NIH Public Workshop on Catalyzing Development and Use of Novel Alternative Methods
Meeting Summary
Bethesda, Maryland
August 21, 2023

WORKING GROUP MEMBERS

Howard Y. Chang, MD, PhD, *Stanford University (Co-Chair)*
Lyric Jorgenson, PhD, *National Institutes of Health (Co-Chair)*
Antonio Baines, PhD, *NC Central University/University of North Carolina*
Szczepan W. Baran, VMD, MS, *VeriSIM Life*
Wendy Chapman, PhD, *University of Melbourne*
Myrtle Davis, DVM, PhD, ATS, *Bristol Myers Squibb*
Linda Griffith, PhD, *Massachusetts Institute of Technology*
Ranu Jung, PhD, *University of Arkansas*
Arnold Kriegstein, MD, PhD, *University of California, San Francisco*
Nancy E. Lane, MD, *University of California, Davis*
Kelly Metcalf Pate, DVM, PhD, *Massachusetts Institute of Technology*
Sergiu Pasca, MD, *Stanford University*
Gordana Vunjak-Novakovic, PhD, *Columbia University*
Danilo A. Tagle, PhD, *National Institutes of Health* (*Ex Officio)*
Maureen Gwinn, PhD, *Environmental Protection Agency* (*Ex Officio)*
Namandje Bumpus, PhD, *Food and Drug Administration* (*Ex Officio)*
Brittany Chao, DPhil, *National Institutes of Health* (*Executive Secretary)*
Jessica Creery, PhD, *National Institutes of Health* (*Executive Secretary)*

AGENDA

9:00 AM Welcome

Howard Chang, MD, PhD & Lyric Jorgenson, PhD – ACD NAMs Working Group Co-Chairs

9:15 AM The Opportunities and Challenges for NAMs in Biomedical Research

Successful deployment of NAMs, whether for conducting basic research, uncovering disease mechanisms, or translating knowledge into products or practice, relies on bringing together disciplines, technologies, and data. This session focuses on research areas for which NAMs have been impactful to identify best practices for leveraging these approaches.

Moderator: Nancy Lane, MD – ACD NAMs Working Group

Presenters:

- Nathan Price, PhD – Thorne HealthTech *(novel mechanisms)*
- Thomas Hartung, MD – John Hopkins University *(translation/product development)*
- Nicole Kleinstreuer, PhD – U.S. National Institutes of Health *(regulatory processes)*
- Chirag Patel, PhD – Harvard Medical School *(inter-individual differences)*

10:30 AM BREAK

10:45 AM Cross Sector Approaches for Driving NAMs Use and Development

Each sector within the biomedical research enterprise has a role to play in catalyzing the development and use of NAMs. This session focuses on the unique and complementary efforts underway to identify synergies and potential gaps in needed collaboration.

Moderator: Danilo Tagle, PhD – ACD NAMs Working Group

Presenters:

- Alex Carlisle, PhD – National Alliance Against Disparities in Patient Health *(nonprofit)*
- Elijah Petersen, PhD – U.S. National Institute of Standards and Technology *(government)*
- Yvette Seger, PhD – Federation of American Societies for Experimental Biology *(scientific society)*

12:00 PM BREAK

DISCUSSION OF HIGH PRIORITY NEEDS

The following sessions focus on identifying potential high priority needs for catalyzing NAMs

use and development with human applicability to (1) advance progress into understanding specific biological processes or states or (2) augment the tools and capabilities for biomedical research to complement and/or potentially replace traditional models. Specific goals will include identifying incentives for integrating efforts and barriers to success.

1:00 PM Developing Integrated and Multi-System Models

Moderator: Szczepan Baran, VMD – ACD NAMs Working Group

Discussants:

- Graça Almeida-Porada, MD, PhD – Wake Forest University
- Blanca Rodriguez, PhD – University of Oxford
- Roser Vento-Tormo, PhD – Wellcome Sanger Institute
- Terry Van Vleet, PhD – AbbVie

2:00 PM BREAK

2:15 PM Leveraging Diverse Datasets for Maximally Useful NAMs

Moderator: Gordana Vunjak-Novakovic, PhD – ACD NAMs Working Group

Discussants:

- John Burke, PhD – Applied Biomath
- Anne Gourmelon – Organisation for Economic Co-operation and Development
- Donna Mendrick, PhD – U.S. Food and Drug Administration
- Ivan Rusyn, MD, PhD – Texas A&M University
- James Zou, PhD – Stanford University

3:15 PM BREAK

3:30 PM Equitably Deploying Robust and Reliable NAMs into Practice

Moderator: Antonio Baines, PhD – ACD NAMs Working Group

Discussants:

- Jessie Carder U.S. Department of Agriculture, Animal Welfare Information Center
- Megan LaFollette, PhD – The 3Rs Collaborative
- Michael Moore, PhD – Tulane University; AxoSim
- Manu Platt, PhD – U.S. National Institutes of Health
- Nicholas Tatonetti, PhD – Columbia University

4:30 PM DISCUSSION AND NEXT STEPS

Howard Chang, MD, PhD & Lyric Jorgenson, PhD – ACD NAMs Working Group Co-Chairs

5:00 PM ADJOURN

Welcome

Howard Chang, M.D., Ph.D., began the meeting at 9:00 a.m. ET. The meeting took place with hybrid virtual attendance and was webcast live. Dr. Chang conveyed regrets that the co-chair of the working group, Lyric Jorgenson, Ph.D, was unable to attend the workshop due to illness.

Achieving scientific breakthroughs requires continuous development, validation, and adoption of innovative technology and techniques. Although animal research is necessary for understanding human disease and development of treatments, novel alternative methods (NAMs) hold promise for increasing tools available to achieve NIH's mission to understand biology and advance human health. With this technology rapidly changing, it is important for researchers to consider NAMs use in an equitable and accessible way.

A broad range of tools fall under the category of NAMs, but the workshop narrowed its scope to rapidly changing techniques and methods complementary to animal models: *in chemico*, *in vitro*, and *in silico* methods. NIH values methods like invertebrate animal models and clinical trials; however, addressing the depth these topics require would not be possible within this workshop.

To identify and prioritize areas for support within these methods, Lawrence Tabak, Ph.D., D.D.S., Acting Director of NIH, appointed Dr. Chang and Dr. Jorgenson as cochairs of the Advisory Committee to the Director (ACD) Working Group (WG) on Catalyzing the Development and Use of Novel Alternative Methods to Advance Biomedical Research. In January 2023, Dr. Tabak charged the WG with assessing the landscape of NAMs for biological research, including challenges and opportunities for different methods and different applications.

The WG convened experts from diverse backgrounds, including industry, academia, and sister federal agencies, to identify high-priority areas for NIH to invest in and catalyze the use of NAMs to inform biomedical research. The WG provided progress updates on these meetings to the ACD in June 2022. The WG then published a Request For Information (RFI) on June 12, 2023, to invite public comments. The deadline for comments was extended from August 16, 2023, to September 5, 2023. Public and expert comments, and the WG recommendations for NAMs will be presented in a final report to the ACD on December 14–15, 2023.

