agreed that the Center would have to go beyond statistical analyses. Dr. Stokes said they need hypotheses to test. Dr. Daston commented that the statistical analyses will suggest the hypotheses that should be tested. Dr. Quackenbush stated that the Center cannot develop the sophisticated statistical analyses until it develops some hypotheses. It is the chicken and egg dilemma. The Center should develop a list of questions to address (e.g., What key pathways are activated in hepatotoxicity?) and start to partition the data into pieces.

Dr. Mumtaz commented that the Center should construct the database to answer the questions. He agreed that NCCT needs a list of questions in the planning phase.

Question #4: Does the work offer to significantly improve environmental health impacts and is the path toward regulatory acceptance and utilization apparent?

Dr. Daston stated that both the DSSTox and ToxRefDB are in demand and being used. He was not certain about their regulatory acceptance but he did know that the program office was using them.

Question #5: Have appropriate data management and analysis tools been incorporated into the project?

Dr. Quackenbush thought the Center had a good strategy for each data type to feed into a central warehouse. The Center, however, needs to give some thought to the questions that will be asked of the data and the analysis tools that will be used. He acknowledged that the tools to be used relates to the questions that will be posed. Dr. Quackenbush said that the Center had accomplished a great deal in 18 months and probably will be able to do a lot more once the data are available. Dr. Quackenbush stated that he would like to hear tomorrow's presentations before finalizing his thoughts. The Bioinformatics Centers may communicate with EPA but it is not clear that they communicate with each other. He asked about the RFA that led to the Bioinformatics Centers. Ms. Kowalski indicated that the Subcommittee can request information about the RFA from NCCT.

Question #6: How would you assess the outreach to other groups in executing the project?

Dr. Daston commented that the Center's outreach, both inside and outside the Agency, is stellar. He was disappointed, however, that the program offices were not at the meeting. NCCT needs to develop a plan for how it will interact and use the efforts of the STAR Bioinformatics Centers. The Subcommittee may want to recommend that the Center take steps to institutionalize the interactions.

Dr. Daston thanked everyone for their input and recessed the meeting until Tuesday morning at 8:30 a.m.

Tuesday, December 18, 2007

Recap of Day 1
Dr. George Daston, Subcommittee Chair

Dr. Daston called the meeting to order at 8:40 a.m. He noted that the time for public comment was scheduled for 10:15-10:30 a.m., and indicated that a short break would be added at some point during the morning session. He then provided a recap of Monday's meeting.

It is clear that ToxCast was always intended to be an early deliverable of the Center and it is exciting that this project is on the verge of making significant contributions. The Subcommittee also heard yesterday about information management/information technology. NCCT appears to have made great progress on this front and it was good to see how the projects dovetail with one another. The Subcommittee was pleased to hear that OPP is using NCCT's database for accessing its own data. Dr. Daston then asked Dr. Kavlock to proceed with today's presentations.

Dr. Kavlock stated that yesterday focused on the early efforts of the Center; today will focus on NCCT's thoughts about how the Center may be able to implement, in the next 10-20 years, the thinking in the NAS report on toxicity testing in the 21st century. Today's presentations include the Virtual Liver Project, Virtual Embryo Project, and Arsenic BBDR Model.

Virtual Liver Project

Dr. Imran Shah, NCCT, ORD, EPA

Dr. Shah indicated that this project involves NCCT, NCEA, and others. The objective is to develop a biologically-based tissue model that predicts injury due to environmental chemicals across doses, chemicals, genders, life stages, and susceptible populations. This model would be designed to aid risk assessment. The Virtual Liver would help answer the following questions: What are the mechanisms of toxicity? Who is susceptible? At what dose is toxicity observed? The challenge is that extrapolating animal toxicity to humans is difficult. Biologically-based modeling can aid with this extrapolation. Dr. Shah mentioned that novel approaches, such as this one, were emphasized by the National Research Council (NRC) vision for toxicity testing.

The hypothesis of the Virtual Liver project is that normal and adverse tissue function can be modeled *in silico* as (i) a cellular network adapting to xenobiotic and nutrient gradients, (ii) cells responding to stimuli by integrating outputs of (iii) molecular network modules perturbed by endogenous/exogenous factors.

The criteria used for the proof of concept are:

1. EPA priorities for protecting public health
    i. Organ most relevant for regulation?
    ii. Target-organ toxicity?
    iii. Chemicals causing this type of injury?
2. Scientifically challenging problem.
3. Sufficient information for model development and evaluation.

Among the chemicals regulated by EPA, the liver is the most common site of earliest injury for non-cancer critical effects according to data from IRIS. EPA regulates pesticides on their liver toxicity because more pesticides show toxic effects in the liver than in any other organ. With respect to liver toxicity, cellular alteration and steatosis are the most common endpoints for subchronic 90-day exposure and adenoma and carcinoma are the most common for chronic 2-year exposure. Dr. Shah pointed out that non-genotoxic liver cancer is a frequent adverse effect.

A number of HPV chemicals cause non-genotoxic liver cancer in rodents. A subset of these chemicals shares a common early mechanism—nuclear receptor (NR) activation (constitutive androstane receptor [CAR], pregnane X receptor [PXR], peroxisome proliferator-activated receptor alpha [PPAR-α]). The relevant chemical classes are: pesticides (pyrethroids), fungicides (triazoles), flame retardants (polybrominated biphenyl ethers), and persistent toxics (polychlorinated biphenyls, polyfluoroalkylacrylate). Dr. Shah commented that the Center is working to narrow down the doses of chemicals that lead to toxic effects.

The focus of the Virtual Liver Project is NR-mediated non-genotoxic liver cancer. Chronic stimulation of these relevant chemical classes leads to NR-activation, which leads to xenobiotic metabolism and cell proliferation, followed by hyperplasia, preneoplastic nodules, and adenoma/carcinoma. Dr. Shah pointed out that the mechanisms are unknown and the data are gleaned from a large number of publications and datasets.

Dr. Shah described the 5-year research plan for the project. In the short-term (1 to 2 years), the Center will gather knowledge on the MoAs, gather relevant data on selected chemicals for model evaluation, model NR-mediated molecular network, and develop the initial physiological tissue model. In the long-term (3 to 5 years), the Center will expand mechanistic knowledge to model cell proliferation and develop a multiscale tissue model to predict hyperplasia.

Dr. Shah explained that it is complicated to determine how exposure leads to toxic effects. Another challenge is taking the qualitative descriptions that are available and translating those into a quantitative model. The first component of the strategy focuses on the liver knowledgebase (KB) and includes development of a qualitative model of MoA (for species), integration of complex physiologic tissue information, and development of logic-based semantic representation. The second component of the strategy is virtual tissue simulation, which includes a quantitative dynamic model of MoA, model construction using KB, and dynamic simulation at multiple scales for physiologic outcomes.

For KB development, the Center must identify tissue components and behavior. The entities include molecules, organelles, cells, tissue, and organs. The types of molecules include proteins, genes, metabolites, and chemicals; the types of organelles include mitrochondria, lysosomes, and microsomes; the types of cells include hepatocyte, Kupffer, endothelial, stellate, and NK; the types of tissue include liver lobule, vasculature, and cell population; and the types of organs include liver and lobe. The behaviors include networks, pathways, interactions, biological processes, cell-cell interactions, cell fate, and normal and abnormal.

Dr. Shah presented a graph of events/processes at multiple scales for the MoA model. This graph proceeded from chemicals (e.g., pyrethroids, PCBs) to molecular perturbations (e.g., xenobiotic metabolism, NR activation) to cellular perturbations (e.g., macrophage activation, mitrochondrial alterations) to tissue injury (e.g., hyperplasia, carcinoma). It showed the changes that occur following chemical exposure. Dr. Shah then provided a diagram that depicted the translation of MoA to a qualitative model for the molecular network. This diagram included input (xenobiotic), mechanism (NR signaling), and output (active NR-complex). A second diagram for the intracellular network progressed from input (xenobiotic) to mechanism (xenobiotic metabolism, oxidative stress, NR signaling, MAPK signaling) to output (survival, apoptosis, proliferation).

The next slide presented by Dr. Shah indicated the complexity of the ontology. NCCT has talked to staff at the National Center for Biomedical Ontology (NCBO), which is a consortium of leading biologists, clinicians, informaticians, and ontologists who develop innovative technology and methods that allow scientists to create, disseminate, and manage biomedical information and knowledge in machine-processable form. NCBO's vision is that all biomedical knowledge and data are disseminated on the Internet using principled ontologies, such that the knowledge and data are semantically interoperable and useful for furthering biomedical science and clinical care.

The Center will work with NCBO to use what has been developed to represent relevant physiological entities and properties. NCCT also will use the open source tool Protégé.

