Report of Considerations of the National Toxicology Program (NTP) Vision by NTP Participating Agencies and Institutes

June 7, 2004

Marilyn Wind, Michael Babich- U.S. Consumer Product Safety Commission (CPSC)
Bill Allaben, Paul Howard- Food and Drug Administration (FDA)
Chris DeRosa- Agency for Toxic Substances and Disease Registry (ATSDR)
Tom Sinks- National Center for Environmental Health (NCEH)
John Howard, Mark Toraason- National Institute for Occupational Safety and Health (NIOSH)
J.Carl Barrett, David Longfellow, Michelle Bennett- National Cancer Institute (NCI)
Amanda Edens- Occupational Safety and Health Administration (OSHA)
Jack Snyder- National Library of Medicine (NLM)
Bill Farland, Hal Zenick- Environmental Protection Agency (EPA)
John Bucher, Scott Masten- National Institute of Environmental Health Sciences (NIEHS)

The NTP Vision for the 21st Century:

To move toxicology from a predominantly observational science at the level of disease-specific models to a predominantly predictive science focused upon a broad inclusion of target-specific, mechanism-based, biological observations

The activities of the National Toxicology Program (NTP) are focused in three broad areas. These are:

1.) Generating toxicology data for substances or agents of potential public health concern
2.) Evaluating health effects data and bringing substances of concern to public attention through the NTP Report on Carcinogens and the monographs of the Center for Evaluation of Risks to Human Reproduction (CERHR), and
3.) Developing and evaluating and/or validating new methods to serve the first two activities and to assist Regulatory agencies in adopting new toxicological methods through the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

From the standpoint of the health regulatory and research agencies that provide oversight to the NTP through the NTP Executive Committee (ATSDR, CPSC, EPA, FDA, NCEH/CDC, NCI/NIH, NIEHS/NIH, NIH, NIOSH/CDC, OSHA) the first function is clearly of critical importance. The NTP is the primary Federal entity created to provide comprehensive toxicological evaluations of substances of public health concern, and funded to carry out basic toxicology research to fill regulatory or scientific data gaps in situations where industry cannot be compelled to do so. Over the 25 years of its existence, the NTP has helped to set the standards for the design, performance, evaluation and interpretation of toxicology studies. These are critical functions of the program and the need for these capabilities is not likely to diminish in the future.

Currently and historically, the three activities outlined above have all been oriented towards a particular disease endpoint, be it cancer, impaired fertility, dysmorphogenesis, hypersensitivity, or some other affliction of interest. Each substance is studied for its potential to produce disease in a model organism using a test protocol developed for that purpose. Recently the program has attempted to provide more complete kinetic information concerning how substances are absorbed, distributed, metabolized and excreted (ADME) in the test species. Based on this collective information, an informal, or in rare cases formal and quantitative characterization of human hazard is developed. As science progresses and interests expand, new disease endpoints e.g. developmental neurotoxicity, endocrine disruption, or windows of disease vulnerability (sensitive subpopulations), require new testing paradigms and often reinvestigation of relatively well studied substances.

This disease-specific approach differs little from the approach to toxicology developed decades ago and adopted by the NTP at its inception in 1978. It is a comfortable approach. It has proven to serve the needs of most US Regulatory Agencies and has with some exceptions been accepted as an adequate basis for public health regulations by the Judicial Branch. This approach operates within the prevailing assumption that most substances in the environment and commerce are benign, and is designed to reactively respond to study substances or agents that have in some manner generated a suspicion of hazard. This focus on studying chemicals likely to pose a hazard may be the most efficient way to identify true hazards; however, it also may in some instances impede the discovery of unexpected results because of adherence to a current scientific bias.