Today's workshop featured open presentations and discussion panels to highlight areas for NAMs use. A subsequent closed session would be used to distill ideas presented in the workshop for inclusion in the WG's final report.

Session 1: The Opportunities and Challenges for NAMs in Biomedical Research

Nancy Lane, M.D., Chair

This session highlighted research areas where NAMs have been impactful. Presentations and discussions were meant to identify areas for leveraging these approaches for basic research, uncovering disease mechanisms, and translating knowledge into products or practice.

Nathan Price, Ph.D., Thorne HealthTech and Institute for Systems Biology

Dr. Price's research is focused on precision health and the use of digital twins as an alternative approach to clinical trial design.

Over the past decade or so, researchers have begun to leverage dense data to understand wellness, disease, and disease prevention. One intriguing area of study is preventing disease by predicting the effects of nonacute toxins.

Dr. Price provided the study of Alzheimer's disease (AD) as an example. Regrowing lost synapses is difficult, but preventing AD may be simpler. To this end, *The Lancet* published recommendations around modifiable risk factors of AD in 2020. Dr. Price suggested that studies using digital twins may help add to these recommendations.

Over the last three years, Dr. Price and his team have developed a digital twin model of how the brain maintains health. The ordinary differential equation (ODE)–based model with a Bayesian network overlay leverages data from approximately 1,000 papers in AD literature. This model allows researchers to layer different information to see how AD progresses. The model also represents complex testable hypotheses of how to maintain health. These capabilities could allow researchers to approach clinical trials in a new way.

Dr. Price demonstrated that digital twin simulations could be performed on both a population and individual-patient basis. Individual risk factors from genetics, lifestyle, and physiological measurements can be entered into the model to build a probabilistic future for the age of a dementia diagnosis. The model can also simulate the effects of different interventions (lifestyle changes and therapeutic treatments) to demonstrate their potential effects, both on their own and in combination. The model can produce personalized predictions for the age of diagnosis and recommend interventions, and it can produce simulations of health and cognition for entire populations.

To validate the model, Dr. Price's team compared data from genetic and known hazard ratios taken from literature and clinical trials with simulated data from 10 million digital twins. The team demonstrated through a common complex hypothesis that the model showed good explanatory power on when people were diagnosed with dementia and what factors were related to diagnosis.

Digital twins and NAMs could therefore be used to inform recommendations like those published by *The Lancet*. For example, observational data has demonstrated that low levels of vitamin D are associated with increased risk for AD. However, clinical trials have been inconclusive in determining this relationship, causing vitamin D to be excluded as a modifiable risk factor. Dr. Price has used the digital twin model to simulate clinical trials around vitamin D. These simulations have shown that vitamin D plays a subtle role in how cholesterol is trafficked in the brain. High variability in cholesterol trafficking, combined with other conditions, can then play a role in cell death. Higher levels of vitamin D are therefore related to lower variability in cholesterol trafficking and ultimately a 25% reduction in AD risk. These findings suggest that traditional clinical trials may not be set up to demonstrate these kinds of effects, causing information to be overlooked.

The digital twin model can also be used for personalized multifactor recommendations, in which different interventions are tested for different people within the model. The model can simulate and determine personalized combinations of interventions, which can be more economically viable and test interventions in a different way. Ultimately, digital twins could allow researchers to take a different approach to clinical trials that better inform preventive measures against disease.

In response to a question, Dr. Price clarified that the model can be used to decrease variability within treatment mechanisms, as opposed to reducing variability of a population. In the vitamin D example, the mechanism of the vitamin's effect on neuronal death can be simulated—even though the effect is small—to generate the hazard ratio and demonstrate subtle effects over time.

Arnold Kriegstein, M.D., Ph.D., asked how the model disentangles different combinations of interventions. Dr. Price explained that the model is not additive but that mechanisms are instead all simulated and built into the model. Referring to the vitamin D example, he said the model can simulate the sum of neurons maintaining a positive energy balance and the effect of vitamin D on that balance. The model can simulate conditions in which low vitamin D over decades of life increases the variability of cholesterol managing neuronal energy balance, and the model also simulates conditions that can result in neuronal death and eventual AD.

In response to another question, Dr. Price explained that the model can be validated and iterated with empirical data. The model incorporates clinical trials and molecular data across different types of interventions. The empirical hazard ratios can then be compared with those generated by the model for both prediction and calibration.

Thomas Hartung, M.D., Ph.D., Johns Hopkins University

Dr. Hartung is an international expert on translational product development, especially around microphysiological systems (MPS).

Approximately 50% of Americans and 60% of Europeans object to animal testing, especially cosmetic testing. NAMs provide an opportunity to reduce animal experiments and are already used in toxicology. But wider adoption of alternative methods is needed, especially given the issues in translating results between animals and humans.

For example, traditional toxicology studies should be the best possible use of animal models because these studies are done under high quality standards, with high doses, and require young, healthy animals rather than those modeling a disease state. However, analyses of 2,839 chemicals with approximately 100 repeat studies demonstrated only 81% general reproducibility—and just 69% reproducibility for the actual identification of toxic chemicals. More complex toxicities for the study of cancer or reproductive toxicity show only 60% correspondence between different animal species.

Cell culture studies show similar reproducibility issues. In academic research, 25% of cell lines are misidentified by species, organ, or sex; 15% to 25% are infected with mycoplasma; and genetic instability and culture artifacts can further interfere with results.

However, through stem cells and bioengineering, MPS can keep cells alive to interrogate and demonstrate systems' functionality. This research area has grown rapidly over the past 10 years. Dr. Hartung and his team held workshops with leaders in MPS to discuss the opportunities for technological development for both science and business. A key area of interest was the use of human MPS for drug development. Dr. Hartung, along with Suzie Fitzpatrick, Ph.D., and Don Ingber, M.D., Ph.D., then hosted the first MPS World Summit, which took place in New Orleans from May 30 to June 3, 2022. The summit received strong engagement from FDA and methods support from the National Center for Advancing Translational Sciences (NCATS). The summit had 665 registrants from 52 organizations. The second MPS World Summit took place in Berlin June 26–30, 2023. Dr. Hartung encouraged those who were interested to register for the third summit, to be held in Seattle next June.

As this research area continues to grow, the quality of MPS is crucial for replacing traditional methods. Dr. Hartung and his team will be publishing a manuscript on *in vitro* reporting standards that will include MPS, and the group is also thinking about ways to validate these systems for regulatory use.

Dr. Hartung used his team's research on brain organoids to illustrate MPS applications. His lab uses standardized human 3D systems based on pluripotent stem cells. MPS produce approximately 1,200 cells on well plates, making them cost-effective. Different genetic backgrounds and disease phenotypes can be introduced to allow researchers to study risk genes of interest. MPS can be further modified with reporter genes for developing different cell types. Finally, the organoid includes neurons, glial cells, oligodendrocytes, astrocytes, and microglia to mimic physiology and immunocompetency. MPS can be used to study infection, cancer grafts, toxicity, neurodegeneration, and gene-environment studies. The team has recently used the organoid to study risk genes and exposures related to autism, receiving funding from NIH.