The Center is using SIMILE's (Semantic Interoperability of Metadata and Information in unLike Environments) Welkin, which is an open source graph-based resource description framework (RDF) visualizer (http://simile.mit.edu/welkin). Dr. Shah used human PPAR-α as an example to illustrate the biological interactions.

NCCT is populating the knowledgebase with biomolecular entities and interactions. The Center's efforts include manual curation of information on organelles, cells, tissue, cell-cell communications, etc. NCCT is expanding the knowledgebase by retrieving information using relevant terms and curating useful facts from the literature manually. The Center also is collaborating with others doing text mining as a means of populating the knowledgebase.

The Liver Knowledgebase will provide users with tools for simple and complex queries. The Center plans to educate the programmatic users and then disseminate the knowledgebase for their use.

Dr. Stokes asked if this essentially yielded a large diagram and Dr. Shah confirmed that it did, adding that the visualization helps the Center see connections. It is a compatible representation computable on a symbolic level.

Dr. Shah presented a diagram that illustrated that cell injury is the result of dynamic processes. He explained that when a cell is exposed to toxic insult, if that insult is not severe, the cell may return to its normal state. If the insult is severe, the cell may return to baseline or go to a new trajectory that is slightly altered. If the injury is too severe for the cell to handle, it leads to cell death or necrosis.

There are several approaches for predicting tissue injury. The Center's goal is to build empirical models. Exposure to environmental chemicals leads to molecular perturbations that lead to cellular perturbations that lead to tissue injury. QSAR, molecular dynamics, QBAR omics, QBAR cytomics, molecular network dynamics, and cellular network dynamics are used to predict injury in this model. Dr. Shah noted that it would be much more effective (in terms of cost and time) to predict injury. The goal is to build a system that has three components—molecular networks, cell response, and tissue outcomes. The molecular networks are being built to bridge between what is observed in rodents and in humans. It should show how NR signaling and xenobiotic metabolism affect cell fate. The knowledgebase provides NR-perturbed signaling and genetic regulatory network. The Center has compiled information about interactions and changes in xenobiotic metabolizing enzymes (P-450s in Phase I, transferases in Phase II, and transporters in Phase III). Dr. Shah commented that most of the literature is focused on one marker for activation; the Center needs to expand and look at others. Three areas of future expansion include: (1) quantitative dose-dependent activation of NRs and xenobiotic metabolizing enzymes (XMEs), (2) network topology and function/stability, and (3) NR-signaling cross-talk with other networks.

Dr. Shah described one project with the goal of using a PBPK models to predict levels of *in vivo* tri-iodothyronine (T3)/tetra-iodothyronine (T4) in plasma. The liver XMEs play a role in regulating hormones (e.g., testosterone, thyroxine). Therefore, NR-activators can disrupt normal

hormone levels leading to reproductive and thyroid effects. A quantitative model of PXR-mediated CYP3A4 transcriptional regulation is being developed. This effort is looking at first order (mass action) kinetics, gathering *in vitro* data, and using parameter values from the literature (four unknowns).

Dr. Shah's next slide addressed Boolean networks. Probabilistic Boolean Networks (PBNs) model molecular interactions as logical operations and can use expression data. The Center is modeling NR-mediated perturbations discretely to understand dynamics. NR molecular interactions are expressed as a truth-table and system trajectory by successive rule application. Staff search for "attractors"/stable states and determine their biological relevance. Genomic data can be used to evaluate or extend models. Dr. Shah explained that the Center is using the network to try to understand complex molecular interactions that lead to outcomes. The topology of the network provides an understanding of the system. He noted that perturbations switch the cell from one attractor to another (e.g., division/death).

The key events that the Center is trying to include in the model are NR activation, xenobiotic metabolism, mitochondrial alterations, cell cycle, apoptosis, hepatocyte proliferation, and hyperplasia. NCCT is doing profiling to determine if the staff can look at the topology to build the molecular network models. NCCT and NHEERL have gene expression data on NR-activators. The Center also is initiating collaborations with industry that has large expression databases. NCCT has established industry and academic partnerships to analyze and generate metabolite profiles. HTS data are available from ToxCast (*in vitro* dose-dependent data on molecular and cellular perturbations).

Dr. Shah stated that cells respond to xenobiotic stress in discrete ways and result in different endpoints (e.g., necrosis, hyperplasia, carcinoma). To understand the dynamic balance between cell death and division, the Center is modeling hepatocyte state using data on key events in the MoA. In the long term, NCCT will model cell state with dynamic molecular networks.

The Center is focusing on key cell-cell interactions among hepatocytes, Kupffer cells, endothelial cells, stellate cells, and pit cells as a starting point. Specific data on hepatocytes at the cellular level are required to calibrate the model. The project staff is collaborating with ToxCast and external groups to develop KC-Hep co-culture for rats and humans.

The hepatic lobule has a heterogeneous structure with five cell types organized in a network around sinusoids. The zones are functionally different (different physiology and different metabolic processes) and injury can vary depending on the zone. One approach for looking at the biology is to model the overall gradient.

Agent-based tissue modeling requirements include:

- ✧ Qualitative rules and quantitative equations.
- ✧ Heterogeneous spatial information: zones, cell density, etc.
- ✧ Cell networks in nutrient/xenobiotic gradients; cell death/division or movement.
- ✧ Molecular interaction networks in cells.

Dr. Shah noted that the Center is not the first group to develop a liver model; he mentioned the work by Yan and Hunt on agent-based liver PK modeling and Anderson's work on agent-based

cancer modeling. Anderson is looking at molecular properties and asking questions about what genes are targets for therapy. Dr. Shah stated that the Center wants to be able to calibrate the molecular models and cell-based models and build the models in a way that mimics what the staff expects to observe.

Dr. Shah presented a timeline for the project. Prior to Year 2, the liver knowledgebase will be deployed, *in vitro* data gathered, the NR model developed, the Hep-KC model developed, and the initial tissue model developed. In Year 2, the liver knowledgebase will be expanded and the molecular network model developed. In Year 3, *in vivo* data will be gathered and the cell-state model will be developed. In Year 4, the multiscale tissue model will be developed.

The Virtual Liver Project team is a multidisciplinary team from across EPA and ORD and external organizations. The team includes NCCT, Office of Prevention, Pesticides, and Toxic Substances (OPPTS), NCEA, NERL, NHEERL, the STAR Bioinformatics Centers, NCBO, Physiome, HepatoSys, National Institute of Environmental Health Sciences (NIEHS), Department of Energy (DOE), and the Hamner Institutes for Health Sciences.

Dr. Shah stated that it is difficult to extrapolate chemical-induced animal toxicity to estimate human risk. Biologically-based virtual tissues can extrapolate injury between species, chemicals, and doses. Liver toxicity is the most frequent critical effect in rodents and NR-mediated non-genotoxic cancer challenge. MoAs can be modeled using the Virtual Liver. Dr. Shah concluded by stating that if the project is successful, it can be applied broadly for predicting target-organ toxicity.

Panel Discussion

Dr. Clark said that he was overwhelmed with the complexity of the project. Development of a product that will model all of the processes in the liver would be amazing. Dr. Shah commented that the Center will select one MoA and one toxicity endpoint before doing two or three to determine whether this is possible. Dr. Rory Conolly stated that the Center is attempting to define a framework for liver biology and how liver is perturbed by stressors in the environment. The framework should incorporate the information in a logical manner.

Dr. Daston commented that Dr. Shah has described a great framework but the starting point is not clear. Where will NCCT start in terms of the many processes and how will the Center determine if it is following the ultimate process? Dr. Shah responded that it is very complicated but the plan is to take a stepwise approach looking at NR-activation. Staff members will look at specific genes and consider what they expect to observe. Dr. Daston commented that XMEs may not be driving hyperplasia. Dr. Shah replied that the staff cannot ignore them just because some researchers hold this opinion; he noted that XMEs may help the Center understand other things so it is not a wasted effort. An unbiased narrow view of metabolites and proteins could guide the staff to alternative models. He noted that there are reports of relationships between regulators and MAP kinases.

One Center staff member stated that a number of PBPK models focus on XMEs and use rat *in vivo* data. If the model can describe XME induction in the rat then that could be used for a scaling factor for humans. Some of the ToxCast work would come into play here. Dr. Daston pointed out that if an XME is not driving the response then the scaling factor may be incorrect.

Dr. Conolly mentioned the links between signaling biology. The question is whether there is an apparent empirical correlation between the xenobiotic enzyme and the outcome. Dr. Hugh Barton said that Dr. Daston's point is well taken. He sees the liver as the key interface; PBPK modeling is further along so it is easier for the staff to describe what will be done on that front. He mentioned a new paper by Frank Gonzalez that demonstrated that the PPAR-α humanized mouse model provides an *in vivo* platform to investigate the species difference mediated by PPAR-α. The paper concluded that this also is an ideal model for human risk assessment peroxisome proliferators exposure. Dr. Barton said that the Center is cognizant of Dr. Daston's concern; therefore, the staff will have to model other paths. He explained that there was little detail on other paths in the presentation because that work is newer to the Center and the staff has not developed the detail yet.

