Summary Minutes

Scientific Advisory Committee on Alternative Toxicological Methods Meeting

September 2-3, 2020

ZoomGov Virtual Meeting

I. Contents

II. Location of Background Materials and Presentations

Background materials and presentations for the 2020 Scientific Advisory Committee on Alternative Toxicological Methods meeting are available on the National Toxicology Program Past SACATM Meetings page (https://ntp.niehs.nih.gov/events/past/index.html?type=SACATM)

III. Frequently Used Abbreviations

3Rs	replacement, reduction, and refinement of animal use
API	application programming interface
CATMoS	Collaborative Acute Toxicity Modeling Suite
cHTS	curated high-throughput screening
DoD	U.S. Department of Defense
EPA	U.S. Environmental Protection Agency
EU	European Union
EURL ECVAM	European Union Reference Laboratory for Alternatives to Animal Testing
FDA	U.S. Food and Drug Administration
GAO	U.S. Government Accountability Office
GHS	Globally Harmonized System for Classification and Labelling of Chemicals (United Nations)
HSLF	Humane Society Legislative Fund
HSUS	Humane Society of the United States
ICATM	International Cooperation on Alternative Test Methods
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ICE	Integrated Chemical Environment
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ILS	Integrated Laboratory Systems, Inc.
iPSC	induced pluripotent stem cell
IVIVE	in vitro to in vivo extrapolation

LLNA	murine local lymph node assay
MPS	microphysiological systems
MWG	ICCVAM Metrics Workgroup
NAMs	new approach methodologies
NICEATM	NTP Interagency Center for the Evaluation of Alternative Toxicological Methods
NIEHS	National Institute of Environmental Health Sciences
NIH	National Institutes of Health
NTP	National Toxicology Program
OECD	Organisation for Economic Co-operation and Development
OPERA	Open Structure-activity/property Relationship App
PCRM	Physicians Committee for Responsible Medicine
QSAR	quantitative structure-activity relationship
SACATM	Scientific Advisory Committee on Alternative Toxicological Methods
VOCs	volatile organic compounds

IV. Attendance

The Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) met virtually on September 2 and 3, 2020, via ZoomGov. The following individuals participated in the meeting. In addition to participants named below, over 200 people viewed the meeting via ZoomGov.

SACATM Members

Michael Bolger, PhD, Simulations Plus, Inc.

Joseph Charest, PhD, The Charles Stark Draper Laboratory, Inc.

Amy Clippinger, PhD, PETA International Science Consortium Ltd.

Kelly Coleman, PhD, DABT, RAC, Medtronic PLC

K. Nadira De Abrew, PhD, The Procter & Gamble Company (chair)

Sean Gehen, PhD, DABT, Corteva Agriscience

ClarLynda Williams-Devane, PhD, North Carolina Department of Health and Human Services

Ad Hoc SACATM Members

Szczepan Baran, VMD, MS, Novartis Institute for BioMedical Research

Denis Fourches, PhD, North Carolina State University[1]

Sue Leary, MS, Alternatives Research and Development Foundation

Kathryn Page, PhD, DABT, The Clorox Company

Priyanka Sura, DVM, MS, DABT, ANGUS Chemical Company

Tamara Tal, PhD, Helmholz-Centre for Environmental Research UFZ

Misti Ushio, PhD, TARA Biosystems, Inc.

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Principal Representatives

Brian Berridge, DVM, PhD, DACVP, National Institute of Environmental Health Sciences (NIEHS)

Suzanne Fitzpatrick, PhD, DABT, U.S. Food and Drug Administration (FDA) Center for Food Safety and Applied Nutrition

John Gordon, PhD, U.S. Consumer Product Safety Commission

Anna Lowit, PhD, U.S. Environmental Protection Agency (EPA), ICCVAM Co-chair

Emily Reinke, PhD, U.S. Department of Defense (acting principal agency representative), ICCVAM Co-chair

Richard Probst, DVM, MPH, DACLAM, National Institute for Occupational Safety and Health

Other ICCVAM Representatives

Paul Brown, PhD, FDA Center for Drug Evaluation and Research

Janet Carter, Occupational Safety and Health Administration

Warren Casey, PhD, DABT, NIEHS

Nicole Kleinstreuer, PhD, NIEHS

National Institute of Environmental Health Sciences Staff

Dori Germolec, PhD

John Maruca (Image Associates, NIEHS support contractor)

[1] Current affiliation: Oerth Bio™.

Elizabeth Maull, PhD, Designated Federal Official

Kyle Messier, PhD

Nathan Mitchiner (NETE, NIEHS support contractor)

Mary Wolfe, PhD

Rick Woychik, PhD

Integrated Laboratory Systems, Inc. (NICEATM support contractor) Staff

David Allen, PhD

Shannon Bell, PhD

Kamel Mansouri, PhD

Steven Morefield, MD

Catherine Sprankle, MS

Public

Elizabeth Baker, Esq., Physicians Committee for Responsible Medicine

Robert Patton, PhD, U.S. Department of Energy Oak Ridge National Laboratory

Katie Paul-Friedman, PhD, EPA Office of Research and Development

Sachdev Sidhu, PhD, University of Toronto

Kristie Sullivan, MPH, Physicians Committee for Responsible Medicine

Joseph Wu, MD, PhD, Stanford University School of Medicine

September 2, 2020

V. Welcome and Opening Remarks

Dr. Nadira De Abrew, The Procter & Gamble Company, chair of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM), called the meeting to order at 10:02 a.m. on September 2. SACATM members and ad hoc participants introduced themselves. Dr. Elizabeth Maull, National Toxicology Program (NTP), the SACATM Designated Federal Official, read the conflict of interest statement and reviewed meeting logistics.

In welcoming remarks, Dr. Rick Woychik, National Institute of Environmental Health Sciences (NIEHS) and NTP Director, noted the 20th anniversary of ICCVAM in 2020. He encouraged SACATM members to challenge ICCVAM to embrace advances in innovation and asked all the meeting participants to bring innovative suggestions to

the discussions. He introduced the International Cooperation on Alternative Test Methods (ICATM) representatives and recognized the SACATM members whose terms were ending. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) co-chairs Dr. Anna Lowit, U.S. Environmental Protection Agency (EPA), and Dr. Emily Reinke, U.S. Department of Defense (DoD), and acting NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) Director Dr. Nicole Kleinstreuer each welcomed the committee and thanked them for their attendance.

VI. ICCVAM – Past, Present, and Future

20 Years of Scientific Accomplishments

Dr. Reinke reviewed a timeline of ICCVAM accomplishments. Highlights included the establishment of ICATM in 2009; the reinvention of ICCVAM in 2013, which began an era of greater public engagement; and the addition of the National Institute of Standards and Technology as a member agency in 2016. ICCVAM, originally an ad hoc committee, was established as a permanent committee by the ICCVAM Authorization Act of 2000[2]. Dr. Reinke noted that most of ICCVAM's work is done by workgroups[3], and reviewed the focus areas of all the workgroups through ICCVAM's history. The direction and vision of ICCVAM was influenced by the 2007 National Research Council report[4] and the 2013 reinvention of ICCVAM reflected the direction articulated by the 2007 report. The number of ICCVAM and NICEATM abstracts and peer-reviewed publications has increased since 2013. The most recent edition[5] of the ICCVAM Biennial Progress Report, mandated by the ICCVAM Authorization Act, was published in July and summarizes activities during 2018-2019. In closing, Dr. Reinke reiterated that ICCVAM activities continue to be guided by the 2018 Strategic Roadmap, the three key goals of which are connecting end users with developers of alternative methods, establishing new validation approaches that are more flexible and efficient, and ensuring adoption and use of new methods by both regulators and industry.

Clarifying questions and comments: There were no clarifying questions for this presentation.

Implementing the Strategic Roadmap

Dr. Kleinstreuer reviewed activities undertaken by ICCVAM agencies to address the

[2] Available at https://ntp.niehs.nih.gov/iccvam/docs/about_docs/pl106545.pdf.
[3] For more information on ICCVAM workgroups, visit https://ntp.niehs.nih.gov/go/iccvam-wg.
[4] "Toxicity Testing in the 21st Century: A Vision and a Strategy," available at https://www.nap.edu/catalog/11970/toxicity-testing-in-the-21st-century-a-vision-and-a.
[5] Available at https://ntp.niehs.nih.gov/go/2019iccvamreport.

goals of the Strategic Roadmap. The success of the approach articulated in the roadmap has been evidenced by the guidances and policies issued by ICCVAM agencies such as the EPA and the U.S. Food and Drug Administration (FDA)[6]. She summarized a broad array of ongoing NICEATM efforts covering a range of endpoints, methods evaluations, data curation, tools development, and modeling activities, and noted that several of these topics would be covered in detail later in the agenda.

