UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

August 28, 2020

EPA-SAB-20-011

The Honorable Andrew R. Wheeler
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: Consultation on New Approach Methods and Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing

Dear Administrator Wheeler:

EPA's Science Advisory Board held a public meeting on June 23 - 24, 2020, and conducted a consultation with EPA staff on new approach methods and reducing the use of laboratory animals for chronic and carcinogenicity testing. Members of the Science Advisory Board's Chemical Assessment Advisory Committee also participated in the consultation.

The Science Advisory Board Staff Office has developed the consultation as a mechanism to provide individual expert comments for the EPA's consideration early in the implementation of a project or action. A consultation is conducted under the normal requirements of the Federal Advisory Committee Act (FACA), as amended (5 U.S.C., App.), which include advance notice of the public meeting in the Federal Register.

No consensus report is provided to the EPA because no consensus advice is given. Individual written comments were requested from all members of the Science Advisory Board and the Science Advisory Board Chemical Assessment Advisory Committee. The EPA's charge questions for the consultation are provided in Enclosure A. The individual written comments that were received from EPA Science Advisory Board members are provided in Enclosure B, and the individual comments that were received from members of the Science Advisory Board's Chemical Assessment Advisory Committee are provided in Enclosure C.

We thank the EPA for the opportunity to provide early advice on new approach methods and reducing the use of laboratory animals in chronic and carcinogenicity testing.

Sincerely,

/s/

Dr. Michael Honeycutt, Chair
EPA Science Advisory Board

/s/

Dr. Hugh A. Barton, Chair
SAB Chemical Assessment Advisory Committee

Enclosures

NOTICE

This report has been written as part of the activities of the EPA Science Advisory Board (SAB), a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The SAB is structured to provide balanced, expert assessment of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names of commercial products constitute a recommendation for use. Reports of the SAB are posted on the EPA Web site at http://www.epa.gov/sab.

U.S. Environmental Protection Agency
Science Advisory Board

CHAIR
Dr. Michael Honeycutt, Division Director, Toxicology Division, Texas Commission on Environmental Quality, Austin, TX

MEMBERS
Dr. Rodney Andrews, Director, Center for Applied Energy Research, UK Research, University of Kentucky, Lexington, KY

Dr. Hugh A. Barton, Independent Consultant, Mystic, CT

Dr. Barbara Beck, Principal, Gradient, Boston, MA

Dr. Deborah Hall Bennett, Professor, Environmental and Occupational Health Division, Department of Public Health Sciences, School of Medicine, University of California, Davis, Davis, CA

Dr. Frederick Bernthal, President Emeritus and Senior Advisor to the Board of Trustees, Universities Research Association, Washington, DC

Dr. Bob Blanz, Associate Director, Office of Water Quality, Division of Environmental Quality, Arkansas Department of Energy and Environment, North Little Rock, AR

Dr. Todd Brewer, Senior Manager, Grants, Education, and Utility Programs, American Water Works Association, Denver, CO

Dr. Joel G. Burken, Curator's Professor and Chair, Civil, Architectural, and Environmental Engineering, College of Engineering and Computing, Missouri University of Science and Technology, Rolla, MO

Dr. Janice E. Chambers, William L. Giles Distinguished Professor and Director, Center for Environmental Health and Sciences, College of Veterinary Medicine, Mississippi State University, Mississippi State, MS

Dr. John R. Christy, Distinguished Professor of Atmospheric Science and Director of Earth System Science Center, University of Alabama in Huntsville, Huntsville, AL

Dr. Samuel Cohen, Professor, Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE

Dr. Louis Anthony (Tony) Cox, Jr., President, Cox Associates, Denver, CO

Dr. Alison C. Cullen, Interim Dean and Professor, Daniel J. Evans School of Public Policy and Governance, University of Washington, Seattle, WA

Dr. Otto C. Doering III, Emeritus Professor, Department of Agricultural Economics, Purdue University, W. Lafayette, IN

Dr. Susan P. Felter, Research Fellow, Global Product Stewardship, Procter & Gamble, Mason, OH

Dr. Joseph A. Gardella, SUNY Distinguished Professor of Chemistry, Department of Chemistry, College of Arts and Sciences, University at Buffalo, Buffalo, NY

Dr. John D. Graham, Professor, School of Public and Environmental Affairs, Indiana University, Bloomington, IN

Dr. John Guckenheimer, Professor Emeritus, Cornell University, Ithaca, NY

Dr. Margaret M. MacDonell, Department Head, Argonne National Laboratory, Lemont, IL

Dr. Robert E. Mace, The Meadows Center for Water and the Environment, Texas State University, San Marcos, TX

Dr. Clyde F. Martin, Horn Professor of Mathematics, Emeritus, Department of Mathematics and Statistics, Texas Tech University, Crofton, MD

Dr. Sue Marty, Senior Toxicology Leader, Toxicology & Environmental Research, The Dow Chemical Company, Midland, MI

Mr. Robert W. Merritt, Independent Consultant, Houston, TX

Dr. Larry Monroe, Independent Consultant, Braselton, GA

Dr. Thomas F. Parkerton, Senior Environmental Scientist, Toxicology & Environmental Science Division, ExxonMobil Biomedical Science, Spring, TX

Dr. Robert Phalen, Professor, Air Pollution Health Effects Laboratory, School of Medicine, University of California-Irvine, Irvine, CA

Dr. Kenneth M. Portier, Independent Consultant, Athens, GA

Dr. Robert Puls, Owner/Principal, Robert Puls Environmental Consulting, Bluffton, SC

Dr. Tara L. Sabo-Attwood, Professor and Chair, Department of Environmental and Global Health, College of Public Health and Health Professionals, University of Florida, Gainesville, FL

Dr. Mara Seeley, Unit Chief – Exposure Assessment, Environmental Toxicology Program, Bureau of Environmental Health, Massachusetts Department of Public Health, Boston, MA

Dr. Anne Smith, Managing Director, NERA Economic Consulting, Washington, DC

Dr. Richard Smith, Professor, Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC

Dr. Jay Turner, Professor and Vice Dean for Education, Department of Energy, Environmental and Chemical Engineering, McKelvey School of Engineering, Washington University, St. Louis, MO

Dr. Brant Ulsh, Principal Health Physicist, M.H. Chew & Associates, Cincinnati, OH

Dr. Donald van der Vaart, Senior Fellow, John Locke Foundation, Raleigh, NC

Mrs. Carrie Vollmer-Sanders, Director, Agriculture Engagement Strategy, Efroymson Conservation Center, The Nature Conservancy, Indianapolis, IN

Dr. Kimberly White, Senior Director, Chemical Products and Technology Division, American Chemistry Council, Washington, DC

Dr. Mark Wiesner, Professor, Department of Civil and Environmental Engineering, Director, Center for the Environmental Implications of NanoTechnology (CEINT), Pratt School of Engineering, Nicholas School of the Environment, Duke University, Durham, NC

Dr. Peter J. Wilcoxen, Laura J. and L. Douglas Meredith Professor for Teaching Excellence, Director, Center for Environmental Policy and Administration, The Maxwell School, Syracuse University, Syracuse, NY

Dr. Richard A. Williams, Retired Economist (Food and Drug Administration), Independent Consultant, McLean, VA

Dr. S. Stanley Young, Chief Executive Officer, CGStat, Raleigh, NC

Dr. Matthew Zwiernik, Professor, Department of Animal Science, Institute for Integrative Toxicology, Michigan State University, East Lansing, MI

SCIENCE ADVISORY BOARD STAFF
Dr. Thomas Armitage, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC

U.S. Environmental Protection Agency
Science Advisory Board
Chemical Assessment Advisory Committee

CHAIR
Dr. Hugh A. Barton, Independent Consultant, Mystic, CT

MEMBERS
Dr. Richard Belzer, Independent Consultant, Mt. Vernon, VA

Dr. Tiffany Bredfeldt, Senior Toxicologist, Toxicology, Texas Commission on Environmental Quality, Austin, TX

Dr. Karen Chou, Associate Professor, Department of Animal Science and Environmental Science & Policy Program, College of Agriculture and Natural Resources, Michigan State University, East Lansing, MI

Dr. Harvey Clewell, Principal Consultant, Ramboll Environment and Health, Research Triangle Park, NC

Dr. Joanne English, Independent Consultant, Menlo Park, CA

Dr. David G. Hoel, Distinguished University Professor, Department of Biometry and Epidemiology, Medical University of South Carolina, Charleston, SC

Dr. Jacqueline Hughes-Oliver, Professor, Statistics Department, North Carolina State University, Raleigh, NC

Dr. Michael Jayjock, Exposure/Risk Assessor, Jayjock Associates LLC, Langhorne, PA

Dr. Wayne Landis, Professor and Director, Institute of Environmental Toxicology, Huxley College of the Environment, Western Washington University, Bellingham, WA

Mr. Charles McConnell, Director, Center for Carbon Management, Office of the Chancellor, University of Houston, Houston, TX

Dr. Dennis Paustenbach, President, Paustenbach and Associates, Jackson, WY

Dr. Isaac Pessah, Professor, Molecular Biosciences, School of Veterinary Medicine , University of California, Davis, Davis, CA

Dr. Joann Brooks Powell, Assistant Professor, Center for Cancer Research and Therapeutic Development (CCRTD), Clark Atlanta University, Atlanta, GA

Dr. Thomas Rosol, Chair and Professor, Biomedical Sciences, Ohio Musculoskeletal & Neurological Institute, Ohio University, Athens, OH

Dr. Ted Simon, Toxicologist, Ted Simon, LLC, Winston, GA

Dr. Eric Smith, Professor, Statistics, College of Science, Virginia Tech, Blacksburg, VA

Dr. Chandrashekhar Sonwane, Scientist, Masten Space Systems, Chatsworth, CA

Dr. Laura Vandenberg, Associate Professor, Environmental Health Sciences, School of Public Health and Health Sciences, University of Massachusetts - Amherst, Amherst, MA

SCIENCE ADVISORY BOARD STAFF
Dr. Suhair Shallal, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC

Enclosure A

The EPA's Charge Questions

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

In accordance with the September, 2019 directive from EPA Administrator Andrew Wheeler, EPA's Office of Pesticide Programs (OPP) and Office of Pollution Prevention and Toxics (OPPT) are working to reduce the number of laboratory animal studies requested or required for pesticides and industrial chemicals. Beyond the ethical issues associated with animal use, new approach methods (NAMs) are expected to improve the scientific foundation of risk assessments by providing human-relevant information that is more efficient and less costly. In collaboration with the Office of Research and Development and multiple stakeholders, EPA's Office of Chemical Safety and Pollution Prevention (OCSPP) has developed a draft white paper highlighting three projects that are improving the science used in risk assessment for chronic/carcinogenicity testing. These activities are organized by the 3Rs principles for laboratory animal testing-- *reduce, replace, refine* as originally proposed by Russell and Burch. Because of the complexities in biology and toxicology, there will not be a "one-size-fits-all" solution to improving chronic/carcinogenicity testing. As such, EPA and its collaborators are taking a multifaceted approach that advances several areas simultaneously. The agency requests the SAB provide comment on the following charge questions.

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.

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

[1] https://www.epa.gov/sites/production/files/2016-01/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

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.

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?

Enclosure B

Individual Comments from Members of the EPA Science Advisory Board on New Approach Methods and Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing

Dr. Hugh Barton

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 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. Janice Chambers

Chronic/Carcinogenicity Waivers:

Question 1:

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 attachment seems to cover most of the information that would need to be analyzed to approve a chronic/carcinogenicity test waiver, but the statements are quite brief and do not provide any guidance to EPA staff regarding their analysis of these data. A WOE approach is reasonable and could definitely be science-based, but could be subject to judgment calls if guidance and rationales are not carefully presented. More detail would be useful.

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 makes a well-documented WOE case for waiving the chronic/carcinogenicity tests based on knowledge of rodent physiology and toxicology with other compounds and read-across as possible.

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

The shorter-term gene expression approach to predicting carcinogenic potential has considerable merit, especially if some well documented commonalities are found among toxicants resulting in the common endpoint considered here, i.e., liver tumors. However there are many uncertainties involved in extrapolating from short term changes in gene expression (which may or may not be related to the mechanism leading to tumors) and the EPA staff indicated that it was not essential to know that the short term changes had any role in causing tumors. This leads to a lot of uncertainty in using this as an indicator of carcinogenic potential unless the data sets are very well correlated with essentially no exceptions to the correlations on the short term measures. Considerable analysis and consideration of the physiology reflected by these changes in gene expression must be made before accepting short-term changes as a surrogate for long-term effects.

3. 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?

The case study is very interesting and well described. It makes a very good case for a KMD which assesses the biology behind the kinetic results and makes an effective case for the KMD instead of the MTD as the high dose in a study. This new approach has considerable potential for more science-based risk assessments on carcinogenicity compared to the earlier approach that indicates the adverse results seen at excessive dose levels must be considered in the risk assessment. One example that the SAB considered was the ETBE/tBA review that the SAB did a couple of years ago. One of the very contentious issues that came up was the tumors induced at only the high dose that probably exceeded the MTD. A KMD approach may have generated better data that could have been easier to interpret.

General comment: There was much to consider on these new approaches and very little time to consider them. Because these would rewrite the long-standing guidance that EPA has been following for risk assessments, these new approaches should be considered in a fashion that is not rushed and allows the SAB to provide more thoughtful advice than is currently possible.

Dr. Samuel Cohen

New Approach Methods in reducing the Use of Laboratory Animals for Chronic Carcinogenicity Testing.

I strongly encourage EPA not to use the approach referred to as the 10 key characteristics of carcinogens. This approach does not specifically address the science related to carcinogenesis and does not address mode of action analysis. It has become essentially a check box exercise that does not incorporate the best scientific evaluation with regard to risk assessment. For example, Simvastatin, one of the statins that has been used by literally millions of individuals around the world to reduce cholesterol, produces liver tumors in males and females, rats and mice, and yet it is widely used in human medicine. It has seven of the listed key characteristics of carcinogens. Utilizing a mode of action/human relevance analysis as described by EPA, Health Canada, and IPCS shows that the rodent liver tumors are not relevant to human cancer risk. Based on such an analysis, Simvastatin was approved by FDA and subsequently by agencies around the world. If the key characteristic approach had been taken, Simvastatin would be classified as a human carcinogen based on the strongly positive animal studies and the number of key characteristics present. This, despite the fact that mode of action shows that it is not relevant to humans and the epidemiology strongly indicates that it is not a risk. This is only one of many examples where the key characteristics approach leads to inappropriate conclusions.

