Peer Review Comments on the *Strategic Plan for the Future of Toxicity Testing at the U.S. Environmental Protection Agency*

September 2008

Review by John R. Bucher, Ph.D.
Associate Director, National Toxicology Program, NIEHS

General Comments: This is a very nice, concise overview of an approach to transform EPA's current model of toxicity testing, and is generally consistent with the NRC report recommendations (NRC, 2007). It is clearly written, organized and well thought out. The plan is appropriately exceedingly ambitious and necessarily vague in many places, and does not try to hide this fact. The complexity of the current problem facing regulatory toxicology, and toxicology in general is, if anything, understated in the introductory materials. Toxicological testing methods in animals, in particular those outlined in regulatory guidelines, are being shown to be increasingly inadequate as research into the molecular and cellular impacts of chemicals on biological systems advances.

Specific comments:
At several points in the document (e.g. p 4, p 7, p 14, etc) the concepts of toxicity pathways, key events and mode of action are discussed in an integrated fashion. While these are clearly interrelated, it is unlikely that a close linkage of these concepts is going to be helpful in the initial interpretation of HTS data. In fact, the application of key events and MOA analysis to HTS output may hinder appreciation of the full range of information coming from these approaches. As a participant in the development and testing of MOA frameworks for experimental animal and human information, I believe our understanding of the various MOAs is far too rudimentary to encompass the 100 or so toxicity pathways as projected by the NRC Report. It is likely that "unsupervised" analyses will be far superior at identifying associations of endpoints that change in a manner that may be coordinated or linked to various phenotypes, and this information, rather than being interpreted in light of what we know, will more likely give us a better appreciation of pathways we haven't yet identified.

On p 6 the following statement appears: "Essential to this iterative process will be the demonstration that the predictive nature of these new approaches is superior to that of our current practices for toxicity testing and risk assessment." I suggest that this statement either be eliminated or expanded to better deal with the fact that the very nature of the information generated, certainly in the initial phases of the new approach, will be completely different from what we are used to. I would suggest that we have at best only a theoretical sense of the predictive nature of our current practices with regard to human health protection, and I think that the potential trade off between gleaning some predictive information for thousands of chemicals vs. what can be done under the current testing regimen cannot be taken into account if this requirement remains.

On p 10 the "current" paradigm for toxicity testing is characterized as including a genomic component and verification of gene expression changes with phenotypic

expression. While this would be optimal, it overstates the degree to which molecular changes are currently contributing to, and providing context for, understanding disease outcomes in traditional *in vivo* toxicology models. But this does point out a place where the concept of utilizing genomic methods in *in vivo* models could provide valuable information to improve the predictive power of the HTS methods. One of the often-unstated assumptions of the HTS approach is that immediate perturbations of toxicity pathways in cells *in vitro* will lead to changes that can be linked to chronic diseases. Our experience with developing predictive models relying only on genomic approaches suggests that initial changes in transcriptomics in *in vivo* studies are not predictive of longer-term outcomes, but that predictive models are much improved if one doses animals for 3 months before profiling genomic transcripts. A better understanding of the linkage between early and late events in relation to gene expression changes may provide information that would improve our ability to make sense out of HTS data.

On p 13, following the very nice section on virtual organs there is an opportunity to discuss the promise for development of HTS systems utilizing combined cell types to both accommodate intracellular signaling as well as deficiencies in metabolic capability of many individual isolated or cultured cells.

On p 13, in contrast to the virtual tissue section, the mention of biomarkers is vague and receives only passing mention of their someday being integrated into risk assessment. This is somewhat of a tangential issue and could be handled in subsequent iterations of the strategic plan if warranted by further developments.

You may want to add reference to the Japanese Center for Validation of Alternative Methods (JaCVAM) to the Appendix.

Edits
P2, para 1- The NIEHS Roadmap for the Future, should be the NTP Roadmap.
P 20, para 1, first line- change “the” to “their”.