While implementation of a generic plan may seem like a stepwise process, Dr. Kleinstreuer noted that, in reality, all the activities inform one another. Specific implementation plans have been developed for acute endpoints (systemic toxicity, skin and eye irritation, and skin sensitization) that reflect EPA's goal of eliminating animal testing for the six-pack. Documents outlining U.S. agency information needs in these areas have been published. Dr. Kleinstreuer noted that waiver guidance issued by EPA has paved the way to eliminate requirements for dermal lethality testing. While still works in progress, great advances have been made in predicting toxicity by computational methods and implementing alternatives, as shown through the brief summaries of ongoing efforts in predicting acute systemic toxicity and eye and skin irritation.

Clarifying questions and comments: In response to a question from Dr. Kelly Coleman, Medtronic PLC, Dr. Kleinstreuer indicated that Dr. Warren Casey, NIEHS, is leading an NTP initiative to modernize carcinogenicity testing, of which genotoxicity is a critical component. Dr. Denis Fourches, North Carolina State University, asked if NICEATM's integrated approach for eye irritation testing of agrochemical formulations included in silico predictors. Dr. Kleinstreuer replied that current efforts are focused on combining in vitro test results but NICEATM could consider if adding in silico models adds value. Other activities on this project will expand the test substance set to include more formulations in the mild to moderate irritation range, and to better relate the test systems to human biology. When asked by Dr. Tamara Tal, Helmholz-Centre for Environmental Research UFZ about prioritization of activity areas, Dr. Kleinstreuer indicated that NICEATM's activities are driven by ICCVAM agency priorities, which, in turn are driven by the regulatory needs, number of tests being conducted, the limitations of existing tests, and the status of available alternatives. Prompted by a comment from Dr. Kathryn Page, The Clorox Company, Dr. Kleinstreuer agreed that the EPA 2015 guidance on eye irritation testing for antimicrobial cleaning products is a good example of how agencies can inform stakeholders of opportunities to use alternatives in lieu of animal tests, and it deserves to be more broadly communicated.

[6] More information at https://ntp.niehs.nih.gov/go/3rs-apps.

NTP Approaches to Assessment of Dermal Hypersensitivity: Using Alternative Methods to Predict Skin Sensitization

Dr. Dori Germolec, NIEHS, discussing NTP studies using alternative methods to predict skin sensitization, commented that great progress has been made in regulatory acceptance of alternative methods in this area. Skin sensitization has a well-defined adverse outcome pathway with a number of in vitro methods mapping to key events within the pathway. Current NTP activities focus on three of these methods: the direct peptide reactivity assay, the KeratinoSens™ assay, and the human cell line activation test. Individually, both the human cell line activation test and the KeratinoSens assay are more accurate in predicting human skin sensitization hazard than the widely used murine local lymph node assay (LLNA). Furthermore, most nonanimal testing strategies that combine multiple in vitro assays and in silico predictions perform better than the LLNA for predicting both human skin sensitization hazard and potency. NTP is testing over 200 substances to expand the applicability domain for these three assays. Substances being tested include mixtures and difficult-to-characterize chemicals nominated by EPA, the U.S. Consumer Product Safety Commission, FDA, and NTP.

As an example of how this can be applied in a regulatory framework, Dr. Germolec commented on an evaluation of six isothiazolinones using a defined approach that included an artificial neural network model. EPA used the results of this analysis for a draft risk assessment published in May, the first use of such information in a regulatory risk assessment. In summary, in vitro data can be, and is, used for hazard identification and risk assessment for this endpoint.

Clarifying questions and comments: Drs. Nadira De Abrew, Priyanka Sura (ANGUS Chemical Company), Joseph Charest (The Charles Stark Draper Laboratory, Inc.), Sean Gehen (Coreva Agriscience), and Tamara Tal had clarifying questions for Dr. Germolec. Dr. Germolec clarified that there are three key events in the skin sensitization adverse outcome pathway: binding to skin proteins, mobilization of dendritic cells, and response of keratinocytes. Several in vitro assays measure these events. The chemicals selected for testing were required to have LLNA data available. Others were recommended because they were particularly challenging. The ultimate goal was to evaluate and potentially expand the applicability domains of the assays. Criteria for determining accuracy of the in vitro method considered both human and LLNA data. Availability of the human patch test data is driving replacement of animal data with human data as the reference for this endpoint. When asked about gene-expression technologies as potential one-to-one replacements for animal tests, Dr. Germolec responded that she was not convinced that any single test will be able to replace the animal assay. Asked about efforts to harmonize acceptance criteria within the federal agencies, Dr. Germolec

indicated that each agency has its own set of criteria based on their information needs. Dr. Kleinstreuer added that each agency has different domains of chemicals that they deal with as well as different regulatory contexts. Internationally, ICCVAM is working for adoption of an Organisation for Economic Co-operation and Development (OECD) test guideline for a defined approach for skin sensitization testing. That will help with harmonization, but the information needed will still be up to the individual agencies.

Measuring Success of 3R Initiatives

Dr. Suzanne Fitzpatrick, FDA, and chair of the FDA Alternative Methods Working Group, provided an update on FDA activities, including the launch of the working group's webpage[7] meant to keep FDA stakeholders informed on the agency's activities in this area.

Dr. Fitzpatrick continued by introducing the ICCVAM Metrics Workgroup (MWG). In response to a recommendation in the U.S. Government Accountability Office (GAO) report "Animal Use in Research,"[8] ICCVAM established a workgroup to develop or implement metrics for federal agencies that demonstrate how acceptance of alternative methods affect animal use. The size of the MWG (19 members) representing eight of the ICCVAM agencies demonstrates the importance of this activity. The workgroup is compiling a report defining metrics relevant to each agency. Dr. Fitzpatrick emphasized that the focus of this report will be on animal use for toxicity testing, not research, and that the report findings will be communicated as recommendations to federal agencies, as ICCVAM does not have any authority over agencies.

Clarifying questions and comments: Responding to a question from Dr. Coleman, Dr. Fitzpatrick indicated that FDA's animal use is for testing; other agencies may also have different activities, including research.

The Strategic Roadmap: What Lies on the Horizon?

Dr. Casey provided an overview of key technologies that hold promise for the future along with challenges to their use:

- Machine learning, a promising technology that should be applied as much as possible, requires large volumes of high-quality, machine-readable data in a standardized format.
- High-throughput transcriptomics, an established technology that provides a broad spectrum of information on physiology and pathophysiology, needs an

[7] Available at https://www.fda.gov/science-research/about-science-research-fda/advancing-alternative-methods-fda.
[8] Available at https://www.gao.gov/products/GAO-19-629.

established knowledge base to capitalize on past work.

- Dynamic models of biological systems that allow researchers to interact with data to identify the tipping points in pathways rather than relying on existing categorical models that focus on a single parameter.
- Microphysiological systems (MPS) are starting to justify to the attention that has been given to them. The National Center for Advancing Translational Sciences, FDA, and NTP are investing in MPS; FDA is bringing MPS in-house. NTP and other agencies are supporting characterization of MPS for a range of use contexts through the TexVal consortium. NTP is also working with the National Centre for the 3Rs in England to develop MPS for cancer hazard assessment.
- Error-corrected DNA sequencing, a very sensitive technology with many potential applications, including human diagnostics and toxicity testing, is under evaluation by NTP to determine the usefulness of this technology to assess genotoxicity and other cancer-related endpoints.

Areas representing opportunities for advancement include developmental neurotoxicity, cardiovascular safety, cancer, and biological products. Dr. Casey shared some data from a European Union (EU) report on animal use for scientific purposes[9]. Referencing Dr. Coleman's question from the previous presentation, Dr. Casey noted that many more animals are used for basic and applied research than for regulatory testing in the EU. Within the regulatory context, animal use for efficacy testing of biological products including batch potency, far exceeds all other regulatory applications. These tests also tend to cause pain and distress. These endpoints are more tractable for decreasing animal use than, for example, predicting carcinogenicity; ICCVAM welcomes ideas from stakeholders on how to have a greater impact in this area. Dr. Casey noted the need for ICCVAM to continually evolve. As agencies put more effort into developing and advancing methods that are relevant to their own needs, ICCVAM needs to focus less on evaluating methods and more on coordinating the sharing of information.