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.

Dr. Tony Cox

General Comments

Following is a thought on how to minimize animal testing while preserving it as a last resort option if needed. This is motivated in part by Dr. Dennis Paustenbach's observation that the goal of minimizing animal testing can be pursued within a scientific framework that considers the value of information that such testing produces in the context of information available from other sources.

A data-driven approach to minimize animal testing without eliminating it in the (probably few) cases where it might be justified by high value of information for human health and lack of alternatives is as follows. The main idea is to use optimization to design least-cost decision/classification (CART-type) trees in which a sequence of costly tests is performed until it is optimal to stop and draw a conclusion. Animal tests might appear in a few places in that tree. A new element compared to existing CART algorithms is to consider (a) the costs of collecting different types of information needed to progress through the tree; and (b) costs of errors in the final conclusions or decisions reached. An optimization algorithm can be used to design trees that minimize expected total cost (the sum of information costs and decision error costs).

Increasing the cost of animal testing to reflect ethical concerns would shift the optimized trees toward using less animal testing. In this approach, one could quantify the tradeoff between cost of animal testing and expected cost of error, taking into account the availability of other sources of information (and their costs and values of information). Such an optimization approach would allow for optimized testing, with animal testing as an allowed component, but one that is only used when use is justified by the information it produces and the lack of better alternatives.

I think we could make such an approach rigorous as a way to minimize use of animal testing while recognizing that it may sometimes have value in the context of limited alternatives.

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.

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.**

- 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?

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).**

    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. Susan Felter

General Comments:

(1) I am very supportive of EPA's research and efforts to reduce the use of laboratory animals in toxicity testing, as well as to reconsider approaches for dose-setting (i.e., the KMD research). That said, this should still be recognized as a goal and not a mandate.

(2) One over-arching principle should be that the goal is not to replicate the findings of a (high-dose) rodent bioassay (or any other high-dose toxicity study). This makes it challenging since there is a need to 'validate' NAMS against existing data. We have a long history in toxicology of testing at the highest dose we can, whether this is the MTD in a rodent cancer bioassay, or a cytotoxicity (lethality)-based limit in an in vitro genotoxicity study. One of the rationales for this was to ensure that we didn't miss something of concern (hazard identification). One of the consequences of this is that we continued to generate data at very high doses – sometimes doses that are impossible to achieve in humans – and then had to figure out what to do with those data. As we consider NAMS, a strong emphasis needs to be put on appropriate dosing that will provide meaningful data to help ensure our ability to protect human health while not generating data that has no relevance to human health and can actually have unintended consequences ranging from perceived safety concerns by the public to possible bans. The need to consider appropriate dosing in NAMs applies to in vivo studies (including short-term toxicogenomic studies) to in vitro studies. Information regarding phys/chem properties, toxicokinetics, and human exposure should all be taken into account.

(3) In general for NAMs, it is important that a precedent does not get set that would lead to any change (e.g., in transcriptomics) being defined as adverse. Results from in vivo studies reflect complex physiology, including multiple compensatory mechanisms and feedback loops that are critical to maintaining homeostasis and avoiding toxicity. It will be important to consider how these translate to results in short-term (toxicogenomics) studies or even in vitro assays designed to detect activity.

(4) It is great to see collaborations between EPA/OPP and a number of external partners in NAMs-related research to move away from current requirements for chronic rodent bioassays. It is also important to ensure collaborations across EPA Offices to ensure that the Agency does not end up in a situation where approaches taken by one Office are in conflict with those taken by another Office.

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.

I am supportive of a WOE approach for waiving chronic/ carcinogenicity studies and overall find the framework to be helpful. Specific suggestions on the information needed in a study waiver request include:

1. Use pattern & exposure scenario: I am very glad to see this highlighted as the first step. It is not clear the extent to which quantitative exposure estimates are needed – to the extent possible, these should be included.

3 (ADME/TK)

- The bullet on metabolism indicates that metabolites formed in the environment should also be included. While it's important to evaluate this, it might be best to add a separate section related to environmental fate, particularly since the description for Section 3 states that the information in this section "should be presented as it relates to the potential for chronic toxicity."
- The second bullet under "Toxicokinetics" refers to nonlinearities as reflected by a lack of proportionality of the AUC vs dose. This should be tied into the discussion of the KMD.

4.4: Consider changing to 'Immunotoxicity' rather than limiting this to immune suppression.

4.5 (Genetic toxicity): Suggest expanding the 2nd bullet to indicate that, in case of positive genotoxicity data, a rationale should be provided as to why the overall WOE is still that the chemical does not pose a genetox concern.

4.6 (Special studies and endpoints): The first bullet in this section is for "Mode of action studies." Given the importance of MOA, it seems this should have its own section, and then Sections 4.3 (hormone perturbation), 4.4 (immune suppression), and 4.5 (genetic toxicity) would be bullets within that section. For each of these, human relevance should be discussed/addressed. The MOA section should also draw upon what is already known about pesticides and MOAs that have been evaluated and accepted by OPP (the White Paper indicates this is > 50 pesticides) to support a nonlinear dose-response assessment.

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.

---

[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

It is very helpful to have a draft case study to illustrate how the WOE analysis would be performed and how the available data can be used to support a waiver. Additional case studies with chemicals with different MOAs and differing amounts of data will be helpful to further 'test' the framework, including cases where a waiver might not be supported. This might lead to further questions of whether additional studies are needed to fill data gaps (vs. making a decision that a chronic/carcinogenicity study is needed).

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.**

These three projects (referred to below as "DNTP", "HESI POD" and "Gene Expression") are each focused on gaining a deeper understanding of aspects of the cancer risk assessment process that will ultimately lead to improvements. A few comments on each follow:

- DNTP: I am fully supportive of the new directions being taken by NTP under the HEI program. I do have a concern about one of the next steps, described as "evaluating the association between the Key Characteristics of Cancer and outcomes in NTP rodent bioassays for chemicals in the Tox21 inventory." While the KCC approach can be helpful as a framework to understand the MOA for chemical carcinogenesis, it is also raises concerns about mis-use for hazard identification. It is critically important that this evaluation be extended to chemicals that are negative in a rodent bioassay as well.

- HESI POD: The use of NAMs (e.g., toxicogenomics) to develop PODs for use in risk assessment is a critically important area – not only does it offer the potential for significant reduction in use of animals, but it also offers a relatively fast way to generate data on an untested chemical that can be critical to our understanding and ability to establish acceptable exposure limits. That said, we are still in the early stages and a lot of work remains to answer some critical questions. The White paper recognizes the challenge that a transcriptomic approach to derive a POD will be just that (a transcriptomic POD), and we will not have an apical endpoint associated with it. I am not particularly concerned about that (we often don't see concordance between a rat and a mouse), but do have concerns relating to how we will ensure that the organs/cells tested to generate the transcriptomic data represent the breadth of sensitivity of different organs. On the flip side, how do we ensure that the approach is not overly conservative given that a 5-day study might show the initial transcriptomic changes associated with chemical exposure, but these will not necessarily translate to any adverse effect?

- Gene expression: This work will help link an apical endpoint (rat liver tumors) and its MOA (genotoxicity, cytotoxicity, and activation of AhR, CAR, ER, or PPARα) with the underlying changes in gene expression. In addition to the qualitative understanding of these connections, the ability to identify biological thresholds for events associated with the tumorigenic responses will significantly advance our understanding of what is happening at lower doses and may help support the move away from testing at high doses because of the historical desire to make sure we aren't missing something that could be of concern for human health (e.g., tumors associated with high-dose testing).

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?

I am fully supportive of EPA's KMD-related research. Using the KMD will support EPA's commitment to reducing the use of animals in research because it will avoid generating data that are not relevant to humans and yet often require follow-up studies to explain the high-dose effects. It will also help avoid a situation where a chemical is 'labelled' (i.e., hazard identification / classification) with a hazard (e.g., "carcinogen") when, in fact, there is no hazard under human-relevant exposure conditions. The example provided (1,3-dichloropropene) is helpful in illustrating both the concept and the advantage to using the KMD. A few comments:

- It should be clear that sometimes the highest dose will be the KMD and sometimes it will be the MTD, and/or should be informed by what we know of human exposure (e.g., set the high dose at a level xx-fold higher than human exposures).

- This concept should also be applied to short-term (toxicogenomics studies) and other repeat-dose toxicity studies, and even in vitro assays.

Dr. Sue Marty

General Comments:

This reviewer agrees with EPA's statement, "Because of the complexities in biology and toxicology, there will not be a "one-size-fits-all" solution to improving chronic/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.**

      The EPA has defined a number of parameters that are considered during these cancer assessments (e.g., physical-chemical properties, ADME and toxicokinetics, toxicity, and conclusions via a weight-of-evidence (WoE) approach, but it is not clear how the decision on waiving the chronic/carcinogenicity study will be made. Clarification on the decision matrix and rationale are needed.

      During the call in June, the Chartered SAB discussed the importance of validation (i.e., against positive bioassays). However, it also would be valuable for the EPA to develop some case studies where waiving the chronic/carcinogenicity study is not appropriate and thereby, demonstrate to stakeholders that such compounds will be identified during the assessment process and will not receive chronic/carcinogenicity study waivers. While chronic/carcinogenicity study waivers may be relatively rare initially, this approach may be leveraged more often as agrochemical registrants gain experience, so it will be important to establish that all compounds will not make it through the waiver process.

Framework template document (Attachment #1) appears to be reasonably complete. Mode-of-action data may be particularly relevant as EPA evaluates eligibility for waiving the chronic/carcinogenicity study. More specific comments on the Framework template document appear below:

II. Study Waiver Requests:

- Section 3 - ADME and Toxicokinetics: Parent compound and metabolites are required in mammalian species. Is it worth specifying human comparative metabolism? This is

[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

frequently done as part of global registrations and would be important to confirm relevance of animal species tested.

- Section 4.3 – Evidence of Hormone Perturbation: For clarification, consider making the last 3 bullets in this section 'sub-bullets' under "Report evidence of hormonal perturbations…". Clarify 'weight change' – Is this weight change of endocrine organs (presumably not body weight)? Many endocrine organs are impacted by body weight change. Is this absolute organ weight or relative organ weight? For serum hormone levels, glucocorticoids are not required in any regulatory test guidelines.
- Section 4.4 Evidence of Immune Suppression: As with the endocrine section, add sub-bullets for clarity. Changes in thymus weight may not indicate effects on immune function as thymic involution is common with age, stress, etc. Also, in this section, it is not clear how often' increased incidence of infections' is relevant for most test animals, which are kept in well controlled environments.
- Conclusion allows the statement: "Based on a WOE approach, the registrant requests that the chronic/carcinogenicity studies [be/not be (as appropriate)] required at this time…" Are all potential registrants submitting this data package? If not, the fact that the form is submitted to the EPA means that the registrant considers it possible to waive these studies. It's not clear that an option is required for this statement.

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 ReCAAP draft case study is helpful, but it may benefit from some revisions. It is important to illustrate the expected level of detail in this document so that other registrants will understand the expectation and presumably provide sufficient detail in future submissions. On p. 7, Table 4, 58 mg/kg/day (males) and 70 mg/kg/day (females) are NOAELs; however, there is an apparent increase in the incidence of collecting duct hyperplasia (not dose-related in males); a footnote to Table 4 would easily clarify the import of these data. Correct "damns" on p. 8. On p. 9, Table 5, "Rearing Index" may not be a commonly reported reproductive parameter for some labs. Consider defining this in a footnote to Table 5. Why is the rearing index in the F1 generation controls so low? Again, a footnote might be helpful. Lastly, late in the document, the terminology changes from 'calculi' to 'stone'; consider making this consistent. Overall, the case study is clear and concise and the conclusion is well supported, but these clarifications will allow the document to better stand alone.

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 concept of using quantitative transcriptomic data from short-term toxicity studies for benchmark dose analysis and point of departure (PoD) determinations is an interesting concept as it would allow for the identification of a safe dose that is not expected to result in toxicity, including carcinogenesis; thus, it may help to address concerns of some stakeholders who advocate that chemicals are not sufficiently tested/regulated. Using transcriptomics to identify these PoDs has been the subject in the scientific literature in recent years (e.g., Farmahin et al., 2017) and have generally shown that transcriptomic PoDs are similar to traditional PoDs. Based on the potential advantages, this approach merits consideration and further development. The presentation highlighted the use of PoD for liver carcinogens, which is a good starting point as there is a reasonable understanding of liver carcinogenesis MOAs in rodents. As this approach is applied more broadly, the EPA will need to consider how many organs need to be assessed in these short-term studies to ensure that other organs are not at risk for tumor formation/toxicity or to ensure that a potentially lower PoD may be identified. In addition, there may be age-related differences in gene expression that need to be considered as part of this assessment. Additional validation will be important as the EPA looks to implement this approach in different risk contexts.

Perhaps 'off topic' for this example, but one additional point with respect to the use of non-animal alternative methods (NAM) data…the EPA should think carefully about how to use NAM data, presumably to support mode-of-action (MOA) and weight of evidence. This includes ToxCast/Tox21 data, which may be used to assess the carcinogenic potential of chemicals. Kleinstreuer et al. (2017) proposed that ToxCast high throughput (HTP) assays could be used to predict rodent carcinogenicity using a model based on cancer hallmark genes and pathways. IARC (e.g., IARC, 2017) has proposed that ToxCast/Tox21 assays can be aligned with key characteristics of carcinogens (Smith et al., 2016) to identify carcinogenic potential of test compounds. However, to date, this approach has been used on a case-by-case basis. Recently, researchers have examined a broader application of ToxCast/Tox21 data for cancer-related bioactivities and determined that these assays have limited predictive capability for distinguishing "probable/likely" carcinogens from "not likely" carcinogens (Hill et al., 2017) or were "no better than chance" at differentiating carcinogens from non-carcinogens (Becker et al., 2017). This question is important and requires a carefully controlled, systematic review to determine the relevance of these assays to identify carcinogenic potential and whether the key characteristics of carcinogens (as currently assessed) add valuable information to identification of carcinogens. Currently, these high throughput assay results appear to have the greatest utility when included as potential supporting data in an overall mode-of-action assessment for carcinogens.

