Compilation of Preliminary Individual Comments from Members of the Chartered SAB and SAB Chemical Assessment Advisory Committee on New Approach Methods and Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing

(As of June 20, 2020)

COMMENTS FROM CHARTERED SAB MEMBERS 2

Dr. Hugh Barton **2**

Dr. Samuel Cohen **6**

Dr. Tony Cox **8**

Dr. Kimberly White **11**

COMMENTS FROM SAB CHEMICAL ASSESSMENT ADVISORY COMMITTEE MEMBERS 13

Dr. Dennis Paustenbach **13**

Dr. Ted Simon **16**

Dr. Eric Smith **18**

Comments from Chartered SAB Members

Dr. Hugh Barton

New Approach Methods and Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing

Hugh A. Barton

June 18, 2020

1. EPA-OPP is participating in the Rethinking Carcinogenicity Assessment for Agrochemicals Project (ReCAAP) with government, non-governmental organization, and industry stakeholders (Section 2). ReCAAP is developing a risk-based weight of evidence (WOE) approach for waiving chronic and carcinogenicity studies. This proposed approach is consistent with existing guidance and current practice for other types of toxicology studies.
    a. Please comment on the draft risk-based WOE approach for waiving chronic/ carcinogenicity studies (Attachment 1). Please include in your comments a discussion of the clarity and completeness of the proposal.

The draft risk-based WOE approach lays out reporting for a variety of information to consider for determining whether to waive chronic/carcinogenicity studies. It is essentially a format to be followed for preparing such a request. What it does not provide is information about how such a decision would be made other than a sentence here or there. Section 4.4 on Evidence of Immune Suppression begins with the statement "In the absence of genotoxicity, hormonal effects, or liver enzyme induction, indications of immunosuppression could raise concern for potential tumor formation." This is useful information, but nothing similar is said in the sections on genotoxicity or hormonal effects. The statement in Section 4.5 Genetic Toxicity is completely confusing – "If the chemical is mutagenic, then the evaluation is complete and no further documentation is needed (the pesticide would not undergo a cancer bioassay if the chemical is mutagenic, so no need to write a full carcinogenicity waiver)." Mutagenic chemicals would generally be of great concern, so it is unclear what this sentence means.

    b. Please comment on the draft case study provided (Attachment 2). Please include in your comments a discussion of the clarity and completeness of the proposal.

The case study is useful but seems such a simple example that it provides limited guidance for making decisions when results are more complex. It is noted in the white paper that additional case studies are in development for trying out the WOE approach; this would be valuable.

2. EPA is collaborating with Division of the National Toxicology Program (DNTP) of National Institute of Environmental Health Sciences (NIEHS) and HESI to consider NAM-based approaches to begin to replace the chronic/carcinogenicity testing in mammals. In addition, EPA-OPP and ORD are working together to collect quantitative gene expression data from short-term *in vivo* rat studies for a selected set of pesticides that cause liver

tumors in rodent with known modes of action. All of these efforts are in the early stages of development and would benefit from expert and public input.

a. Please comment on the direction and scope of the three collaborative projects described in Section 3 of the draft white paper.

The three collaborative projects appear to be valuable efforts to move the disciplines of toxicology and risk assessment as well as the organizations dependent upon them, such as EPA, towards new alternative approaches to assessing and protecting human health. These projects are useful and deserving of sufficient support to allow them to make a difference in the next few years.

3. EPA is working with HESI, NTP, and other government and industry stakeholders to accelerate the incorporation of kinetically-derived maximum doses (KMD) into repeat dosing studies like the chronic/carcinogenicity study as an alternative to the traditional maximum tolerated dose (MTD). The KMD approach is consistent with numerous guidance documents developed by EPA, OECD and other international organizations as a more humane and human relevant approach to dose selection. One KMD study has been provided to the SAB along with the description and agenda of an upcoming workshop and the scope/charge of a workgroup at HESI to develop additional case study and a best practices document.

   a. Please comment on EPA's current KMD-related activities as described in Section 4 and Attachments 3 and 4. Does the SAB have additional activities that EPA could consider?

Choices of the doses to use in toxicity studies have long posed a challenge. A kinetically maximum dose has been proposed as an additional selection criteria. The general concern is that toxicities observed at very high doses compared to human exposure may not be as indicative of potential human health effects as toxicities observed at lower doses. The materials provided do not make a strong case for use of kinetics as a determinant, but rather text in the white paper and the provided example argue for the importance of toxicodynamics or mode of action in making determinations about dose selections.