Dr. Daston commented that it will be valuable to have the knowledgebase and the sophisticated model, but the Center needs to figure out how to use these tools to quickly test new hypotheses and discard them if they are not appropriate.

Dr. Stokes said she was impressed by the breadth of thinking that has gone into this project. To move from the molecular level all the way to the tissue level is a huge undertaking. She was glad to see that the project was focused initially on one area. Dr. Stokes then stated that she did not see how the cell injury slides fit with the first part of the presentation. Dr. Shah responded that the key change for causing carcinogenicity is a change in cell state; therefore, an understanding at the cell level is key to understanding proliferation. The Center is implementing the project in stages. The hope is to develop a two-stage model to understand how cells get perturbed. Dr. Stokes encouraged the staff to bring in the big picture modeling early. This would give the connection, from both pharmacodynamic and biological standpoints, to the animal endpoint earlier. Dr. Shah replied that when the project began, the staff thought to focus on the molecular scale. The Center has recruited a post-doc to work on the tissue level model beginning in January 2008. He mentioned that Physiome Sciences took an interesting approach, but could not show tissue level physiological changes—the endpoint was missing. The Center staff realizes the tissue scale work must be initiated sooner to get to the endpoint earlier. The project is working from the middle out. Dr. Stokes commented that from the modeling standpoint, it is from end to start. What happens once the cell decides to grow or die is significant. Dr. Shah commented that the differences between species are important. Dr. Stokes stressed the importance of keeping a very specific endpoint in mind.

Dr. Quackenbush thought the Center had done a good job of laying out a broad vision; he was glad to see that the staff was focusing on one type of mechanism. He expressed some concern about how NCCT was going to link all of the models together. Dr. Shah said the staff has been giving that considerable thought, adding that Physiome Sciences and others have been thinking about this for some time. He noted that people usually focus on the physiological or on the cellular but not in between.

Dr. Quackenbush applauded NCCT for attempting to put together the knowledgebase, but he was concerned that it might consume all of the project's resources. What is the fall back position? Dr. Shah responded that the fall back position would be to not conduct the work in house but to stimulate opportunities for academics to contribute. The Center is interested in identifying the relevant papers and concepts, but staff time is limited. Dr. Quackenbush mentioned that there is a CRADA with Ingenuity that has focused on this issue. Dr. Shah replied that the Center is

working with Ingenuity as well as others to see if NCCT can leverage commercial resources. If a knowledgebase exists, the Center will use it.

Dr. Mumtaz said it was an excellent presentation. He was pleased to see that the Center is working with other collaborators. He asked if the model is being developed for use by risk assessors. If so, he cautioned that risk assessors often are confused by simple models and this one is quite sophisticated.

Public Comment

At 10:15 a.m., Dr. Daston called for public comments. He asked that the comments be limited to 5 minutes each. He noted that public comments are useful to the Subcommittee but the members are not obligated to respond to them.

Dr. Stan Young from the National Institute of Statistical Sciences commented on reproducibility and credibility. He referred to a paper authored by Dr. John Ioannidis that was published in the *Journal of the American Medical Association* in 2005. This paper points to a real problem: observational studies do not replicate. Five out of six highly cited claims from observational studies failed to replicate. Dr. Young noted that the 80 percent failure to replicate figure is holding up very well.

Dr. Young made the following points: (1) ask many questions and you will receive p-values less than 0.05; (2) missing middle, five groups ignore the middle three; (3) crazy corrections, 20 confounders (pick your favorite 3); (4) fun with ratios; (5) up or down? You pick, I can do either; (6) subgroup selection bias, and (7) put it all together—always $p<0.05$. Averages can hide important variation. Dr. Young used the example of ozone regulation to prove his point. EPA is regulating ozone the same across the entire United States even though there are significant variations in the percent rise in mortality per 10 ppb rise in 24-hour ozone across the country.

Dr. Young was concerned that investigators are cheating on their results to get p values suitable for publication. To address this problem, he proposed the following: (1) acknowledge the problem, (2) share data sets, (3) training and holdout set, (4) multiplicity adjusted p-values, (5) demand replication, and (6) criticize poor statistical practice. Dr. Young closed his comments with the following questions: Is the data set available? Under what conditions? How many questions are posted on the data set? Is there multiple testing adjustment? He stated that if the Risk Ratio is < 2.0, ignore it.

Dr. John Dale Dunn submitted the following comment. He is a civilian emergency medicine faculty member at Carl R. Darnall Army Medical Center, Fort Hood, Texas. He has been a physician for 35 years and a non-practicing attorney for 28 years. He submitted for the Subcommittee's consideration the chapters on epidemiology and toxicology of the Federal Judicial Center's Reference *Manual on Scientific Evidence* (2nd Edition, 2000, West Publishing). The chapter on epidemiology was written by Michael Green, D. Mical Freedman, and Leon Gordis, and it is unequivocal on the relative risk proof causation in observational population studies.

The chapter on toxicology by Bernard Goldstein and Mary Sue Henifin and the *amicus curiae* brief on behalf of 30 internationally prominent and distinguished scientists both discuss the rules

of threshold in toxicology. Neither source supports the EPA use of liner modeling for low dose toxic claims from high dose toxin experiments. He posed the question: What is the BOSC or the Computational Toxicology Subcommittee or any other BOSC subcommittee doing to keep the EPA from violating the rules on relative risk and linear modeling? EPA toxicology looks to dose response to save small effects research. Where is biological plausibility when thresholds are discarded? Does this new small effects dose/response doctrine encourage EPA-sponsored researchers to data dredge?

These toxicology research problems are compounded by multiple testing at some EPA and NIH funded research mills that churn through surveys and cohort information for new toxic effects. Journal editors and the Academy are complicit in accepting this new scientific canon of small effects, linear modeling, one-hit toxicology, and multiple testing without adjustment. The Subcommittee knows the problems. Currently, EPA is embarked on pseudo heroics, chasing phantom toxic effects for what benefit and at what expense to the public.

Failures to maintain scientific integrity in EPA can be traced to size, budget, and influence of the Agency. No researcher is rewarded for dissenting from the EPA hymnal on toxicology, even when there are major scientific integrity issues in epidemiology and toxicology. Dr. Dunn thought that perhaps a federal judge will change that someday soon by opening up the Reference Manual.

The BOSC should commence a Karl Popper concept of science reassessment of EPA toxicology. For the last 2 years, there has been no public comment in the BOSC meetings. The BOSC mission is assuring scientific integrity at EPA. Vigorous debate and skeptical inquiry are the basis for scientific integrity. The BOSC and its subcommittees must focus on the mission.

Dr. Daston thought that these two comments were better suited to a different discussion than the one that is taking place at this meeting. He agreed to share these comments with the BOSC Executive Committee. He then asked if there were any additional public comments and there were none so the panel discussion resumed.

Panel Discussion (Continued)

Dr. Barton said that the Virtual Liver Project was intended to be a visionary project—it was not directly applicable to risk assessment. On the other hand, the project has been designed so that it can have an impact. EPA's current processes for criteria pollutants, IRIS, etc., are designed to create hardcopy documents. The Agency needs a different paradigm to create knowledgebases that can be analyzed by many people and in different ways. He hopes that this project will help the Agency shift its thinking toward the electronic world rather than the publication world.

Dr. Barton mentioned Dr. Dix's statement yesterday that the pharmaceutical industry has about 500 chemicals for which it has animal and human data; the Center does not have human data on the ToxCast 300 chemicals. The NAS report was clear that EPA needs to be able to predict human and not animal toxicity. He mentioned a presentation at a meeting on evidence-based toxicology. There was a wonderful talk on a Pfizer antibiotic. One in 14,000 people ended up with serious liver injury to the point of death or transplant. This is a 1 in 10^{-4} risk level. They regulated nuclear receptors so there is potential to make linkages between animals and humans,

but the real challenge is to make the comparison to human toxicology. It would be useful for the pharmaceutical companies to share more information, perhaps through ACToR.

Dr. Conolly commented that risk assessment has involved statistical modeling of data and the movement to biologically-based models to get a more accurate picture is just starting. This project is well ahead of what risk assessors currently are doing, but that is okay because the Agency wants to use robust science in its risk assessments. This difference, however, causes a tension between the risk assessors and those that work on these speculative models. As soon as the researchers think the model can be trusted, they want the risk assessors to use it, but there is skepticism among the risk assessors to trust the model. How do we get regulatory groups to work with Center staff and others to minimize this tension? Dr. Mumtaz suggested the integration of exposure and effects and to consider dose when modeling because it may help simplify the model.