References

Becker, R.A., Dreier, D.A., Manibusan, M.K., Cox, L.A., Simon, T.W., Bus, J.S. (2017). How well can carcinogenicity be predicted by high throughput "characteristics of carcinogens" mechanistic data? Regul. Toxicol. Pharmacol. 90, 185-196.

Farmahin, R., Williams, A., Kuo, B., Chepelev, N.L., Thomas, R.S., Barton-Maclaren, T.S., Curran, I.H., Nong, A., Wade, M.G., Yauk, C.L. (2017). Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment. Arch. Toxicol. 91, 2045-2065.

Hill 3rd, T., Nelms, M.D., Edwards, S.W., Martin, M., Judson, R., Corton, J.C., Wood, C.E. (2017). Editor's highlight: negative predictors of carcinogenicity for environmental chemicals. Toxicol. Sci. 155, 157e169.

IARC (International Agency for Research on Cancer) (2017). IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, me 112. Some Organophosphate Insecticides and Herbicides: Diazinon, Glyphosate, Malathion, Parathion, and Tetrachlorvinphos. http://monographs.iarc.fr/ENG/Monographs/vol112/index.php. (Accessed 10 July 2020).

Kleinstreuer, N.C., Dix, D.J., Houck, K.A., Kavlock, R.J., Knudsen, T.B., Martin, M.T., Paul, K.B., Reif, D.M., Crofton, K.M., Hamilton, K. , Hunter, R., Shah, I., Judson, R.S. (2013). In vitro perturbations of targets in cancer hallmark processes predict rodent chemical carcinogenesis. Toxicol. Sci. 131, 40e55.

Smith, M.T., Guyton, K.Z., Gibbons, C.F., Fritz, J.M., Portier, C.J., Rusyn, I., DeMarini, D.M., Caldwell, J.C., Kavlock, R.J., Lambert, P.F., Hecht, S.S., Bucher, J.R., Stewart, B.W., Baan, R.A., Cogliano, V.J., Straif, K.. (2016). Key Characteristics of Carcinogens as a basis for organizing data on mechanisms of carcinogenesis. Environ. Health Perspect. 124, 713-721.

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?

For this reviewer, kinetically derived maximum dose (KMD) refers to doses at which the internal dosimetry (i.e., blood Area-Under-the-Curve [AUC]) is no longer proportional to dose. If a 1X dose generates a specific blood AUC, a 2X dose generates 2X AUC, and a 3X dose generates 12X AUC, it is clear that something kinetically different is happening at the 3X dose. However, it is more difficult to identify and agree on a threshold for non-dose-proportional toxicokinetics, which depends on numerous factors (e.g., robust

data sets, dose spacing, variability). It would be beneficial to achieve some scientific consensus at the October workshop on 'what constitutes an appropriate data set(s) for the application of KMD'.

The issue of identifying and using a KMD is contentious. Recent papers by Slob et al. (2020) and Heringa et al. (2020) have questioned the ability to accurately identify KMD and have articulated concerns about using KMD for dose level selection in toxicity studies. However, while the threshold for KMD may be difficult to agree upon, non-dose-proportional AUCs clearly indicate a change in toxicokinetics. If KMD doses identified in toxicity studies are markedly above human exposures (i.e., humans would not be expected to exhibit non-dose-proportional toxicokinetics), the relevance of the dose producing this non-dose-proportional AUC and toxicity seen at these doses is questionable. This may occur with studies using the maximum tolerated dose (MTD), resulting in questionable hazard predictions. Hazard identification above the KMD needlessly increases public alarm on the use of chemicals and may result in discontinued use of chemicals that offer advantages to society. Some additional observations:

- EPA decision to hold a workshop on KMD in October 2020 is well supported. One issue with the application of KMD is to understand what constitutes sufficient data for reasonable determination of KMD. Interanimal variability in toxicokinetic outcomes will need to be considered. In some cases, an understanding of toxicokinetics at various life stages (e.g., pregnancy/lactation, immature offspring) may be needed.
- KMD can only be used if exposure is adequately understood such that there is a large margin of exposure between actual human exposure levels and KMD levels based on non-dose-proportional AUCs. What information/confidence is required in exposure assessments/
- What statistical models or combination of statistical models will help to discern a reasonable determination of KMD?

References

Heringa MB, Cnubben NHP, Slob W, Pronk MEJ, Muller A, Woutersen M, Hakkert BC. (2020). Regul Use of the kinetically-derived maximum dose concept in selection of top doses for toxicity studies hampers proper hazard assessment and risk management. Regul. Toxicol Pharmacol. doi: 10.1016/j.yrtph.2020.

Slob, W., Zeilmaker, M.J., Hoogenveen, R.T. (2020). The relationship between internal and external dose: Some general results based on a generic compartmental model. Toxicol Sci. https://doi.org/10.1093/toxsci/kfaa082

Dr. Thomas Parkerton

General Comments

EPA correctly acknowledges some of the shortcomings with the present approach used for determining the need for chronic toxicity and carcinogenic data and should be commended for developing alternative approaches. The Office of Pesticide Programs has led efforts to provide a more flexible and fit for purpose approach in these assessments given the activities, actions and projects described. However, it is encouraged this initiative includes outreach and engagement across EPA. In addition to the Office of Research & Development, the Office of Pollution Prevention and Toxics (OPPT) should be viewed as an essential partner in actions that are planned consistent with the direction identified in the update of TSCA. Regrettably, OPPT does not appear involved in the Rethinking Carcinogenicity Assessment for Agrochemicals (ReCAAP) Project which would help ensure alignment on the development of a risk-based weight of evidence (WOE) approach for waiving chronic and carcinogenicity studies. Further, OPPT scientists should contribute to the HESI eSTAR program to ensure perspectives and concerns are addressed if they are not already active participants. As this initiative is progressed, efforts should be made to ensure consultation and alignment for implementation with other relevant EPA program offices such as the Office of Superfund Remediation and Technology Innovation, Office of Water, Office of Air and Radiation, and the National Center for Environmental Assessment.

Response to Charge Questions

1. **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.**

   RESPONSE: The draft WOE template appears like a positive step forward. Two questions for further consideration:

   How can this document be extended beyond agricultural chemicals?
   Read across is mentioned in step 5 in relation to carcinogenicity assessment. Can read across principles be extended to other toxicity endpoints if confirmative short terms assays indicate similar biological activity (specifically, subchronic and mode of action studies) ?

2. **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.**

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

4. **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?**

   RESPONSE: It is recommended that kinetically-derived maximum dose (KMD) concept for determining dose ranges proposed for carcinogenicity testing be extended to all repeat

dose toxicity testing. OPP is using KMD for pesticides and OPPT should be encouraged to apply this approach for industrial chemicals. Further, as this approach is further adopted and more broadly implemented across the EPA, will there a process or effort to reassess prior decisions and evaluation with regard to carcinogenic hazard/risk? More specifically, how will prior substance evaluations for which hazard thresholds were derived above the KMD and for which NAM data, MOAs, human relevance were not available or considered be reevaluated in light of updated assessment paradigms and new data?

Dr. Robert Phalen

Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing:

The EPA should consider whether or not its mandate legitimately includes "working to reduce the number of laboratory animal studies." This is a centuries old controversial issue that could dilute EPA's attention to protecting the environment and human health. There is sustained progress on protecting laboratory animals through the work of Institutional Care and Use Committees, and countless other entities. Vast expenditures, university programs, pharmaceutical and other industries, nongovernmental organizations, governmental regulatory bodies, and individuals work for animal welfare, so it's not an area for which there is a need for additional governmental expenditures. The Nuremburg Trial principles that defined crimes against humanity included "The experiment must be based on animal experimentation and other prior knowledge." Those principles are a cornerstone in ethics.

The SAB was told by a program speaker at the June 23-24 teleconference that "a rat is not a man." But it is important to recognize the inadequacies of non-whole animal models. For example as a chair of our IRB, I saw a report where a pharmaceutical company kept a group of dogs on their anti-depressant drug (on a chronic low dose administration protocol) after their drug passed all FDA requirements and was apparently in wide human patients use. The dogs underwent thorough periodic veterinary examinations. When they saw subtle unexpected changes in the dogs' retinas, they immediately pulled the drug from distribution. That action may have prevented blindness in countless people. This is just one of many examples where animal studies are irreplaceable. Mammals are complex beyond what we toxicologists can comprehend.

Dr. Mara Seeley

Charge Questions: NAMS/Reducing Use of Laboratory Animals for Chronic and Carcinogenicity Testing

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

- The WOE approach seems reasonable and well-thought out, as a framework.
- EPA should consider incorporating evidence of additional pathways or findings potentially associated with an increased likelihood of tumor formation, such as suppression of programmed cell death, inflammation/cytotoxicity/regenerative cell proliferation.

1b. Draft case study

- The draft case study is comprehensive and clearly presented.
- The case study should include references to testing protocols or studies relied on for the toxicity testing.

2. Direction and scope of collaborative NAMs projects

- The direction and scope of the collaborative NAMs projects should advance the science of risk assessment for carcinogens away from reliance predominantly on tumor findings from animal bioassays, which can be both resource and time intensive; towards alternatives that may ultimately be able to provide more timely information for a greater number of environmental exposures.

3. Current KMD-related activities

- The KMD seems like a reasonable alternative to the MTD, and thus developing best practices for KMD analysis would be a worthwhile effort for EPA.
- Several scientists at the Netherlands' RIVM and Centre for Safety of Substances and Products have expressed concerns related to use of the KMD (Heringa et al., 2020; Woutersen et al., 2020)[1] including for example that internal/external dose relationships may not actually exhibit a true inflection point; and that test doses may not be sufficiently high to observe more subtle effects, or effects that occur at a relatively low incidence. In developing best practices for KMD analysis, EPA should consider the concerns expressed by Haringa et al. and Woutersen et al.

[1] Heringa, MB, NHP Cnubben, W Slob, MEJ Pronk, A Muller, M Woutersen, BC Hakkert. 2020. Use of the kinetically-derived maximum dose concept in selection of top doses for toxicity studies hampers proper hazard assessment and risk management. Regul Toxicol Pharmacol. 114; Woutersen, M, A Muller, MEJ Pronk, NHP Cnubben, BC Hakkert. 2020. Regulating human safety: How dose selection in toxicity studies impacts human health hazard assessment and subsequent risk management options. Regul Toxicol Pharmacol. 114.

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?

Dr. S. Stanley Young

I agree with Dr. Sam Cohen. Given that long-term rodent studies are unlikely to be predictive of human cancer, it makes no sense to continue running them. Experts have known of the problems from the beginning. (See Meyers et al., 1985).[1]

There is solid evidence that the science literature in general and the environmental epidemiology literature specifically is not reliable (see references below). These comments written for journal editors should be considered by EPA.

It is critical that decision makers understand multiple testing and multiple modeling, MTMM.

i. **Experts** all agree that false positives can be due to MTMM.
ii. **Theory:** A statistical test can be positive by chance 5% of the time. If you test 100 questions you expect 5 or so statistically significant findings even if there is nothing going on.
iii. In addition to A and B, it is critical to get **hands on experienc**e with MTMM. Perform your own experiment and see the false positives. Get 5 coins and cast them 100 times and fill in the table (provided as an example below).

a. Editors of scientific journals

1. Standards for exploratory/confirmatory research

Editors need to develop and publish standards for naming an observational study or trial as exploratory or confirmatory. For example, if a study does not have a pre-study written protocol, asks many questions (implying the conduct of many statistical tests) without dealing with multiple testing and multiple modeling, does not provide access to the data set, and does not provide the analysis code, then the word "Exploratory study" should be appended to the title of the paper. Conversely, if these aspects are done, then the words "Confirmatory study" should be attached to the paper title. [SSY: the EPA should denote each study used in considering a regulation as either Exploratory or Confirmatory.]

2. Require data and analysis code

Editors should require the analysis data set and analysis code for any published paper to be made public. [SSY: Any EPA funded study should make public: a. protocol, b. analysis data set, c. analysis code. It is the responsibility of the researcher to devise means that data can be made public. Micro aggregation usually protects personal identity.]

3. Require treatment of MTM

Editors should require that researchers deal with multiple testing and multiple modeling. [SSY: the EPA should require control of MTMM on any study considered Confirmatory.]

---

[1] Meyers, D.B., S.S. Young, and C.L. Gries. Design of Cancer Assays for Pharmaceutical Agents. Journal of the National Cancer Institute 74(5):1151. https://doi.org/10.1093/jnci/74.5.1151
[Available at: https://academic.oup.com/jnci/article-abstract/74/5/1151/882438?redirectedFrom=fulltext]

4. Require treatment of MTMM

Epidemiology journals and journal editors should formalize editorial policies and peer review practices around making decisions to publish based on issues of quality and logical reasoning by researchers and not on novelty or the direction (i.e., positive vs. negative results) and strength of study results. [SSY: The EPA should maintain a list of negative studies.]

b. Funders of research

"He who pays the piper calls the tune." Those funding a study should specify the data to be used, the analysis methods to be employed, and the transparency and completeness of the reporting. The NIH has taken a step in this direction. The EPA should follow their lead. The NIH will provide funding only if the researcher provides a data access plan.

1. Fund data building and data analysis separately

Human nature is such that researchers will so align their work to enhance its acceptability for publication. They might do some data gardening. They might adjust their claims to the results found in the data. They might cite supporting literature and omit reference to contrary results. One way that funding agencies might prevent this opportunistic behavior is to fund 'data building' separately from 'data analysis.' The resulting data sets should be made public. They could fund multiple researchers to analyze the resulting data. Any researcher could apply their analysis methods and write a paper.