The development of functional endpoints has allowed for meaningful assays using MPS beyond standard toxicity studies, including assays of learning and memory using organoid intelligence (OI). This area of study combines artificial intelligence (AI) and *in vitro* methods, allowing MPS to demonstrate short- and long-term learning effects. Dr. Hartung and his team recently developed electroencephalography signals from brain organoids to show opportunities for new MPS endpoints, an effort that has been met with significant press coverage and funding from the National Science Foundation.

Dr. Hartung acknowledged the controversy currently surrounding AI but noted its popularity in this research area for its power in computation and data production. Dr. Hartung also demonstrated in 2018 that even simple forms of AI could be used for toxicology prediction with better precision than animal studies provide. Through its computational power to catalogue exposomic impacts on human disease etiology, AI also has potential for use in toxicology studies within the Human Exposome Project and for advancing an evidence-based toxicology movement that has developed over the past 20 years.

In summary, the traditional method of moving from hypothesis to causation and proof is complemented by both evidence-based approaches and machine learning (ML), which will allow researchers to arrive at evidence from a different point of view. Successfully integrating

technology will require open access publishing, fair and accessible data, explainable AI, mechanistic validation, and study of the human exposome.

Nicole Kleinstreuer, Ph.D., National Institute of Environmental Health Sciences (NIEHS)

Dr. Kleinstreuer is the Director of the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Executive Director of the Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM). These organizations are federal resources for alternatives to animal testing, providing scientific and operational support to a congressionally mandated committee of 17 federal regulatory and research agencies. These agencies require (or consider) chemical safety data and are interested in moving toward rapid, efficient, human-relevant testing approaches.

NICEATM and ICCVAM develop datasets, models, and NAMs that follow a cyclical process in which iterative, mutually informative approaches use big data, predictive models, experimental design, and mechanistic models. The ultimate goal of this work is to generate human-relevant insights into disease processes that support effective environmental research and protect sensitive and susceptible populations. For the purposes of the workshop, NICEATM and ICCVAM's validation work in the field of toxicology can have an impact in the biomedical research space.

Five years ago, ICCVAM published a strategic roadmap representing the 17 agencies. The roadmap explained that NAMs that are being tailored to meet regulatory needs must:

1) Help end users guide development of new methods.
2) Use efficient and flexible approaches to establish confidence in new methods.
3) Encourage adoption of new methods by federal agencies and regulated industries.

During development of this roadmap, it became clear that there was no one-size-fits-all approach to validation. As a follow-up to the roadmap, ICCVAM has been working on a report called "Validation, Qualification, and Regulatory Acceptance of New Approach Methodologies," which is open for public comment until September 5, 2023. The report presents considerations for the development and implementation of NAMs, including their context of use (e.g., biological relevance, technical characterization of methods, data integrity, and information transparency), and independent review.

Biological and mechanistic relevance is especially crucial for supporting regulatory translation and application. These concepts, while they seem intuitive, reflect a fundamental shift in thought for validation and establishing scientific confidence in NAMs as scientists move away from exclusive reliance on existing historical reference animal data and instead begin to consider methods based on human biology.

Dr. Kleinstreuer presented the study of skin sensitization as an area of research where these validation principles have been applied and where use of NAMs has gained regulatory acceptance. The Organization for Economic Cooperation and Development (OECD) has published guidance on the adverse outcome pathway (AOP) framework, which can help develop testing strategies for the study of how chemical exposures interact on a molecular level to result

in organismal adverse effects. The framework can be used to map initiating events, correlated biological processes, and different test methods (including animal models) to different stages of the AOP. In the case of skin sensitization, individual NAMs have been mapped to the AOP according to OECD guidelines to cover the entirety of this process's biology and mechanism, allowing researchers to use these methods for prediction of adverse skin sensitization outcomes. The decades of establishing confidence in these combined methods have additionally shown that the combination of *in vitro* and *in silico* methods based on human mechanistic understanding have outperformed animal tests when comparing results to human reference data.

NICEATM is continuing to apply principles of the validation report by partnering with federal agencies to optimize protocols that require special instrumentation. The hope is that applying higher throughput to such protocols will increase accessibility and encourage their broader uptake. The results of this validation study are currently being written up and will be submitted to the OECD later in 2023.

NICEATM is additionally working with the U.S. Environmental Protection Agency (EPA) to apply AI and integrate NAM data to derive points of departure for quantitative risk assessment and regulatory decision making. In a recent assessment of isothiazolinones, EPA determined that *in vitro* and *in chemico* studies provided data that were more reliable, reproducible, and human-relevant than existing animal tests. This has been the first use of such information in regulatory risk assessment.

Dr. Kleinstreuer's group has taken this work further by using Bayesian probabilistic modeling to study human-relevant points of departure in variability when studying susceptibility across populations. This model can be used to facilitate decision making on consumer safety by inferring human-relevant metrics of human potency across populations. The team has also partnered with Unilever to expand the database and make data open source. This work is featured in the OECD work plan for inclusion in the defined approach guidelines.

In closing, there was mention that a 2019 report from the European Union stated that the vast majority of animals are used in basic and regulatory research. Dr. Kleinstreuer advised the WG to consider this use when drafting their recommendations. She also suggested that the fields of toxicology and biology could learn from each other to develop methods that answer critical questions about health and environment for effective regulatory decision making. NAMs will be fundamental to this work.

Chirag Patel, Ph.D., Harvard Medical School

Dr. Patel's research uses informatics to integrate data sources from large-scale biobanks. Combining data modalities allows Dr. Patel to dissect interindividual differences and capture comprehensive clinical experience. Studies are observational, nonrandomized, convenient, and single-cohort. Researchers are able to start with any hypothesis. They can also access participant location to study exposures at the individual level.

The development of biobanks has been a welcome addition from candidate gene studies of the past. However, a key challenge in their use is addressing variability of multiple ancestry groups and using these data in a translational capacity for predictive ability. Addressing this challenge

involves integrating data across the genome and exposome. Researchers are working to compile exposures in a single dataset to analyze them systematically and contextualize how these factors play a role in disease. This work requires developing an output of causal factors and biological responses that are connected to clinical outcomes. Dr. Patel presented his research demonstrating how analysis of biobank samples could be used to relate the exposome to biology and disease.

Broadly, Dr. Patel's work models exposomic factors and their relationship to disease, accounting for individual variation in gene-environment interactions. His team has put together a large twin and sibling dataset using biobank samples to examine the shared versus nonshared exposome. These data revealed variation yet to be explained by genetics or shared environment. Correlation studies of different metabolites and nicotine exposure for adverse health outcomes also showed significant variability across a dense correlational globe of attributes, demonstrating a need to study exposome-wide variation.

Data variability and heterogeneity in data modalities (e.g., mass spectrometry, geospatial markers, and self-report questionnaires) can be addressed by developing a polyexposure risk score (PXS) that incorporates all of these measures. This score has been validated in UK Biobank and the Personalized Environment and Genes Study cohort. PXS scores are also complementary to other risk scores such as polygenic risk scores (PGS). PXS may therefore be another tool for assessing clinical risk for disease, especially in undiagnosed populations where PGS has not been deployed.