Developmental Systems Biology: The Virtual Embryo
Dr. Tom Knudsen, NCCT, ORD, EPA

Dr. Knudsen explained that his presentation would cover the scope of the problem (developmental toxicity and EPA context, hypothesis, and research prototype), knowledge discovery (information infrastructure and analytical text mining), *in silico* morphogenesis (developmental ontologies and gene networks, hybrid cellular automata [CA] simulators, and case examples), and a summary of the project and potential collaborators.

Developmental toxicity is the potential of a chemical to cause adverse effects in the embryo or fetus and it is an important health consideration. A number of risk assessments used critical effects based on childhood or gender susceptibility (http://www.epa.gov/iris). Developmental endpoints include fetal weight reduction, malformations, resorptions, and functional deficits. Data usually are derived from experimental studies dosing pregnant animals and evaluating fetuses at term.

Developmental endpoints may be critically sensitive to low-dose exposure. Many chemicals have data gaps following early life exposure. Reproductive-development testing is the largest consumer of animal resources so it would be prudent to develop methods of predicting exposure outcomes. One place where NCCT could make a real difference is predicting toxicity after developmental exposure.

The goal of the Virtual Embryo Project (v-Embryo) is to build capacity for predictive modeling and regulatory analysis of developmental processes and toxicities. The hypothesis driving the research is that critical effects of environmental agents can be captured from computer simulations of normal development that structure: (1) the flow of molecular regulatory information, (2) cell-autonomous responses to genetic and environmental cues, and (3) emergent properties associated with collective cellular behaviors.

There are two approaches of computational techniques—knowledge discovery and computer simulation. Knowledge discovery involves data collection, data mining, and data analysis. The challenge is that framework data are highly dispersed in literature, databases, and paper records. Computer simulation involves representation, a computational model, and a visualization model.

The challenge is to develop a computational toolbox to integrate data into biological networks and model emergent properties of a developing system.

Dr. Knudsen described vertebrate eye development as the research prototype. The attributes are:

- ✧ Good literature base
- ✧ Human relevance
- ✧ Reciprocal tissue interactions
- ✧ Morphogenetic cell death
- ✧ Range of ocular phenotypes
- ✧ Often non-syndromic
- ✧ Genetic susceptibility
- ✧ Environmental sensitivity
- ✧ Developmental ontology
- ✧ Gene regulatory networks (how the control genes work together is unknown)
- ✧ Conservation of cell signaling
- ✧ Mitochondrial dependency (early eye development is sensitive to the mitochondria)
- ✧ Cellular interaction networks (assume they are expressed in eye development and there is a need to formalize these networks).

The short-term (1 to 2 years) research objective of the v-Embryo is to build a directed computer animation and accompanying knowledge management system of vertebrate eye development that is biologically-based and mathematically-driven. This will help researchers understand how the eye forms from the cell up. The long-term (3 to 5 years) objective is to validate/refine the prototype against new and existing data from standardized tests and begin to consider other embryological systems and concepts to expand its scientific capability and regulatory utility. In the next 20 years, Dr. Knudsen expects to build models of embryo systems to help predict which organisms would be susceptible.

The knowledge discovery component involves the information infrastructure for modeling and analysis. The Birth Defects Systems Manager (BDSM) will be connected to the CompTox API layer. The Center will build a searchable document repository capturing information from various sources (ToxRefDB, ToxNET). The logic-based text mining rules are based on the development toxicology thesaurus (standardized ontology). NCCT will build a translational Regulatory Toxicology Support module to assist stakeholders using the information for risk assessments and education.

Dr. Knudsen presented a diagram of the NCCT application framework; he commented that the Center was not sure how to bring v-Embryo and BDSM together but the staff is working on it. The project is using a shallow parser for text mining (SearchPERL.pl). A keyword can be entered and various toxicological sources searched. He mentioned that 950 non-redundant terms were used to identify one effect.

IRIS was searched for eye malformations in core risk assessment studies. The search revealed that 68 of 544 (12.5%) chemicals caused eye malformations, mostly “small eye.” The core OPP tests indicated that 28 of 268 (10.4%) pesticides caused eye defects in rat chronic studies.

Dr. Knudsen presented a diagram that mapped eye malformations to ToxRefDB chemicals. The staff searched 4,618 DERs for 37 developmental toxicity terms for eye defects. The DERs of 160 chemicals mentioned 6 conditions. A search of PubMed for the 160 chemicals yielded 13,027 PubMed Identifiers (PMIDs) of which 2,282 refer to these eye conditions.

Natural Language Processing (NLP) is being used to identify relevant information, for example, to identify relevant gene expression studies in PubMed.

The three-dimensional computer visualization model is being developed using Blender, which is a free open source 3-D content creation suite used by video game producers (http://www.blender.org).

Dr. Knudsen presented a series of scanning electron microscope (SEM) images that depicted early eye development in the mouse (Theiler stages 12-18, 72 hours in mouse [gestation days 8-11] and about 20-37 days human gestation). The growth curves from mouse eye development are a simplistic model for eye development that gives a framework for applying trajectories. As the Center moves to developing a systems model, it looks at the sequence of events as the lens and retina form. The staff is starting to extract information from the ontology map to build networks controlling early eye development. Dr. Knudsen noted that metabolic interactions are important for understanding eye development.

A search of EMAGE, a free database of 2D and 3D gene expression patterns during mouse embryo development, yields 65 genes (TS 12-18) and captures 2,164 PMIDs. One hundred sixty-five genes (overlapping in mouse and rat) are developmentally regulated in optic pit, optic vesicle, and optic cup. The Center is looking at these genes to identify which ones show the same development relevance. Dr. Knudsen presented data on genetic response networks in mouse embryos 3 hours after maternal alcohol on gestation day 8. The C57Bl/6N mice showed low susceptibility and the C57BL/6J mice showed high susceptibility. Looking at pathways that are up-regulated and down-regulated, he pointed out that certain pathway responses only occurred in the insensitive system; susceptibility should be viewed as a lack of response rather than a response. Although only a moderate number of genes are being examined, there is still an astronomical number of calculations.

Dr. Knudsen noted that receptor-mediated cell adhesion genes increase in the more susceptible mouse and decrease in the less susceptible mouse. Over time, the more adhesive cells are surrounded by the less adhesive cells.

Quoting Butler's paper published in *Nature* in 2001, Dr. Knudsen said that "our ability to create mathematical models describing the function of biological networks will become just as important as traditional lab skills and thinking." He also quoted Schnell, et al., "Molecular biology took Humpty Dumpty apart…mathematical modeling is required to put him back together again" (*American Scientist* 2007;95:134).

CompuCell3D (CC3D) is a cell-driven model of chick limb development. This is a nice piece of software for cell-based computing. This model simulates the process leading to three digits. The cells come together and start bone development. It is a cell-based model of morphogenesis using cell-biological cues. It considers the radius and ulna as being formed separately but it now is known that they occur together. There are deficiencies to this model, e.g., it does not consider

cell growth as a possible mechanism. Therefore, it is a plug in module for cell sorting. The Center can use this and other tools to build more complex models.

Hybrid CA simulators may be helpful in developing the v-Embryo. It is a computational modeling method to simulate the collective cellular behavior and emergent properties of a complex system. The generalized "cell" is a connected domain of lattice nodes (pixels) with the same value (spin); spin flips depend on parameter set (type, state) deduced from local gene networks, core processes, and environment. The automaton advances in discrete steps; hybrid simulators overlay continuous agents (morphogens, signals, metabolites, chemicals). CC3D computes energy Hamiltonian (H) from inputs, agents, and ΔE and then evolves simulation based on the principle of energy minimization.

To build something similar for eye development, the Center would focus on the cell-based processing driving the natural system—patterning (cell interaction), morphogenesis (cytoskeletal remodeling), selective growth (proliferation and apoptosis), and cell differentiation (crystallins, photoreceptor genes)—as well as the cell-based processes driving the formal system—network logic (information flow), cellular automata (cellular Potts model), biochemical fields (Reaction-Diffusion equation), and state transitions (cell type and state module).

Lack of oxygen can stop lens development and cause changes in the retina. Using simulators, the Center can model eye defects. Catastrophe theory can be used to project phenotype. There are stable and unstable positions between state A and state B. Catastrophe theory provides a mathematical formula to calculate the critical point.

The Center has proposed the v-Embryo for predictive modeling and regulatory analysis of developmental processes and toxicities. Cell-based simulators of morphogenesis compute phenotype outcome associated with collective cell behavior. The Center can begin to think about systematic parameter sweeps to identify lever points in the energy, control, and robustness of a system. The model can help define associative relationships in the developing system across dose, species, and stage.