2. No data, no funding

The funding agency should provide funding only if a researcher agrees to make the analysis data set public.

3. Have a "holdout" data set

Even more robust methods should be considered. The funder can require two data sets, a public data set, and a holdout data set. The researcher, knowing that there is a holdout data set, might then carefully address the multiple testing and multiple modeling problem, among other analysis aspects so as not to be embarrassed if the holdout set does not support his claims.

References:

Long-term rodent studies:
Meyers DB, Young SS, Gries CL. (1985) Design of cancer assays for pharmaceutical agents. *J Natl Cancer Inst*. 74,1151-1152.

General reliability of science.
Young SS, Karr A. 2011. Deming, data and observational studies. Significance 8, 116–120. [SSY: 52 claims from observational studies failed to replicate in RCTs.]

Begley CG, Ellis LM. 2012. Raise standards for preclinical cancer research. Nature 483, 531–533. [SSY: 47/53 experimental biology studies could not be replicated.]

Baker M. 2016. 1,500 scientists lift the lid on reproducibility. Nature 533, 452−454. [SSY: 90% of surveyed scientists say there is a serious (52%) or real (38%) crisis of science claims failing to replicate.]

Reliability of Environmental Epidemiology

Young SS, Smith RL, Lopiano KK. 2017. Air quality and acute deaths in California, 2000-2012. Regulatory Toxicology and Pharmacology 88, 173-184. [SSY: there is no association of all-cause, respiratory, or cardiovascular deaths in California with PM2.5 or ozone. There are several separate confirmation California studies.]

Young SS. 2017. Air quality environmental epidemiology studies are unreliable. Regulatory Toxicology and Pharmacology 88, 177-180. [SSY: Each of eight studies appearing in Environmental Health Perspectives examined massive numbers of possible claims without any correction for multiple testing or multiple modeling, a fatal analysis flaw.]

Young SS, Kindzierski KB. 2019. Evaluation of a meta-analysis of air quality and heart attacks, a case study. Crit. Rev. Toxicol. 49(1), 85−94. [SSY: 34 studies were used in this meta-analysis. The air components studies were the standard six, PM2.5, PM10, NO_2, SO_2, CO and ozone. There were studies with small and large p-values. One interpretation is that multiple testing and multiple modeling were the cause of the small p-values.]

Example

The coin experiment – Compounds are tested in long-term rodent studies, rats and mice, males and females for two years, and at the end of the studies hundreds of tumors were assessed by pathologists. The tumor counts are subjected to statistical analysis, each tumor separately. Think of the whole process and of the MTMM problem. Will you have a clean test?

Cast five coins 100 times (representing 25 tumors x 4 rodent population groupings). In each cell write down the number of heads for each cast. The probability of obtaining 5 heads in a single cast is $(1/2)^5$=0.03125.

Run the experiment (table below). Keep in mind that this is simply an experiment of chance and 5 heads showing up in a single cast is nothing more than a chance finding; it does not have real meaning. Also, there is no difference between this experiment of chance and the toxicology researchers statistically analyzing 100 different possible rodent-tumor combinations.

	Male	Female	Male	Female
Tumor	**Rat**	**Rat**	**Mouse**	**Mouse**
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				

Enclosure C

Individual Comments from Members of the EPA Science Advisory Board Chemical Assessment Advisory Committee on New Approach Methods and Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing

Dr. Richard Belzer

Comments on SAB/CAAC Review of 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 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).[1] Please include in your comments a discussion of the clarity and completeness of the proposal.**

This draft weight-of-evidence (WoE) framework provides a well-organized list of categories of information that a registrant would be expected to provide. I see a few potential problems, however:

i. § II is ambiguous with respect to the framework's actual information requirements.

How would a registrant know that a petition is adequately supported? This section includes references to *potential* dietary and non-dietary exposure (§ II.1), bioaccumulation (§II.2), chronic exposure (§ II.1), chronic toxicity (§ II.3), tumor formation (§ IV.4), molecules in the same chemical class (§ IV.5), and human exposure (§ IV.7). Outside of physics, *potential* has no scientific meaning. Thus, § II appears to be less about weight-of- *evidence* (WoE) than weight-of-*worry* (WoW).

ii. "Specific criteria to determine when not to require the Combined Chronic Toxicity/Carcinogenicity studies" are not absent; rather, they have been ignored.

Briefing materials for the public meeting include a "problem statement" for the HESI eStar POD:

> There are no specific criteria to determine when not to require the Combined Chronic Toxicity/Carcinogenicity studies (OECD 453; 451), or how to determine appropriate POD for chronic risk assessments for pesticides based on available toxicological and exposure data in the absence of chronic toxicity studies…there is a movement to transition away from a routine 'check-box' approach towards a more scientifically sound weight of evidence (WOE) carcinogenicity assessment for non-genotoxic food-use pesticides.[2]

[1] U.S. Environmental Protection Agency (2020a).

[2] Hilton and Akerman (2020, Slide 12).

Specific criteria do in fact exist, however. They are found in the Paperwork Reduction Act (PRA) and its accompanying regulations.[3] All toxicological data requirements are covered "collections of information."[4] Before any federal agency can seek information from the public, whether mandatory or voluntary, it must seek and obtain prior approval from the Office of Management and Budget (OMB).[5] Failure to do so renders the information collection unenforceable.[6]

To be approvable by OMB, the agency must show *inter alia* that its information request "[i]s the least burdensome necessary for the proper performance of the agency's functions to comply with legal requirements and achieve program objectives."[7] Given the persuasive evidence reported by here that U.S. EPA's toxicological data requests often do not satisfy this standard, and that OMB may be legally obligated to disapprove them, one might think that these criteria are important enough to at least consider.

The PRA criteria can be easily understood as an application of the value-of-information (VOI) principle. Data requirements that satisfy VOI are statutorily approvable; data requirements that do not are not. U.S. EPA is certainly fortunate that OMB has either been too busy with higher priorities to disapprove these collections, or perhaps it has chosen to look the other way. But there is no guarantee that forbearance will continue indefinitely.

iii. § II.6 concerns *safety* assessment, not *risk* assessment, and as such it cannot produce outputs compatible with regulatory benefit-cost analysis.

Each of the items requested here is part of safety assessment. There is no scientific definition of *safe*, so transparency requires that this be properly characterized as risk *management.* I realize that safety assessment is the dominant form of what federal and state governments do under the rubric of *risk* assessment. Nonetheless, the outputs of safety assessment are inherently subjective and not refutable, and hence they are nonscientific. They cannot meet established principles for information quality, including substantive and presentational objectivity.[8] And because they do not estimate expected value risk, they cannot be used as inputs to regulatory benefit-cost analysis.

Further, § II.7 reinforces the inference that the registrant would be asked to propose a risk *management* decision with respect to whether the default requirement to conduct additional studies should be waived. For these reasons, the second half of the title of § II.6 should be labeled as Proposed/Prospective Risk *Management* Decision. If U.S. EPA does not want to invite the registrant to opine on risk management, then this subsection should be revised accordingly.

---

[3] 44 U.S.C. § 3501 *et seq*., 5 C.F.R. Part 1320.
[4] 44 U.S.C. § 3502(3) and 5 C.F.R. § 1320.3(c).
[5] 44 U.S.C. § 3507 and 5 C.F.R. § 1320.8-12.
[6] 44 U.S.C. § 3512, 5 C.F.R. § 1320.6.
[7] 44 U.S.C. § 3504(c)(4), 5 C.F.R. § 1320.5(d)(d)(ii)
[8] Office of Management and Budget (2002); U.S. Environmental Protection Agency (2002).

Of course, if the Agency were to follow this VOI-based framework, it could ask registrants to estimate the practical utility and burden of performing a toxicological study for which they want a waiver. Burden (i.e., costs) would be estimated in part the conventional way – i.e., by adding up the outlays required to plan, conduct, and report results from the test. But costs are properly measured in terms of the benefits foregone resulting from the expenditure of scarce resources. These foregone benefits (i.e., "opportunity costs") include the costs from delayed decision-making, an inherent and unavoidable consequence of deciding to obtain more data. Opportunity costs may include human health risks not prevented, such as would arise of the pesticide proposed for registration is likely to be risky than the pesticide it would replace. Opportunity cost also includes the value of animal lives not sacrificed.[9] Practical utility (i.e., benefits) would be measured in units of human health risk reduction provided by the information, preferably monetized using WTP methods.

This framework converts the problem from one of *safety* assessment to *risk* assessment. A waiver is justified if the burden of the test exceeds its practical utility. No policy judgments are required concerning such things as the choice of a point of departure, a margin of exposure, and the value of uncertainty factors.

iv. This proposed WoE framework is missing a coherent model and the weights that would be used to run it.

Is U.S. EPA inviting registrants to propose which (WoE? WoW?) model the Agency should use, and what weights it should apply? To be clear, this is a question for clarification and not a criticism. Registrants may be better positioned to propose a choice of model and weights, and document these proposed choices in a way that is transparent and reproducible. Note that the VOI framework proposed in subsection (iii) above, which uses the PRA as its statutory foundation, would eliminate debate about the choice of WoE framework and weights to be applied.[10]

While it seems logical to me that VOI principles should guide decisions to acquire information, these principles seem nowhere in evidence.[11] Clearly, the case for waiver is

[9] The proper way to value animal lives not sacrificed relies on the same willingness- to-pay (WTP) methods used for valuing reductions in human health risk and premature mortality. This WTP is presumably is positive, but it cannot be infinite and additional research may be necessary to develop objective estimates because, to date, U.S. EPA has never placed a positive value on avoiding animal sacrifice.

[10] This is not to say that stakeholders won't differ with respect to estimates of practical utility and burden. But the estimation of these quantities is subject to objective standards. Intentional bias, whether for policy reasons or some other purpose, is simply impermissible in economic analysis.

[11] VOI principles are hinted at in OPP's Data Requirement Guiding Principles: "These guiding principles for data requirements will enable OPP staff to focus on the *information most relevant* to the assessment. The goal is to ensure there is sufficient information to reliably support registration decisions that are protective of public health and the environment while *avoiding the generation and evaluation of data that does not materially influence the scientific certainty of a regulatory decision.* It is important to *only require data that adequately inform regulatory decision making* and thereby *avoid unnecessary use of time and resources, data generation costs, and animal testing*" (emphasis added). See U.S. Environmental Protection Agency Office of Pesticide Programs (2013, p. 1). The Data Requirements Principles even give a welcome (if halting) nod to the principle of opportunity costs: "Delayed regulatory decisions affect the delivery of health and environmental protections and access to benefits

strongest when burden vastly exceeds practical utility and is weakest when burden exceeds practical utility by only a small amount.[12] Information of this type would inform the balancing test necessary to support rational choice. It also would help U.S. EPA document the basis for its decisions, and defend them if (when) challenged.

If a VOI-based framework is not used, registration decisions will remain mired in conflict. The choice of WoE (WoW?) model and weights may be (correctly) criticized as subjective and perhaps self-serving. These choices are not matters of science but policy, and policy debates (whether explicit or disguised as science) have been shown to be difficult to resolve when various stakeholders (including U.S. EPA) have different values and preferences.

As an alternative, a registrant could offer to describe what results a study proposed for waiver would have to yield to alter the default decision set forth in § II.7. It may be reasonable for the registrant to describe how extreme the results of a study proposed for waiver would have to be, and how likely it is that such extreme results would be obtained given existing scientific knowledge. The case for waiver is strongest if results must be extraordinarily extreme or unlikely to materially change the Agency's decision. Of course, registrants might need guidance from the Agency concerning specifically what results equate to what the hypothetical implementation of the WoW framework in Attachment 2 calls a "potential for concern" (there's that word, again).[13] Still, this approach would use well understood VOI principles to inform the decision to waive or retain default information collection burdens, As a result, Agency decisions could be substantially more reproducible, and more easily defended if challenged.

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.**

i. The case study in Attachment 2[14] invents, but does not define, new terms of art that are not in Attachment 1.[15]

At the top of this list is "potential for concern" in § I ("Purpose of the Analysis"), a phrase that is ambiguous in both nouns. The case study's elaboration on this purpose is unhelpful. Apparently, the analysis is supposed to be informed by "the *potential* for long-term exposure from dietary sources" when "*possible* total chronic exposure is *very low*" (emphasis added). Is *potential* different from *possible*? If so, how? If not, why are different terms used? How low is *very* low? Ambiguity is partially relieved by text that seems to implicitly invoke VOI principles, but the lack of any concrete expression of those principles

such as pest management tools and safer products." The Consolidated Guideline project would be greatly enhanced if VOI principles and opportunity cost were given leading roles.

[12] Where practical utility demonstrably exceeds burden, there would be no justification for waiver.

[13] U.S. Environmental Protection Agency (2020b, p. 3).

[14] U.S. Environmental Protection Agency (2020b).

[15] U.S. Environmental Protection Agency (2020a).

raises doubt as to whether that inference is correct, or whether a registrant interpreting the text this way might later be blindsided to learn it's not.

ii. Some text in Attachment 2 is a source of potential … worry.

In my comments on the proposed WoE framework in response to Question 1(a), I said that text looked more like a weight-of-worry (WoW) framework. Attachment 2 reinforces this concern. *Potential*, an adjective that is ambiguous without bound, appears 18 times in the case study. *Possible* appears another six times. In none of these uses, however, is the adjective defined to reveal, for example, an associated probability – particularly the probability sufficient to trigger enough worry to deny a waiver request.

The stated facts in this case study make the waiver decision easy, so it's superfluous *in this case* to unpack the meaning of *potential.* But that does not illuminate how the framework would be implemented if the facts were less one-sided.

iii. Where simple declarative scientific statements are made, they are sometimes undermined by ambiguity.

§ II.1 makes a factual statement that contains substantial probative information establishing a strong case for waiver given the precautionary character of HED's *levels of concern*:

> All exposure scenarios, including dietary (food and water), residential, aggregate, and occupation, are reported to be below the level of concern for EPA's Health Effects Division (HED).[16]

If *all* exposure scenarios are below HED's worry threshold, that ought to be sufficient. Why, then, does the text follow with a second sentence that undermines the first with multiple sources of ambiguity?