But as these methods are used, validation methods may also need to be revisited. Dr. Patel noted that the Bradford Hill criteria have largely been used to guide decision making at the federal and policy levels. But studies of the exposome feature individual exposures with variable levels of risk, and 1:1 study of exposures and phenotypes may not be practical. Other areas of Bradford Hill criteria may still be relevant for NAMs and biobanks. For example, biobanks can easily allow for replication of studies. However, other Bradford Hill criteria of temporality, biological gradients, biological plausibility, and coherence are still being examined.

Researchers are also suggesting the need to reassess reproducibility by retesting under different assumptions. For example, researchers have suggested using genetic interindividual variability to simulate a "randomized trial" examining correlations between genotype-exposure relationships and adverse health outcomes.

Dr. Patel closed his talk by suggesting that real-world exposures and correlated health outcomes be used to guide biomedical research, and that biobanks be used to examine those relationships and interindividual variability. The scientific community will need to consider guiding questions such as the relevance of current heuristics, how multiomic readouts are clinically and biologically relevant to disease, and how multiple correlated exposures can be mapped to real life with disease-relevant NAMs.

Discussion

Dr. Lane asked what lessons could be learned regarding tech readiness and choice of hypothesis when developing a system that would be scalable to NAMs for human susceptibility. Dr. Kleinstreuer said researchers need to start from a human biology-based framework to cover key

aspects of cellular, molecular, and tissue-level events that contribute to adverse outcomes. NAMs can also represent human variability and susceptibility across populations in ways that are not scalable or feasible with traditional mammalian models. There is exciting work being done using NAMs, and NIH can step in to provide an influx of resources that support their continued development.

Dr. Hartung said stem cells provide opportunities in studying human diversity.

Dr. Patel reiterated that biobanks are a first step, noting their ability to recall participants for further testing if needed. Dr. Price added that the ability to follow participants opens up opportunities to study long-term effects that have previously been invisible to biomedical researchers.

Dr. Lane asked Dr. Hartung how integrating technology advanced his projects and how his team addressed challenges in this work. Dr. Hartung said bridging expertise among different areas is a significant challenge. He works to recruit students and researchers from different areas and encourages standards in communication among different specializations. Workshops have also been used to share fundamental knowledge and to build common vocabulary and methodologies. Dr. Patel also works to recruit students from multiple research areas. He added that students need experience at the hospital bedside and experience with data early in their training.

Dr. Kleinstreuer said building data ecosystems can be done in partnership with AI to generate testable hypotheses and also iteratively improve AI models. Investing in data infrastructure in a way that produces unstructured and structured datasets will be fundamental to developing better models. Dr. Lane noted untapped data from clinical trials and biobanks that could be added to such infrastructure.

Dr. Lane asked Dr. Price what infrastructure was needed to address challenges in NAMs. Dr. Price said standards can encourage interoperability, but they also need to avoid stifling innovation. In addition, researchers need an infrastructure for representing mechanisms. Modeling a virtual human is not yet possible, so researchers will need to understand the parameters of effective systems modeling. Finally, the public needs to be able to understand the work being done with NAMs. Dr. Price suggested large language models could be effective in generating such communication to a public audience.

Dr. Hartung reiterated the importance of open access publishing. Dr. Kleinstreuer agreed. She added that federated data sharing and model building approaches could be used to address potential issues with proprietary data. Encryption technology and federated model building could be shared externally, even if the data used to train the models remain proprietary. NIH could additionally play an important role in building public-private consortiums.

Session 2: Cross-Sector Approaches for Driving NAMs Use and Development

Danilo Tagle, Ph.D., Chair

This session focused on unique and complementary approaches in different sectors for the purpose of identifying synergies and needs for collaboration. Dr. Chang noted that ACD WG

member Myrtle Davis, D.V.M., Ph.D., would, in place of a formal speaker, provide perspectives from industry.

Alex Carlisle, Ph.D., National Alliance Against Disparities in Patient Health (NADPH)

Dr. Carlisle is the Founder and Chair of NADPH. He presented his research experience and current work to encourage the need for broad efficacy and usage of NAMs in an ethical and equitable fashion.

Dr. Carlisle contributed to the development of first-generation microarrays at the National Cancer Institute and National Institute of Neurological Disorders and Stroke. He has further improved on genomic interrogation platforms by helping to develop early bioinformatic profiles. This work faced computational challenges in storing and analyzing data because of the top-down nature of technology generation: Subject matter experts developed the technology, but high levels of expertise were then required to understand and use it. Technology is then widely shared but still not made to be understandable, resulting in a situation in which academic institutions and the public face these same challenges.

Dr. Carlisle said that scientists developing technology based on patient samples and data have a responsibility to translate their work back to patient communities. Scientists must also ensure that the technology being developed is informed, robust, ethical, and equitable. This requires scientists to work with human-centered design experts who are able to engage with the customer.

NADPH is a research and technology development organization that works to translate technology and ensure its application is serving those who have been marginalized, with the ultimate goal of closing the gap in precision health care. Pharmaceutical and technology companies frequently do not invest in qualitative assessment of the community they are interested in. One example of NADPH's work includes partnership with the Robert Wood Johnson Foundation (RWJF) on the Rising Equitable Community Data Ecosystems (RECoDE) project. NADPH worked to understand the barriers facing technology's development and adoption. In addition, to develop principles for how new technology should be applied, they assessed its potential for good and harm.

Dr. Carlisle is also lead multiple principal investigator on the Infrastructure Core for NIH's Artificial Intelligence/Machine Learning Consortium to Advance Health Equity and Researcher Diversity (AIM-AHEAD) program, an initiative meant to accelerate innovation of computational tools for use in health disparities research by engaging across stakeholder communities. By building capacity for technology and investing in its use within underrepresented communities, AIM-AHEAD hopes to add talent and diversity of thought to the field of technology development while also delivering technology that is relevant and beneficial to those communities. Dr. Carlisle and his colleagues are applying lessons learned from both of these lines of work and collaborating to engage communities and develop principles that will guide technology development.

As NAMs are developed, the entirety of the development life cycle needs to be considered, which means engaging with communities throughout that cycle. Researchers need to be fair,

authentic, and genuine in their willingness to engage with the broader community. They additionally need to be aware of power imbalances when seeking community input.

Dr. Carlisle closed his talk by noting that no group is monolithic, and engagement also needs to be conducted with heterogeneity in mind. From gathering community opinions and identifying themes, Dr. Carlisle and his colleagues have found that the needs of a group may be different from the needs of a researcher, and those needs may also differ between and within communities.

Elijah Petersen, Ph.D., M.S., National Institute of Standards and Technology (NIST)

Dr. Petersen is a research scientist in the Cell Science Group at NIST. His research is on the comprehensive evaluation of *in vitro* assays to improve repeatability and interlaboratory agreement of assay results. He shared a multiagency government effort to gain greater assurance in NAMs and how researchers can maintain high-quality data for use in these new models and methods.

This project brought together researchers from the U.S. Consumer Product Safety Commission (CPSC), NICEATM, the U.S. Department of Defense (DoD), NIST, and Empa, the Swiss Federal Laboratories for Materials Science and Technology. The different agencies in this effort have been valuable for their distinct viewpoints and expertise, ranging from measurement to regulation to the development of new technology and understanding of bench science.