There are a number of potential collaborators on the v-Embryo Project, including:

- ✧ OPPTS—animal studies subchronic/chronic
- ✧ NCEA—risk assessment and IRIS database
- ✧ Environmental Monitoring and Visualization Laboratory (EMVL)—environmental modeling and visualization
- ✧ NHEERL—embryonic stem cells, embryo culture, genomics, cytomics, and mitochondria
- ✧ Commercial—Merck Research Labs, Numira Biosciences, and Zygogen
- ✧ Organizations—International Life Sciences Institute (ILSI), DEVTOX, Teratology Society, and EMAP-United Kingdom
- ✧ Government—National Institute of Child Health and Human Development, NIEHS, National Institute of Alcohol Abuse and Alcoholism, and RIVM-Netherlands
- ✧ Academia—University of Louisville, University of Notre Dame, and University of North Carolina.

Panel Discussion

Dr. Daston asked what is needed to test the hypothesis. Dr. Knudsen responded that there has to be a good deal of interaction with NHEERL to provide information at the genetic and cellular levels. He noted that there are some sources of data available already. Embryonic stem cells could be an important approach as well as use of whole embryo culture systems. The Center has an ongoing collaboration with RIVM-Netherlands looking at chemical effects in embryo culture. NCCT will focus more on the cellular level and understanding mitochondrial effects and stage changes.

Dr. Daston commented that the Virtual Liver Project is expending considerable effort on creating a framework, but the v-Embryo Project appears to be relying more on what others have developed (Bard, et al.). Does the framework for v-Embryo already exist? Dr. Knudsen said that Bard is not seeking to develop a model for testing the effects of chemical exposure. He mentioned that the Jackson lab has a phenotypic database on mouse knock-outs and knock-ins and many of those affect the eye. Therefore, there are data available for the project. Teratologists have been capturing data but there is a need to apply good biomarkers to follow the cells. The zebrafish researchers have been doing a good job of working out the circuitry and the Center can use that information. Dr. Shah commented that this project is approaching organ simulation from a physiological modeling standpoint. By looking at the cellular level, the phenotype outcomes can be seen. The Virtual Liver Project sounded more de novo because there is no paradigm for looking at tissue injury. There is no mechanistic approach to linking perturbations to outcomes. The Center plans to use everything available for both the v-Embryo and Virtual Liver projects. There will be a lot of synergy between these two efforts.

Dr. Di Giulio said he thought it was odd that the Center chose to focus on the eye. What about the nervous system? Is the eye predictive of what is going on in the brain? Dr. Knudsen responded that the nervous system is more morphologically sensitive but right after the optical nerve forms, it is very morphologically sensitive. Some thought was given to this criticism, but the decision was a practical one given the knowledge and expertise of the Center staff. He thinks the critical issue is to avoid isolating the Center from the neurotoxicology community. Mitrochondria are important for eye development and they are critical for neurotoxicologists as well.

Dr. Di Giulio said that his colleagues at Duke are very excited about REACH. Is the Center trying to take advantage of the data that will be available because of REACH? Dr. Dix mentioned that the Center has a contract with Phylonix. There have been some delays in getting started, but that work may be applicable to the v-Embryo Project. That could be a vehicle for generating data for the zebrafish model for v-Embryo. He also mentioned some collaborative work with Zygogen.

Dr. Quackenbush asked Dr. Knudsen to comment on the integration of this project and ToxCast. Dr. Knudsen answered that ToxCast was not designed with development endpoints in mind. It can be moved into a developmental system by looking at cell level information using the same technologies as the other data in ToxCast. At some point, more resources will be required for institutional collaborations. At the present, there is a fair amount of data available from literature and databases that can be captured for use by the project.

Dr. Stokes wondered what questions would be asked of the v-Embryo model if it were available today. Dr. Knudsen replied that the model could be used to study the MoA of fluorouracil. It is an antiproliferative agent that inhibits cell cycle progression; when examining tissues, it induces cell death. When we look closer at this process, we may see all kinds of changes. We do not know the primary path of action. The model would allow us to look at just inhibiting cell proliferation for an amount of time. We can see how well the model does at predicting those associations. What are the molecular networks that control the system? Can we use those as targets to mechanistically better understand how exposure leads to effects? Dr. Stokes asked which of those questions would be priorities at EPA. How much priority does EPA place on understanding mechanisms versus determining whether a chemical is toxic or not? What is the purpose of the modeling? Is the Center building the model to predict outcomes?

Dr. Barton commented that EPA has multiple priorities. One priority is the large number of chemicals for which the Agency has limited information. Are there better ways to prioritize them for testing? EPA also spends a lot of money developing risk assessments for chemicals. The Center is trying to improve the Agency's ability to better estimate risk. This is another EPA priority. The balance among the various priorities fluctuates.

Dr. Daston said his understanding of EPA's mission is that the Agency is to ensure that chemicals are not present in the environment at levels that will impact public health. A better understanding of the mechanisms and levels that affect public health helps the Agency to accomplish this mission. The Center needs to apply this project to the problems of dose response and chemical prioritization. Dr. Kavlock pointed out that this research will be applicable to risk assessments conducted in 2027, not those being done today. There are many questions to answer with this model. The model can extrapolate dose by time from mouse to human. This is an important issue and can yield a lot of payback. He commented that in 2005, $1000 of computing power bought a computer about the size of an insect brain, in 2010, it will buy a computer the size of a mouse brain, in 2030, a computer the size of a human brain, and in 2060, it will buy a computer the size of all the minds on the planet. Over time, the computing power available will be tremendous, and the Center's work could lead to the development of some very powerful tools.

Dr. Shah stated that there are no data on chemicals at environmentally relevant doses; therefore, the Center is using models to understand what is happening at low doses and to predict the effects of exposure to such low doses.

Dr. Clark commented that the Center will need to find partners to bring in the data as this project grows. NCCT will have to entice partners to become involved. He also noted that people will want answers to questions on extrapolation.

Dr. Kavlock said that the Center wants to do more but these projects are resource intensive. Perhaps there will be an increase in the investment in this area in the next few years. The Center needs to publicize its success stories to attract collaborators. This is something that the staff will discuss internally.

Working Lunch

Computational Toxicology Subcommittee Members

Dr. Daston asked the members to provide their initial responses to the charge questions with respect to the projects presented.

Question #1: Does the scope and involvement of expertise in the project reflect activities consistent with the function of a Center?

Dr. Clark stated that the Center has the people necessary to create the vision and has developed an approach for attracting partners. Dr. Stokes said that she was concerned about continuity. Would different staff members build a different model? Dr. Quackenbush thought the Virtual Liver Project looked like a young investigator's first NIH grant because it is so ambitious. Each part of the project could be a 5-year effort. Dr. Daston noted that this touches on goals and milestones, which is the next question.

Question #2: Are the goals and milestones suitably described, ambitious, and innovative?

Dr. Daston stated that the vision is well cast but more specific milestones may be needed. Dr. Stokes thought the Center should develop some specific questions that would be asked of the model. What will the Center do with the model? A list of questions also might garner the interest of post-docs. Dr. Daston said the Subcommittee may endorse updating the implementation plan. Dr. Quackenbush did not think the timeline was realistic. Dr. Daston commented that the Center will need to be pragmatic about identifying realistic milestones.

Dr. Quackenbush thought Dr. Shah has done a great job of thinking about the Virtual Liver and its elements.

Question #3: Are there significant gaps in the approach that can be pointed out at this point in the evolution of the project?

Dr. Daston asked if there were any significant gaps that the Subcommittee should point out. He acknowledged that these projects are so new that the Subcommittee may only be able to suggest additional communications, collaborators, etc.

Dr. Stokes asked how the different pieces will fit together. It is one thing to make the mathematical equations fit together but what about the science and the thinking behind the parts. The assumptions and boundaries must be compatible. That is one gap that was not addressed. Dr. Quackenbush liked Dr. Stokes' suggestion that the Center do screening to get empirical data. He commented that Dr. Knudsen has not been at EPA long enough to figure out how this project fits into the other efforts. Dr. Daston added that the v-Embryo Project also needs collaborators. Dr. Mumtaz thought the report should acknowledge that Dr. Knudsen's short tenure at EPA has made it difficult for the Center to fully integrate the project and enlist collaborators. Dr. Quackenbush thought Dr. Knudsen would play an important role in Phase II of ToxCast. His contribution will add an important element.

Question #4: Does the work offer to significantly improve environmental health impacts and is the path toward regulatory acceptance and utilization apparent?

Dr. Daston said it is probably too early to answer this question, but the Center needs to keep an eye on this for the future. Dr. Stokes agreed that it is important to work to improve risk assessments in 2020, but it would be great to find a practical application in the short term. Dr. Quackenbush thought some of the milestones for the Virtual Liver Project could be completed earlier, perhaps by 2010, and used to evaluate chemicals.

Question #5: Have appropriate data management and analysis tools been incorporated into the project?