Review by George Daston, Ph.D.
Research Fellow, P&G

I found this document to be a succinct, readable plan that clearly outlines the strategic direction EPA intends to take in exploring improvements in chemical risk assessment. I found the title to be a little too restrictive, though. While I understand the desire to tie the document to the "Tox Testing the 21st Century" report, the plan is beyond tox testing and really addresses the process of chemical risk assessment. I would recommend changing the title to reflect that fact. Perhaps it could be called "...Toxicity Testing and Risk Assessment Approaches..." or something like that. It isn't as pithy, but it is more accurate.

The actual meat of the plan is entirely consistent with the mission and goals of the Comp Tox program. In one way this is good, because the Comp Tox program should be leading the Agency in this effort. However, I am worried that it is almost exclusively Comp Tox-focused. I think it is important to stress the fact that the Comp Tox program is not only a research center but is also a hub in the EPA science network. Bring in examples in which other parts of the Agency, both inside and outside ORD, are adding value to this effort. Examples could be the expertise in in vivo models from the pesticides program in helping to design and interpret the first phases of ToxCast; the efforts in chemi-informatics, SAR, and QSAR from other parts of ORD and Toxics that contribute, via a Community of Practice, to the CompTox strategic goal of chemical prioritization; etc. I don't think this list needs to be comprehensive, but it needs to be enough to show 1) that this is EPA's plan, not just NCCT's plan; and 2) that the rest of the Agency is an active participant, not just the recipient of the Institutional Transition at some undetermined date.

The strategic plan has been well thought through. It clearly demonstrates that EPA got the core message from the NRC report and is already making great progress in translating the report's ideas into action. I believe that the strategy that is outlined here has already been well-reviewed and incorporates the thoughts of lots of smart people inside and outside the Agency. I don't have any substantive criticisms of the goals or approaches outlined in the document. I do have some specific comments, however, that may help clarify the message:

P. 4, section 2.2: I was left with a lot of questions as to exactly how pathway-based risk assessment was going to be done, and what the purpose was. Is it to use the new approaches to reduce uncertainties inherent in the currently used tox testing methods (i.e., are we still going to be doing the standard battery of in vivo tests, and the new tests are going to help us with extrapolations)? Or, are we going to be using the new methods to tailor the testing to better characterized the outcomes predicted by knowing the tox pathway targets? Or, is this an entirely new, alternative approach in which we won't be doing the traditional testing methods?

P. 7, lines 10-17 of the second paragraph: Please indicate that these critical components include understanding the **quantitative** relationships between molecular events and the higher order changes at which toxicity is manifested. Also in that paragraph, I think it would be useful to indicate that it may be possible to use this information to select more appropriate models for toxicity testing (e.g., those that express disease states similar to humans when a particular tox pathway is perturbed, or a model that is sensitized (or humanized) for a particular gene/ pathway that is representative of a sensitive subpopulation, etc.).

P. 8, first paragraph: I like this paragraph, and would also like to see a similar paragraph indicating that ToxCast data may be used to support the development of better chemi-informatic-based toxicity predictions.

P. 11, first line in section 4.1: I think there is a good chance for the word "postulates" to be misunderstood. Maybe it's just that I am helping my son too much with his geometry homework these days, but I think many readers would misconstrue postulates as something proven. Instead, try, "guesses" or "estimates" or provides the opinion".

P. 11, last paragraph: A key question that is worth mentioning here is that we still need to determine how far back in the progression from initial interaction of chemical with molecular target to toxicity (Fig. 2, e.g.) we will need to go to effectively define, detect, and model these toxicity pathways. If we only need to go back a step or two from toxic manifestation before we come to common pathogenetic steps (bottlenecks), the emphasis on in vitro or in vivo model development will be greater; if we need to go back several steps, the emphasis on chemi-informatics, 'omics, etc., may be greater. It would be useful to acknowledge this here, and indicate that these are questions that can be addressed through Actor and the virtual systems models.

P. 14, second paragraph in section 4.3: This paragraph ends with the sentence that dealing with extrapolation of non-homologous effects will be on a case-by-case basis, but I think this paragraph would be helped with a few examples. E.g., in instances in which the MOA is expected to be identical but pathology different, one would rely on MOA to predict human outcome.