The project applied basic NAMs to conceptual analyses. The group measured these results and built statistical models considering relevant uncertainties within the method to increase confidence in that method's use. While this analysis ensures confidence, it can also increase costs and throughput power of the assay. It is therefore more practical to assess the extent of controls needed for a given assay based on its context of use.

The group developed a framework to aid in this decision making process, based on considerations of the NAM's biological relevance and the testing capabilities of the assay. The framework can guide researchers checking for sources of variability and can also serve as a reference point in cases when researchers need to revisit earlier steps or take a nonlinear approach to assessment.

Dr. Petersen provided examples of this conceptualization through different diagrams. Fishbone diagrams can be used to represent the expected outcome and different results and conditions factoring into the results from different labs using the assay. This visualization of the assay's use allows researchers to identify key sources of variability. Flowcharts can also display the key steps of an assay, along with control measure that address each step. Variability can also be easily displayed through the use of scatter plots and histograms, while also displaying variability within the specific variables being used within the assay. These visualization tools allow researchers to encode quality into each step of a NAM so that it can be easily run across different labs.

In summary, government agencies can provide foundational pieces that organizations can build on to drive innovation. Employing standard steps can build confidence in NAMs use, especially as they grow in complexity.

Yvette Seger, Ph.D., Federation of American Societies for Experimental Biology (FASEB)

Dr. Seger is the Director of the Office of Science Policy, Deputy Director of the Office of Public Affairs, Director of Strategic Scientific Program Advancement, and a member of the Board of Directors at the Center for Open Science at FASEB. Since 1912, FASEB has been a federation of scientific societies that advance health and well-being by promoting research and education through collaboration among its societies and members. The organization is considered the largest coalition and policy voice of biomedical researchers, with 26 member societies and more than 10,000 scientist members.

Dr. Seger shared FASEB's comments submitted to the WG RFI and posted on FASEB's website. Successful advancement of NAMs depends on flexible and iterative approaches for decision making, including establishing uniform guidelines and standards. Potential barriers include lack of funding mechanisms and lack of mechanisms to foster collaboration between animal researchers and NAM developers. Dr. Seger noted that university researchers will engage in research in which they have funding, which can come from NIH or other partnerships. Dr. Seger suggested that addressing this barrier could be key to greater interest in NAMs.

Dr. Seger reiterated some best practices for NAMs, including raising community awareness. Although there is high awareness of NAMs in the field of toxicology, biomedical communities still have gaps in their understanding and knowledge of NAMs. Building awareness of NAMs and their requirements for rigor and reproducibility is especially critical in this community as it looks to transition from animal models to these methods.

Dr. Seger suggested that scientific societies can help by organizing scientists through their stakeholder meetings. These include individual society convenings, public forums, and closed-door FASEB meetings. In addition, many scientists in FASEB could share their perspectives on how to advance science using NAMs.

Scientific societies can also amplify communications on resources, opportunities, and collaborative partnerships; disseminate guidelines; distill information; and raise awareness of notices, including the WG RFI. Dr. Seger noted that scientists often do not consider something pertinent until they hear from their peers. FASEB can expand engagement of NAMs among researchers from many different career levels and among the general public.

Discussion

Dr. Davis said that she would provide perspective on how her colleagues in the pharmaceutical industry work to provide safe, effective medications to patients and how this work interfaces with the WG's goals for NAMs. She noted that scientific curiosity and the expectation that animals would be used only when necessary were common across the different sectors at the workshop, and she reiterated the group's interest in developing confidence in these methods.

Dr. Tagle asked the panel to reemphasize what each of their sectors have in terms of strengths in advancing NAMs and what cross-sector collaborations are needed for the greatest impact. Dr. Carlisle said NADPH had strong qualitative and quantitative research skills. The organization is able to engage with communities with trust and sincerity so that they understand how technology can affect them, while also developing structures for data analytics and AI models. He reemphasized the need for scientists to be aware that certain communities distrust scientific research and the need for policy to ensure that ethics and equity are embedded in the practice of data collection and use. Scientists must also think about commoditization and incentives for participating in research. Finally, Dr. Carlisle reemphasized the need for improved metrics and standardizations as organizations are brought together.

Dr. Petersen reiterated the government's ability to convene different sectors and the government's role in improving infrastructure for evaluation outside of any technological competition. Dr. Seger said that scientific societies could help bring scientists together for the purpose of standardization. She suggested that having scientific communities involved in these processes could encourage broader adoption of and enthusiasm for building and validating NAMs.

Dr. Davis said that the pharmaceutical industry understands the importance of establishing basic disease mechanisms, which the WG must consider alongside technology investment. Pharmaceutical rigor in disease mechanisms combined with its fast pace and knowledge of the patient experience all feed into the industry's goal of producing medications that improve patient health with minimal side effects.

Dr. Tagle asked whether there were barriers in addition to a lack of funding and whether there were additional incentives that could encourage collaboration. Dr. Seger reiterated that funding can be a limiting factor but added that scientists will want to have the right partnerships and know how to properly collaborate with different sectors. Additional resources will therefore be needed to ensure interested parties understand how to think of science holistically in terms of project management. Dr. Carlisle also recommended having mechanisms that exist outside of competition, to allow for collaboration without losing out on important revenue or grant sources.

Dr. Petersen said government regulatory agencies are trying to remove barriers to NAMs' validation by updating guidance, some of which is 25 years old. In terms of funding, Dr. Petersen felt there was more money available for innovation than understanding sources of variability.

Dr. Tagle asked Dr. Petersen whether there were best practices he could recommend for collaborating across federal agencies, especially in developing NAMs for human relevance. Dr. Petersen said that in addition to fostering dialogue among different sectors to understand the needs of different agencies, transparency around what is needed for regulators has been very important.

Dr. Tagle asked Dr. Seger whether there were distinct ways she approached different societies within FASEB and how scientists are trained. Dr. Seger said that she first determines the needs of a society and whether there are ways they can be addressed. She also recommended that the scientific community approach collaboration with humility and openness about where they do

and do not have expertise, while aiming to communicate without jargon. In regard to training, conferences and society-driven funding or workshops can be important resources for developing skills.

Dr. Carlisle recommended developing relationships with community colleges and providing internships to under-resourced communities. These pathways could be useful for NAMs validation. On internships, Dr. Davis added that Bristol Myers Squibb (BMS) hosts summer interns who learn to develop new assays and rigorously test them. BMS has also partnered with the Society of Toxicology to fund a graduate student award. The award supports the student, who then teaches BMS about their academic research. Dr. Petersen said internships are also available at NIST.

Session 3: Developing Integrated and Multisystem Models

Szczepan Baran, D.V.M., Chair

Dr. Baran had the panelists introduce themselves before moderating a discussion on integrating technology and NAMs for impact on drug development and scientific discovery.

Graça Almeida-Porada, M.D., Ph.D., Wake Forest University

Dr. Almeida-Porada is a professor of Regenerative Medicine at Wake Forest University. There is a high-priority need to develop multisystem approaches in human biology. Dr. Almeida-Porada's research develops cell and gene therapy platforms to treat patients with genetic disorders. Her lab uses animal, *in silico*, and *in vitro* models, and collaborates with experts to simulate mutations in sickle cell patients and create an animal model of sickle cell disease. These systems are used with different organoids and serotypes to address aspects of gene therapy and predict toxicity. Dr. Almeida-Porada said that she hoped this workshop could establish common ground to develop technology that will be priceless in finding better solutions for patients and finding better answers than those that have been found using animal models.