> While there is the *potential* for long-term exposure from dietary sources, the *possible* total chronic exposure is *very low*.[17]

Text in §§ II.2, II.3, II.4.5, and II.5 provide better models. § II.2 explains, in just two sentences, everything that is needed to infer that Herbicide1 should not trigger concern (worry?) about bioaccumulation. § II.3 makes the same point in a few short paragraphs, all with the same message. The paragraph on excretion includes a profoundly powerful scientific inference: " "*no* bioaccumulation *potential*" (emphasis added). § II.4.5 clearly states, "All [genotoxicity] studies were negative, *clearly eliminating* any concern for tumor formation via genotoxic mechanisms" (emphasis added). § II.5 reports that Herbicide 1 "does not contain" elements "demonstrated to be associated [not necessarily *causally*] with sensitization

[16] U.S. Environmental Protection Agency (2020b, p. 3).
[17] U.S. Environmental Protection Agency (2020b, p. 3 [emphasis added]).

reactions" and "showed *no effects* on the thyroid" (emphasis added).[18] Scientifically valid declarative statements that cut off a risk pathway should be encouraged precisely because they are useful for reducing uncertainty.

Taking these texts at face value, the burden of collecting additional information based on what is scientifically known about Hertbicide1 almost certainly exceeds the practical utility of the information which the studies proposed for waiver would supply. It would be helpful if U.S. EPA invited registrants to characterize the available science in such VOI terms.

iv. WoE/WoW inferences are not obvious from the evidence presented.

§ II.4.2 says continuous administration of Herbicide1 leads to calculi-induced hyperplasia in dogs at specified (high) doses. These effects are said to be reversible. Why, then, does this evidence "strongly support that the threshold of concern for toxicity for Herbicide1 is the exposure that leads to calculi formation *regardless of the duration of the exposure*" (emphasis added)? A rational basis for dose-independence seems to be missing. Does reversibility not matter? Is there truly no difference between exposure for a lifetime versus exposure for a single day? Absent a rational basis, WoE and WoW frameworks are indistinguishable.

v. WoE/WoW inferences cannot be interpreted in welfare economic terms.

The benefits of avoiding a risk depend on WTP. An economist asked to estimate WTP to avoid reversible calculi-induced hyperplasia would be stymied. Valuation requires that a risk be well-defined, but this biological effect is not easily understood by nonexperts. Moreover, any WTP would be severely attenuated precisely because effects are reversible. Given the information provided, my best estimate of the value of avoiding this effect is zero. And if that is so, the proposed WoE framework clearly has strayed into WoW territory. The authors need to make a stronger case why reversible calculi-induced hyperplasia is meaningful to actual people, particularly if doses are below the threshold whereby calculi formation can act as urinary bladder carcinogens. Similar concerns arise for hormone perturbation as a presumptively adverse effect . It is likely that perturbation within normal boundaries would not be construed by actual people as an adverse effect. And actual people might interpret some hormone perturbations (or even permanent changes) as beneficial. A biological effect, whose interpretative meaning to actual people is unclear in both magnitude and sign, cannot be presumed to be adverse just because toxicologists are able to observe it in laboratory experiments.

The case study description of chronic urinary tract toxicity suggests that information essential for valuation is missing (or perhaps was not reported). The summary inference – "*with sufficient exposure*, crystalluria and the accompanying toxicity might be expected for

[18] Qualifying such statements with dubious probability statements is unhelpful. After declaring the absence of thyroid effects, stating that "Herbicide1 is *highly unlikely* to result in thyroid tumor formation" (emphasis added) creates contradictory ambiguity.

Herbicide1"[19] begs the question what exposure is *sufficient*, a nonscientific term. If that exposure level exceeds background environmental concentrations, then WTP to avoid it is likely to be zero and the social benefit from denying the waiver would be negligible at best. Only if the level of exposure exceeds background environmental concentrations could there be any practical utility in acquiring the information. The quantification of practice utility depends on the extent it exceeds background environmental concentrations and whether such concentrations pose human actual human health risk.

The achievement of safety-assessment type dose indicators (e.g., NOAELs) also cannot be valued, which makes them dubious health-based targets. Setting aside for the moment the fact that purported adverse effects understood only by experts are not well- defined goods, a monetary value in principle can be estimated for the avoidance of any response in a causal dose-response relationship. But dose indicators that are not located on a causal dose-response relationship are much more challenging to value. They consist of two parts: (1) WTP to avoid the nearest dose causally located, and (2) WTP for reducing exposure below this causally located dose, though without any information concerning the magnitude of additional risk avoided. The second part is purely speculative, and the greater the implied safety factor the more speculative it is. WTP for the latter part is likely to be zero, as would be the WTP for any noncausal dose-response.

The value of waiving default studies on value-of-information grounds is diminished when risk management adjustments are made without their burdens being accounted for in the analysis. Despite the strong case given for why chronic toxicity in humans is impossible at environmentally relevant doses, the case study proposes to add an additional 10x factor for database uncertainty (see Table 7). The proper question is the net practical utility of a waiver under both scenarios. Only if net practical utility is higher under the scenario in which the 10x database uncertainty factor is added is it possible to justify adding that factor on value-of-information grounds. Otherwise, it is a just a risk management concession.

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.**

The description of the NAM project does not explicitly say so, but it appears to be grounded on the realization that traditional rodent bioassays, which for decades were characterized as the "gold standard" in toxicology, have significant information quality deficiencies, most notably "imperfect translational relevance to human health."[20] In addition, rodent bioassays have been assailed for lacking practical utility sufficient to justify their

[19] U.S. Environmental Protection Agency (2020b, p. 10 [emphasis added]).
[20] U.S. Environmental Protection Agency (2020b, p 2 list).

burden.[21] Practical utility limitations are self-evident from this acknowledged information quality deficiency, and burden includes both cost, delay, and low throughput. A relatively recent addition to these burdens is rising salience of ethical doubt about animal sacrifice, particularly when it has little or no information value.

I am concerned that the "3R principles" summarized in § 1.2 of the Cancer NAMs White Paper[22] may lack a proper respect for VOI principles. There are optimal levels of reduction, replacement, and refinement, each of which depends on the result of VOI analysis.[23] But I cannot find any discussion in the White Paper that properly addresses tradeoffs. There are a few hints (e.g., "there will not be a 'one-size fits all' solution," p. 4), but nowhere are tradeoffs transparently discussed and analyzed. Even this limited acknowledgement of tradeoffs is subsequently discarded in favor of on-size fits all approaches to KMD:

> There is an immediate need to *standardize* these approaches for broad regulatory use and facilitate *global harmonization*[24].

But "standardization" is the *sine qua non* of every one-size-fits-all solution, and "global harmonization" makes clear that "one-size" is not an approximation.

Further, the "risk-based weight of evidence analysis framework" does not include VOI principles or recognize tradeoffs. If it is true that "NAMs are expected to improve the scientific foundation of risk assessments by providing more human-relevant information that is more efficient and less costly,"[25] then why has the technology of toxicology been mired for decades in tools that generate low-value information in a manner that is inefficient and expensive?[26]

---

[21] Pitot III and Dragan (2001, p. 293-294) were not the first to call the chronic 2-year bioassay the "gold standard," but they offered a prophetic warning about value of information in these studies: "Because so many research dollars go into carcinogenicity testing and the data resulting from such studies are expected to be useful not only in hazard identification but in risk estimation, an acceptable scientific protocol with quality assurance must be followed to produce scientifically and statistically valid data." They then discuss an array of issues, but note that "[d]espite these criticisms and problems, the chronic 2-year bioassay continues to be the major basis for regulatory action in this country and in many countries throughout the world."

[22] U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development (2020, pp. 4-5).

[23] If the concern about animal sacrifice were truly a matter of ethics, such experimental would be simply banned instead of reduced, replaced, and/or refined.

[24] U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development (2020, p. 5).

[25] U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development (2020, pp. 4-5).

[26] Similarly, if a kinetically-derived maximum dose approach "not only lessens or avoids unnecessary pain and distress in animals, but also generates data that are relevant and more predictive of human health risks" (U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development 2020, p. 5), why have conventional methods that callously imposed pain and distress in laboratory animals without producing useful information been used for so long?

Safety assessment wages war against scientific risk assessment and regulatory benefit-cost analysis, and at least one of the proposed reforms in "Replacement" would continue this war in a new guise. Kudos to HESI for convening "a multi-sector and multi- disciplinary working group," but doing so to "build and implement [yet another] framework" for "health protective point[s] of departure"[27] that are by definition *not* part of risk assessment and are fundamentally incompatible with VOI principles, information quality principles, and regulatory benefit-cost analysis, is not a productive step forward.

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

I am concerned that a key attribute of these three projects is the preservation of certain institutional controls over the path of reform, and that these controls will operate in ways that are incompatible with objective risk assessment and regulatory benefit-cost analysis.[28]

An early sign of trouble can be found in the simultaneous adoption of the "fit for purpose" concept and the goals of "characteriz[ing] the *potential* for environmental exposures to cause *or contribute to* the development of cancer in humans…"[29] These goals are incompatible with regulatory benefit-cost analysis, a key *purpose* of risk assessment, in which *potential* exposures (probability unknown) that *contribute to* risk (in some amorphous way) actually are risk management thumbs on the risk assessment scale. The replacement of one institutionalized incompatible toxicology technology with another is not reform in any meaningful sense. Perhaps the development of a "translational toxicology pipeline" would be useful, but it won't be if it substitutes "carcinogenic *potential*" for risk.[30] Adverse Outcome Pathways do not necessarily bring us closer to objectively estimating cancer *risk*. Rather, they appear likely to be used to justify the substitution of ever more arcane non- or pre-cancer endpoints – endpoints that actual people cannot understand or value, and thus are inherently incapable of valuation.[31]

---

[27] U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development (2020, p. 5 [bullet 2]).

[28] It was NTP that institutionalized the 2-year rodent bioassay and for decades protected it from criticism, and the short description in U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development (2020, Sec. 3.1) suggests that NTP seeks to retain this role. This suggestion is magnified by Casey (2020), who promotes the NTP's library of chronic 2-year bioassays (slide 25) while discreetly acknowledging their limited informational value (slide 26).

[29] U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development (2020, p. 8 [emphasis added]).

[30] U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development (2020, p. 8 [bullet 1]). That the plan would retain the chronic 2-year rodent bioassay as "an option" is similarly worrisome. These bioassays should be justified only by rigorous application of VOI principles and analysis. They should never be treated as acceptable defaults in which the burden of proof rests with an alternative technology.

[31] It is far from clear that toxicologists can consistently *rank* alternative points along an AOP. If they cannot, the purported AOP lacks sufficient scientific content to be used for estimating risk or benefits to human health.

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?**

The observed historical difference in how KMD has been used in pharmacology and chemical safety assessments[32] deserves more thoughtful explanation (though some is presented in the proposed workshop summarized in Attachment 3). If this difference is attributable primarily to science, a concerted application of science may be useful. But if it is primarily attributable to institutional factors, the public interest argues against any governmental entity (or group of such entities) holding reform hostage to the preservation of institutional control.[33] Some support for the former is supplied early in Attachment 3, but unfortunately the structure of the project, and thus constraints imposed upon it, support the latter. To oversimplify, the principle value of KMD appears to be obtaining information at dose ranges closer to human exposure and determining where and why the decades-long assumption of linearity breaks down. Why, then, is the proposed project so focused on "best practices"?

"Best practices" are, of course, generally preferred; who, after all, is against what's "best"? Problems can and do arise when, under the guise of "best practices," stakeholders (including government agencies) seek to establish their preferred models and/or outcomes as "best," to the exclusion of their scientific or institutional competitors. Especially at this stage, therefore, it may be premature to establish the goal of seeking a consensus on "best practices," as the HESI project, summarized in Attachment 4, appears trying to do. A better approach may

[32] U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention and U.S. Environmental Protection Agency Office of Research and Development (2020, p. 13): "While KMD is routinely considered in preclinical tests to provide perspective on the relevance of study results to human safety assessment for drugs, KMD is rarely used in chemical safety assessment."

[33] By "institutional factors" I am, of course, alluding to Public Choice theory. See, e.g., Buchanan and Tullock (1962) and Downs (1967). For a historical perspective on today's moment of reform in toxicology, see Downs (1972). The second stage in Down's "issue- attention cycle" ("alarmed discovery and euphoric enthusiasm") is best captured by the publication of National Research Council (2007). It's a matter for debate whether reform is now in stage 3, 4 or 5. Evidence for each can be found in the White Paper.

be to seek reasonable minimum performance standards that would exclude only demonstrably inferior approaches.[34]

The most salient concern to me is that this workshop emphasize the need for outputs that are compatible with objective risk assessment and regulatory benefit-cost analysis. Outputs that, like the longstanding toxicological tools that would be replaced, are inconsistent with these purposes are not helpful and offer no credible reform. Clarity that objective risk assessment is intended would substantially ameliorate this concern.

REFERENCES

Buchanan, James M., and Gordon Tullock. 1962. *The Calculus of Consent.* University of Michigan Press: Ann Arbor MI.

Casey, Warren. 2020. "Advancing Carcinogenicity Assessment at the National Toxicology Program; Presentation to the Science Advisory Board." Slides 21-29. NTP: Washington DC.

Downs, Anthony. 1967. *Inside Bureaucracy.* Little, Brown: Boston MA.
———. 1972. "Up and Down with Ecology: The 'Issue–Attention Cycle'," *The Public Interest*, 28: 38-50.

Hilton, Gina, and Gregory Akerman. 2020. "ReCAAP: Rethinking Carcinogenicity Assessment for Agrochemicals Project; Presentation to the Science Advisory Board." Slides 8-20. PETA International Science Consortium, Ltd. and U.S. EPA: Washington DC.

National Research Council. 2007. *Toxicity Testing in the 21st Century: A Vision and a Strategy.* National Academies Press: Washington DC.

Office of Management and Budget. 2002. "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Notice; Republication," *Federal Register*, 67: 8452-60.