As higher internal concentrations of chemicals or drugs are obtained in toxicity studies with higher administered doses, several things may occur. There may be changes in how the body handles a chemical (pharmacokinetics or toxicokinetics) and/or what the chemical does to the body (pharmacodynamics or toxicodynamics also referred to as mode of action). Changes in pharmacodynamic processes at higher doses may result in adverse effects that would not occur at lower doses, so these effects have little or no relevance for human health risk assessment with low enough exposures. Some changes in pharmacodynamics, such as depletion of the scavenger molecule glutathione, can also change pharmacokinetics impacting whether appropriate extrapolations to lower doses are feasible, as in the example provided. Saturation of pharmacokinetic clearance processes (e.g., metabolism, urinary excretion) result in changes in the relationship between the external administered dose and the internal concentration time course of parent compound and metabolites. This may effect appropriate extrapolation to lower doses and make pharmacokinetic modeling desirable, but does not by itself determine the human relevance of

the findings. Higher internal concentrations in the absence of changes in toxicodynamic processes can result in observable effects that are unobservable at lower doses given the number of animals in the study and the duration of exposure and observations. Detection of such effects would generally be considered advantageous, since the biological processes occurring are the same as those occurring at lower concentrations but would require much larger numbers of animals to be observable at lower concentrations.

One challenge for using saturation of pharmacokinetics as a characteristic to limit the highest dose in a toxicity study is that pharmacokinetics alone does not identify when observed toxicities are not predictive for humans. The example provided (EPA-HQ-OPP-2013-0154-0104) describes decisions about a KMD for dichloropropene. It argues for selection of a KMD based upon nonlinear kinetics assessed in a steady state study in mice (p33) and then argues that the benign lung adenomas arose only at doses exceeding the KMD. However, the key issue is that glutathione depletion is observed (section 5B p35). Glutathione depletion impacts the pharmacokinetics because the major route of metabolism is glutathione conjugation (section 5A p 35), but importantly glutathione depletion is a major protective pathway in the toxicodynamics. With depletion it is expected that toxicodynamic processes would occur that would be much less likely with adequate glutathione present leading to the lung adenomas. Glutathione depletion, like excessive body weight loss, should be considered a criterion that would limit use of such high doses in toxicity studies or, if the study were already done as is the case for dichloropropene, argue that the data from that dose are too confounded to be appropriate to extrapolate to lower doses or to humans. Thus, the changes in toxicokinetics are here being used as an indicator of when changes are occurring in toxicodynamics. It would be preferable to use the data on glutathione depletion directly.

Another issue with the proposals around the KMD is how the shift from linear pharmacokinetics to saturation tends to be described and then the statistical methods applied. Saturable processes are most frequently well described mathematically by a rectangular hyperbola, often referred to as a Michaelis-Menten equation, referring to a common formulation in enzymology. This is a smooth curve that at low concentrations, well below the mathematical parameter describing the concentration giving half-maximal activity (called Km), behaves essentially linearly within measurement error. As one goes to higher concentrations, the curve increasingly deviates from the low dose linear behavior, effectively bending over from a straight line to result in a near plateau at the maximal activity. There is no "inflection" point as would be created at the point where two straight lines with different slopes join (hockey stick model), but rather a steadily increasing difference between a straight line and a hyperbola that bends over.

Thus, the question is how much of a deviation from the low dose linearity is considered “too much” deviation so one should not use such a dose; this appears a matter of choice. A widely used rule of thumb for experimental designs to determine the parameters of the Michaelis-Menten equation or equivalently equations for specific binding or pharmacological effects (Emax equations), is that one needs data spanning a range from 10-fold above the half maximal parameter (i.e., Km, or Kd, or EC50) to 10-fold below. The 100-fold range centered on the half maximal parameter covers the approximately linear center portion of the sigmoid shaped curve observed when the x-axis is plotted on the log scale. The differences between the curve shapes on the normal and log scales for the x-axis leads to confusion about what are approximately linear portions of the curve. Fitting sparse pharmacokinetic data to obtain an “inflection” point seems a questionable approach.

One avenue for consideration in the various ongoing KMD activities, is whether the issues might be better addressed by focusing on the activities moving towards use of NAMs or transcriptomic PODs. Concerns over how to select doses in a traditional toxicity study, in particular a chronic/cancer study, should hopefully become a lower priority as one moves towards newer methods.

Dr. Samuel Cohen

New Approach Methods in reducing the Use of Laboratory Animals

for Chronic Carcinogenicity Testing.