Overall, I think this is a strong, well written, effective document. My comments are pretty minor. I am happy to expand or clarify any of my comments if you or any of the other authors have questions.

Review by Daniel Krewski, Ph.D., MHA
Professor and Director, McLaughlin Centre for Population Health Risk Assessment, University of Ottawa

As general comments, I found the Strategic Plan to be extremely well written. It not only captures the essence of the NRC vision (which appears to form the foundation for the plan) very nicely, it also extends the scientific thinking underlying the NRC vision in several areas.

The Strategic Plan is both comprehensive and complete, touching on organizational, infrastructure issues that will need to be addressed in moving forward with the plan, in addition to the core scientific issues.

I am attaching copies two articles that members of the NRC Committee on Toxicity Testing and Assessment of Environmental Agents have written since the publication of the two NRC committee reports:

D. Krewski, M. E. Andersen, E. Mantus, and L. Zeise. Toxicity testing in the 21st century: Implications for human health risk assessment. Risk Analysis. Accepted (likely to appear early in 2009).

M. E. Andersen and D Krewski. Toxicity testing in the 21st century: Bringing the vision to life. Toxicological Sciences, Submitted.

These articles also extend the thinking underlying the NRC Committee reports, but do not necessarily represent the views of the full committee. The first article will appear in Risk Analysis, along with a number of invited commentaries; pending final acceptance, the second article will appear in Toxicological Sciences with invited commentaries to be invited at a later date.

I'm also attaching a reprint of a series of opinions on the implications of the NRC vision for regulatory risk assessment, which appeared in the Environmental Forum earlier this year.

Thank you for the opportunity to comment on the EPA Strategic Plan; it was a please to review such a well-written document, which takes as its foundation our work through the National Research Council's Committee on Toxicity Testing and Assessment of Environmental Agents.

Specific comments on the plan

(Not all comments necessarily require a response; in some cases, I have highlighted what I consider to be particular strengths of the plan.)

Page iii: On page 2 of the report, the Future of Toxicity Testing Workgroup is stated to be cross-agency - would at least one co-chair from another agency be useful?

Page 1, paragraph 1: One of the most important challenges is the need for broader coverage of the perhaps 100,000 environmental agents to which people are potentially exposed, only a small minority of which have been adequately evaluated to date.

Page 2, 2nd full paragraph: I was unaware that such a group had been established - this is a natural, and probably necessary, step towards implementing the NRC vision at the national level

Table 1: Would it be too bold to say "Need to screen as many as 100,000 environmental agents. . ."?

Page 7, paragraph 2: Would it be better to say 'agent' rather than 'pollutant'? (Not all environmental chemicals will be pollutants per se.)

Page 8, line 2: This may be my single most important comment. The most common criticism of the NRC vision is the lack of in vivo integration using in vitro tests for pathway perturbations. However, if all critical perturbations are prevented, the non-additive effect of multiple perturbations will also be prevented. If the authors feel this notion has scientific merit, it might be worth mentioning to pre-empt some of the criticism that is certain to occur when a predominantly in vitro approach is proposed.

Page 8, “Simple and reliable screening models are needed that predict exposures to chemicals so that information from the full source-to-outcome continuum is brought into consideration in the evaluation of chemicals.”: Nice phrase for what I have been calling the 'exposure-response continuum'.

Page 8, figure 2: This is a very effective way of depicting the notion of pathway perturbations, which is central to the NRC report (no change suggested).

Page 9, paragraph 1: Post-marketing surveillance for adverse drug reactions will be of some use, but for a narrowly defined class of adverse health effects. However, this example is important, since the drug development process may provide additional information on the pathways by which ADRs observed under real world conditions are induced.

In my view, molecular and genetic epidemiology offer even greater opportunities for the identification of toxicity pathway perturbations directly in humans; such studies may also be useful in mapping toxicity pathways themselves, which is a prerequisite for implementing the vision.