Clarifying questions and comments: Drs. Kelly Coleman, Clarlynda Williams-Devane (North Carolina Department of Health and Human Services), Szczepan Baran (Novartis Institute for BioMedical Research), and Sean Gehan posed clarifying questions on Dr. Casey's presentation. In addressing the technologies covered, Dr. Casey indicated that NTP has several mutational signatures projects ongoing, including one that is comparing mutational signatures observed in animals to those found in humans. He recommended using legacy data and leveraging ongoing studies to advance high-

[9] Available at https://ec.europa.eu/environment/chemicals/lab_animals/reports_en.htm.

throughput transcriptomics. Those data can be used to develop a dynamic understanding of transcriptional networks and how they relate to organ toxicity. Dr. Casey suggested initially focusing on developing dynamic networks. The availability of these would then make the task of incorporating quantitative structure-activity relationship (QSAR) models to generate chemical activity predictions easier to address. In response to questions on sharing agency performance criteria, Dr. Casey commented that EPA publicized its accepted methods and FDA asked for input on performance criteria for MPS. While strategies for adopting technologies are agency specific, ICCVAM will be looking for input on this; however, overcoming policy issues and entrenched attitudes continues to be a challenge. As a next step to moving MPS towards regulatory applicability, the TexVal consortium data will be made publicly available. Any concerns or ideas for the consortium can be directed to NICEATM.

Public Comments

One written public comment was submitted for this section, on behalf of the Humane Society of the United States (HSUS) and Human Society Legislative Fund (HSLF).[10]

Oral Public Comments

Ms. Kristie Sullivan, representing the Physicians Committee for Responsible Medicine (PCRM), noted the increase in activities and impact of NICEATM and ICCVAM in recent years, especially the increase of transparency and engagement with the scientific community. She appreciated EPA's recent acceptance and utilization of alternatives for acute systemic toxicity and skin sensitization and hoped that other agencies would follow EPA's example in communications. The activities of the MWG have the potential to inform where limited resources could be applied most effectively. PCRM recommends SACATM encourage agencies to adopt data reporting practices that include all studies submitted with animal use information for each. A PCRM review of new drug applications submitted to FDA from 2015-2018 indicates that animal testing is still being widely used for acute systemic toxicity, irritation, and skin sensitization.

Comments from Designated SACATM Discussants

Dr. Coleman, Dr. Gehen, Ms. Sue Leary, Alternatives Research and Development Foundation, and Dr. Kathryn Page, The Clorox Company were asked to discuss a series of questions in context of the initial set of presentations. Ms. Leary commented that the 20th anniversary of ICCVAM is a significant milestone and it is important to recognize the progress that has been made. The work on the six-pack represents a good model for how to move forward. In response to the question on other useful data or information

[10] Written public comments for all topics are available at https://ntp.niehs.nih.gov/go/meeting; select Event Type "SACATM" and click on "Meeting Materials."

that would support implementation of the identified approaches for acute toxicity testing, Dr. Coleman offered information from a Medtronic data mining project that includes over 2000 tested chemicals, including irritants and sensitizers, to ensure coverage of the medical devices chemical space in CATMoS. Medtronic's data set could also be applied to an evaluation of publicly available QSAR programs to predict irritation, sensitization, and genotoxicity. Dr. Gehen supported addition of new data to ensure broad coverage of both chemical space and modes of action, especially for highly toxic chemicals and suggested that combining CATMoS predictions with additivity models would be a useful approach to estimate toxicity of mixtures. He also recommended generating CATMoS predictions in parallel with ongoing testing. Dr. Page also supported increasing the chemical space by including antimicrobials and pesticides in the validation of alternatives for skin irritation tests. Dr. Page questioned the use of the LLNA data as reference data rather than the Buehler test and asked if the formulations tested in vivo were identical to the ones used in the in vitro tests. She recommended broadly sharing data from a collaboration between Clorox, PCRM, EPA, and NICEATM on an evaluation of skin irritation models that could address the possible contributions of dermal absorption to overprediction.

Dr. Gehen, Ms. Leary, and Dr. Page provided some practical considerations. Dr. Gehen stated that getting past animal data as the gold standard remains a big challenge and suggested that implementing more human relevant mechanistic approaches will be important in overcoming the challenge. The lack of global harmonization remains a significant barrier. Dr. Gehen recommended defining, up front, what "success" will look like for using other reference data relative to highly variable animal data. Greater adoption of NAMs would be supported by incentivizing their use and reducing the risk associate with that use. Dr. Page recommended sharing information on how NAMs are currently being used to support international acceptance of products to get a better understanding of the poor adoption of available NAMs. Ms. Leary asked the agencies to be mindful of the impact of budget cuts on the advancement of NAMs and asked them to protect against those effects.

When asked for other specific technologies or NAMs to consider, Dr. Gehen supported the inclusion of MPS and transcriptomics and Dr. Coleman recommended that ICCVAM consider the Skimune® testing system. Ms. Leary added her commendation for engaging with National Centre for the 3Rs in the area of carcinogenicity.

Drs. Coleman and Gehen recommended genotoxicity and pyrogenicity, and nongenotoxic carcinogenicity, respectively, from both a risk assessment and hazard identification perspective, as potential endpoints for future consideration. Dr. Gehen added that having in vivo to in vitro extrapolation (IVIVE) information will facilitate the

use of in vitro methods in risk assessment applications. Ms. Leary appreciated that ICCVAM recognized the importance of reducing animal use in vaccine and carcinogenicity testing.

All the discussants had recommendations for how agencies could account for implementation of NAMs and its impact on animal use. Dr. Coleman suggested that agencies broadly communicate their efforts through NAM web portals that list both accepted test methods and resulting estimates of reduced animal use. He recommended that these web portals include an interactive component so that stakeholders could track NAM success and identify opportunities. Dr. Gehen added that having more information about actual animal use will help identify priority areas for advancing alternatives. Knowing when a submission used NAMs rather than traditional animal testing would help non-federal stakeholders identify opportunities to advance NAMs. Dr. Page commented on the progress made by EPA in advancing alternatives by partnering with stakeholders; this information needs to be more broadly publicized and training made available. The impact of waivers and bridging on reducing animal use should not be overlooked and should also be publicized. She noted that several agencies regulate very similar products; it might be helpful to identify the differences in agency requirements. Ms. Leary, recognizing the importance of the MWG, hoped that by developing metrics for toxicity testing, strategies for reducing animal use in research could be developed, as the GAO report was focused on reducing animal use in both testing and research.

Additional SACATM Comments

Dr. Kleinstreuer noted that recent FDA guidance[11] states that they no longer recommend use of the LLNA, but will consider a battery of in silico, in chemico, and in vitro studies that have been shown to adequately predict human skin sensitization with an accuracy similar to existing in vivo methods. She welcomed offers made to share data that would help NICEATM improve its computational models. NICEATM is working on evaluating the United Nations Globally Harmonized System for Classification and Labelling of Chemicals (GHS) mixtures equation for predicting toxicity of mixtures, and she agreed that it would be useful to use computational predictions as a part of this. It would be valuable for NICEATM to partner with industry stakeholders on case studies to obtain structures for some proprietary substances.

Dr. De Abrew asked Dr. Casey if there were any efforts to normalize the differences in what the various MPS are measuring. Dr. Casey responded that this is a question that

[11] Available at https://www.fda.gov/media/135312/download.

the tissue chip testing consortium is trying to address.

VII. Fostering International Partnerships

International Partner Updates

Dr. Lowit provided an update of ongoing activities among ICCVAM international partners. ICATM was initially established among the United States, EU, Canada, and Japan, with South Korea later added to the formal agreement. Brazil, China, Singapore, Taiwan, and OECD have also been participating. ICATM has been an effective international forum to advance alternatives. For example, a 2016 skin sensitization workshop focused almost entirely on regulators and enabled a hands-on discussion to advance elimination of animal testing for skin sensitization.

Dr. Lowit then turned to updates from ICATM participants.