1. The format outlined for the study waiver request presents the needed information in a clear and comprehensive manner. I strongly support the agency in its attempts to no longer require the two-year bioassay in mice and rats for carcinogenicity testing. Not only is this wasteful of resources, including the use of too many animals, it has little predictive value for human carcinogenic risk, particularly for chemicals that are nongenotoxic. I strongly support the concept of using reference doses for carcinogenicity assessment for nongenotoxic chemicals with appropriate safety factors, since the toxicities that occur prior to development of cancer are necessary for ultimate development of the cancer. Since these are precursor lesions and are noncancerous endpoints, they can be handled the same way as other noncancer toxicities. Focusing on genetic toxicity, immunosuppression and hormone perturbation are appropriate, especially for toxicity assessment. However, for cancer assessments, the only hormonal perturbation known to be predictive of human cancer risk is estrogenic activity related to breast, endometrial, and to a much lesser extent, liver and possibly ovarian cancer. Other hormonal perturbations, although of consideration regarding toxicity, do not predict carcinogenic activity in humans. Probably the best example of nonrelevance are the numerous chemicals involved in producing thyroid follicular tumors in rats, all of which ultimately are related to hypothyroid induction with increased TSH stimulation and thyroid follicular proliferation. There are innumerable epidemiology studies indicating that hypothyroidism is not a risk factor for human thyroid cancer. The focus on immunosuppression is particularly important, since many of the immunosuppressive agents used in clinical medicine are actually negative in the two-year bioassay. Immunosuppression of any kind, whether inherited, secondary to therapy for transplantation, cancer, or autoimmune diseases, or due to AIDS, all increase the risk of virally related cancers and a few others. A two-year bioassay for such agents is a waste of resources. A few specific details could also be included in the listing. Some mention should be made about the quality of studies, although for pesticide evaluations these will often be GLP. In evaluation of genotoxicity, it should be stated that appropriate negative *in vivo* studies will usually negate positive findings in the *in vitro* studies.

The examples that were provided are quite good, particularly the bladder and calculus example for cancer assessment and the use of KMD in addition to MTD for assessment of dose. My only comment regarding the bladder cancer example is that there actually are examples of carcinogenicity testing of sulfonamides, dating back to the 1960s. Although these were not GLP studies, they clearly showed an increased incidence of bladder tumors in mice. The chemical was 4-ethylsulphonylnaphthalene-1-sulphonamide (ENS) (Br. J. Urol., 26 :26-34, 1964), and was the example for which David Clayson ultimately hypothesized that the tumors induced by chemicals related to calculi were due to the calculi and not to the chemical itself, this dating back to 1974 (Clayson, JNCI, 52: 1685-1689, 1974). The conclusion is that this is quantitatively not relevant to humans given the differences in exposure. However, one could also argue that qualitatively this is not relevant to human exposures either (Cohen, Toxicol. Res., 7: 565-575, 2018). For EPA and its focus on risk-based assessment, this distinction is not important. However, for agencies, such as the European Chemical Agency, which is hazard based, this can become a significant issue for classification.

2. The efforts of the EPA's collaborations with NTP and HESI to consider NAM-based approaches is strongly supported. For now, a combination of short-term *in vivo* studies with appropriate *in vitro* studies can readily accomplish a screening for carcinogenicity. I would recommend that the studies in mice for carcinogenicity be completely eliminated since they are of no predictive value for humans beyond that which is provided from studies in rats. The EPA might reexamine the basis for the non-use of mice in European pharmaceutical evaluations in the 1980s and 1990s, prior to ICH. They only incorporated the mouse into the overall evaluation when the international harmonization efforts were undertaken due to the requirements by both the United States and Japan for two species. Some of the significant issues that still need to be addressed in these collaborative efforts is the focus on screening for tissues other than the liver. Most of the studies so far have investigated predictive values of genomic and other methods utilizing liver carcinogens. However, this will not screen for carcinogens for a number of other tissues, keeping in mind that there are very few chemicals actually associated with human liver cancer (ethanol, aflatoxin, estrogen). How does the agency propose to screen for carcinogens with target organs beside the liver? Rodent assays will be of little value since the major cancers in humans frequently are not affected in the rodent bioassay, or the rodent model is not predictive of changes in humans. For example, colon, stomach, and pancreas ductal adenocarcinomas are rarely produced in the rodent bioassay with any chemicals, particularly not with nongenotoxic chemicals. Likewise, the prostate in the rodent is not similar to the human prostate either morphologically or endocrinologically. The list could go on extensively. I would encourage the collaborations to develop human cell-based assays that would be more predictive of human risk. However, for nongenotoxic chemicals, it is the toxicity in short-term assays, whether in human cells or in animals, that could provide the basis for a risk assessment since protecting for the short-term toxicities will also protect for the ultimate development of cancers. (For further details, see Cohen, Toxicol. Res., 7: 565-575, 2018; Toxicol. Sci., 80: 225-229, 2004; Toxicol. Pathol., 38: 487-501, 2010; Cohen et al., Reg. Toxicol. Pharmacol., 103: 100-105, 2019)

3. I strongly support the efforts by the EPA and its collaborators in further developing the KMD-related evaluations. The example that was provided for dichloropropene highlights the advantages of such an approach to assess overall risk. Clearly, as long as exposures in humans will be below the KMD, studies performed at doses higher than the KMD are irrelevant to human risk. Combining this with the efforts in reducing the number of animals, a short-term evaluation for KMD in addition to MTD dose determinations could be used in establishing appropriate doses for the short term studies, and completely eliminating the two-year bioassay as described above.