Pitot III, Henry C., and Yvonne P. Dragan. 2001. "Chemical Carcinogenesis.' In Curtis D. Klaassen (ed.), *Casarett & Doull's Toxicology: The Basic Science of Poisons*. McGraw- Hill Medical Publishing Division: New York NY.

U.S. Environmental Protection Agency. 2002. "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency." *EPA/260R-02-008*. U.S. EPA: Washington DC.

[34] The stated objectives include both exclusivity ("Selecting *the* appropriate PK parameter to examine dose proportionality," emphasis added) and inclusivity ("Defining a *minimum* dataset considered necessary to set a KMD," emphasis added).

———. 2020a. "Attachment 1: Draft Risk-Based Weight of Evidence Framework for Chronic/ Carcinogenicity Studies with Agrochemicals." U.S. EPA/SAB: Washington DC.

———. 2020b. "Attachment 2: Rethinking Carcinogenicity Assessment for Agrochemicals Project (ReCAAP); Herbicide1 – DRAFT Waiver based on Risk21 Approaches for Chronic/Carcinogenicity Studies." U.S. EPA/SAB: Washington DC.

U.S. Environmental Protection Agency Office of Chemical Safety and Pollution Prevention, and U.S. Environmental Protection Agency Office of Research and Development. 2020. "New Approach Methods and Reducing the Use of Laboratory Animals for Chronic and Carcinogenicity Testing." U.S. EPA/OCSPP & ORD: Washington DC.

U.S. Environmental Protection Agency Office of Pesticide Programs. 2013. "Guiding Principles for Data Requirements." U.S.EPA/OPP: Washington DC.

Dr. Tiffany Bredfeldt

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.**

        i. Under Section II item 1, the U.S. EPA should consider adding a section to the framework where study waiver applicants can report any known biomonitoring data or any form of historical exposure data if/when available.

        ii. Under Section II item 2 of the framework, the applicants are requested to report what is known about metabolism in mammals and which metabolites are formed in the environment. Metabolites formed in the environment would be better placed in a section describing environmental transport and fete of the chemical in questions to avoid confusion. This section could be referred to as needed throughout the document.

        iii. In item 2 of Section II, the framework document asks applicants to describe results from available toxicokinetic studies. The second bullet describing repeated dose evaluations is worded in a confusing manner. Please consider breaking these instructions into more sentences or rewording for clarity.

        iv. A MOA section should be included as a stand-alone section. Then, following MOA section, more emphasis could be placed on specific MOA where it is important to the exposure duration in each section. This just seems like MOA is being included as too much of an afterthought in subsections when it should be critical in the event chronic studies are being waived.

        v. Overall, the framework for this waiver are clear and well-written. However, they read like instructions. What is not clear is what WOE will be specifically required to waive long-term cancer bioassays. Though, the

[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

WOE is likely considered on a case-by-case basis, the guidelines need to consider in their scope required WOE (or do so in additional case studies). The framework could be more widely used and should be written in a manner that is more comprehensive when it comes to required WOE along with examples.

vi. It is clear that the central MOAs for carcinogenicity, i.e., genotoxicity, endocrine disruption, and immunosuppression, are being addressed in the framework. Additional data to be considered should more clearly requested, e.g., receptor-mediated assays, instead of adding such data into a vague "additional" data section, which is where this data was placed in the example provided in Attachment 2.

vii. The human-relevance of the proposed MOA should be discussed in the MOA section(s) or as a separate section. This evidence is a critical point of discussion that may be best emphasized in a MOA section rather than scattered throughout the document.

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.

i. For a variety of places in the document, figures (e.g., metabolism and MOA) would greatly improve clarity and ease of reading. It is valuable to encourage applicants to produce figures in these documents for ease of reading and transparency.

ii. The EPA provided a very thorough WOE analysis for Herbicide1. The overall report is clear and convincing. The available data support the choice to not require cancer bioassays for this chemical based upon what is known of the MOA and the available toxicity database. However, it would be nice to have another example where the chemical-specific database is less data rich. Such an example may provide clearer guidance for when EPA can apply this waiver and when they cannot. The criteria for waiving or not waiving data is not clear from the available framework. To have consistent application of the waiver, it may be necessary to discuss required WOE a bit more. This might require the production of more prescriptive guidelines or additional WOE case studies, which are apparently underway, to be made available.

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. The scope of the three focus areas for the TTP is appropriate and timely. The strong focus on AOP or MOA framework early in the process appears to be a wise choice for this collaborative project.

2. The scope and direction of the future goals for 2020 are appropriate. While initial work seems focused on collecting methods or models, there is significant emphasis on the mutational comparison between rodents and humans. Other cancer MOA should also be included in these early stages, so that they may matriculate in tandem. I would consider comparison of rodent versus human endocrine disruptors and immunosuppressors also important. I have concern that too much focus on comparing rodent to human data may simply be an act of comparative biology if contextualization is not well integrated throughout these early stages.

3. The structure and key elements of the TTP will be critical for its success. At this time, it is unclear what that structure will be of the TTP. A figure or scheme of that structure would benefit the clarity and make component stages of the pipeline more transparent in regard to their use and scope. It would seem that the obvious and most simple stages that make up the TTP are a development and growth stage for establishing assays, a stage for increased use in the scientific and risk assessment communities for developing and shaping application of assays or methods within these communities. Finally, the mature, characterized assays in final stages should be made available for broad and required use in the regulated community.

4. In addition to clarifying TTP stages, it will be important for inclusion and exclusion criteria to be developed within the context of an AOP/MOA framework for the application of new tools developed or refined in the TTP. Importantly, human relevance should be integrated into throughout the framework to enable it to be truly fit for purpose.

5. The Health and Environmental Sciences Institute (HESI) Point of Departure Program Overview project is of high value and importance. It is critical that we learn to derive POD from evolving

methods in vivo and in vitro. Some impressive progress has been made on these fronts.

6. The EPA and collaborators are right to begin their POD finding studies in rodent cell lines for comparison against in vivo studies. However, the challenge of human relevance is not necessarily met in this approach because studies in human cell lines are not being referred to in a manner that indicates they will be a critical part of POD finding studies. The EPA should consider the pathway forward that moves away from animals and animal cell lines and how to conduct risk assessments in human cell lines. It is apparent that EPA has considered these endpoints per the studies they have published in the past. However, it is unclear how EPA will incorporate human cells for comparison to rodent model systems in the described research projects.

7. In the Gene Expression Evaluation of Pesticides with Established Liver Tumor Modes of Action Project, the EPA is fortunate to have a collection of guideline studies and mechanistic information for many pesticides and I agree that these studies and provide a logical bridge between traditional studies and NAMS. These studies are critical for the risk assessment community to build confidence in NAMs.

8. The EPA chose 6 MIEs that are logical for the investigation of liver cancer in rodents. They at least establish a foundation upon which to build. The EPA indicates that the methods were derived using Affymetrix microarray data of rats exposed to pharmaceutical compounds. The following objective is a bit unclear. Does the EPA intend to generate a novel or new profiling platform or a custom one? More information is needed in this section to provide clarity upon what the EPA intends to do.

9. The primary mention of dose in section 3.3 seems to indicate that studies are being conducted with tumorigenic doses. It is unclear if the EPA intends to conduct studies with a range of doses for liver tumorigenic agents. It would seem logical for the EPA to not only include tumorigenic doses but lower doses where tumorigenesis may not be expected to determine the behavior of genes responding early in the process or whether there are transitions in the toxicity in or among AOP. It is unclear how the EPA intends to conduct these studies and clarity is needed.

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?

i. The EPA's interest in reducing animal research is commendable. However, it is unclear how the KMD-related activities will reduce animal use given that kinetic studies must be conducted to find the KMD.

ii. "KMD refers to the highest dose at, or slightly above the point of departure from dose proportionality, or PK linearity". The KMD approach is interesting, but it appears applicability will be highly chemical-dependent. There are a variety of reasons non-linear kinetics may occur: saturation in metabolism, saturation of absorption, and saturation of excretion. These cases may give rise to different toxicities depending on what step and chemical/metabolite is driving toxicity. With that in mind, the concept of KMD is complex. In cases where the KMD is achieved by saturation that results in accumulation of a less toxic form of a given chemical the saturation kinetics may not drive a maximumly tolerated dose or a meaningful maximal dose. Clearly, MOA must inform the decision to utilize KMD. Such a policy may make the application of KMD variable and, as such, inconsistent, an issue that EPA is aware of as mentioned in the Cancer NAMs White Paper. Further, too low dose selection during toxicity testing may give rise to misleading interpretations of toxicity data. It seems that KMD doses are higher than human exposure. Thus, if a KMD is to be used, it will be important to define human exposure when that data is available to contextualize the use of the KMD.

iii. The concept of inflection point where dose proportionality is lost may be flawed when one considers that saturation, by whatever mechanism, is generally not a rapid change. Thus, it may be a gradual event since saturation may not appear as a simple threshold or inflection point on the dose response curve. If we assume this could happen, at least in some instances, the KMD concept will be flawed. It is unclear how EPA may deal with this issue.

iv. The EPA should consider benchmark dose approach (BMD) as an alternative to the KMD approach. In this approach, it may be possible to use more dose groups to characterize the dose-response curve. In cases

where animal suffering appears significant, high dose studies could be halted.

v. Please see:

1. Harringa MB, Cnubben NHP, Slob W, Pronk MEJ, Muller A, Woutersen M, Hakkert BC. 2020. 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. 114:104659. https://doi.org/10.101106/j.yrtph.2020.104659

2. Woutersen M, Muller A, Pronk MEJ, Cnubben NHP, Hakkert BC. 2020. Regulating human safety: how dose selection in toxicity studies impacts human health hazard assessment and subsequent risk management options. Regulatory Toxicology and Pharmacology. 114:104660. https://doi.org/10.1016/j.yrtph.2020.104660

3. Kalantari F, Ringblom J, Sand S, Oberg M. 2017. Influence of distribution of animals between dose groups on estimated benchmark dose and animal distress for quantal responses. Risk Analysis:37:1716-1728. https://doi.org/10.1111/risa.12741

Dr. Karen Chou

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.**

    The outline, in general, provides a simple and useful outline, but can be improved with the following modifications:

    1) Under section II. 1. Use pattern & exposure scenarios: It may be helpful to state that the exposure profile should include information on both environmental exposure pathways and the routes of exposure. Statements should be made on each of the pathway- or route-specific exposure scenarios, so that it does not appear to be an omission or a lack of consideration. When there is no existing knowledge on a given scenario or when a given exposure scenario is not expected, this should be stated accordingly.

    2) Item II. 2. Physical-chemical Properties, the fifth Bullet Point: Although it is appropriate to list "potential for bioaccumulation" under physical-chemical properties, it is not appropriate to list "chronic toxicity" under this Item. Chronic toxicity should be under Item II.3. ADME & Toxicokinetics. Rationale for this recommendation: Although physical-chemical properties, such as octanol-water partition coefficient, may be used to predict bioaccumulation, the estimation of potential of chronic toxicity must take ADME and many other factors into consideration.

    3) Item II STUDY WAIVER REQUESTS: Add an item for "depletion of micronutrients and endogenous molecules that are important for the removal of reactive oxidants and electrophiles". The mechanisms of toxicity could be different between short-term/subchronic and chronic

[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

exposure. For example, chronic exposure may result in repeated and frequent depletion of micronutrients and endogenous molecules, which may not be observed or predictable by physiological, cellular, or sub-cellular (including genomic or proteomic) measurements after short-term or sub-chronic exposure. Micronutrient depletion that contributes to chronic toxicity can be caused by a long-term presence of pro-oxidants, cytokines, or reactive oxygen species. In addition, the magnification of the reactive oxygen species over a time course can be independent from bioaccumulation of the test substance. In addition, frequent and repeated depletion of the micronutrients (vitamins A, E, and C, beta-carotene, iron, zinc, selenium manganese, copper, etc.) and hydrophilic endogenous molecules required for Phase II biotransformation (glycine, taurine, glutamine, sulfate, glutathione, glucuronide, donors of acetyl and methyl groups, etc.), can result in adverse effects that may not be observable after short-term or subchronic exposure. The potential depletion of these pathologically important micronutrients and endogenous molecules, therefore, should be examined, reported and considered in the WOE assessment, and reviewed for the waiver decision.

4) Item II. 7. Conclusion, under "Clearly Summarize the following points": Add summary of assessment strength, weakness, and uncertainty. For rationale, see Comments under next charge question.

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:

The draft case study is clearly constructed according to the outline shown in Attachment 1. Appendix G presents additional evidence and arguments for the waiver request. The level of confidence of the waiver decision lies in the inductive and deductive reasoning, (a) extending the knowledge from short-term effects to long-term effects and (b) applying read across approaches to fill data gaps. Comparing with the method of hypothesis testing, i.e. actually testing potential chronic effects, both approaches applied in the waiver request are associated with higher uncertainty values. To communicate and share the waiver decision with other risk assessors and the public at large, and to ensure the scientific values of the assessment (i.e. objectivity and reproducibility) the strength, weakness, and uncertainties in each case-specific waiver decision should be explicitly stated in the summary (This comment is presented as a recommendation under the previous charge question).

Recommendations:

1) Section II.1, Paragraph Exposure.: The exposure profile should include statements for all possible environmental exposure pathways and the routes of exposure in humans. In the example presented in Attachment 2, the routes of dermal and inhalation exposure are not mentioned. If some of the routes of exposure are unlikely based on the proposed use of Herbicide 1, it should be stated so to provide clarity.
2) P. 4, Absorption: The absorption of all potential routes of exposure to Herbicide 1 should be reported. If data are not available for certain routes, it should be stated so to provide clarity. The information provide for the route of oral exposure is clear, condensed, and informative.
3) P. 5, first paragraph: Explain/define first label and second label. Otherwise, the paragraph is well done.
4) Section 3. The ranges of test doses in the studies reported under the subsections of Absorption, Distributing, Metabolism, and Excretion need to be reported, for comparing with the amounts of likely exposure of normal use, comparing with the results from other studies, and evaluating potential saturation of ADME pathways.
5) P. 7, "The NOAEL in rat was 58 mg/kg/day and 70 mg/kg/day for males and females respectively": Based on the data presented in Table 4, the reviewer finds the selection of NOAEL values questionable.