- Canada is advancing 3Rs (replacement, reduction, and refinement of animal use) in both the research and regulatory arenas. Canada has ongoing activities in the areas of endocrine disruption, genotoxicity, applications of omics methods, and zebrafish models. The Canadian Centre for Alternatives to Animal Methods is developing in vitro disease models and nonanimal antibodies; they also have ongoing academic and regulatory activities.
- Recent EU activities include publications on animal use data, recommendations on use of nonanimal-derived antibodies, and leadership of activities within OECD and the United Nations. The European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) issues an annual status report summarizing their activities.
- Recent Japanese Center for the Validation of Alternative Methods efforts have focused on alternative methods for skin sensitization, skin irritation, and skin sensitization. Validation studies have been completed on a multi-immunotoxicity assay and a layer-by-layer 3D model skin irritation test method. A validation study is in progress for the epidermal sensitization assay.
- The Korean Center for the Validation of Alternative Methods has also been working on eye and skin irritation and skin sensitization. The MCTT HCE™ eye irritation test was accepted by OECD last year, and the KeraSkin skin irritation test is being considered within a revision of Test Guideline 439. Validation of the spectrophometric direct peptide reactivity assay for skin sensitization is planned for next year. Legislation under consideration in Korea would require a strategy to encourage development and utilization of alternative methods every five years

and codify the center's coordinating role in alternative methods development.

- The Brazilian Center for the Validation of Alternative Methods recommends test methods to the National Council for the Control of Animal Experimentation, which oversees regulatory adoption of validated test methods in Brazil. Most recently, the center recommended the monocyte activation test for pyrogen testing.
- The Taiwan National Health Research Institute is increasing awareness of alternatives to animal testing in Taiwan. The institute organized an international symposium on alternatives to animal testing in 2019 and established a website. Research projects are ongoing in developmental and reproductive toxicity, cardiotoxicity, neurological disease, and immunotoxicity.
- OECD plays an important role in global harmonization, and ICATM partners contribute to OECD guidelines. For example, the ICATM skin sensitization workshop prompted the development of a test guideline for defined approaches for skin sensitization expected to be adopted early 2021. Dr. Lowit summarized other OECD activities in the areas of eye irritation, developmental neurotoxicity, and acute fish toxicity. New policy and guidance efforts are ongoing for computational data, stem cells, in vitro approaches for developmental immunotoxicity, and mechanisms of the retinoid signaling pathway.

Clarifying questions and comments: There were no clarifying questions for this presentation.

FDA, ICH, and the 3Rs

Dr. Paul Brown, FDA Center for Drug Evaluation and Research, described how FDA's interactions with the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) impact the 3Rs. ICH, which includes both regulators and pharmaceutical manufacturers, reduces duplication of preclinical testing, clinical testing, and postmarketing evaluation by publishing technical guidelines that are implemented by regulatory authorities. Other ICH participants include a diverse group of governments, industry trade groups, and nongovernmental organizations.

ICH has issued over 100 guidelines, which are categorized as focusing on quality (Q), safety (S), efficacy (E), or multidisciplinary (M) topics. Dr. Brown summarized the ICH procedure for developing guidances and noted that participating regulators are committed to implement these once they are issued. He emphasized that ICH guidances focus on the later stages of drug development; regulators aren't involved in the discovery phase of drug development.

Multiple ICH guidances address recommendations for nonclinical studies, with ICH M3,

"Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals" being the primary guidance. ICH M3 specifies that in vitro methods should be considered and used when applicable in these studies. Dr. Brown described how nonclinical studies are conducted in parallel with clinical studies in each phase of the drug development timeline, and nonclinical studies are only done to support clinical studies that are actually conducted.

Dr. Brown described other ICH guidelines that impact animal use. The approach taken by ICH for evaluation of methods purposely reflects the approach used by ICCVAM and OECD for their evaluations.

Clarifying questions and comments: Drs. De Abrew and Baran asked several clarifying questions of Dr. Brown. Dr. Brown commented that ICH's harmonization activities have been driven by the pharmaceutical industry and he could not speak as to why this had not happened in other sectors. Asked about areas of opportunity to increase harmonization to further advance the 3Rs, Dr. Brown described several ongoing efforts including possibly waiving two-year rat carcinogenicity studies based on the totality of available information and comparing carcinogenicity predictions from drug companies to the actual outcomes of carcinogenicity studies. Ongoing work in immunotoxicity aims to advance the use of alternatives for skin sensitization. OECD adoption of the test guideline for defined approaches will hopefully drive the development of guidance on this topic.

Nonanimal Test Methods for Hazard Classification – Update on UN GHS Activities

Ms. Janet Carter, Occupational Safety and Health Administration, presented[12] an overview of the activities of the GHS nonanimal testing workgroup and ICCVAM's role in advancing acceptance of nonanimal methods in the context of the GHS. The GHS was established to standardize health and safety information internationally. It has 72 participating countries. Everything is done by consensus to increase the impact of their non-mandatory recommendations. The first GHS document, completed in 2001, encompassed classification of physical hazards, classification of health and environmental hazards, and hazard communication. The GHS is updated every two years to incorporate new data and information. It includes 10 health endpoints, some of which have specific classification criteria and others that are more subjective.

In 2015, the Netherlands and the United Kingdom proposed that an informal working group be established to facilitate use of data from nonanimal methods in the GHS. The working group focused first on skin corrosion and irritation, with the goal of developing

[12] Due to technical difficulties. Ms. Carter gave her presentation on Thursday, September 3.

recommendations based on existing data that would be sufficiently robust to accommodate new data. Skin irritation classifications are based on a tiered approach to testing and evaluation of existing information. The nonanimal working group recommended moving in vitro/ex vivo data up to the second tier in the process, to precede animal data. The working group is currently evaluating classification criteria for eye corrosion and irritation. These criteria consider skin irritation data and data from defined approaches. The working group will evaluate skin sensitization next.

Clarifying questions and comments: In response to a question posed by Dr. Kleinstreuer, Dr. Carter indicated that GHS specified that Tier 1 is based on existing data; there is a paragraph in the text discouraging prospective animal testing.

Asked by Dr. Sura for her thoughts on countries with their own classification systems and preferred testing methods, Dr. Carter indicated that the GHS is meant to be test-method neutral. It doesn't prescribe the use of specific tests, only some qualities of the test and data. Some countries will give preference to certain tests. Because of the expert judgment involved in some of these tests, there is going to be variability, and the GHS has some case studies that confirm this.

Public Comments

Public comments and SACATM discussion for this topic were combined with those from the following topic.

VIII. Moving Away from Animal-based Antibodies

Introduction

Dr. Casey reviewed current activities to promote nonanimal-derived antibodies. These are driven by an evaluation by the EURL ECVAM Scientific Advisory Committee, which found that nonanimal-derived antibodies offer significant advantages over animal-derived monoclonal antibodies. Europe is currently studying how to implement these findings, and NICEATM and ICCVAM are exploring how to support this goal in the United States. NICEATM co-organized a workshop in 2019 that recommended worldwide elimination of the ascites method, converting production of known antibodies to nonanimal means, and increasing education and communication. The 2020 ICCVAM Communities of Practice webinar focused on the use of animal-free affinity reagents[13]; NICEATM is currently collaborating on a webinar series[14] with its partners. ICCVAM is seeking advice from SACATM on how to move this forward with this topic as it is a

[13] For more information, visit https://ntp.niehs.nih.gov/go/commprac-2020.
[14] Webinar series information available at https://www.piscltd.org.uk/antibody-webinars/.

controversial area. How do we develop the necessary confidence to take the next steps? How do we gain support from antibody manufacturers and, most critically, the National Institutes of Health (NIH)?

Clarifying questions and comments: Responding to a question from Dr. Misti Ushio, TARA Biosystems, Inc., Dr. Casey indicated that educational efforts need to focus on both regulators and industry. Specific areas of concern include technical feasibility, cost, and availability. If educational efforts can address the scientific issues, then it will be easier to address others. He noted that false starts in the past have created lingering concerns with the technology. Dr. Tal asked if animal-based antibodies are being used mostly for research or regulatory applications. Dr. Casey, while acknowledging that research activities would be outside of ICCVAM's remit, considered the topic of broad interest and relevant to ICCVAM in that good research is needed to support good regulatory decisions. NICEATM is engaging in recombinant antibody activities in support of NTP interests. Dr. Tal commented that providing incentives for manufacturers and subsidies for users could help advance the nonanimal technologies. Dr. Casey indicated that was one recommendation resulting from the 2019 workshop. However, NIH is the only agency in the position to do that, reinforcing the need to effectively make the case to them for this technology.