Samuel M. Cohen, MD, PhD
Professor, Department of Pathology and Microbiology
University of Nebraska Medical Center

Dr. Tony Cox

Preliminary comments in response to the charge questions for the SAB consultation on New Approach Methods and Reducing the use of Laboratory Animals for Chronic and Carcinogenicity Testing

1a. Please comment on the draft risk-based WOE approach for waiving chronic/ carcinogenicity studies (Attachment 1). Please include in your comments a discussion of the clarity and completeness of the proposal.
1b. Please comment on the draft case study provided (Attachment 2). Please include in your comments a discussion of the clarity and completeness of the proposal.

The proposed WoE approach is mostly clearly explained, and the motivation and rationale provided seem to me to be compelling and admirable. However, the proposal is not fully clear and complete in the following areas:

- Section 4 (Toxicity).
    - Attachment 1 says that "Indications of immunosuppression could raise concern for potential tumor formation." How should it be determined whether such concern is warranted by the evidence? Does a feeling of concern have relevance for risk assessment, or is the risk assessment to be driven by data independent of feelings of concern that it might not or might not engender?
    - Some immune response-related endpoints may be extremely sensitive to even low levels of exposure to some chemicals, and yet be irrelevant for risk assessment. (For example, some peripheral blood lymphocytes may respond to very low concentrations of benzene, and yet have no relevance for risk of acute myeloid leukemia.) How should the relevance of such sensitive responses for risk assessment be determined?
    - A change that "could raise concerns for potential tumor formation" in a non-specialist might not do so in a specialist who understands why the change is irrelevant for tumor formation. In such cases, is the concern itself to be used as a basis for decision-making, or should the possibility of concern instead be discussed by the registrant, and reasons for sharing or rejecting the possible concern be explained?

- Section 5 (Evidence of Chronic Toxicity from Related Chemicals). Attachment 1 says "As outlined above, providing a rationale for why the indicator molecule(s) were chosen as the best comparators to the candidate molecule is an important element of this section." How is "best" defined here? Might the "best comparators" still not be very good? Are there objective tests or principles that should be applied to select indicator molecule(s) as comparators?

- Section 6 (Proposed Points of Departure and Prospective Risk Assessments). Attachment 1 says "Calculate estimates for cancer risk (the Margin of Exposure) – by

linear or non-linear cancer risk assessment methods as appropriate for the molecule." How should the registrant (and EPA) determine which methods are "appropriate for the molecule"? Might several different methods be appropriate? If there is uncertainty about which methods are most appropriate, how should that uncertainty be addressed?

- Section 7 (Conclusion). How is a conclusion supposed to be derived (and supported) from the data considered? The template in Attachment 1 says "**Based on a WOE approach, the registrant requests that the chronic/carcinogenicity toxicity studies [be/not be (*as appropriate*)] required at this time for [Chemical X].** This approach considered all of the available hazard and exposure information for [Chemical X], including: [provide a summary of why studies should not be required]." This leaves unclear exactly how the registrant (and the EPA) should get from "all of the available hazard and exposure information" to a specific request.

    - Might the same body of information lead to different requests by different parties?
    - On what basis should it be decided whether a request should be granted based on the data provided? In other words, this final step, of getting from data considered to the action requested, needs to be explained further.

4. *EPA is collaborating with Division of the National Toxicology Program (DNTP) of National Institute of Environmental Health Sciences (NIEHS) and HESI to consider NAM-based approaches to begin to replace the chronic/carcinogenicity testing in mammals. In addition, EPA-OPP and ORD are working together to collect quantitative gene expression data from short-term in vivo rat studies for a selected set of pesticides that cause liver tumors in rodent with known modes of action. All of these efforts are in the early stages of development and would benefit from expert and public input.*

    a. *Please comment on the direction and scope of the three collaborative projects described in Section 3 of the draft white paper.*

- Characterizing the predictive validity (rather than just the descriptive validity) of NAM-based approaches using gene expression data is critically important for assessing the practical value of this approach. What are its false-positive and false-negative rates (and how should these be defined and estimated from realistically limited data)?

- How does the value of information (VoI) provided by the NAM-based approaches compare to the VoI from traditional approaches? (Animal testing at relatively high concentrations has, arguably, not been very informative about human responses at realistic concentrations, although the decades-old debate on this point continues. Are NAM-based approaches demonstrably better, or at least not demonstrably worse?)

- Is it worth considering that some chemicals might be carcinogenic at toxic concentrations but not at relevant environmental concentrations?

5. *EPA is working with HESI, NTP, and other government and industry stakeholders to accelerate the incorporation of kinetically-derived maximum doses (KMD) into repeat dosing studies like the chronic/carcinogenicity study as an alternative to the traditional maximum tolerated dose (MTD).*

    a. *Please comment on EPA's current KMD-related activities as described in Section 4 and Attachments 3 and 4. Does the SAB have additional activities that EPA could consider?*

- Great idea! This is important work and appears to be well thought-out. An additional activity to consider might be to carefully examine the effects of *timings* of repeated exposures (as well as their concentrations), as they relate to non-proportional responses. For example, dose fractionation and stop-exposure experiments on various chemicals have shown that the same cumulative exposure per unit time (e.g., per week) can have large or small toxic effects, depending on how it is distributed over time. This may require considering pharmacodynamics and well as pharmacokinetics. (See e.g., Figures 3 and 4 of "Implications of nonlinearity, confounding, and interactions for estimating exposure concentration-response functions in quantitative risk analysis www.sciencedirect.com/science/article/pii/S0013935120305314.)