Scientific and societal pressures on this existing approach to toxicology are growing and are pulling the field in different directions. The EPA through its High Production Volume (HPV) chemical initiative, and the European Union through it Registration, Evaluation, and Authorisation of Chemicals (REACH) program are promoting the creation of large publicly available databases of chemical toxicity information that are populated not based on the suspicion of hazard, but based on high production volumes. The NTP is itself in the process of making all of the data from its studies available in an interactive, searchable, electronic mode. This will allow these databases to be probed in ways not before possible, to potentially reveal new associations between toxicity endpoints and perhaps new understanding of what these tests are telling us. In an inverse sense, the new technologies of toxicogenomics and proteomics offer a similar opportunity to provide data unbiased by the model selected to study the disease, through the collection of vast amounts of information on gene and protein expression representing the global response of biological systems exposed to chemicals. For the first time it may be possible to apply computational, or systems biology analysis approaches to toxicology databases of sufficient magnitude to discern meaningful patterns or profiles of toxic responses. Efficiency in the collection, dissemination and use of this information will become more and more important in light of our constant obligation to use animal (and other) resources judiciously.

The NTP vision calls for the Program to provide leadership to move the field of toxicology from a science based largely on the observation of disease to a science focused on the prediction of disease through collection of a broad array of mechanism-based, biological observations. This vision is based in part on the belief that many environmentally influenced disease processes have an underlying similarity in their basic causal mechanisms (such as mitochondrial dysfunction, altered signal transduction, receptor activation, DNA repair inhibition, etc), and that cellular and organ system responses to maintain homeostasis in the face of chemical stressors are discrete and limited in number.

Predictive toxicology is not new, and indeed a major NTP predictive effort in the 1980s involved winnowing a large number of genetic toxicity assays down to only a few that seemed to be most predictive of cancer in rodents. Two subsequent predictive exercises followed in which a prospective challenge was issued to the scientific community to use whatever tools were at its disposal to predict the outcome of ongoing NTP rodent cancer studies. These challenges served to illustrate just how difficult it is to create a body of knowledge sufficient to allow one to predict an outcome as complicated as a disease endpoint. In general, predictive systems that relied solely on decision rules derived from toxicology and cancer data sets existing at that time, performed more poorly than those that relied on decision rules modified by human judgment. This suggests that the predictive models were not yet of sufficient sophistication to capture all the kinds of information needed to succeed. It is conceivable that the extent of information obtainable through genomic and proteomic analyses, high throughput mechanistic toxicology screens, and better toxicokinetic and metabolism information may ultimately strengthen these predictive knowledge bases and allow them to be used to predict rodent carcinogens, dysmorphogens, etc. But predicting rodent disease outcomes is not our primary goal. A serious limitation of all experimental animal-based procedures for predicting human responses is the influence of species- and strain-specific factors that affect either the qualitative or quantitative nature of the disease outcome. Moving from a disease-based science to a mechanism-based science could conceivably bypass these limitations.

It is assumed that the most promising predictive toxicology screens and databases will focus on general mechanisms of disease etiology common to both laboratory animals and humans. It is likely that these tools might be most effectively developed with an eye toward ones that have the capacity to assess a toxicological response in humans or in human tissues as well as in rodents or rodent tissues. It is probably unrealistic to expect that predictive systems, in a broad sense, will ever be developed to the point that they entirely replace in vivo toxicology and cancer studies. There will always be situations of exposure that occur in animals, but cannot be replicated in vitro. For example, continuous exposure of an animal during a two-year study provides exposures during all stages of development of a preneoplastic lesion to a fully developed neoplasm. This would be difficult to replicate in isolated, cell-based systems.

What would it take to move toxicology from a disease-based to a mechanism-based science? Clearly there would have to be broad scientific support for a set of biological observations that are considered to represent an adverse, and perhaps irreversible event. This will not be an easy task. The selection of predictive assays for adverse biological observations may benefit from an evaluation of the types of ancillary mechanistic information found to currently influence decisions in evaluating agents for listing in the Report on Carcinogens, or evaluating reproductive/developmental hazards through the CERHR evaluation process. An analogous discussion takes place currently for regulatory decisions that require distinguishing NOELs from NOAELs, but there is no doubt that

agreement on predictive endpoints of concern will be a major obstacle to achieving the NTP vision. There will also have to be a regulatory framework that allows the use of mechanistic information as an adequate basis for regulation. Recent examples of the use of mechanistic information in listing agents in the NTP Report on Carcinogens have withstood legal challenge and provide support to the notion that this is feasible. However, these documents do not directly support regulatory actions. Acceptance of mechanism-based observations in place of traditional disease endpoints will require that scientists gradually reorient their thinking to place a higher value on early biochemical or gene or protein expression changes than on traditional disease endpoints. It is likely that some laws will need to be rewritten if we are ever able to take full advantage of scientific advancements in predictive toxicology.