COVID-19 Therapeutic Development with Synthetic Antibody Technology

Dr. Sachdev Sidhu, University of Toronto, presented a case study on the application of nonanimal-derived antibodies to drug development. Drawbacks to animal-derived antibodies include lack of control over the development and production process, which affects specificity, affinity, potency, and precision. These drawbacks can be addressed through a synthetic antibody platform that better defines and controls production. Dr. Sidhu noted that therapeutic antibodies have been very successful against a breadth of targets and are more highly validated than other types of therapeutics. He also noted that synthetic methods of antibody production are easier to scale up than natural antibody production.

Turning to the application of nonanimal-derived antibodies to COVID-19 research, Dr. Sidhu reviewed how antibodies act against the spike protein of the virus to block infection, forming the basis of an antibody-based therapeutic. Experience with Ebola and convalescent plasma therapy support the validity of this approach. Applying their platform to the manufacture of COVID-19 therapeutics, Dr. Sidhu's group first produced a variety of different antibodies to the COVID-19 spike protein. They identified three good candidates that had a similar profile to Herceptin, suggesting that they would be suitable for manufacture. The next steps convert the candidate from an immunoglobulin G to a multivalent antibody, which increases the affinity to the target. The converted

candidate is now being manufactured and prepared for clinical trials, and Dr. Sidhu emphasized that this was accomplished in less than five months. The next step is to target different epitopes on the spike protein, which could potentially lead to hundreds of lead antibodies and thousands of cocktails. Testing all these in animals would not be practical, so the group is exploring the use of tissue chips to evaluate the lead antibodies.

Clarifying questions and comments: Dr. Sidhu, responding to a question posed by Dr. Charest, indicated that they can produce about a hundred micrograms each of thousands of antibodies in a week. Dr. Ushio noted that this shows the advantage of the nonanimal-based platform and asked about roadblocks to broader adoption of this approach. Dr. Sidhu commented that the main roadblocks were political and institutional rather than technical. It's a matter of getting people past the platforms they're used to. The conceptual differences between these methods and hybridomas also require training. Dr. Kleinstreuer commented that lung tissue chip models are being used to study SARS-CoV-2 infection. She asked what the prerequisite would be for going into human clinical trials with these reagents. Dr. Sidhu replied that efficacy can be demonstrated in vitro, and because antibody therapeutics are very safe, safety studies should go quickly. The current candidates are good but there is opportunity for improvement. For that, studies in animals or advanced organoid tissue models may be important.

Public Comments

One written public comment was submitted for this section, on behalf of HSUS/HSLF.

Oral Public Comments

Ms. Elizabeth Baker, PCRM, noted that progress made in the United States on reducing animal use can have the most impact if international harmonization is achieved, and she encouraged ICCVAM to pursue this. In particular, she asked that ICCVAM support efforts to encourage China to join the OECD Mutual Acceptance of Data agreement. She commended FDA on launch of the new methods webinar series and the alternative methods workgroup. FDA's intention to provide flexibility to use nonanimal methods was reflected in today's presentation, but additional steps can be taken. She noted that ICH M3R2 states that the applicant "can use an alternative approach if the approach satisfies the requirements of the statutes and regulations," but the relevant regulations still require animal data. Regulations need to be updated to make them method neutral.

On the topic of antibodies, Ms. Baker noted that an influential 1999 National Academy of Sciences report on monoclonal antibody production needs to be updated to reflect current technologies. Because a wealth of research has undermined the justification of

use of animal-derived antibodies, she requested that NIH take the following actions:

- Stop funding development and use of ascites-derived antibodies.
- Develop and implement a roadmap for phaseout of other types of animal-derived antibodies by January 2021.
- Allocate funding to antibody producers that support production of nonanimal-derived antibodies.
- Add incentives and contingencies to research funding to push researchers to adopt nonanimal-derived antibodies.
- Consider the feasibility of producing a public library or partial library of recombinant antibodies to make them publicly available to researchers.

Comments from Designated SACATM Discussants: International Interactions

Discussants for "Fostering International Partnerships" were asked to identify areas of interest or specific projects that could be considered to help further international harmonization through international organizations. Amy Clippinger, PETA International Science Consortium Ltd., commented on a need for more consistent application of well-established and more sensitive nonanimal methods to shellfish toxin testing. Considering a recent publication[15] disputing the concerns over material-mediated pyrogens, she also noted an opportunity to advance the monocyte activation test for pyrogen detection. Dr. Sura cited a need for global education on available NAMs, including case studies detailing successful implementation of alternative approaches, and how to use data from alternative methods in risk assessment for the regulators. Ms. Leary commented that the use of animals in vaccine testing should be addressed.

Dr. Clippinger and Ms. Leary appreciated the impressive summary of ICH activities provided by Dr. Brown. Ms. Leary found that this was a great example of collaboration that can have worldwide impact. Dr. Clippinger suggested that ICH consider adding guidance on the use of nonanimal methods for skin sensitization evaluation, as found in the FDA draft guidance document, "Nonclinical Safety Evaluation of the Immunotoxic Potential of Drugs and Biologics," made available for public comment[16] in February 2020.

Dr. Sura commented that, while the chemical industry would welcome worldwide harmonization, this is challenging as some countries use a risk-based approach and others rely on hazard-based approaches. The amount of repeat testing needed to

[15] Available at https://www.altex.org/index.php/altex/article/view/1027.
[16] Federal Register, https://www.federalregister.gov/d/2020-03426.

satisfy the requirements of individual countries is problematic. She also noted that clarity on the translatability and relevance of NAMs would encourage industry-wide implementation and qualification.

Ms. Leary valued the presentation on ICATM accomplishments highlighting how partners working together can advance specific issues. Dr. Clippinger encouraged increased inclusion of China in OECD efforts.

Comments from Designated SACATM Discussants: Moving Away from Animal-based Antibodies

Discussants for "Moving Away from Animal-based Antibodies" were asked to consider activities to increase the availability and awareness of nonanimal-derived antibodies. Dr. Charest, the first discussant, suggested that there are ways to increase the awareness of nonanimal-derived antibodies, including marketing their utility and performance, inclusion in vendor catalogs, use by other scientists through shared protocols, and publication of use across a broad range of fields. Including antibody sequence in publications will reduce the need for revalidation. Producing entities should also have some ownership in the process, such as patentability.

Dr. Charest noted that standardized conditions for assessing specificity and affinity, as well as validation models, will enhance broader usage. Dr. Baran added that stakeholder involvement, including the regulators, in establishing validation criteria could also support broader use.

Dr. Misti Ushio, TARA Biosystems, commented that the points being made reflect a broader issue: to encourage adoption of new technologies, you need to identify key stakeholders (e.g., manufacturers and commercial and academic users), specific attributes of concern (e.g., quality, equivalency, cost, and infrastructure), and early adopters and their information needs.

Dr. Baran thought that the literature to establish the efficacy of these reagents already exists. There is a need to educate the stakeholders, relying on data and available science.

Addressing the perception of the high costs associated with the use of nonanimal-derived antibodies, Dr. Charest pointed out that the cost estimates do not account for the costs of failed or repeated experiments required by using animal-derived antibodies. Quantifying those costs could drive demand for nonanimal-derived reagents. Automating production could also drive down costs.

Dr. Sidhu commented that, from his experience, using nonanimal-derived antibodies is not necessarily more expensive and he thought that the assumptions used to come to

this conclusion may be outdated.

Additional SACATM Comments

Dr. Clippinger noted the advantages of nonanimal-derived antibodies. She encouraged NICEATM to compile an independent expert report on the state-of-the-science and the practical next steps for advancing reproducible recombinant antibodies. ICCVAM agencies should partner with international counterparts to advance use of these reagents. She also praised the use of lung-chip models for assessing antibody-based COVID-19 therapeutics and the effects of other inhaled substances.

Dr. Kleinstreuer asked if there was anyone on SACATM that did not support the idea of ICCVAM convening an independent peer review of the use of nonanimal-derived antibodies. There were no voiced objections.

Dr. De Abrew thanked the day's presenters and discussants. Dr. Maull asked participants to plan to provide feedback tomorrow afternoon on the platform used for this year's meeting. Dr. De Abrew adjourned the meeting for the day at 3:50 p.m.

September 3, 2020

Dr. De Abrew called the second day of the meeting to order at 10:06 a.m. SACATM members, ad hoc members, key NTP staff, and the ICCVAM co-chairs introduced themselves. Dr. Maull reviewed meeting logistics and read the conflict of interest statement.