Dr. Kimberly White

Charge questions for the SAB consultation on new approach methods and reducing the use of laboratory animals for chronic and carcinogenicity testing.

1. Question: Please comment on the draft risk-based WOE approach for waiving chronic/ carcinogenicity studies (Attachment 1). Please include in your comments a discussion of the clarity and completeness of the proposal.

Answer: The "Draft Risk-Based Weight of Evidence Framework for Chronic/Carcinogenicity Studies with Agrochemicals" is clear regarding what information would be required of a registrant submitting a waiver request. However, it does not provide information on the criteria that the agency may use to review the information or support granting the waiver (i.e. is there a base level set of information that would be required for the waiver to be considered complete). Perhaps the approach could include an additional section between I and II that addresses this component. Additionally, section II.4 Toxicity, requests the registrant to "summarize how available studies can be used to inform chronic outcomes." This section should also consider including whether this information is specific to the chemical being assessed or is data also permitted for a surrogate chemical that is anticipated to act similarly to the chemical under review.

While there is a separate section II.5 which discussed read across information it was unclear if this information also applied to section II.4. In several sections of the draft risk-based WOE approach it requests information on mode of action data and associated key events. The document should consider highlighting specific mode of action frameworks that have already been accepted by the agency. As noted in the draft whitepaper provided to the SAB, the "CARC has evaluated tumor mode of action data for > 60 pesticides and has accepted the proposed MOA and conducted non-linear dose-response assessment for > 50 pesticides." Additionally, it is unclear how or when information about study quality should be presented throughout section II or if full study reports or manuscripts would also be requested as supporting information.

2. Question: Please comment on the draft case study provided (Attachment 2). Please include in your comments a discussion of the clarity and completeness of the proposal.

Answer: Some additional areas for clarity in the draft case study include the following: In Section 1. Use and Exposure Profile – additional information regarding other potential exposure routes (e.g. inhalation, dermal). The current draft notes "All exposure scenarios, including dietary (food and water), residential, aggregate, and occupation" but doesn't identify what those exposure routes are. In Section 4. Toxicity – Acute toxicity information is summarized in paragraph form and then a summary table in appendix B. Does this information need to include more specific study data and associated references?

Another potential consideration for the case study is whether it should include a summary table of the weight-of-evidence in the Section 7, Conclusion, that includes the lines of evidence available, relevancy of that data to the studies being requested to waive, the strength of that

evidence and the reliability/uncertainty associated with this evidence

3. Question: Please comment on the direction and scope of the three collaborative projects described in Section 3 of the draft white paper.

Answer: The direction and scope of the three collaborative projects to help support efforts to replace the use of animals in chronic/carcinogenicity testing appears reasonable. Some thoughts are provided below for consideration:

- DNTP Efforts to Improve Carcinogenic Assessment of Environmental Substances

– An important component of this project as noted in focus area 1 will be "Developing a translational toxicology pipeline (TTP) of capacities to characterize the potential for environmental substances to cause or contribute to the development of cancer." Having an understanding of what translational changes and at what level they represent an adverse impact will be important. Additionally, understanding the role of reversibility of any identified change and impacts to understanding the development or progression of cancer will also be important.

- Health and Environmental Sciences Institute (HESI) Point of Departure Program Overview

– Case study examples demonstrating how the PODs can be established will be important, perhaps this will be included in the 2021 manuscript. As well, addressing the issue of non-correlation of adverse effects with mode of action information and if that impacts the confidence of the established POD.

- Gene Expression Evaluation of Pesticides with Established Liver Tumor Modes of Action

– The agency should provide more details about the specific deliverables of this project.

4. Question: Please comment on EPA's current KMD-related activities as described in Section 4 and Attachments 3 and 4. Does the SAB have additional activities that EPA could consider?

Answer: The project description should be updated to include specific deliverables, outcomes and overall timeline for the effort. For example, will the case studies that will be identified result in a publication or whitepaper; will the planned workshop result in a publication; are there communication or education aspects of this project that should be included; are there any specific challenges to implementation that have been identified?

Comments from SAB Chemical Assessment Advisory Committee Members

Dr. Dennis Paustenbach

New Approach Methods and Reducing The Use of Laboratory Animals for Chronic and Carcinogenicity Testing

My Overview on The Document: I have been interested, as a toxicologist, in reducing the number of animals involved in testing new chemicals and I have encouraged the profession to move away from the more painful ones for the animals. My post-doc in 1986 focused on how to use physiologically-based pharmacokinetic approaches to select doses for chronic bioassays. Interestingly, it has taken about 30 years for those thoughts to take hold.