What would be lost by such a move? Probably the biggest gap in the regulatory process resulting from such a move would be deficiencies in the area of estimating and extrapolating quantitative risks to humans for specific diseases. A framework would need to be established for dealing with mechanism-based predictive data, in conjunction with toxicokinetic and other ADME data from short-term rodent studies, to establish an endpoint for quantifying an adverse outcome. To address the need for exposure assessment information for humans it will be necessary to view biological effects as a reflection of exposure and develop mechanism-based fingerprints from outcome data that can serve as surrogates of exposure.

Another major challenge to the establishment of a hazard identification process based on predictive, mechanism-based, biological observations is the requirement that such a system weigh positive findings of potential adverse events more heavily than the lack of such events. This is an inherent weakness in any predictive or reductionist system that relies on the presumed scientific understanding of a mechanism of adverse action. Reliance on a negative finding to establish a lack of hazard potential requires that the predictive system be designed such that it could and would, without failure, detect all relevant adverse events or mechanisms in all situations. The design of such a system would be essentially impossible. However, it may be possible to estimate a rough false negative error rate if a sufficiently large number of chemicals were assayed in tests with endpoints that were not generally associated with known adverse actions of those chemicals.

Acceptable performance characteristics of predictive systems is also an area on which agreement must be reached within the scientific community. All predictive systems will have inherent error rates, as do the current *in vivo* toxicology models used for predicting human responses. An important assumption in the NTP vision is that predictive systems designed to detect mechanism-based adverse biological outcomes will have inherently more power to predict similar human responses than do current *in vivo* or rodent disease-based models to predict human diseases. This has important implications with regard to

the approach one would take toward the design of any process that would evaluate and validate mechanism-based predictive systems for their public health value.

What effect would implementation of the NTP vision have on the structure and function of the NTP over the foreseeable future? Clearly implementation of the NTP vision will require a coordinated process of scientific and regulatory changes. ***Any movement would by necessity be incremental such that the products of NTP studies would remain relevant for scientific and regulatory needs, while also providing the science base to justify change.*** Thus, we envision a program that continues to provide data from high quality rodent studies and continues to provide scientifically sound evaluations of its own and other published literature, while fulfilling its mandate to establish new test methods through the development of predictive, mechanism-based analysis tools. Utilization of this new predictive toxicology information may be most effective for agents that already have a fairly comprehensive database, such as is found in the ATSDR Toxicology Profiles. Thus, as predictive tools are developed, it will be important to utilize chemicals with a rich existing toxicology dataset as well as "unknown" chemicals studied in a prospective manner.

Why is the NTP in the best position to lead this process? The NTP is neither a regulatory agency nor regulated industry. Thus, it has the latitude to adopt a more flexible approach to answering scientific questions than perhaps these aforementioned entities. The NTP is primarily located within the National Institute of Environmental Health Sciences, which has a strong basic research program focused on fundamental mechanisms of health and disease. The NTP also has a strong mandate to develop the field of toxicology through new methods and through utilization of the ICCVAM process, which provides a mechanism to facilitate the coordinated interagency evaluation and adoption of new toxicology methods. Careful consideration will need to be given to the role of the ICCVAM in implementation of the NTP vision. A clear understanding of where individual mechanistic toxicology screens would fit within a regulatory framework is necessary before any single test or battery of tests can be evaluated for their predictive value. Clearly the NTP has the obligation to provide the scientific justification for the use of new or alternative toxicology test methods within existing regulatory frameworks.

Implementation of the NTP vision will be a difficult and challenging activity. In the short term, success will be measured by the degree to which mechanism-based assays can be developed that are robust and reliable and that provide data accepted by the scientific community as valid for scientific and regulatory purposes. In the longer term, success will be measured by improvements in public health brought about by the prevention of exposures to agents that might have never been tested using our current low capacity, high cost methods.