IX. Curating and Characterizing Data for Alternative Methods Use

Incorporating Variability in Animal Studies into Regulatory Frameworks and NAM Assessment

Dr. Kleinstreuer commented that NICEATM devotes considerable effort to curating and characterizing reference data. These data can be used both to validate new approaches and to characterize the inherent uncertainty in the standard guideline test methods. Quality and variability of these data need to be characterized to provide a fair evaluation of new methods. NICEATM has found that reproducibility of reference data for binary hazard identification of various endpoints ranges from about 70-85%. Reproducibility rates for potency classifications of eye and skin irritation are lower. Examination of physicochemical properties showed no clear differences between chemicals that have reproducible skin irritation classifications and those that have variable classifications. Similarly, examination of chemotype data or physicochemical properties in an oral LD50

data set failed to indicate that chemical properties were causing the variability. Analysis of these data suggests that confidence intervals can be derived to represent the inherent variability of the animal tests.

Dr. Kleinstreuer described how QSAR toxicity predictions can be applied to reviews of reference databases to reveal data entry errors and unit curation issues. For example, discrepancies between predictions from the CATMoS model and reference data were found to be due to transcription errors or unit errors in the original in vivo data sets. NICEATM is currently extracting rabbit and rat data from NTP and European Chemicals Agency repeat dose developmental toxicity studies and mapping the primary data to controlled vocabularies and ontologies to facilitate computational analyses. NICEATM has automated this process and it can be adapted to different study types. The process is now being adapted and used by the EPA Integrated Risk Information System.

Dr. Kleinstreuer noted that while data curation is challenging, it is important for setting reasonable expectations for performance of new approaches. She also noted the importance of continuing to move toward greater use of human data and adverse outcome pathway-based mechanisms as references for developing and establishing confidence in new methods.

Clarifying questions and comments: Drs. Page, Williams-Devane, Fourches, and Coleman posed clarifying questions to Dr. Kleinstreuer. She indicated that while chemical properties affect acute toxicity predictions, they do not appear to affect variability of acute toxicity classification. Applying predictions across endpoints is an area of active research. NICEATM and collaborators at the National Center for Advancing Translational Sciences are applying machine learning methods to predicting acute toxicity using results from different species. Asked about efforts to automate processes, Dr. Kleinstreuer indicated that NICEATM's goal is to have a comprehensive semi-automated pipeline for identification of high-quality papers, extraction, curation, and annotation that will reduce both the time required and the potential for error. Dr. Patton will address other automation issues in his presentation. In response to a question on erroneous data found through public portals, Dr. Kleinstreuer commented on the importance of version control and communication to the data users and providers; NICEATM is exploring the most effective means to do this. Dr. Coleman, noting that rabbit tests are known to be highly variable, found it interesting that in the NICEATM data set, variability did not appear to be related to physicochemical properties and asked if NICEATM did those analyses with in vitro data. Dr. Kleinstreuer responded that variability analysis is built into the test method validation process, so the variability of these methods tends to be lower. Characterization of applicability domain also helps reduce variability. She agreed, however, that an examination of an actual data set would

be of interest. Dr. Coleman further noted that benchmarking against highly variable in vivo methods is a weakness of the current validation process. Dr. Kleinstreuer agreed and suggested that a better alternative might be a human biology-based validation approach that considers the mechanism of injury. Another approach is to use chemicals with highly reproducible results in validation efforts where animal data are being used as the reference.

Quantitative Variability in Repeat Dose Toxicity Studies: Implications for Scientific Confidence in New Approach Methodologies

Dr. Katie Paul-Friedman, EPA Office of Research and Development, noted that EPA requires that NAMs used for regulatory risk assessment be as good or better than animal methods. However, reproducibility of a new method, by definition, can't improve on the reproducibility of the method used to validate it. She presented EPA studies of quantitative variability for in vivo methods that measured how widely values within a data set ranged from the average. These studies considered the range of possible systemic values in replicate studies, the maximal accuracy that could be expected from a new model that predicts systemic effects, and the degree to which sources of variance could be identified. Data used were from ToxRefDB v2.0 and represented over 5000 studies submitted to the EPA Office of Pesticide Programs and other sources, mostly repeat-dose experiments in rat, mouse, and dog. The study divided observed variance into the variance that could be explained by study parameters and a remainder of unexplained variance that could be due to biology or parameters that weren't captured in study metadata. Two statistical approaches, multilinear regression and augmented cell means, were used to evaluate the variances. The two approaches have different characteristics such as stringency levels and ways to account for experimental factors. Analysis[17] suggested that a specific level of variance in systemic toxicity effect values can be explained, and the percent explained variance is stable across statistical models. This enables definition of minimum prediction intervals within which NAMs should be able to provide valid results. Previous work in this area has yielded similar results and suggests that 55-70% of in vivo data variability can be explained. This work has important implications for use of in vivo data for evaluating NAMs.

Clarifying questions and comments: Responding to questions from Drs. De Abrew and Tal, Dr. Paul-Friedman indicated that stratifying the data used in the EPA analysis by potency would have made the dataset too small, though that approach could be used with a larger dataset. It would be possible to apply the approach for quantitative points of departure to classification, but different statistical methods would be called for. Dr. Tal

[17] Pham et al. 2020. Computational Toxicology. https://doi.org/10.1016/j.comtox.2020.100126.

followed up with a question on the value of experiments to understand the source of unexplained variability. Dr. Paul-Friedman commented that, while recording data on easily characterized factors such as feed would be helpful, it was not sufficient rationale for conducting new animal experiments. However, these factors need to be considered when planning experiments with models becoming more widely used such as zebrafish.

Machine Learning in Toxicology: Towards Intelligent Access to the Content of Research

Dr. Robert Patton, Oak Ridge National Laboratory, described the Oak Ridge National Laboratory Computational Data Analytics Group's collaboration with NTP on automation of literature data collection for systematic review. They are automating the process to determine if an experiment meets minimum defined criteria for a study type. These criteria can encompass characteristics such as number of animals per group, animal model, administration route, and so on. Dr. Patton described the approaches used: a supervised approach, which developed classification criteria and applied them to extracting learned patterns and an unsupervised approach, which extracted text segments relevant to criteria descriptions and then classified the extracted segments. The NICEATM uterotrophic database[18] was used as a training dataset for the algorithm. Each sentence in a document was examined for information relevant to the minimum criteria and the top five sentences were extracted. Approaches used to improve the algorithm include applying context and targeting analysis to specific sections of papers. information delivered in non-text sections (e.g., tables or figures) remain a challenge. Current efforts focused on automatically extracting data from tables use a deep neural network developed to differentiate paragraph text from table text. Once the data extraction process is optimized, the group will look at millions of documents to make connections between documents that might not necessarily appear related. This is done by identifying concepts and finding the shortest paths between nodes representing papers. One study showed that new connections created with the addition of recent data can be a means for predicting new discoveries.

Clarifying questions and comments: Responding to a question from Dr. Charest, Dr. Patton stated that it is challenging to visualized different sections of older imaged-based PDFs; optical character recognition processing depends on the quality of the image.

Stem Cells and Genomics for Precision Cardiovascular Medicine

Dr. Joseph Wu, Stanford University School of Medicine, described how human induced pluripotent stem cells (iPSCs) can be applied to toxicological questions. His laboratory at

[18] Kleinstreuer et al. 2016. Environ Health Perspect 124(5):556-62. https://www.doi.org/10.1289/ehp.1510183.

Stanford derives iPSCs from individual patients and differentiates them into different types of cardiac cells. These can then be used to create engineered organoids for drug testing. They are using these resources in three areas of study: cardiovascular disease mechanisms, chemotherapy-induced cardiotoxicity, and screening new drugs for cardiotoxicity. Tissues generated from families with specific disease types are used to study mechanisms underlying these diseases to suggest appropriate therapies. Studies of chemotherapy-induced cardiotoxicity allow better understanding of mechanisms of toxicity and inform chemotherapy decisions and cardiac therapy. These data are also used to develop a "cardiac safety index" for new chemotherapy drugs. Additionally, tissues from diverse patients are being used to conduct "clinical trials in a dish" both to identify patients that are the most likely to benefit from new cardiac drugs and identify possible cardiotoxic effects of new drugs. These approaches could improve efficiency of clinical trials. Stanford has a biobank of samples from over 1000 patients that are available to other investigators for use. This work represents a human-based approach to drug testing that represents human genetic diversity in a way that's not possible in animal studies.