Blanca Rodriguez, Ph.D., M.Sc., University of Oxford

Dr. Rodriguez is a professor of Computational Medicine at the University of Oxford, specializing in the development of computational methods for understanding the heart. The goal of her research is to develop paradigms in clinical trials and digital twins for novel therapy development. Dr. Rodriguez is also developing human-based *in silico* technology to simulate clinical trials. Her team works with clinicians, scientists, industry, and regulatory agencies to build virtual organs and validate simulation results to ultimately replace animal models.

Roser Vento-Tormo, Ph.D., Wellcome Sanger Institute

Dr. Vento-Tormo is a group leader at the Wellcome Sanger Institute. Her team uses and develops technology in the fields of genomics, bioinformatics, and bioengineering in the interest of regenerating function in mucosal tissues. The team has used NAMs since its start, especially when animal models are not applicable to tissues of interest. One area where the team has found NAMs especially useful has been in the study of the uterus. The team has taken measurements of

individual cells across the menstrual cycle. These measurements and bioinformatics are then used to integrate data and build an *in vitro* model based on the cells. Multimodal information has also been used to build bioinformatic tools for the study of cell-cell communication. In creating these models, the team has been able to make reference maps of tissue and improve other *in vitro* models, such as organoids. The team has also compared *in vivo* and *in vitro* models to show that new methods can perform similarly in recapitulating disease.

Terry Van Vleet, Ph.D., AbbVie

Dr. Van Vleet serves as the head of Investigative Toxicology at AbbVie. The company works with discovery to build fit-for-purpose screening profiles for safety and *in vitro* assays. AbbVie is also working to bring in technology to allow for complex *in vitro* models that characterize disease. This work has given Dr. Van Vleet experience in building and maintaining models and determining their strengths and weaknesses. Dr. Van Vleet has also served as Chair for the IQ MPS Affiliate to help develop broader industry perspective on how pharmaceutical industries are applying NAMs in safety assessment. In this role, he has developed a toxicology group that assesses data structure in the application of advanced analytics.

Discussion

Dr. Baran asked what the challenges were in developing mixed models and what incentives were needed to overcome them. Dr. Van Vleet noted there were different types of data, with inconsistent methods of maintaining the data in the proper format. Maintaining data in proper formats and proper databases will be an important area for future investment. These issues are prevalent in preclinical data, which can feature different terminology depending on the sector the data come from. These inconsistencies can bring challenges to translating results to the clinic. However, through opportunities to share data anonymously and greater investment in preclinical and clinical data, the challenges could be met. Federated databases could also provide integrative models for proper data maintenance.

Dr. Rodriguez said the challenges included management's willingness to invest in the uptake of new technology. Management will need a clear understanding of the technology's benefits and trust in a new system to support and fund it. She said that she shared Dr. Van Vleet's concern about having data of high quality and quantity to construct effective new methods. Teams developing new methods will also need the right expertise and training to work together with new approaches, which management will need to understand. Incentives will need to be clear, and they will need to come from government or regulatory agencies to fight skepticism and unwillingness to change.

Dr. Vento-Tormo added that instead of a focus on funding for things that are innovative and new, there needs to be an increase in funding for maintaining databases. As in previous discussions, she suggested more opportunities for validating models and standardizing measurements.

Dr. Baran asked what the limiting factors were for making hybrid NAM models more useful for studying human conditions and deploying them for human advantage. He also asked how to address these challenges and, where possible, examples of ways to do so. Dr. Almeida-Porada said that innovation and qualification are both pathways forward that need equal consideration.

She reiterated that there are limiting factors coming from the multisystems themselves, in that they require experts from different fields to come together and communicate across their areas of study.

Dr. Almeida-Porada also noted that MPS technology is advancing with the use of ML and bioinformatics. However, scientists will need to ensure that these systems are recapitulating proper tissue responses and retaining characteristics of their donors that are important for studying interindividual differences. MPS models with multiple organs will also need to establish whether a single donor will be used and how to address differences among donors within the system. Having an agency that maintained these tissues could help address some of these questions. Wake Forest University has developed a universal medium to allow for multiple organoids, but this is not a perfect solution for toxicological studies. Researchers will also need to consider the benefits of ML and AI for analysis of multiple organoids to establish a big picture and make full use of those data.

Dr. Baran asked Dr. Almeida-Porada whether she has standardized the kinds of metadata being collected. Dr. Almeida-Porada said this is an ongoing, challenging effort.

On cell sources, Dr. Van Vleet agreed that there are often issues with cells being the wrong type or from the wrong species. He agreed that a verified bank of cells was an area for investment and also that universal media and size could be potential issues in multiple organoid systems. Dr. Van Vleet said he felt that the biggest barrier to multisystems was having the confidence to make decisions on the data. This issue could be addressed by having validated data sources and tools, plus corroborating evidence of NAMs' efficacy.

Building on Dr. Almeida-Porada and Dr. Van Vleet's suggestions, Dr. Rodriguez noted that a credible system will still face issues with uptake because systems developed in academia are frequently not transferred to industrial settings. Adjusting funding cycles to support the time needed for validating systems and moving them forward could help address this issue. Funding could also help support agencies that are not well staffed.

Dr. Vento-Tormo added that validation of the system needs to be discussed as it moves from academia to other sectors. Models also need to account for immune responses and complex aspects of biology to recapitulate long-term responses.

Dr. Baran asked Dr. Van Vleet whether his team engages with MPS experts. Dr. Van Vleet said the team was formed to try to leverage historic data. The team is currently reformatting these data and establishing proper analytics but does not yet have enough MPS data to work with. He noted that multisystems are already being used at AbbVie, but greater confidence will need to be established before these systems are ready to replace animal models. Complex systems can last for long periods of time, but because of the length of experiments, it also takes more time to get data.

Dr. Baran asked Dr. Almeida-Porada whether *in silico* tools also need to be qualified. She replied that these tools can be generalized more easily. Her team reaches out to collaborators, and they work together on specific projects. She aims to work with people from different backgrounds so that they can provide information her team does not have. Collaborative teams need to be

prepared for a long investment that is ultimately worthwhile. Dr. Rodriguez commented that *in silico* methods do need to be qualified for specific contexts of use but can face issues because of the speed at which *in silico* methods are updated, and qualification of an updated *in silico* model can often take longer than the actual update.

Dr. Baran asked what challenges researchers will face integrating NAMs with traditional methodologies and how those challenges could be addressed. Dr. Vento-Tormo said that harmonizing measurements and metadata will present a large challenge, especially when working with *in silico* models. She said that she also expected challenges in translating methods between mice and humans, and increasing complexity in recapitulating long-term responses.

Dr. Rodriguez said anticipating implementation at the outset of a project and engaging with those using the new methods could address the challenges of complex methodology. Dr. Almeida-Porada agreed, adding that panels evaluating academic work should be included in this engagement. These changes, along with expanded funding cycles, could encourage integration.