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.

Project 1, the US National Toxicology Program (DNTP) will populate the translational toxicology pipeline (TTP) with existing technologies and database and implementing AOP and IATA approaches, as a part of the effort to expand the use of alternative methods in assessing chemical carcinogenicity. A framework will be developed to link mechanisms to pathways for site-specific cancer. Carci HEI could enhance the communication among federal agencies and develop a communication strategy to deliver clear and actionable information to all stakeholders.

In the second project, led by HESI, alternative approaches will be applied to evaluate omic-technology and non-apical effect-based POD, such as transcriptomic POD, for human health risk assessment. Results from previous

studies, although limited, have demonstrated the possibility that short-term non-apical PODs may be used to predict chronic apical PODs.

The third project is designed to provide evidence to support the replacement of the long-term carcinogenicity testing study with short-term studies, by using (a) the liver-tumor database evaluated by CARC of EPA-OPP, (b) the discovery that liver-cancer biomarkers observed after short-term exposure can accurately identify chemical-dose combinations that cause liver tumors at 2-years post exposure, and (c) the observed threshold in the dose-response relationships of the six biomarkers.

The three projects are likely to provide supporting information for threshold dose-response relationship for liver cancer, justification for using data generated from short-term study for the assessment of potential long-term effects, including carcinogenicity, and potentially replacing some of the required toxicity testing studies, refining human health risk assessment, and reducing the number of animals used for testing studies. Together, they are likely to provide information, tools, and justification for the harmonization of guidelines for cancer and noncancer effects, including the assumption of threshold dose-response relationship. These projects can certainly facilitate new approaches of risk assessment toward the 3R goals.

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?**

    The assumption that an observed inflection point indicates kinetic saturation is possible and likely to be demonstrable using certain datasets generated from toxicity testing studies, but unlikely to be universal. Studies have shown that some fitted toxicokinetic models provide no inflection points (Heringa et al. 2020). In addition, toxicity testing studies are designed for observations that are the outcome of changes in both types, toxicokinetic and toxicodynamic pathways simultaneously. Selectively excluding test doses based on toxicokinetics alone or assuming an observed dataset is a consequence of toxicokinetics only could lead to misuse of the information.

As for the application of the KMD approach in the case of 1,3-dichloropropene, which is cited in the White Paper as an example for the KMD approach, observation of the test dose of 60 ppm was dismissed from the toxicity assessment. The decision was made based on the inflection point of 40 ppm, without deliberating potential human exposure. Furthermore, the toxicological meaning of the difference between the two doses, 40 ppm vs. 60 ppm, is not considered in the judgment of dismissal.

Recommendation for additional topics to be discussed at the KMD Workshop:

1. Define toxicologically sound validation requirements before the KMD concept can be recommended in regulatory testing guidelines.
2. Defining toxicologically-based “unacceptable top doses” when KMD approach is applied in toxicity testing studies.

Dr. Harvey Clewell

Discussion/Charge Questions

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.**

    I found the draft framework to be clearly written and comprehensive. I'm sure it will require modification/expansion as additional case studies are evaluated.

    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 provides an excellent example of the potential application of the framework for a chemical that causes renal toxicity from calculi formation, which is a fairly straightforward case. I would suggest that additional case studies are needed to illustrate the application of the guideline for other cases such as liver toxicity associated with enzyme induction, metabolism-related thyroid toxicity, etc. It would also be useful to included cases where the mutagenicity data is equivocal but read across suggests a non-mutagenic mode of action.

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 OPP efforts described in Section 3, including the collaborative projects with NIEHS and HESI, are important for moving methods of carcinogenic assessment

into the future, increasing the capability for rapidly identifying carcinogenic potential while reducing animal testing requirements, and should be continued.

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?**

    I applaud EPA-OPP for taking a proactive role in developing approaches for evaluating submissions that propose the consideration of a KMD to reduce (or interpret) animal testing. The HESI activities will provide an opportunity to document multiple case studies that illustrate both the value and the limitations of the KMD approach.

    The 1,3-DCP case study provides an example of a well-constructed argument for the application of a KMD to interpret animal toxicity study results (Bartels et al 2020). The argument included evidence of a kinetic nonlinearity together with evidence of significant depletion of glutathione in the target tissue at exposures about the KMD.

    However, in the EPA's evaluation of the carcinogenic potential of 1,3-DCP (EPA-HQ-OPP-2013-0154-0104) the determination to accept the proposed use of a KMD to eliminate consideration of tumor outcomes at high concentrations/doses was supported solely by the observations of kinetic nonlinearities. The crucial evidence of GSH depletion above the KMD is mentioned in the report, but it is not cited as weighing into the decision to apply the KMD. This illustrates the importance of the HESI effort to develop principles for documenting and evaluating KMD decisions.

Dr. David Hoel

Reducing animal testing.

When the Ames' Salmonella mutagenicity testing appeared, everyone was excited about avoiding animal testing for possible cancer effects. To test this idea, we at NIEHS took the data from all NTP cancer studies and correlated the cancer results with the Ames testing results. In this study 63 animal tests (mouse and rat following NTP's test design) were positive and 49 were negative. Of the positive cancer results only 28 were positive in the Ames mutagenicity assay. (Piegorsch W.W. and Hoel D.G. Exploring relationships between mutagenic and carcinogenic potencies. Mutation Research 196: 161-175, 1988.) This result ended the idea of simply using the Ames assay for cancer risk.

What I would like to see is a similar analysis using the more scientifically data sources and the resulting projections. This may have been done but I have not seen the results in a published analysis.

Dr. Wayne Landis

General Comments

The slide presentations were very useful and a more complete description of the program.

It was interesting that one of the EPA program manager felt compelled to describe the various constraints that EPA was under in order to make decisions on new chemicals and in other situations. Many of the panel members are well aware of those constraints having been in applied toxicology for many years, have worked on many EPA programs, and often teach courses on the operations of the various programs. The long comment seemed a pushback on the suggestions from the committee regarding better tools and techniques. While I understand the frustration inherent in the decision-making process as constrained by current law, there are better ways of doing the science as recognized in the slide deck.

It is not clear that EPA has set specific goals for its ability to correctly estimate toxicity of new chemicals or pesticides. What are the false toxic or false safe rates? We are being asked to review a program without a clear set of specifications for its performance.

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 guidance1 and current practice2 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**.

I do not have much to say in support of qualitative WOE approaches. I have seen them misused in arguments with such topics as atrazine effects and others. There are quantitative approaches to WOE published by EPA authors. Below are two citations.

Carriger et al. 2016. Bayesian networks improve causal environmental assessments for evidence-based policy. Environ. Sci. Technol. 2016, 50, 13195−13205. DOI: 10.1021/acs.est.6b03220

Carriger JF, Barron MG. 2016. A practical probabilistic graphical modeling tool for weighing ecological risk-based evidence. Soil and Sediment Contamination: An International Journal, 25:4, 476-487, DOI: 10.1080/15320383.2016.1171293

The approaches do require specific kinds of information to be collected during the WoE process.

    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.**

Seems too simple and not probabilistic. If the goal is to provide information for risk assessment then how can an approach that does address uncertainty with an attempt to be quantitative be the best approach. It is not clear to me how this approach would fit into the more quantitative approach as described by Carriger et al (see citations above).

One of the issues I see with the WOE approach is that it fails to incorporate the adverse outcome pathway as a causative model. While the slide deck does describe AOPs and potential uses, no mention is made that this is also a cause-effect model that can be used to guide the WOE process. What information from various toxicity tests can inform whether or not the various segments within an AOP are confirmed for carcinogenicity or other endpoints? AOPs can be made quantitative (although rarely done) and QSAR evidence can also be used to estimate a probability of a certain type of damage to DNA or the inhibition of a particular enzyme It is not clear that the WOE can take advantage of the AOP approach or that those building AOPs understand the potential utility of the approach in making rapid assessments of the toxicologic potential of a new chemical or pesticide.

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.**

As a former chair of the WWU IACUC and a proponent of the 3Rs I have some very specific questions. I did not really get answers to these questions in the presentations. What are the goals for accuracy and precision? Are they similar in predictive ability as conventional methods and where are the data? How do NAM approaches compare to conventional animal testing for accuracy and prediction of human toxicity? If there are deficits, are they acceptable given the goal of reducing the number of and pain to the animals. How are the extrapolations going to be made from both the NAM and gene-expression studies to human carcinogenicity? Are there better data extrapolation-modeling tools than have been conventional used and will the approach lead to the 3 Rs being realized?

Comment: If there are to be collaborative programs with other laboratories and the use of various datasets in applying WOE approaches the agency needs to upgrade its capability in data analysis, data presentation and experimental design. There are some examples in the slide deck that demonstrate these issues.

Slide 38 presents a regression of the 5-day genomic POD compared to the Apical Potency Value. The R^2 value is presented as 0.89. The upper and lower limits for BEPOD are presented, implying that the value is not a datapoint but an extrapolation from a regression model. The APV is a log range, why? The regression is supposed to present how well the BEPOD predicts the APV. Both axes are in a log format.

So first, the BEPOD is not a data point but the result of a regression. Therefore uncertainty is masked even the upper and lower limits. The limits are highly dependent on the regression model being used along with other factors. It appears that the regression points are used for the linear regression so that any uncertainty is not incorporated into the analysis making the R value appear better than if the uncertainty was incorporated Why use the ½ log for the APV? That is not uncertainty but just a range. So I would critique the graph as being a poor presentation of the relationship between the two variables because of the lack of incorporation of uncertainty. Slides 40 and 41 present a common misleading indication of variability in the estimation of toxicity and reference values. In slide 40 there is the presentation of three dose response curves just to the right of the diagram with the short-term animal study. Instead of plotting the results of each replicate for each dose it appears as if the mean value with a confidence interval for that result is plotted. Then a point is taken from that is supposed to be the BMDL. I have no confidence in that value for predicting toxicity. First, is the regression run on the actual datapoints or on the mean value? If it is run on just the mean value the most likely value (the basic regression line) will be the same as if all the data are incorporated but the confidence interval will be smaller. I also suggest that the predictive interval for a regression be estimated as well. The predictive interval is the bounds for the next observation to be taken, or what would likely happen at each of the concentrations to an individual replicate. Then the BMDL is taken, divided by what seems to be largely arbitrary. Slide 41 has a similar graph but with a wild curve. Second, the supposed experimental design of relatively few treatments but with many replicates is adding uncertainty to the analysis. Given that the BMDL is often at the lower concentrations where the transition is often going from a steeper slope to a very shallow slope it is that part of the curve that needs to be described. The traditional experimental design is not built to focus on this part of the dose-response. A better experimental design for determining the BMDL is to better describe this portion of the dose-response This can be done without an increase in animals. The better design to is have treatments that are not replicates of one exposure but are at a broad range of concentrations focusing on being accurate in that transition part of the dose-response curve.

An even better approach is to not use just one regression model but several approaches and then apply Bayesian averaging to summarize the results. In this manner we would better describe model uncertainty.

I did catch a comment on the use of artificial intelligence to describe the interactions. I do use case learning to derive interactions and Bayesian networks to estimate risk. Both of these tools are considered part of the AI or big data world—also simple ones. I suggest that for these studies we start by doing fundamental statistics better and then build up to AI tools.

Dr. Dennis Paustenbach

All too often, when the Agency brings topics to the public or SAB, the train has gone a long way down the tracks before enough introspection has occurred and the gathering of alternative approaches has been completed. I believe the NAM may fit into that box.

My overarching concern is that policy, not science, is currently the rationale for the moving the NAM approach forward in a fairly aggressive manner. It is among the first times in my 40 year career where an administration says "this is what we want to achieve, please supply the science or creativity to justify….and then make it happen by a particular date." That is almost never a good approach to dealing with any complicated problem. It is fine to be "science forcing" but it is not fine to arrive at a conclusion (with a date for implementation) without evaluating whether another approach could achieve 80% of the desired outcome and at the same time eliminate most of the risk of making the wrong decision. The 80/20 rule is almost always the most effective way to approach any significant challenge.

The desire to use fewer animals and to treat them in an ethical/humane manner is a very important goal. It is definitely overdue that we revamp the approach. We are 20-40 years behind where we should be with respect to optimizing animal testing and using in-vitro testing to inform how many animals are needed, the proper species, the proper dosing regimen, and an evaluation of the need to have animals suffer. These challenges are present for every chemical we need to evaluate. I want to believe that we know more than enough to identify optimize the amount of information that needs to be gathered without using so many animals or having them suffer.

However, I hope it is not necessary and I doubt that it is appropriate to flatly claim "we will no longer be using animals, or not many animals, within 10-15 years." Indeed, it is probably on the verge of irresponsible to approach a problem of this type in this manner. There is a rich history about why certain animal tests have been needed and the circumstances associated with selecting doses. We have known for many years that it is nearly impossible to predict which compensatory mechanisms may come into play, in a significant manner, when the whole animal is challenged by a toxicant at a given dose; and these different mechanisms/responses influence the outcome (sometimes beneficially, sometimes not). For those of us trained in the 1970s and 1980s, one thinks of phenobarbital and what we learned about compensatory mechanisms (and reversibility).

I say nearly irresponsible because there has not much discussion in the past 4-5 years, since this initiative gained steam, about the consequence of "what if we get this wrong?" There are 300,000,000 Americans who expect not to be harmed by chemicals that are present in our lives and they are looking forward to the benefits of the newly developed chemicals that will (and are) being synthesized or discovered in the natural environment. I am not counting the 1 billion others who just assume that if the United States or EU say a chemical at given doses is safe, then they are more than willing to assume it to be true.