Dr. Quackenbush suggested that the Center give more thought to integration. Dr. Daston thought the Center needed more outreach to other groups. The Subcommittee should encourage NCCT to formalize partnerships within EPA and outside the Agency. It may be difficult to get active partners outside ORD in the early years, but it is important to get outside organizations interested in the projects as soon as possible.

BBDR Modeling for the Carcinogenicity of Inorganic Arsenic
Dr. Rory Conolly, NCCT, ORD, EPA

Dr. Conolly's presentation included an overview, EPA's current cancer risk assessment, the definition of a BBDR model, description of a BBDR model for carcinogenicity of arsenic, and first steps in model development.

NCCT's expertise in computational modeling and NHEERL's expertise in arsenic MoAs will be combined to develop a data-driven dose-response assessment model (BBDR model). Center staff members have met with NHEERL several times to work on the BBDR model. This model will be designed for use by regulatory groups such as OW and NCEA to reduce uncertainty in low-dose extrapolation, which will lead to better risk assessment.

The Center is working with NHEERL, OW, and NCEA to identify data needs and NCCT and NHEERL are developing a proposal for research by external collaborators. These efforts will lead to the collection of new data. There is a considerable amount of data on As but not much of the data is useful for development of a BBDR model. NHEERL's existing dosimetry data and MoAs will be used, along with new data to build the BBDR model.

The Safe Drinking Water Act (SDWA) requires review of regulations every 6 years. The deadline for the BBDR model to impact the next review cycle is 2012. Dr. Conolly noted that it is too late to impact the current review cycle. New data from the current planning process are needed by 2010-2011. Computational modeling will be initiated using existing data and new data will be incorporated as they become available.

The level of effort expended on this project is 50 percent of Dr. Conolly's time and a full-time post-doc as well as 19 principal investigators and 15-20 full-time equivalents (FTEs) from NHEERL. This effort from EPA staff along with the external collaborators should really make a difference.

Inorganic arsenic (iAs) is a multi-site human carcinogen (bladder, lung, and skin). The main concern is drinking water exposure in the United States and globally. In India and Bangladesh there are high concentrations of As in the groundwater. The current cancer risk assessment references mortality data from a lung and bladder cancer cohort in Taiwan. It uses linear extrapolation to low doses. The current drinking water standard for Arsenic is 10 µg/L (ppb). The standard was derived from balancing the cost of removing arsenic against the potential health benefits produced from reduced exposures.

Dr. Conolly pointed out that the data do not constrain the dose-response curve at relevant levels of exposure. There are a number of issues with the current assessment. Linear extrapolation is a default and not a consequence of a hypothesis about the MoA of iAs as a human carcinogen. It is probably health protective, but it is unlikely to be accurate, which is problematic for risk-benefit analysis.

A BBDR model is one in which biological mechanisms determine dose-response. It is a predictive tool used to estimate potential human health risks by describing and quantifying the key steps in the cellular, tissue, and organismal responses as a result of chemical exposure. The Guidelines for Carcinogen Risk Assessment state that a "robust, biologically based model" is the preferred approach for dose-response assessment. The As risk assessment did not use BBDR; the Center hopes to be able to actually build in information on cellular and tissue events that are secondary. Dr. Conolly noted that it is important to understand the biochemistry that drives tumor development.

BBDR models can predict time course and severity of a toxicological response. They make maximum use of mechanistic information (PK, PD), and they can be further "tuned" for special cases (e.g., sensitive subgroups, genotypic variation, age-dependence).

Since 2005, the Center has been participating in an ongoing series of meetings with the arsenic investigators at EPA. These meetings have the support of senior management and include representatives from OW and NCEA. A 2-day meeting was held in RTP in October 2007. As a result of that meeting, a proposal for new research is being prepared and an external review is planned for early 2008. The Center is looking for external collaborators to increase coverage of issues relevant to dose-response. The BBDR model is needed by 2012 to impact the mandated 6-year review of the drinking water standard for iAs.

The first step is to look at arsenic metabolism and a human PBPK model for iAs and methylated metabolites. Detailed descriptions of the inorganic arsenic metabolism used in the model are given in a manuscript with references to studies examining each step of the pathways. The sequential reduction/methylation has been shown by many researchers earlier (reference for these pathways is given by Thomas, et al., 2004 and Zakharyan, et al., 2005). The presence of inhibitory reactions of As and MMA on the methylation reaction also has been documented (Kedderis, et al., 2006; Easterling, et al., 2002; Kenyon, et al., 2001; Zakharyan, et al., 1999; and Styblo, et al., 1996). The inhibitory metabolic pathways of inorganic arsenic also were used by Gentry, et al. (2004) in their mouse PBPK model. The ample evidence and support for the metabolic pathways as they were described in the manuscript necessitated the use of these mechanisms so that the quantitative model would be as biologically plausible as shown in literature. The purpose of this model was to quantify these pathways in a model that can be used to predict levels of arsenic and its metabolites in urine and blood. The model was intended to

challenge or examine metabolic pathways that were shown to exist by the use of experimentation as shown from the ample evidence in literature. Support for this mechanism of interaction was recently provided in a publication by Kedderis, et al. (2006).

Developing an earlier As PBPK model for mice, Gentry, et al. (2004) reached a similar conclusion to the one used in the Center's model. Specifically, comparing the kinetics of monomethylarsenate (MMA) and dimethylarsenate (DMA) immediately following administration of MMA^V and iAs, it was observed that one set of kinetic parameters would not adequately describe the formation of DMA from MMA under both situations (Gentry, et al., 2004). Therefore, adjustment to the metabolic mechanism of inorganic arsenic was made to allow for the production of DMA directly from As^{III}. This mechanistic assumption should not be interpreted as instantaneous double-methylation of As^{III} to produce DMA. One possible biochemical basis for this assumption is that a portion of MMA is not immediately released to circulation following the methylation of As^{III}. Some MMA still is present in the proximity of the methylating enzyme to be further metabolized to DMA. These assumptions are justified based on several mechanistic hypotheses regarding the functions of methyltransferases. AS3MT catalyzes both the mono- and dimethylation reactions that produce MMA and DMA from iAs. The second methylation reaction is several hundred times faster than the first methylation reaction when studied using purified enzyme preparations. Hence, concentrations of MMA are typically much lower than those of DMA. In addition, there are analyses which suggest that the rate of transport of MMA out of hepatocytes is slow compared to other arsenicals based on *in vitro* studies in rat hepatocytes (Easterling, et al., 2002).

The planned human PBPK model will describe mechanisms of differential tissue-specific accumulation of metabolites. It looks like four models but it is only one that tracks four different metabolites. It will be modified to facilitate analysis of human biomonitoring data (e.g., skin and toenail compartments). It will describe mechanisms of quantitative variability in metabolism (e.g., polymorphisms identified in human AS3MT).

Dr. Conolly presented a schematic of the overall PBPK model for inorganic arsenic and methylated metabolites. The model accommodates oral exposure to As^{III}, As^V, MMA^V, and DMA^V as represented by block arrows to the GI tract lumen in each of the submodels. Reduction of As^V to As^{III} occurs predominantly in the GI lumen, as well as liver, lung, and kidney. Reduction of MMA^V and DMA^V to their respective trivalent forms occurs in the liver and kidney. The trivalent forms are not described as entering the circulation in the model; rather urinary excretion of the trivalent forms occurs via active transport in only the kidney. Methylation of As^{III} and MMA^{III} occurs in the liver and kidney. Oxidation of As^{III}, MMA^{III}, and DMA^{III} to their respective pentavalent forms occurs in liver, kidney, and lung and is quantitatively minor.

Model simulations are compared to data of total As, As, MMA, and DMA cumulative levels in urine of human volunteers exposed to 100 µg As in the form of sodium arsenate and sodium arsenite. Data were obtained from a doctoral dissertation by Lee (1999). In the slide, data of As, MMA, and DMA levels in collected urine were shown in squares, triangles, and circles, respectively. The simulations were performed to evaluate the model behavior using parameters of previous optimized simulations of Buchet, et al. (1981) data. Best results were obtained when an assumption of reduction of sodium arsenate was assumed in the stomach.

There is motivation for PBPK models for rats and mice. A number of studies report that exposure to As causes bladder cancer. DMA^V causes bladder cancer and trimethylarsine oxide (TMAO) causes liver cancer in Fisher 344 rats. Arsenate is a tumor promoter with dimethylbenzanthracene (DMBA) in the mouse, and arsenite enhances UVR-induced skin cancer in hairless mice. The mouse model will be consistent with the human model. MoA studies could be conducted in support the human BBDR model.

For the ongoing development of the mouse PBPK model, there are four submodels (As^V, As^{III}, $MMA^{V\&III}$, $DMA^{V\&III}$) linked by methylation in liver. There is greater consistency with the human model and additional target tissues are being incorporated (e.g., bladder). Mechanisms to account for tissue-specific accumulation of certain arsenicals are being incorporated (e.g., binding proteins).