Dr. Van Vleet suggested that having animal cell-based models could help in making comparisons and translation. They could also be used as preliminary assays to assess safety.

Dr. Baran asked the panel what they would do for the state of NAMs use if they had unlimited power and funding. Dr. Vento-Tormo called for greater complexity of NAMs. Dr. Rodriguez wanted longer funding cycles. Dr. Almeida-Porada suggested gathering researchers from different sectors to improve NAMs and move the field forward, regardless of whether a team wished to find a clinical response. Finally, Dr. Van Vleet called for large blinded anonymous databases that could increase the power of new analytics.

Session 4: Leveraging Diverse Datasets for Maximally Useful NAMs

Gordana Vunjak-Novakovic, Ph.D., M.S., Chair

Dr. Vunjak-Novakovic had the panelists introduce themselves before discussing the topic of leveraging diverse datasets for maximally useful NAMs.

John Burke, Ph.D., Applied BioMath

Dr. Burke is the cofounder of BioMath and brings two decades of experience in mathematical modeling, life science, biology, and pharmaceutical and biotechnology research to NAMs use and integrating disparate types of data. His work in systems modeling considers the interface and parameters of a model to perform biochemical equations with minimal unknown parameters. The model can then become a repository for data and hypotheses. The model can also incorporate both *in vitro* assays and clinical data to ultimately inform therapies in a more cost-effective way.

The ability to do millions of simulations across different considerations makes it fundamental to consider the data available and whether the size of the model is suitable for the question at hand. Once those elements are established, the model allows for both forward and backward optimization to find clinical uncertainty points and can help establish therapies for which the scale is still unknown.

These features make the potential for NAMs exciting. Researchers can input biology, systems, and timepoints with greater complexity, which may add time to experiments but also allows for the identification of many more therapeutic possibilities. Dr. Burke ended his introduction with a caution to researchers to consider the context of data generation and how a model will be applied.

Anne Gourmelon, Ph.D., M.S., Organization for Economic Cooperation and Development (OECD)

Dr. Gourmelon's work at OECD is to set standards for methodologies and assess chemicals including pesticides, industrial materials, and cosmetics. This work comes with the challenges of having limited data to build standards and reach regulatory acceptance of evaluation methods. Use of NAMs in chemical and safety testing is likely to face similar challenges because of a lack of human-relevant data. Dr. Gourmelon suggests building consensus databases that are structured and can be augmented to generate alternative data. These types of databases are typically used for proprietary data, but making public consensus databases will allow for methodological developers to have equal footing while also increasing competition to produce the best methodologies alongside new technological developments. Databases must also include positive and negative results so that models can predict a wide range of responses.

Dr. Gourmelon said that researchers must also develop a greater understanding of the diversity of mathematical approaches and guide regulators to help them understand new approaches as well. At an OECD meeting in July 2023, member countries reported that they lacked the capacity at the national level to approach methods such as AI with confidence. Dr. Gourmelon suggested meeting these needs and working to leverage diverse databases.

Donna Mendrick, Ph.D., U.S. Food and Drug Administration (FDA)

Dr. Mendrick noted that she was not speaking on behalf of FDA but said that the agency's Alternative Methods Working Group has also worked to bring researchers together from the modeling, commercial, and academic sectors. She encouraged the ACD WG members to visit the FDA WG website.

Dr. Mendrick said there are pros and cons to any method, *in vivo*, *in vitro*, or *in silico*, and it is important to use the best models in any research. Researchers should seek out multiple donors, despite the difficulty, instead of only having one or two. *In silico* approaches also need high-quality data and models that are well trained without being overfit and overpromising.

Building confidence in a new method is also important for bringing in new data and approaches. People, including scientists, are hardwired to not accept change, so winning hearts and minds requires confidence. Building animal systems to bridge the gap between human MPS to human clinical studies could help in this endeavor. At the same time, overpromising must be avoided. While obtaining funding is critical, Dr. Mendrick said, she believes that overpromising leads to the disenfranchising of people and ultimately holds the field back.

Ivan Rusyn, M.D., Ph.D., Texas A&M University

Dr. Rusyn spoke of his approaches to building confidence in NAMs. His lab performs NCATS-supported experiments testing NAMs in many contexts, especially MPS with tissue chips. As part of a government consortium, his lab has also worked to define what methods should be tested each year and how data from those methods should be collected and shared. While transparency of data is not a new problem, it has been an important component of this work.

For NAMs to grow, there also need to be more data. Dr. Rusyn said that he was unsure whether complex NAMs were currently significant data sources, but he noted that NCATS has funded the development of an MPS database that has grown over time. Databases must continue to grow, and researchers should work with developers to develop models around NAMs and bring them into the context of large-scale modeling.

James Zou, Ph.D., Stanford University

Dr. Zou said there were exciting opportunities for *in silico* approaches to generate data that could then be a resource for NAMs. He provided three examples:

1) Language models have been trained to predict protein structure, allowing for researchers to introduce coding mutations and study how they affect the model's prediction on protein structure *in silico*. Dr. Zou noted that this model needs to be validated but could help researchers find mutations in the genome.
2) Generative AI has also been used to generate synthetic data where natural data are hard to collect. One area of this work includes using AI to computationally generate transcriptomics from widely available histology images. These synthetic data can recapitulate what you would get from histology slides and are richer than what is normally obtained.
3) Synthetic data can be used for clinical trials, including combining AI-generated data with electronic health records (EHRs). Instead of running an expensive clinical trial, researchers can use real-world data to develop and compare cohorts with AI developing synthetic control arms to augment them.

Dr. Zou also noted that alongside these examples are opportunities for generative AI to organize data, which could help researchers interface with NAM resources.

Discussion

Dr. Vunjak-Novakovic asked what data, models, and methods, both experimental and analytical, would most effectively drive NAM research, technological development, and translation. Dr. Mendrick said academic labs and small companies both need funding to understand and adopt state-of-the-art techniques, and the conditions best suited to alternative approaches. Researchers also need to have regulatory involvement in the early stages of investigation. Dr. Vunjak-Novakovic noted these recommendations could be tied to earlier remarks.

Dr. Vunjak-Novakovic asked Dr. Zou to expand on analytical approaches and how researchers should approach AI methods. Dr. Zou said the biggest challenges to AI-based methodologies were experimental validation and follow-up. While AI can produce a wide variety of predictions, researchers need to establish when those predictions are more or less reliable. Iterative loops

need to be developed so that AI can make predictions and design experiments that broaden their reliability, which will contribute to wider adoption of AI *in silico* methods.

Dr. Vunjak-Novakovic asked how researchers can best maintain confidence in data. Dr. Rusyn recommended that the ACD WG look at the National Academies of Science, Engineering, and Medicine's report providing solutions on how NAMs could be used. The report's recommendations include bringing developers and users of technology together to establish systems' parameters and encourage transparency about what is being modeled. As Dr. Mendrick recommended, alignment must also avoid overpromising on what can be delivered. Dr. Gourmelon agreed, adding that transparency is critical and should be available to users either upon request or up front. Users should be able to access technology so they can adopt it and understand how it can be useful in a specific context.