Page 10, paragraph 2: Exposure assessment was largely outside the terms of reference of the NRC Committee (which was therefore not constituted to address this component of risk assessment). The inclusion of exposure assessment in the EPA operationalization of the vision is an important expansion of the NRC vision.

Page 10, figure 4: This figure is very elegant - it illustrates the role of cell signalling (a key concept that was perhaps not discussed in sufficient depth in the NRC report) within the sequence of events leading to adverse health outcomes very clearly.

Page 11 “It is important that increased emphasis be placed on examination of exposure concentrations that are more reflective of real-world condition.”: Excellent point! As highly sensitive (and hopefully specific) in silico and in vitro methods are developed, it will be possible to make direct inferences about the likelihood of toxicity pathway perturbations at environmentally relevant exposures.

Page 11, Section 4.1: This is one area in which inter-agency co-operation will be extremely valuable. NTP has a wealth of data that could populate such databases, as does the National Chemical Genomics Center.

It could be useful to involve the private section, particularly the pharmaceutical industry, although the issue of confidential business information will need to be addressed.

Page 13, last paragraph: Biomarkers used in molecular and genetic epidemiology may also provide valuable information for toxicity pathway identification.

Page 14 “Such efforts will help address the question of the extent to which "key events" that are predictive of ultimate endpoints (whether cancer or immunosuppression or kidney disease) must be demonstrated or whether the perturbation of baseline biological processes sufficient to induce substantial cellular level response (*e.g.*, a stress response) should be considered an adequate endpoint for risk assessment.”: Would it be worth also referring to these 'key events' as 'critical perturbations of toxicity pathways' to make the link to the NRC vision more direct?

Page 14, last paragraph: As the vision becomes more fully articulated, the notion of 'mode of action' may become subsumed within the notion of 'toxicity pathways'.

Page 16, last bullet: Interaction with international programs such as REACH will be extremely important in the interests of international harmonization. If the U.S. were to move forward with the NRC vision in isolation from the international community, the progress made towards international harmonization of toxicity testing and risk assessment guidelines, and the establishment of mutual recognition agreements for risk assessments done in different countries, would be compromised. Involvement of the international community in the implementation of EPA's strategic plan should be done at the earliest possible stage.

Page 17, paragraph 1: A collection of 5 or so case studies would be extremely valuable in demonstrating how the NRC vision is likely to be operationalized. This might be a good topic for a workshop, with a proceedings volume containing the detailed case studies, along with a synthesis of the lessons learned from the case studies.

Page 18, second full paragraph: As noted in section 5.1, the academic community will be an important participant in articulating the vision for the future of toxicity testing. Although specifics of how this might be done might may be too detailed for the present strategic plan, it would be worth thinking about specific mechanisms such as targeted collaborative research agreements - motivating the academic community to elaborate components of the NRC vision could prove to a cost-effective way of developing the science base on which the vision rests.

Page 20, paragraph 2: One possible mechanism that could be used for this purpose is to ask the NRC to establish an oversight committee charged with providing specific guidance and ongoing evaluation of progress, not just by EPA but across the scientific community in general. EPA has previously done this through the NRC Committee on Research Priorities for Airborne Particulate Matter, which provided oversight for a 13 year $440 million program in this area. The Committee published 4 reports that provided valuable direction on this important risk issue between 1998 and 2004.

Review by Martin Stephens, Ph.D.

Vice President for Animal Research Issues, The Humane Society of the United States

I commend the Environmental Protection Agency (EPA) for embracing the report of National Research Council (NRC), *Toxicity Testing in the 21st Century: a Vision and a Strategy*, by convening the Future of Toxicity Testing Workgroup, which has prepared the draft *Strategic Plan for the Future of Toxicity Testing at the U.S. Environmental Protection Agency*. As the agency that commissioned the NRC report, it is appropriate for the EPA to take a lead role in its implementation, notwithstanding the interest and activity of sister agencies and other potential partners in working with the EPA to turn the NRC vision into a reality.