Clarifying questions and comments: Dr. Tal posed a question on how to determine if the genetic diversity included in his models captures all possibly toxicities for new drug trials. Dr. Wu commented that typical clinical trials need to be large due to the uncertainty inherent in these data; for example, it can't be determined whether all the participants took the medication as directed. The Stanford lab is currently treating samples from 100 patients under very controlled conditions to determine if this is a sufficiently large study to address the question of genetic diversity as well as that of outliers in clinical trials. They are also incorporating patients with genetic susceptibilities into the biobank. In response to a second question from Dr. Tal, Dr. Wu replied that the iPSC reprogramming process uses defined reagents; fetal bovine serum or other animal-based reagents are not used.

Public Comments

One written public comment was submitted for this section, on behalf of HSUS/HSLF.

Oral Public Comments

Dr. Sullivan, PCRM, commented that these presentations show the innovative approaches that are being used to evaluate data used to validate NAMs. This is crucial to leveling the playing field for new methods. For many endpoints, defining an appropriate reference standard will be challenging. She encouraged ICCVAM to think creatively about how to gain more practical experience using human data in a weight-of-evidence manner. Adverse outcome pathways could provide context for epidemiological

evidence or biomarkers from in vitro or clinical studies. Moving beyond animal data as the reference standard is necessary if we are to achieve the goal articulated in the Lautenberg Chemical Safety Act, which specified that NAMs perform as well or better than animal tests in predicting hazard. In a different context, it's possible that assessing reproducibility of in vivo models could be useful to efforts to evaluate and replace in vivo models being used to develop and assess pharmaceuticals. She closed by encouraging the pharmaceutical industry and ICCVAM to consider solutions to make more human data available for use in the evaluation of new methods for pharmaceuticals.

Comments from Designated SACATM Discussants

Dr. Gehen, first discussant for the session, commented that the variability of the in vivo data needs to be considered in evaluating new methods. Simplifying regulatory classification systems suggests opportunities for considering different approaches to using data. He also thought representing information as a range or probability may possibly be more useful than making discrete predictions. It is also important to understand sources of variability in context of human applicability and to anchor new testing approaches to human biology and mechanisms. Understanding sources of variability could lead to a better understanding of the range of biological responses or perturbations that might be expected from chemical exposures and could help define more relevant exposure levels. Dr. Tal added that the practical effects of data variability could be addressed by reducing the number of hazard categories for some endpoints. She recommended that ICCVAM organize a symposium or develop a white paper on developing reference ranges.

Dr. Charest noted that the greatest variability in hazard classification occurs in mild to moderate categories, possibly due to their narrow boundaries compared to the most or least toxic categories. There might be approaches to measuring toxicity that are more robust relative to sources of variability. He added that variation is not random; it might be worthwhile to consider how experimental details are recorded. Referring to the literature studies, Dr. Charest though that getting quantitative and numerical data whenever possible will be helpful for training artificial intelligence approaches and developing new categories. He commented that bias might be introduced when less readable PDFs are discarded.

Dr. Tal recommended that the approach for identifying errors in reference data described by Dr. Kleinstreuer be automated to identify discordant data. She also asked that the notations be made in the ICE database to enable the user to select the latest, most curated data set for analysis.

Addressing a question specific to stem cells and precision medicine, Dr. Charest

commented that the stem cell concept is a great approach to capturing genetic variation, but it is important to consider how the physiology of the organoids will affect results. He closed by wondering about the practicality of using iPSCs to personalize drug therapies, and whether the technology exists to apply this approach in real time during a patient's treatment. Using iPSCs to set exclusion criteria for clinical trials is a good near-term goal. Dr. Tal suggested that an ICCVAM agency fund a case study using the stem cell platforms described to determine the ability of NAMs to detect toxicity across a range of genetic variability, and how other sources of variability could be captured.

Dr. Ushio, fourth discussant, felt that these talks addressed the challenges inherent to using animals for research. She agreed with Dr. Sullivan that we need to move toward relating new methods to human data. It's important to demonstrate that in vitro human models are human-relevant and look for the best way to bridge that to humans.

Additional SACATM Comments

Dr. De Abrew asked Dr. Kleinstreuer if any factors other than physiocochemical properties were examined with regards to the data variability, such as the impact of expert judgment. Dr. Kleinstreuer replied that although expert judgment wasn't involved in the classification, the subjective nature of the scoring for some of these methods is an important potential source of variability.

Dr. Page commented that implementing NAMs presents an opportunity to make models that are better than animal tests. It's important that we understand the science behind the endpoints of concern. That knowledge can be used to identify cases where nonconcordance is due to differences between animals and humans and justify why new methods might be an improvement on animal models.

X. Computational Resources

Introduction

Dr. Kleinstreuer noted the importance of computational tools to chemical evaluations. This session highlighted tools NICEATM has developed or updated over the past year. She summarized how NICEATM has acted upon previous SACATM feedback to improve these tools. Since this session includes presentations on ICE and CATMoS, Dr. Kleinstreuer gave brief overviews of other tools that NICEATM has participated in developing:

- InterPred[19] uses QSAR models to predict whether a chemical will interfere with

[19] Available at https://sandbox.ntp.niehs.nih.gov/interferences/.

assay readout mechanisms.

- Tox21 BodyMap[20] visualizes where a chemical might produce effects in the body based on the Tox21 HTS data.
- ChemMaps[21] enables visualization of a group of chemicals within the universe of chemical space; the user can specify chemical properties of interest. ChemMaps can also be used to standardize chemical structures and generate molecular descriptors for use in other applications.

Clarifying questions and comments: There were no clarifying questions for this presentation.

Integrated Chemical Environment (ICE) & In Vitro to In Vivo Extrapolation (IVIVE)

Dr. Shannon Bell, Integrated Laboratory Systems (ILS; contractor supporting NICEATM), provided an update on ICE, a resource point for data and tools for NICEATM stakeholders. Because NICEATM stakeholders are a diverse group, it's important that ICE be user-friendly, accessible, and have high-quality data. Dr. Bell reviewed the sources and types of data in ICE as well as what users can do in ICE. Recent and ongoing developments include establishing an advisory group to provide a diversity of user viewpoints, improving integration with the EPA Chemistry Dashboard and the NTP Chemical Effects in Biological Systems database, and increasing visualization and interactivity. Dr. Bell highlighted a recent publication that describes a major ICE[22] update. Otherwise, all updates to ICE are publicized via NICEATM News. Dr. Bell gave a quick overview of some ICE features:

- ICE[23] Chemical Quick Lists, which can be used to populate a search on chemicals of interest.
- Curated high-throughput screening (cHTS) data are Tox21 data that have been curated to remove questionable data (i.e., chemicals which fail analytical quality control) and annotate assays to cell processes and modes of action. The annotation helps users who are not familiar with Tox21 assays but know what biological endpoints they're interested in.
- ICE IVIVE tool now supports upload of user-provided data.
- Chemical Characterization tool has the potential to examine properties of

[20] Available at https://sandbox.ntp.niehs.nih.gov/bodymap/.
[21] Available at https://sandbox.ntp.niehs.nih.gov/chemmaps/.
[22] Bell et al. 2020. Toxicol In Vitro. https://doi.org/10.1016/j.tiv.2020.104916.
[23] Available at https://ice.ntp.niehs.nih.gov/.

chemicals that, for example, perform differently in an assay. This tool now features principal component analysis plots, providing another visualization option for examining different groups of chemicals.

Clarifying questions and comments: Dr. Bell responded to clarifying questions posed through chat and from Drs. William-Devane, Bolger (Simulations Plus, Inc.), and Fourches. Dr. Bell noted that ICE is a public website; however, the IVIVE tool is also available as an R notebook that can be downloaded and run locally. Documentation for renaming a chemical list in the Chemical Characterization tool is found in both the downloadable user guide and under an info button on the website. Dr. Williams-Devane recommended highlighting this useful feature in a video tutorial. Dr. Bell added that an application programming interface (API) is available to use with ICE and all the data in ICE are available for export as a flat file.

Responding to a question from Dr. Bolger, Dr. Bell indicated that the equivalent administered dose used in the IVIVE tool is based on the predicted fraction unbound to plasma protein. NICEATM plans to add a feature that allows the user to control how this is modeled. The IVIVE tool uses fraction unbound and intrinsic clearance as inputs and these parameters are currently only available as predictions from the Open Structure-activity/property Relationship App (OPERA). Experimental data for these parameters will be included in the next ICE update, allowing the user to select either experimental or in silico data for their simulation.