Interestingly, even though I have lobbied to significantly reduce the number of animals tested, I still believe that given our current level of understanding of the chemicals, it is premature for the agency to arbitrary say that there were will no more or very limited testing by 2035. We currently don't have a sufficient scientific bases for doing that and I believe it is implausible that we can rely on in-vitro and in-silico tests alone for safety evaluations.

I wholeheartedly support using kinetics to select the maximum dose in animal testing. MTD has existed for far too long and its use has often led us down the wrong paths.

Using computers, zebra fish, in-vitro assays, gene arrays, and subtle perturbations to guide us to understand the safety in humans seems like a worthy management objective that deals with the concerns of many stakeholders but it is unlikely that these, alone, will allow us to safely bring chemicals to the marketplace.

There is an undertone in toxicology today, and it is reflected in this document, that these in-silico and isolated cell perturbation assays will be effective at performing the hazard identification step in the risk assessment. This is possible and I would support it. However, for these to be used to give us insight about the dose-response relationship in mammals is probably asking for too much. The number of compensatory mechanisms and feed-back looks that can occur in-vivo vs the information gleaned from cellular assays (which don't have compensatory mechanisms) are many. Thus, I would urge that EPA "not" continue to say that they will achieve their goal by a particular date.

If the data are truly convincing and there are not too many exceptions as to when you can extrapolate, then it can be decided what a modern era test program can look like. In my experience, you can learn a vast amount by conducting whole body, in-vivo, assays in a much more limited number of animals; but we need animals.

My impression is that the Agency recognizes that if you declare NOELs in in-vitro assays as the Point of Departure, you will end up with remarkably low regulatory guidelines which, by definition, will be protective of the public. We have seen these inferences made in many dozens of in-vitro and in-silico papers of the past 15 years. However, if we just want to choose very low numbers, using any approach possible, I would dare say that you won't need many toxicologists to do that and the nation will surely decrease the number of new chemicals which can be used in the coming years.

I don't think it should be the role of EPA to stifle technological advancement by setting limits on chemicals that are unnecessarily low just because it can be done. Whenever we have done that, there have been unintentional consequences that prevent significant advancement. With the technological age at its infancy, especially with respect to microdevices and in developing vaccines and antibiotics, EPA would want to be nervous about unintentional "holding back" some important discoveries.

So, as having watched this initiative "get legs" over the past 25 years, and having served as chairman of one of the BOSC committees to look over this work, I would just add a word of caution that EPA's first objective is protect the safety of our citizens and that this should not be compromised by an elected official or an appointed bureaucrat who is well intended (and wanting to be popular among certain stakeholders). The Agency needs to be driven by the facts and it could take many years to really understand how far these modern tests can take us.

Part B

I did not see charge questions so have no comments on that front. I did read the **Rethinking Carcinogenicity Assessment for Agrochemicals** and have some comments.

a) As an engineer, and then a toxicologist, I am very fond of example problems for illustrating the strengths and weaknesses of an approach. So, I thought this exercise was valuable.
b) The objective was to allow the reader to decide if a 2 year bioassay was needed for this new chemical.
c) It was wise to start with the exposure assessment.
d) It was also wise to introduce more toxicologists to the important role of chemical and physical properties. This topic is nearly always understated with respect to its importance.
e) It is unclear how the ADME info was incorporated into the decision making other than to know it did not accumulate in tissues with repeated exposure.
f) The example suffers having a control group of limited size (see subchronic testing). In my experience, in recent years, it is wise to have twice as many controls as you use in the treatment groups.

g) Moving to the conclusions. It is interesting that the risk assessment is driven by data from the in-vivo studies and that were was virtually no reliance on gene arrays or subtle in-vitro tests for a myriad of effects.
h) I am not sure a great argument was made for NOT conducting a 2 year bioassay but, perhaps, that is the direction we are headed.

Overall, this is an interesting analysis but very few of the NAM methods were applied.

Dr. Ted Simon

Charge Questions on White Paper "New Approach Methods and Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing

1a. Risk-based WOE approach for waiving chronic carcinogenicity studies.

Comment 1: An exposure-based triage is missing from this first section. For many chemicals, exposure estimates may be available from far-field studies [2–4]. Any extant exposure estimates should be used and if unavailable, the methods used to develop them could be used within a read-across type inter-chemical extrapolation for exposure [5–10].

These exposure estimates could then be used as a prioritization/triage scheme in the development of PODs (#6, Proposed Points of Departure).

Comment 2: I am in favor of the use of MOE to address cancer risk, which appears to be the statement in the 7th bullet in section 6. The qualification of "…by linear or nonlinear cancer risk assessment methods as appropriate …" is also unclear. How would the choice of methods be determined?

I would propose the use of a WOE for this determination as proposed by several authors in the scientific literature [11; 12]

1b. Case Study with "Herbicide1"

Comment 1: A table of the exposures for all the scenarios would be helpful in "1. Use and Exposure Profile."