I understand from the charge to reviewers that the Strategic Plan is intended to be "concise while providing a 'big picture' strategic view to facilitate the understanding of EPA's direction by its many stakeholders. As such, it does not present the level of detail (e.g., research projects, milestones) that would be articulated in a subsequent implementation plan, nor does the strategy articulate the level of funding that would be necessary to carry out the strategy." I have reviewed the document in this light but look forward to the issuance of an implementation plan and its particulars concerning specific deliverables, collaborations, timeline, and budget. These details will be a more concrete reflection of the agency's level of commitment to this priority, and better reveal how the agency will go about ushering in the tools for 21st century toxicology.

I think the key goals articulated in the strategic plan are well reasoned and very supportable, and reflect positively on the agency's willingness to embrace the latest science and technology in pursuit of its mission to protect human health and the environment against risks posed by environmental chemicals.

I first provide general comments and then address the specific questions posed by the charge to reviewers.

General Comments

Animal data as the reference standard: The pace of progress on protecting human health from exposure to hazardous environmental agents will depend on how quickly the agency's implementation of the NRC vision focuses on modeling the biology of humans (21st century toxicology), not animals (20th century toxicology). This entails, among other things, using human data on adverse effects to evaluate the relevance of the new pathways-based approach. In this light, I am concerned by several statements that imply that data from animals may continue to be regarded as the reference standard, such as the following:

- "As toxicity pathways are identified, relevant in vitro assays can be utilized and their results compared to in vivo studies as appropriate given the need to predict effects in humans or other species" (p. 7).

I welcome the discussion of human data (p. 14). However, even here, undue weight continues to be given to animal data. "Linkage is more complicated for effects observed in animals that may predict human effects that are related, but not identical to, the outcomes in animals (e.g., developmental effects in an animal model may predict developmental effects in humans, but the exact manifestation might be different). Cleary this aspect will need to be addressed on a case-by-case basis as we gain experience." Continuing to regard animal data as the starting point is not consistent with the NRC vision of human-based toxicology.

We agree that “Additional emphasis needs to be placed on toxicities demonstrated to occur in humans”, with efforts made to obtain data from human clinical trials or post-marketing surveillance (p. 9; p. 14).

Is there a role for careful microdosing in humans as a means for generating human data in developing the PBPK modeling data needed for risk assessment?

Minimizing the generation of new animal data: In light of our concern about using animal data as a reference standard, we do not want to see any emphasis placed on generating new animal data to advance the implementation of the vision. Consequently, we are concerned about statements such as the following:

- "Comparative approaches using samples from a range of species, including rodents (or other species used in toxicity testing), humans, and proposed alternative species (e.g., Danio rerio, zebrafish) could be valuable” (p. 9).
- "Development of these virtual tissue and organ systems will require newly generated data across phylogenetic systems to both fill data gaps identified within the iterative process and test the predictive nature of these virtual systems" (p. 12).
- "This translates to a continued role over the foreseeable future for in vivo systems in the development of this research strategy and implementation" (p. 13).

If EPA doesn't intend for "in vivo systems" to mean "animal tests," this should be clarified.

On a related matter, it is important to realize that the supplemental “targeted testing” envisioned by the NRC vision (p. 14) does not necessarily entail animal testing, especially as the new paradigm is progressively elaborated. Such targeted testing could also be in vitro.

Omics as an extension of animal testing: I am concerned about the following statement: "Following administration of the chemical to the test animal (usually at high

doses), genomic approaches are used to detect alterations in molecular pathways, the data are mined to describe the ensuing cellular alterations...." (p. 10). This would seem to position -omics research as an add-on to new animal studies, which would tie (i) the generation of genomic data to animal experimentation and (ii) data interpretation to rodent (vs. human) biology. This is neither necessary nor appropriate.

Human-derived cells and tissues: As mentioned above, the pace of future progress on the human health side will depend on how quickly the work focuses on modeling the biology of humans (21st century toxicology), not animals (20th century toxicology). Aside from moving swiftly to using human data on adverse effects as the reference standard for evaluating the relevance of the new pathway-based approaches, there is also the issue of the sources of the cells and tissues used as substrates for pathway testing. These should be human-derived, as underscored by the NRC vision. This point should be emphasized in Section 4.3 (p. 14).