Dr. Fourches asked how ICE interacts with the EPA Chemistry Dashboard. Dr. Bell indicated that DSSTox identifiers used in ICE link to the Chemistry Dashboard, allowing the user to explore single chemicals in more detail. ICE also allows the user to export search results in batch mode to the EPA Dashboard.

Collaborative Modeling Project for Predicting Acute Oral Toxicity (CATMoS)

Dr. Kamel Mansouri, ILS (contractor supporting NICEATM), provided an update on the acute oral toxicity CATMoS model. This project followed an approach used by EPA on two earlier projects to predict estrogen receptor[24] and androgen receptor[25] activity. ICCVAM agencies identified five endpoints of interest for acute systemic toxicity: point estimates of LD50, highly toxic chemicals, nontoxic chemicals, and EPA and GHS hazard classifications. Dr. Mansouri reviewed the data provided to develop the CATMoS models, the cleaning and curation procedure, and establishment of the training and evaluation datasets. The evaluation dataset, embedded in the prediction set, contained chemicals of regulatory interest including ToxCast/Tox21 chemicals, EPA Endocrine

[24] Mansouri et al. 2016. Environ Health Perspect. https://doi.org/10.1289/ehp.1510267.
[25] Mansouri et al. 2020. Environ Health Perspect. https://doi.org/10.1289/EHP5580.

Disruptor Screening Program chemicals, and chemicals submitted to EPA under the Toxic Substances Control Act. Nearly 140 models were developed using diverse machine learning approaches; each chemical on the list was predicted by at least 10 models, and most chemicals were predicted by at least 20 models. There was a high degree of concordance among the models, especially for predicting very toxic substances.

Dr. Mansouri reviewed development of the consensus model and the consensus predictions, which used a majority rule weight-of-evidence approach. The CATMoS consensus predictions performed as well as replicate in vivo experiments in predicting oral acute toxicity outcomes. NICEATM is collaborating with 10 offices within a number of ICCVAM agencies to apply CATMoS; these collaborations have generated predictions for over 9000 substances. CATMoS, which is available through OPERA, can be run as a downloadable command-line application or as a web-based graphical user interface. In addition to providing CATMoS predictions, OPERA can predict physicochemical properties, environmental fate properties, absorbance/distribution/metabolism/excretion properties, and other toxicity endpoints. OPERA now has a QSAR-ready standardization tool that improves consistency of predictions. OPERA also provides an assessment of whether the chemical is within the applicability domain of the relevant model, as well as an accuracy assessment of the prediction. OPERA predictions on approximately 800,000 substances are available via ICE. Dr. Mansouri closed by noting that the success of this project was due in large part to the collaboration of a broad range of stakeholders and especially the contribution of regulators.

Clarifying questions and comments: There were no clarifying questions for this presentation.

Toxicokinetic and Toxicological Based Geospatial Risk Mapping

Dr. Kyle Messier, NTP, described his work to provide a human context to communicate and visualize toxicological data. His current project arose out of an interest in relating modeled air pollution concentrations to biological activity and using these relationships to characterize geographic risk. Data needs for the project included:

- Spatiotemporal models of exposure and population.
- Toxicokinetic information.
- Toxicological dose-response information.

A good monitoring network for volatile organic compounds (VOCs), from which spaciotemporal exposure data can be obtained, is available in the United States. This project focused on 46 VOCs for which high-quality toxicological data and toxicokinetic

models were available. The advantages of this approach over other risk assessment approaches include:

- Incorporation of spatial and population context.
- Utilization of toxicological data rather than epidemiological data.
- Flexibility in considering in vitro or in vivo dose-response.
- Consideration of different biological targets.

This approach uses a risk evaluation approach like that used by the pharmaceutical industry, which can help with interpretability. Results of these analyses can be discussed in terms of hazard or risk quotient, population impact, or disease burden.

Clarifying questions and comments: There were no clarifying questions for this presentation.

Public Comments

One written public comment was submitted for this section, on behalf of HSUS/HSLF.

Oral Public Comments

Dr. Kristie Sullivan, PCRM, indicated that her organization supports NICEATM's efforts to make user-friendly models available. PCRM has worked with NICEATM to offer training. She encouraged companies to share data to support tools such as CATMoS, and ICCVAM agencies to utilize and encourage the use of these tools to replace in vivo testing wherever possible.

Comments from Designated SACATM Discussants

Discussants for "Computational Resources" were unanimous in praising the progress made over the past year. Dr. Page commented that NICEATM listened to and acted on the feedback provided in the past year, and especially liked the improvements made to the user guides. Dr. Williams-Devane noted the improvement in the use of unrestricted data that maximizes access by academic and other users. Dr. Bolger thought the tooltips were a useful addition and agreed that the user guides were very helpful. Dr. Fourches added his praise for including in silico predictions in the ICE database and the increased interactivity with the EPA Chemistry Dashboard. He commented that the added ability for users to upload their data was in response to a request made at last year's SACATM meeting. Dr. Fourches also commented on OPERA's great success, noting that it integrates well into other resources. The graphical user interface has facilitated wider use.

In addressing a question regarding future uses for ICE, Dr. Bolger thought that ICE

could be useful for screening large numbers of substances to get a rank order for toxicity, but further testing would be needed to obtain specific equivalent administered doses. Dr. Fourches suggested development of hybrid models that use both chemical descriptors and in vitro data to make toxicity predictions. Dr. Page would like to see expansion of ICE to mixtures analyses and to enable exploration of interactions.

Discussants identified some areas for improvement. Dr. Williams-Devane thought it would useful to identify if the hazard classifications for specific datasets were based on regulatory guidelines or internal curation. The use of the DSSTox identifiers with ICE allows data to be integrated with resources such as gene expression and other data and tools from the National Center for Biotechnology Information. Additional APIs to support these interactions would be useful. Dr. Williams-Devane appreciated the interactivity of the ICE Search function with the EPA Chemicals Dashboard and the ability to search results; however, the ability to use the search box in the filtering dialog to create a custom filter was not intuitive. She suggested adding visual cues to improve this.

Commenting on his personal experiences, Dr. Bolger noted some errors in downloading ICE data sets to Excel format; some of the column headers in the downloaded files were not clear. There was also an apparent conflict between some of the data in the inhalation dataset. Dr. Bolger was also unsure of the accuracy of some of the calculated properties used in the IVIVE tool and how corrections for fraction unbound were made; he had some questions about values he had observed for pKa and intrinsic clearance. Values that are experimental or (predicted)? properties should be clearly identified. Dr. Bolger asked if there was a custom download for this.

Dr. Fourches commented on the utility of the principal component analysis plots; linking this to ChemMaps and providing more opportunities to generate “paper-ready” figures would encourage use of ICE. He recommended integrating more visualization tools such as ToxPi diagrams for chemical structures and providing additional information about functional groups. Enhancing ICE to support stereochemical variants would be very useful, as this affects toxicity predictions. It is important to be aware that neutral compounds and analogous salts can have different CAS numbers, a system of identification that ICE relies on. This should be made clear to users. In the ICE Search, Dr. Fourches was not clear about the option to “add chemical with identical QSAR structure” function. Dr. Fourches also commented that the OPERA predictions based on consensus models should include some indication of the range of contributing predictions.

Dr. Williams-Devane considered making ICE user-friendly a priority as advancing the use of computational tools for predicting toxicity depends on engaging users without a

toxicology background.

Additional SACATM Comments

Dr. Kleinstreuer responded to points made in the discussion. NICEATM has been studying how ICE can better support analysis of mixtures. In response to Dr. Fourches' question, she noted that QSAR-ready structure tool allows Search to return information on both neutral and salt forms of a chemical. Clearly the documentation on this feature needs to be improved. She expressed appreciation for Dr. Bolger's detailed comments and stated that NICEATM would follow up with him offline to discuss them further.

XI. Adjournment

Dr. De Abrew thanked the organizers and staff supporting the meeting and expressed his appreciation for the SACATM members' participation. Dr. Brian Berridge, NTP, stated that he was impressed by the progress realized by NICEATM and ICCVAM over the last year. Dr. Lowit thanked the public commenters for their participation and noted the high attendance over the two days of the meeting.

Dr. Maull reminded attendees that slides will be made available on the NTP website when they meet government accessibility guidelines. Attendees will be notified when slides and minutes are available.

Dr. De Abrew adjourned the meeting at 3:52 p.m.

(signature redacted)

K. Nadira De Abrew, PhD
SACATM Chair
Date: January 26th 2021