The DMA^V PBPK model for the rat builds on the mouse DMA^V PBPK model (Evans, et al.). It accounts for binding of DMA^{III} to rat hemoglobin and extensive metabolism to TMAO. The rat model tests the hypothesis that difference in bladder cancer response in mice versus rats is a function of kinetics rather than dynamics. The challenges include limited PK data in the rat and the data are not collected for purposes of model development.

Dr. Conolly stated that As is clearly a human carcinogen, but what is the dose response at doses present in the environment? There is a large amount of *in vitro* studies of MoA. There are less epidemiological cancer dose-response data available, and there are very few data from laboratory animal *in vivo* dose-response and time-course studies. The *in vitro* studies of MoA generally lack the dose-response and time course data needed for BBDR. In addition, there are relevance issues (i.e., transformed cell lines). The other two types of data (epidemiological cancer dose-response and lab animal *in vivo* dose-response and time-course data are very relevant to BBDR.

Dr. Conolly mentioned a March 2007 paper by Xie, et al., entitled "Aberrant DNA methylation and gene expression in livers of newborn mice transplacentally exposed to a hepatocarcinogenic dose of inorganic arsenic." This study involved one applied dose of 85 ppm arsenic in drinking water. The data reported in that paper cannot be used in BBDR modeling. He mentioned another study by Lam, et al., entitled "Transcriptome kinetics of arsenic-induced adaptive response in zebrafish liver" (*Physiological Genomics* 2006;27:351-361). One applied dose of arsenic of 15 ppm in drinking water was used in this study. What is the human relevance?

Another study by Liu at the University of New Mexico Health Sciences Center, was conducted in the HaCaT human keratinocyte transformed cell line. These data were more useful for BBDR model development because there was a nice dose response.

Dr. Conolly then presented data on formaldehyde—dose-time response surface for regenerative cellular proliferation in nasal epithelium of the F344 rat. He noted that these data are very useful for BBDR model development. Rather than a single dose at a single time, there are different doses at different times.

Overall dose-response and time-course is built up from the key event relationships from dosimetry to the regulatory endpoint. The Center needs new data to identify the path(s) that lead to the effect. Dr. Conolly identified several plausible pathways, noting that data are needed to determine which pathway is correct. What is the dose and MoA that leads to bladder cancer,

skin cancer, and lung cancer? The questions driving the effort are: What are the responsible metabolites? What are the key event sequences? What are the tissue similarities or differences?

Dr. Conolly listed the following characteristics of experiments that support BBDR modeling:

- ✧ Identify sequences of key events.
- ✧ Emphasize dose-response and time-course.
- ✧ Coordinate among labs.
- ✧ Evaluate *in vitro* to *in vivo* extrapolation for *in vitro* studies.
- ✧ Evaluate interspecies extrapolation for animal studies.
- ✧ Consider conducting fewer, more elaborate studies.

One of the first steps in model development is linking tissue doses to tumor outcomes. There has been significant use of PBPK modeling in risk assessment during the past 20 years. The influence of PK non-linearities on tumor dose-response will be identified and compared with linear low-dose extrapolation. PBPK predictions will be validated against human biomarker data. The next step is to simply fit the clonal growth model to epidemiological dose-response data. Parameter changes associated with observed tumor responses will be identified. What are the effects on cellular division and death rates? What are the effects on mutation rates? What are the normal cell effects? What are the effects on preneoplastic lesions? The Center wants to build in mechanistic information to understand the process from normal cells to initiated cells to cancer cell to tumor.

The charge questions for the As BBDR project are:

- ✧ Does the proposed computational model have the potential to identify and reduce uncertainties with the risk assessment process?

- ✧ Will the model be able to help identify susceptible populations and compare potential risks in those populations with less susceptible groups?

- ✧ Is coordination between model development and associated data collection sufficient to avoid problems with the model being either over- or under-determined?

Panel Discussion

Dr. Daston said he was not certain who would judge the model in terms of its risk assessment value. At what point will they be involved in the project? Dr. Conolly replied that OW and NCEA have been involved with the project and OPP has been briefed and is aware of the Center's work. Dr. Doug Wolf commented that the lead from OW participates in the project's regular conference calls. Dr. Daston was concerned that if the BBDR model does not explain all of the toxicity of As, OW may not recognize it and continue to use the linear default in the risk assessment. Dr. Conolly responded that no model can do everything. Its adoption will depend on the level of sophistication of the regulators. ORD always runs the risk of the regulators not using a model; the Center will have to actively involve them in the process to increase their understanding of the model. NCCT is trying to be comprehensive in candidate MoAs, which will be very useful to the regulators.

Dr. Daston asked if the Center would consider creating an advisory committee for the project from a subset of individuals who attended the February workshop. Dr. Conolly said he liked that idea.

Dr. Stokes asked Dr. Conolly to elaborate on the types of approaches that will be used. Dr. Conolly responded that he envisions it as a differential equation based model. Dr. Di Giulio asked if Dr. Conolly had a sense of the shape of the curve at low doses. Dr. Conolly answered that many people think it is non-linear but unless there are compelling data to support that assumption, it will be difficult to convince others. From an engineer's perspective of cellular biology, it is difficult to imagine that it is not non-linear.

Dr. Daston asked if the model will include non-cancer effects. Dr. Conolly agreed that there are many effects associated with As exposure, such as diabetes and cardiovascular disease; the Center thought about including those in the model, but decided to focus on cancer endpoints. Dr. Wolf stated that the cancer guidelines specifically mention biologic MoA; the default linear approach must be used unless the data can prove otherwise. Therefore, the Center is dealing with cancer MoA first. This is being driven by the project office and NCEA—target tissues, endpoints, etc., were selected based on their input.

Dr. Daston was concerned that by focusing only on cancer the Center could open itself up for criticism. For example, what if the model indicates that As at high doses would affect only a limited number of people; however, if As at low dose levels causes cardiovascular disease, the model will not be able to predict this outcome. Dr. Conolly replied that if the cancer model is successful, the Center will have achieved a great deal. Once the cancer model is successful, NCCT can expand to cardiovascular and other effects using the same process.

Dr. Quackenbush asked if Dr. Conolly had talked to Dr. Dix about incorporating this into ToxCast. Dr. Conolly answered that Dr. Dix has talked to him about integration; he mentioned that dimethylarsenic is in ToxCast. Dr. Dix said that he is thinking about adding three or four arsenicals in Phase II of ToxCast. Dr. Conolly asked if there was flexibility in ToxCast to collect information on more lines. Dr. Dix replied that there may be some other assays (skin, bladder, lung) that might be useful.

Dr. Mumtaz asked if any parameters had been calculated. Is the Center thinking about getting data to link to tissue doses rather than urinary excretion? A member of the panel responded that they only optimized parameters that were found to be sensitive—basically urine data. That makes this model unique from previous efforts. Dr. Mumtaz asked if dermal exposure was an issue. Dr. Conolly answered that the program office is more concerned about other exposures than skin. Dr. Daston said there is a need for large coordinated studies. Is the program office willing to support such studies? Dr. Conolly replied that there will be a need for a lesion pathogenesis study and without such a study it will be difficult to go beyond PBPK modeling. Dr. Hal Zenick commented that one way to deal with data generation is to strengthen modeling. This work falls under the Human Health Research Program. There are many different testing approaches and ORD is committed to improving the modeling to reduce uncertainty in risk assessment.

Dr. Daston thought this project was a good example of the Center using its expertise to tie into a large, important ORD program to leverage much more than NCCT could do alone. Arsenic is clearly one of the most important public health issues that needs to be resolved.

Dr. Daston thanked the panel for answering the Subcommittee's questions. He then thanked Dr. Kavlock and all the staff members who presented at the meeting. The Subcommittee's report will provide the Center with comments and recommendations to improve the program. It is clear that NCCT is making great progress and the Center's expertise is being sought out by others in the Agency. The Subcommittee is pleased to see that the Center has developed tools that already are being used by the program offices. Dr. Kavlock said he really appreciated the advice and feedback the Center has been getting from the BOSC.

Subcommittee Working Time

Computational Toxicology Subcommittee

Dr. Daston said that Question #4 had a specific reference to the BBDR model. There are three subquestions under Question #4. Dr. Mumtaz asked about the focus of the research proposal that was developed by NCCT and NHEERL. One member replied that the proposal will focus on filling data gaps.

Dr. Daston commented that the Center has a huge potential to reduce the uncertainty of the As risk assessment. He did not know whether the 10 ppb standard was the most protective level. There is a pragmatic limit to what can be done to control As. It was not clear that the model will identify susceptible populations.