Comment 2: Essentially, this section used the exposure estimate for infants as the highest. Hence, consistent with my comment above, what's missing from the first section is this infant dose.

2a. Direction and Scope of Section 3 in the white paper

Comment 1: I disagree with the use of the key characteristics of cancer (KCCs). The KCCs have no better predictive ability than random chance, as demonstrated with a set of chemicals identified by EPA-OPP CARC as carcinogenic or not [13].

3a. Use of the KMD

Comment 1: I agree with proposals outlined in attachments 3 and 4.

References

1. Carli G, Cecchi L, Stebbing J, Parronchi P, Farsi A (2020) Is asthma protective against COVID-19.

Allergy
2. Wambaugh JF, Wang A, Dionisio KL, Frame A, Egeghy P, Judson R et al. (2014) High throughput heuristics for prioritizing human exposure to environmental chemicals. Environ Sci Technol 48: 12760-12767.
3. Wambaugh JF, Wetmore BA, Pearce R, Strope C, Goldsmith R, Sluka JP et al. (2015) Toxicokinetic Triage for Environmental Chemicals. Toxicol Sci 147: 55-67.
4. Wambaugh JF, Setzer RW, Reif DM, Gangwal S, Mitchell-Blackwood J, Arnot JA et al. (2013) High-throughput models for exposure-based chemical prioritization in the ExpoCast project. Environ Sci Technol 47: 8479-8488.
5. Sipes NS, Martin MT, Kothiya P, Reif DM, Judson RS, Richard AM et al. (2013) Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays. Chem Res Toxicol 26: 878-895.
6. Patlewicz G, Ball N, Booth ED, Hulzebos E, Zvinavashe E, Hennes C (2013) Use of category approaches, read-across and (Q)SAR: general considerations. Regul Toxicol Pharmacol 67: 1-12.
7. Patlewicz G, Roberts DW, Aptula A, Blackburn K, Hubesch B (2013) Workshop: use of "readacross" for chemical safety assessment under REACH. Regul Toxicol Pharmacol 65: 226-228.
8. Patlewicz G, Ball N, Becker RA, Booth ED, Cronin MT, Kroese D et al. (2014) Read-across approaches--misconceptions, promises and challenges ahead. ALTEX 31: 387-396.
9. Hartung T (2016) Making big sense from big data in toxicology by read-across. ALTEX 33: 83-93.
10. Shah I, Liu J, Judson RS, Thomas RS, Patlewicz G (2016) Systematically evaluating read-across prediction and performance using a local validity approach characterized by chemical structure and bioactivity information. Regul Toxicol Pharmacol 79: 12-24.
11. Becker RA, Dellarco V, Seed J, Kronenberg JM, Meek B, Foreman J et al. (2017) Quantitative weight of evidence to assess confidence in potential modes of action. Regul Toxicol Pharmacol 86: 205-220.
12. Dekant W, Bridges J, Scialli AR (2017) A quantitative weight of evidence assessment of confidence in modes-of-action and their human relevance. Regul Toxicol Pharmacol 90: 51-71.
13. Becker RA, Dreier DA, Manibusan MK, Cox LAT, Simon TW, Bus JS (2017) How well can carcinogenicity be predicted by high throughput "characteristics of carcinogens" mechanistic data? Regul Toxicol Pharmacol 90: 185-196.

Dr. Eric Smith

Comments from EP Smith - New Approach Methods and Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing

1. EPA-OPP is participating in the Rethinking Carcinogenicity Assessment for Agrochemicals Project (ReCAAP) with government, non-governmental organization, and industry stakeholders (Section 2). ReCAAP is developing a risk-based weight of evidence (WOE) approach for waiving chronic and carcinogenicity studies. This proposed approach is consistent with existing guidance[1] and current practice[2] for other types of toxicology studies.
    a. Please comment on the draft risk-based WOE approach for waiving chronic/ carcinogenicity studies (Attachment 1). Please include in your comments a discussion of the clarity and completeness of the proposal.
    b. Please comment on the draft case study provided (Attachment 2). Please include in your comments a discussion of the clarity and completeness of the proposal.

Comments on: Attachment 2 June 2020 SAB ReCAAP Waiver_Draft Final
Information in the tables related to statistical summarization or testing is lacking. I would suggest requiring that information. In particular

1. When there is an interval what is the confidence level, or are the values mean plus or minus standard deviation or standard error?
2. When p-values are reported, what is being tested? It is not always clear.
3. Samples sizes should be included. Typically, these studied are balanced so the sample sizes should be the same for each dose. If not the same, there should be a comment as to why not.