I would dare say that if it isn't too late, it would be wise to go back and, perhaps, change the question, and try another approach. I want to believe that could we identify a goal and a

timeline, then ask if science can achieve the goal (or get close) of the elected or appointed officials. Would it not be more reasonable to challenge the toxicology/pharmacology community to achieve the following (simultaneously):

A) To cut back on the number of animals that need to be tested by 30% in 5 years, 50% in 8-9 years, and 75% within 15 years? It is even possible that we could have as a goal to use no more than 10% of the animals we use today, conducting studies on only the most important of endpoints, by 2040. This would leave open the door to use animal tests, only conducted humanely, to insure that the health of large populations of persons not be jeopardized.

B) To nearly eliminate animal testing for those agents (and for those animals) for which we are 90% certain of the outcome, before we even begin the test.

C) Apply the best knowledge today, based on 40 years of "read across analyses with animals", new knowledge of what can be learned in-vitro, what we have learned about chemical properties with respect to irritation and toxicity, insights from SAR and in-silico work, to identify those toxicological endpoint which need to have some level, not a massive level, of animal tests to give us confidence in our decision making.

D) Convene a panel to arrive within a very short time frame, at a statistically correct, but minimal, number of animals for conducting the most common (but necessary) animal tests.

E) Immediately convene a panel to identify methodologies (or for identifying doses for animal testing) that insure that the animals do not suffer during the study. Period. I believe that this can be achieved by thoughtful persons who have decades of experience in animal testing. It is not that difficult to imagine having as few as 5 animals have key blood parameters and visual examination tell us if the animal is seriously stressed or suffering; then come up with an appropriate protocol. In a large percent of studies, maybe all of them, doses can be selected which will yield adequate insight yet not cause pain to the animals.

F) Other "charge questions" can be identified but I hope the concept is apparent.

I have believed for that for more than 30 years (due to information gathered during work in PB-PK modeling, a recognition about the importance of chemical/physical properties)the toxicology, and in-vitro testing) that regulatory community could have cut back on the number of animals needed to gain approvals by more than 50%.

The continued "mining" of so-called "big data" surely will continue to help us find optimal approaches. I am aware that EPA and the pharmaceutical companies have invested heavily in answering some questions using the massive amount of data collected since 1960, but perhaps

not enough of those efforts were directed to the topics which this panel has been asked to address.

Like most things in life, it is nearly always better to work toward a goal incrementally rather than abruptly or by edict. I well remember the "dioxin wars" of about 1985-2005. In the U.S., we conducted hundreds of millions of research, assigned perhaps 400 or more senior scientists in the U.S. (perhaps 100+ within EPA) to focus on what we perceived as an urgent problem. There were many panels and many law suits over the "right number" for regulating the concentration of PCDD/PCDF in air, water, soil, sediments, fish, meat, consumer products, etc. After 20 years, the "science" and debate were not resolved.

In the EU, the same debates were happening. However, a couple of countries said "this is extremely complicated so let's just do something now and work toward a positive outcome." A couple of countries said "We want 25% lesser aerial and water emissions within 3-5 years. We then want to have 50-75% fewer emissions, compared to today, within 8-10 years. After 10 years, we will see if more work is needed." That initiative was accepted and implemented within 18 months. Within 5-7 years, aerial and water emissions were down 90%+ (way ahead of schedule). I learned a lot watching the differences between the "all or none" approach used in the U.S. and the wisdom of having "the long view" (for which the EU, China and Japan are well known).

I see the challenge of reducing the reliance on animal testing in the same manner. I urge that the scientists who have been brought together to listen to the Agency scientists and who now have been tasked to answer these "charge questions," to give some thought to another approach.

I have said for the past 10-15 years that convening panels like this, then constraining them to answer only the questions that the Agency "thinks" are important to answer for the Agency to move forward, is a wasted opportunity to draw on the creativity and wisdom of the incredible talent whom SAB has historically identified for their panels.

The first day of these SAB meetings, by my way of thinking, is to ask the committee, "are we looking at this problem in the right way?" And, perhaps state, "The Agency had this challenge and this is how our team, and our contractor, thought it should be tackled. Here are the results." Did we ask the right question or begin the process in the right way?

It is a shame, in my view, to gather together a team of advisors who collectively have nearly 800 man-years of experience, wisdom and finely tuned scientific minds, and who have the ability to "think out of the box" (not be constrained by what their supervisor or the administrator has said is the best way to tackle the problem), and then close out their engagement by asking them to focus on a specific set of "charge questions."

Having served on many panels over the years, like everyone else who have been tasked for this engagement, I would much rather the EPA have convened us BEFORE the Agency went very far down the road of figuring out "the answer." Time and again, as I feel certain regarding this work on New Approach Methods, this group would have tackled it in a slightly different or a significantly different manner.

Given the importance to the public health and well-being of Americans of demanding "we will generally be out of business of animal testing, or we will be using 80% less animals in 15 years because we believe that is the right thing to do," I would rather be absolutely certain that we had explored all the options. As noted previously, I am confident that 75-80% of animal testing can be eliminated if we draw on the best minds and am equally confident that we can learn what we need to know without putting the animals through uncomfortable or painful tests.

My comments on the "charge questions" are not substantial. The preliminary comments shared by Dr. Armitage on June 22, about 150 pages of them, are exceptional. And, they present a wide variety of solid ideas for the Agency to address.

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.**

    Comment: I was not that impressed with the draft risk-based WOE approach. It did not seem sufficiently thoughtful to convince me that the chronic study should be waived. However, like other panel members, I much prefer having illustrative examples of precisely how the Agency is thinking rather than read vague descriptions. So, please continue to add MANY more examples of how the Agency is thinking in future proposals.

    I believe the comments that you have received thus far cover all of the key themes that deserve to be addressed.

    One factor which did not receive discussion, and I sort of expected Dr. Belzer to mention it, is that the magnitude of testing should be proportional to the number of persons expected to be exposed and the anticipated daily dose. For decades, we have placed the bar much higher for understanding the public health hazard for pesticides or food additives to which the entire population (which has a range of susceptibilities) might be exposed vs. a new chemical for which only 100 workers might be exposed. The PMN process is illustrative of the approach.

---

[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

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.

Comment: Nothing to add beyond that which has been mentioned.

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.

This was a very good charge question. The comments you have received, for right now, are more than enough to digest.

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?

I agree with others that it is unclear to me that the current KMD approach is much better than relying on NOELs and the BMD. Perhaps I am missing something but this might deserve more discussion.

Dr. Isaac Pessah

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**.

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.**

    CONCERNS ABOUT THE APPROACH: I raised a fundamental concern about the lack of face validity and predictive value of the use of results from 5-day exposure transcriptomic data as an alternative to long-term carcinogenicity models. The point I raised hinges on the significant body of high quality peer reviewed literature compiled over the last 10 -15 years indicating that the patterns of mRNA expression levels are highly dependent on: (1) the life-stage at which they are measured (epigenomic modifiers), (2) the overall genomic background of the individual (leading to significant background gene influences that modify disease outcome- onset and severity; especially for non-monogenic diseases such as cancer and age-related neuropathology) and (3) lifestyle factors- especially in midlife and late-life onset diseases (multiple environmental factors that influence epigenetic marks on DNA that influence transcription). A five-day exposure model in an inbred mouse or rat strain will likely provide little predictive value

[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

of long-term outcomes in humans. Worse, it may lead to inaccurate, possibly misleading predictions. I would be willing to provide literature references, but have not because I was not part of the original review committee on this topic.

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

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?**

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 non-linear 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.

Comment 2: The KMD may prove to be most useful in determining the concentrations used in high throughput *in vitro* studies that are part of the NAMs. The dose range for assay can be informed using the KMD. Then, effects at kinetically impossible doses

observed *in vitro* could not be used to support incorrect hazard identifications as I have seen from IARC.

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 "read-across" 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 on: 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.**

    The document provides a useful approach for waiving studies. It would be useful to provide a "roadmap" for what is acceptable and also how decisions would be made. There are a wide variety of WOE approaches so some advice would be helpful here as well as some examples (perhaps they will be given in case studies).

    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 interesting and provides a good example. Hopefully there will be a variety of these (and perhaps a repository of cases that have been submitted). It would be useful to have a couple of studies where there is a failure.

    Some comments related to presentation and interpretation:

1. When there is an interval reported it is important to report the type of interval. It is not clear in examples if the values presented are mean plus or minus standard deviation or standard error or a confidence interval.
2. When p-values are reported it is important to identify what is being tested and to give additional information, for example effect size and sample size.

[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. 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.

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?

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 material on designing experiments, especially computer aided experiments? For example, what is a good design for determination of dose proportionality? Should the potential effect of interactions be considered?

Design of studies seems especially important and the EPA should be able to give good guidance about design. For example, if a poor design can lead to high estimates of KMD then study design needs to be explored and defense of the design should be part of the study results. Case study 7: Statistical tests to determine KMD from sparse data points seems important. 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.

These authors 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.

These authors argue that KMD is ill-advised for top-dose estimation. The following is noted 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."

There is 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.

These papers indicate the importance of the design and analysis of studies, as well as the interpretation and suggest that these topics should be considered in the workshop and guidance documents. Perhaps a part of the workshop and document would include discussion of criticisms of the approach.

The EPA should require researchers to make available (to the extent possible) all data used in KMD studies as well as the computer code used to analyze the data. Researchers should be encouraged (required) to provide comments to the code so that

others may use the code in their analyses. There are a variety of "best practices" articles that are relevant, for example Wilson et al. (2014).

Wilson G., Aruliah, D.A., Brown, C.T., Chue Hong, N.P., Davis, M., Guy, R.T., et al. (2014). Best Practices for Scientific Computing. PLoS Biol 12(1):e1001745. Https://doi.org/10.1371/journal.pbio.1001745

Dr. Laura Vandenberg

The agency requests the SAB provide comment on the following charge questions.

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.**

The problem that I see with use patterns and exposure assessments is that they are always limited to a single snapshot in time. I presume that this waiver would be requested prior to any human exposures. Thus, exposure assessments are going to be based on models and data from other agrochemicals, and not based on known uses. We have seen repeatedly over the years that use data (e.g., raw numbers of applied pesticide, use of pesticides in novel agricultural technologies, use of pesticides as harvest desiccants, etc.) collected for pesticide exposures are almost immediately out of date. Thus, starting with "exposure" as if these data are meaningful, especially over the lifetime of the pesticide, is concerning. Furthermore, questions regarding carcinogenesis and toxicity are not issues of risk, but rather issues of hazard. Thus, it is entirely unclear how exposure data would be used in this determination; even low exposures to carcinogens are expected to cause harm, when harm is evaluated across populations.
As an endocrinologist, I would also note that the three bullet points related to how "hormone perturbation" (I presume you mean endocrine disruption) will be evaluated are insufficient. Please review the numerous expert statements from the Endocrine Society and UNEP/WHO on how guideline outcomes are insufficient to evaluate endocrine outcomes. I have also written extensively on this topic.

Some references:

Bergman et al. The impact of endocrine disruption: A consensus statement on the state of the science. Environmental Health Perspectives. 2013;121:A104 - A6.

[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

Bergman Å, Heindel J, Jobling S, Kidd K, Zoeller R, eds. The State-of-the-Science of Endocrine Disrupting Chemicals – 2012. available from: http://www.who.int/iris/bitstream/10665/78101/1/9789241505031_eng.pdf.

Diamanti-Kandarakis et al. Endocrine-disrupting chemicals: an Endocrine Society scientific statement. Endocr Rev. 2009;30(4):293-342.

Zoeller et al. Endocrine-disrupting chemicals and public health protection: A statement of principles from the Endocrine Society. Endocrinology. 2012;153:4097 - 110.

Gore et al. EDC-2: The Endocrine Society's Second Scientific Statement on Endocrine-Disrupting Chemicals. Endocr Rev. 2015;36(6):E1-150.

Vandenberg et al. Endocrine disruptors and the future of toxicology testing - lessons from CLARITY-BPA. Nat Rev Endocrinol. 2019;15(6):366-74.

Vandenberg LN. Low dose effects challenge the evaluation of endocrine disrupting chemicals. Trends in Food Science & Technology. 2019;84:58-61.

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 illustrates exactly the concerns I raised above regarding exposure data. All exposure scenarios are below the level of concern by the HED. These are modeled data – how are they used? What do they look like? What happens when uses deviate from those scenarios, as happens for many modern pesticides and agrochemicals? Further, I am not convinced that the HED level of concern is sufficiently protective. How many pesticides/agrochemicals have exposure data below that level of concern, yet have epidemiological evidence suggesting harm in human populations? Frankly, language like "the possible total chronic exposure is very low" is insufficient and unscientific. What is meant by "very low"??

I am also not convinced by the "hormone perturbation" data that are provided. Where are the anogenital distance data for the F1 generation? Where are the confidence intervals/standard deviations for AGD in the F2 generation? Are there in vitro/mechanistic data (e.g., receptor binding assays) available for this chemical? 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.

NTP and NIEHS have been pioneers in the development of in vitro screening tests (e.g., Tox21 assays). What remains unclear is how entirely in vitro-derived data can actually be used by regulatory agencies that have defined an "adverse effect" as occurring in a whole/intact animal. How, on a very practical level, can animals be replaced if the in vitro work is never going to be considered "adverse"? This is a fundamental issue in the use of NAMs and I am not satisfied by the white paper, nor the responses by EPA staff during the public meeting in response to my queries. This seems like a first step that must be tackled prior to investing time and energy into developing assays.

I am also confused by Section 3.2, describing alternatives to the traditional carcinogenesis / chronic toxicity test guidelines including a short-term repeat dose studies (5 days) to derive transcriptomic PODs. These are certainly not NAMs, and it would be inappropriate to classify them as such. Although they may be cost saving, they are not animal sparing. Certainly, should sufficient evidence be provided that these short-term assays are equally predictive and more sensitive than the traditional assays, they should be explored further. But considering the EPA administrator's call to eliminate animal testing, this seems like a poor use of resources, as animal experimentation is not eliminated.

I also am not convinced that transcriptomic analyses are as straightforward as suggestion in this white paper. Perhaps even more than other kinds of data, there are many, many ways to evaluate these large datasets (and, if multiple organs are profiled, there is the potential to create truly "big data"). Is there guidance for which transcriptomic approaches are best? Most sensitive? Most predictive of harm? Most predictive of effects in human populations?

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?**