Dr. Stokes said there was some mention of subpopulation risk but it was not clear how this would be addressed. Dr. Daston suggested seeking input from experts attending the workshops about this issue. Dr. Stokes asked if the Center will be gathering data at lower concentrations. Dr. Daston stated that the idea is to understand MoA so they can predict effects at low concentrations. Dr. Clark pointed out that there is more detail provided in the notebook. Dr. Kavlock commented that it has been difficult to get the staff thinking about biological models.

Dr. Quackenbush asked where the NCCT staff is located. Dr. Kavlock responded that they all are located in RTP, but the staff members occupy space in three different locations. Dr. Quackenbush suggested that there be more interaction among Center staff members; perhaps they could present their work to one another at regular meetings. Dr. Kavlock responded that the Center staff members attend biweekly meetings where they present their work. Dr. Quackenbush asked if these meetings include all Center staff members. Dr. Kavlock confirmed that all staff members attend these meetings. He commented that he did not want everything to be focused on ToxCast, but agreed that there should be integration with ToxCast when it makes sense.

Dr. Daston asked if the Subcommittee members had any questions or comments about Question #4. Dr. Stokes asked if there were any major gaps in the project. Dr. Daston thought it would be difficult to answer that question until the Center develops a detailed research plan. Dr. Clark said NCCT will have to rely on input from stakeholders concerning whether the model addresses the risk assessment uncertainties. Regulators must be confident with the model and its ability to

address their questions. Dr. Daston pointed out that even if the model is not used for risk assessment, it will be useful in meeting the goals of the Center.

Dr. Di Giulio mentioned the As study in Chile, noting that there was a 50-year delay in observance of the effects.

Dr. Daston said that he expected the Subcommittee members to draft the response to their assigned question and send it to Ms. Kowalski. He reminded the Subcommittee members that all items should be sent to Ms. Kowalski; she will send the responses to him and he will consolidate them and send the file to the Subcommittee members. Dr. Daston asked that the drafts be submitted in the next 2 to 3 weeks. A conference call will be scheduled for late January or early February to discuss the draft report. Ms. Kowalski agreed to work with the members to determine their availability and schedule the call.

Dr. Stokes asked Dr. Daston to explain the issue associated with the As standard. Dr. Daston replied that the As level was changed from 50 to 10 ppb. Based on an NAS review of the data, the change was suspended by then Administrator Christie Todd Whitman. A second NAS panel reviewed the data and upheld the 10 ppb standard. Dr. Quackenbush stated that one of the goals of the NCCT is to move beyond linear extrapolation.

Dr. Daston thanked the Subcommittee members for their comments and adjourned the meeting at 2:35 p.m.

Action Items

- ✧ Dr. Clark will draft the response to Question #1 and send it to Ms. Kowalski by January 8, 2008.
- ✧ Dr. Di Giulio will draft the response to Question #2 and send it to Ms. Kowalski by January 8, 2008.
- ✧ Dr. Stokes will draft the response to Question #3 and send it to Ms. Kowalski by January 8, 2008.
- ✧ Dr. Mumtaz will draft the response to Question #4 and send it to Ms. Kowalski by January 8, 2008.
- ✧ Dr. Quackenbush will draft the response to Question #5 and send it to Ms. Kowalski by January 8, 2008.
- ✧ Dr. Daston will draft the response to Question #6 and send it to Ms. Kowalski by January 8, 2008.
- ✧ Ms. Kowalski will send the draft responses to the charge questions to Dr. Daston.
- ✧ Dr. Daston will consolidate the draft responses to the charge questions into one document and send it to Ms. Kowalski who will distribute it to the Subcommittee.

- ✧ NCCT will provide the RFA (or the URL that links to the RFA) for the STAR Bioinformatics Centers to Ms. Kowalski and she will distribute it to the Subcommittee members.

PARTICIPANTS LIST

U.S. EPA Board of Scientific Counselors
Computational Toxicology Subcommittee Meeting
December 17-18, 2007

Subcommittee Members

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

James R. Clark, Ph.D.
Exxon Mobil Research and Engineering Company

Richard T. Di Giulio, Ph.D.
Professor
Nicholas School of the Environment and Earth Sciences, Duke University

M. Moiz Mumtaz, Ph.D.
Science Advisor
Division of Toxicology and Environmental Medicine
Agency for Toxic Substances and Disease Registry

John Quackenbush, Ph.D.
Professor of Biostatistics and Computational Biology
Department of Biostatistics and Computational Biology
Dana-Farber Cancer Institute

Cynthia L. Stokes, Ph.D.

Designated Federal Officer

Ms. Lorelei Kowalski
Office of Research and Development
U.S. Environmental Protection Agency

Presenters

Rory Conolly, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

David Dix, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Richard Judson, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Robert Kavlock, Ph.D.
Director
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Thomas Knudsen, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Imran Shah, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Contractor Support

Beverly Campbell
The Scientific Consulting Group, Inc.

Denise Hoffman
The Scientific Consulting Group, Inc.

Other EPA Participants

Barbara Abbott, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Hugh Barton, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Carl Blackman, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Jerry Blancato, Ph.D.
Deputy Director
National Center for Computational Toxicology

Ray Burton
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Elaine Cohen-Hubal, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Karen Dean
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Stephen Edwards, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Robert Ellis-Hutchings, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Fathi Elloumi, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Hisham El-Masri, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Andrew Geller, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Michael R. Goldsmith, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Keith Houck, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Lora Johnson
National Exposure Research Laboratory
U.S. Environmental Protection Agency

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

Kirk Kitchin, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Melissa Kramer
Office of the Science Advisor
U.S. Environmental Protection Agency

Stephen Little
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Matthew Martin, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Dale Pahl
National Exposure Research Laboratory
U.S. Environmental Protection Agency

Heidi Paulsen
Office of Environmental Information
U.S. Environmental Protection Agency

Jim Rabinowitz, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

David Reif, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Joseph Retzer, Ph.D.
Office of Environmental Information
U.S. Environmental Protection Agency

Ann Richard, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Chester Rodriguez, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Bill Russo
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

R. Woodrow Setzer, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

David Thomas
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

John Wambaugh, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Doug Wolf, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Hal Zenick, Ph.D.
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

Maggie Zhou, Ph.D.
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Other Participants

Zhen Li
UNC STAR Bioinformatics Center

Richard Paules, Ph.D.
National Institute of Environmental Health Sciences

Matt Shipman
Risk Policy Report

Amar V. Singh
Lockheed Martin

Richard Spenser
Lockheed Martin

Ray Tice, Ph.D.
National Institute of Environmental Health Sciences

Tom Transue
Lockheed Martin

Maritja Wolf
Lockheed Martin

Stan Young, Ph.D.
National Institute of Statistical Sciences

B·O·S·C
BOARD OF SCIENTIFIC COUNSELORS

COMPUTATIONAL TOXICOLOGY SUBCOMMITTEE
AGENDA
December 17-18, 2007
U.S. EPA Research Triangle Park (RTP) Campus
EPA Main Building (Room C114 on December 17 and Room C111C on December 18)
109 T.W. Alexander Drive
Research Triangle Park, NC 27711

Monday, December 17, 2007

12:30 p.m. – 1:00 p.m.	Registration	
1:00 p.m. – 1:10 p.m.	Welcome and Introductions - New Subcommittee Member - Draft Charge - Meeting Agenda	Dr. George Daston, Subcommittee Chair
1:10 p.m. – 1:15 p.m.	DFO Remarks	Lori Kowalski, Office of Research and Development (ORD)
1:15 p.m. – 1:45 p.m.	National Center for Computational Toxicology (NCCT) Overview	Dr. Robert Kavlock, Director, NCCT
1:45 p.m. – 3:05 p.m.	ToxCast	Dr. David Dix, NCCT
3:05 p.m. – 3:20 p.m.	Break	
3:20 p.m. – 4:40 p.m.	Information Management/Information Technology (IM/IT) - Informatics	Dr. Richard Judson, NCCT
4:40 p.m. – 5:30 p.m.	Subcommittee Working Time	Comp Tox Subcommittee
5:30 p.m.	Adjourn	

Tuesday, December 18, 2007

8:30 a.m. - 8:45 a.m.	Recap of Day 1	Dr. George Daston, Subcommittee Chair
8:45 a.m. – 10:15 a.m.	Virtual Liver	Dr. Imran Shah, NCCT

10:15 a.m. – 10:30 a.m.	Public Comment	
10:30 a.m. – 12:00 noon	Developmental Systems Biology	Dr. Thomas Knudsen, NCCT
12:00 noon – 1:00 p.m.	Working Lunch	Comp Tox Subcommittee
1:00 p.m. – 2:30 p.m.	Arsenic Biologically-Based Dose Response Model (BBDR)	Dr. Rory Conolly, NCCT
2:30 p.m. – 3:00 p.m.	Subcommittee Working Time	Comp Tox Subcommittee
3:00 p.m	Adjourn	