2. EPA is collaborating with Division of the National Toxicology Program (DNTP) of National Institute of Environmental Health Sciences (NIEHS) and HESI to consider NAM-based approaches to begin to replace the chronic/carcinogenicity testing in mammals. In addition, EPA-OPP and ORD are working together to collect quantitative gene expression data from short-term *in vivo* rat studies for a selected set of pesticides that cause liver tumors in rodent with known modes of action. All of these efforts are in the early stages of development and would benefit from expert and public input.
    a. Please comment on the direction and scope of the three collaborative projects described in **Section 3 of the draft white paper**.

[1] https://www.epa.gov/sites/production/files/2015-04/documents/data-require-guide-principle.pdf ; https://www.epa.gov/sites/production/files/2014-02/documents/part158-tox-data-requirement.pdf
[2] Craig et al (2019) Reducing the Need for Animal Testing While Increasing Efficiency in a Pesticide Regulatory Setting: Lessons From the EPA Office of Pesticide Programs' Hazard and Science Policy Council. Regul Toxicol Pharmacol, 108, 104481 Nov 2019. PMID: 31546018

3. EPA is working with HESI, NTP, and other government and industry stakeholders to accelerate the incorporation of kinetically-derived maximum doses (KMD) into repeat dosing studies like the chronic/carcinogenicity study as an alternative to the traditional maximum tolerated dose (MTD). The KMD approach is consistent with numerous guidance documents developed by EPA, OECD and other international organizations as a more humane and human relevant approach to dose selection. One KMD study has been provided to the SAB along with the description and agenda of an upcoming workshop and the scope/charge of a workgroup at HESI to develop additional case study and a best practices document.

    a. Please comment on EPA's current KMD-related activities as described in Section 4 and Attachments 3 and 4. Does the SAB have additional activities that EPA could consider?

Comments regarding: Agenda and description of the upcoming workshop the on kinetically-derived maximum dose (Attachment 3 and 4)

One of the objectives is Conducting statistical analyses to determine a KMD for interpreting dose-response data which is good. Would there be an opportunity for designing experiments, especially computer aided experiments be useful? For example, what is a good design for determination of dose proportionality?
Should the potential effect of interactions be considered?

Case study 7: Statistical tests to determine KMD from sparse data points – again it may be useful to consider different experimental designs, especially adaptive designs for estimation of KMD.

There is some literature on the problem that should be considered in the workshop:

L.G. McFadden, M.J. Bartels, D.L. Rick, P.S. Price, D.D. Fontaine, S.A. Saghir Statistical methodology to determine kinetically derived maximum tolerated dose in repeat dose toxicity studies Regul. Toxicol. Pharmacol., 63 (2012), pp. 344-351.

They argue that an appropriate method is to fit linear then add a quadratic term and evaluate if it is statistically significant. Design comes important here as dose spacing and number of replicates can affect significance. Guidance on design could be valuable.

The paper below seems relevant.
Minne B. Heringa, Nicole H.P. Cnubben, Wout Slob, Marja E.J. Pronk, Andre Muller, Marjolijn Woutersen, Betty C. Hakkert, Use of the kinetically-derived maximum dose concept in selection of top doses for toxicity studies hampers proper hazard assessment and risk management, Regulatory Toxicology and Pharmacology,Volume 114, 2020, 104659,

https://doi.org/10.1016/j.yrtph.2020.104659.

The authors argue that KMD is ill-advised for top-dose estimation. The following is indicated in their paper

"KMD concept aims at having top doses in toxicity tests below non-existing inflection point. The KMD leads to lower test doses, resulting in less informative or inconclusive data. Testing at too low doses does not meet 3R principle and has regulatory consequences."

Note also the recent paper

Marjolijn Woutersen, Andre Muller, Marja E.J. Pronk, Nicole H.P. Cnubben, Betty C. Hakkert, Regulating human safety: How dose selection in toxicity studies impacts human health hazard assessment and subsequent risk management options, Regulatory Toxicology and Pharmacology, Volume 114, 2020, 104660,https://doi.org/10.1016/j.yrtph.2020.104660.

Both of these papers indicate the need to have the design and analysis of studies, as well as the interpretation, have an important role in the workshop and guidance documents. Perhaps a part of the workshop and document would include discussion of criticisms of the approach.

Comments on: EPA-HQ-OPP-2013-0154-0104: 1,3-Dichloropropene: Report of the Cancer Assessment Review Committee

There are a considerable number of statements of the form: statistically significant ($p>0.05$) that are not supported with additional information. There has been considerable pushback in the statistical and scientific community about the use of p-values and the reliance of statistical significance as a surrogate for scientific significance (Ioannidis 2005, 2019). The American Statistical Association has published two editorials on the topic with some recommendations (Wasserstein and Lazar, 2016, Wasserstein, Schirm, and Lazar, 2019). In particular, the p-value by itself gives little information. When sample sizes ar small, a large effect is needed for statistical significance. If sample sizes are large then a small effect would give statistical significance. Hence without sample size or effect size, there is difficulty in interpreting the p-value. Consider implementing some of the recommendations perhaps requiring effect size, sample sizes and/or confidence intervals rather than just p-values.