Gradualism: The plan states that "advances are likely to be gradual over the next decade or two" (p. 15). However, progress is likely to accelerate as unforeseen technological developments remove bottlenecks, as in the Human Genome Project. In fact, we recommend that the Human Genome Project be cited as an example of a "big biology" project akin to the one envisioned in the NRC report. We further recommend that the present effort be entitled the Human Toxicology Initiative.

Operational transition: The basis of the NRC vision is that adverse effects in whole organisms can be predicted from underlying perturbations to toxicity pathways (the Plan's Figure 1). For most such pathways, it is possible that not all intense perturbations will lead to adverse effects--biology is not chemistry. Thus the new approach, which regulates on the basis of pathways, is likely to have an element of precaution. EPA should anticipate and plan its operational transitions accordingly. Similarly, the transition to regulating on the basis of pathway perturbations versus pathological effects may require changes to some of the operating guidelines and policy documents cited in the report. A paradigm shift should not be expected to conform to all existing standards.

Outreach: I applaud the attention to stakeholder involvement, transparency, and outreach (p. 17, 18). This will promote constructive feedback on the agency's own efforts as well as synergistic partnerships with stakeholders.

Specific Questions to Reviewers

1. *Is the paper written in a clear, concise, and readable manner? If not, please provide detailed comments*.

I believe the paper is written largely in clear, concise, and readable manner, given its scope as a high-level strategic plan to be followed up by an implementation plan. The

EPA should clarify whether "in vivo systems" is being used to refer to "animal tests" or more broadly (see General Comments).

2. *Are there additional activities that you believe EPA should consider to facilitate the transition to a new toxicity testing paradigm? If so, please describe these activities.*

The plan should clarify the extent to which the envisioned research will be conducted in-house or via external contracts, and how external partners could work with EPA to facilitate implementation of the vision.

3. *Is the balance and relative timing of activities related to chemical prioritization, quantitative risk assessment, and regulatory use appropriate?*

There is a danger that too much emphasis will be placed on chemical prioritization and, as a result, the strategic plan's implementation would stall at this stage. If that happened, the EPA would miss the opportunity to create a far-reaching paradigm shift. Prioritization should be seen as a side-benefit off of the main road leading to realization of the NRC vision.

Support for the EPA strategic plan will increase to extent that efforts move swiftly from screening and prioritization (p. 7) to risk assessment (p. 10). If that transition is slow, the plan will be seen largely as a means to tee-up animal testing, not to replace it.

4. *Are there additional issues regarding the transition to a new toxicity testing paradigm that you believe should be noted in the paper? If so, please describe these issues.*

Chemical coverage: The traditional approach to toxicity testing has substantial practical shortcomings, including low throughput (p. 1). These shortcomings would be overcome by the envisioned paradigm shift. This would mean, for example, that thousands of untested chemicals could be practically tested. Consequently, the benefit of the new paradigm is, in part, its ability to generate meaningful data where none exists today. The issue is not only the envisioned superiority of the data generated in the new versus animal data. This advantage should be part of the EPA's outreach communications (p. 17).

Ecological risk assessment: I'm pleased to see EPA's plan to implement the 21st century vision in a manner that encompasses ecological risk assessment as well as human health (p. 9), given the recent moves to markedly expand ecotox data requirements in the pesticide/endocrine areas. However, whereas there is considerable discussion of establishing human relevance on the human health side, the report does not explicitly address creating toxicity pathways based on the biology of the target species of ecotox testing. This should be better addressed in the plan.

Evaluation/validation: The strategic plan has little to say on the key question of validation of the new paradigm; indeed, it seems to avoid use of that term. Proof of concept studies are mentioned (p.16). ICCVAM is listed as a potential partner in this effort (Table 1). The plan should be more explicit on what will be needed for validation when a wholly new paradigm is developed. Surely, pre-existing approaches, if applied, would present an enormous logistical and financial bottleneck to progress.