Dr. Gourmelon also suggested that regulators support peer reviewers, so that their conclusions can be trusted. In this way, scientific societies may be useful by developing training and educational materials on how technology is maintained and data are generated.

Dr. Vunjak-Novakovic asked what approach can be recommended for managing, sharing, and integrating NAMs data. Dr. Burke said transparency is needed at every level, including at the higher-level scale of why a model is set up in a certain way. Scientists without modeling expertise need to be able to look at a model and understand its design. Scientists also need to understand how a model can be integrated with other methods and know when the model is or is not working to guide future experiments. Dr. Burke also recommended that NAMs be used iteratively to prioritize experiments.

Large public databases could help interpret information that can be used with NAMs, but data are often distributed in the form of spreadsheets or supplemental information. Dr. Rusyn encouraged greater and more thoughtful efforts by researchers contributing to large databases and contributing data that have been curated for broad use.

Dr. Vunjak-Novakovic asked the panelists to give a final statement on what they want to see in the next year to advance the use of NAMs. Dr. Mendrick encouraged more meetings like today's workshop, where both the good and bad are discussed. Dr. Gourmelon wanted to see greater consensus with highly curated databases that are publicly available. Dr. Zou suggested more resources and efforts in creating targeted validation of AI. Dr. Rusyn said NAM developers will need to think carefully about existing problems and developing NAMs that address them. Finally, Dr. Burke wished to see NAMs for relevant animal models, to allow for backward and forward translation, and a better understanding of how new methods will translate to humans.

Session 5: Equitably Deploying Robust and Reliable NAMs into Practice

Antonio Baines, Ph.D., Chair

Before moderating the discussion on equitably deploying NAMs, Dr. Baines had the panelists present their opening remarks.

Jessie Carder, M.S., U.S. Department of Agriculture (USDA) and Animal Welfare Information Center (AWIC)

Ms. Carder's work at AWIC promotes and provides guidance on the use and care of research animals. She also helps researchers find information on NAMs in academic literature. She can share insights into the NAMs application process and how they can be more accessible and affordable to universities and institutions outside of the United States.

Megan LaFollette, Ph.D., M.S., The North American 3Rs Collaborative

Dr. LaFollette is the Executive Director of The North American 3Rs Collaborative, a nonprofit that seeks to advance science for both people and animals. NAMs are critical to this goal. Dr. LaFollette and this organization have led an initiative made up of commercial providers with expertise in developing systems and have also started an AI branch with the goal of increasing NAMs' implementation.

Dr. LaFollette is also interested in spanning industries through science-based communication and dissemination. She said that change can be difficult and that it was important to avoid alienating animal researchers by putting down animal research to elevate NAMs.

Michael Moore, Ph.D., Tulane University and AxoSim

Dr. Moore is a professor and Chair of Biomedical Engineering at Tulane University, and the CEO of AxoSim. The company began through peripheral nerve MPS technology that Dr. Moore developed in his lab and is now used in a variety of commercial settings. As a panelist, Dr. Moore would be sharing his perspectives on the different priorities between academia and commercial sectors, and the value of balancing such tensions.

Manu Platt, Ph.D., NIH

Dr. Platt earned his undergraduate degree at Morehouse College, a historically Black college in Georgia, which has encouraged his focus on equity and access to research. Over the course of his career, Dr. Platt has researched racial health disparities, including sickle cell disease and HIV-mediated cardiovascular diseases. He is now the director of the NIH-wide Center for Biomedical Engineering Technology Acceleration (BETA Center). Engineering has the power to make biomedical work simpler, faster, and cheaper, but this technology is not always accessible. Researchers will want to think about how to implement technology in rural communities and how to make science faster, simpler, and more accessible for all.

Nicholas Tatonetti, Ph.D., M.S., Cedars-Sinai

Dr. Tatonetti recently moved from Columbia University to Cedars-Sinai. He uses real-world data for drug toxicity studies. Dr. Tatonetti uses data points collected in hospitals to address unexpected adverse drug reactions and studies how such reactions are distributed unequally across populations. This work can involve reverse translational medicine by using patient data to generate hypotheses for training algorithms, which can then be modeled *in vitro* or in organoids.

These methods allow researchers to discover the effect of a drug in a relevant population. Confidence can then be built by performing a simple experiment that proves causality.

Discussion

Dr. Baines asked the panel whether NAMs really can replace animal research or whether they can supplement these approaches. He also asked panelists how researchers can be comfortable with both methods. Dr. LaFollette said her organization focuses messaging on performing the best science and using the best model for a particular question. She noted that animals will likely be needed at some stage of drug development, but NAMs could be used earlier in the pipeline to identify compounds or determine which model is best to use. There are many areas where NAMs can be complementary to animal models, which could be used in messaging to speed up progress in NAMs' use.

Dr. Platt agreed that animal models have a place, but their use depends on the questions being asked. He suggested researchers focus on questions as opposed to what technology to use. At the same time, he expressed concern over what happens to people's research if they miss a new wave of technology. He recommended taking a measured approach in phasing methods out so that new methods are applied equitably.

Dr. Baines referred to earlier discussion of having different expertise involved in NAMs and asked whether there were any gaps in subject matter representation. Ms. Carder said she felt that developers, engineers, and industry could all become too focused on developing new technology. As a result, they do not think of the intended use of a product and potentially overpromise its benefits. Dr. Moore noted that biologists show a similar lack of engagement. This issue could result in part from biologists' feeling they are being asked to give up models that have been the basis of their careers, but it is also the result of engineers' unwillingness to design technology with biologists in mind. Dr. Moore suggested funding agencies could be creative in incentivizing collaboration, citing the National Institute of Biomedical Imaging and Bioengineering (NIBIB) collaborations among mathematicians, engineers, and biologists. Dr. Tatonetti added that similar issues are seen when translating AI models. When systems are not easy to use, other sectors pass them over.

Ms. Carder said the AWIC is encouraging larger funding agencies to train researchers in NAMs use to increase confidence. Dr. Baines asked how training could affect undergraduates, graduate students, and postdoctoral researchers, and how NAMs themselves could be used in training the next generation of scientists. Dr. Platt suggested training programs that integrate subjects. For example, Georgia Tech integrated engineers into their biomedical programs and brought sociologists in to observe the classes. The right language needs to be implemented across disciplines to incorporate NAMs' use. Dr. Moore added that inclusive language is necessary from the very beginning: When he taught a course geared toward both neuroscience and engineering graduates, at first no neuroscience students attended it, because the course description featured only engineering concepts.

Dr. LaFollette additionally suggested an online program that gives high-level exposure to NAMs in a way that everyone can understand. Ms. Carder noted that earlier discussion raised the issue

of there being no centralized training for NAMs, an omission that could increase the risk that young scientists would become siloed within their specialties.

Dr. Baines asked for examples of NAMs that have been validated, similar to Dr. Kleinstreuer's example on skin sensitization. Dr. Moore said liver and kidney cells from multiple donors have showed toxicities similar to what would be expected. However, the nervous system still has a long way to go toward validation. Finding high-quality cell lines interferes with validation and ties back to previous issues with validation work's being passed over for funding. Dr. Tatonetti said AI systems face issues with validation but have significant potential in modeling complex systems. With greater validation, they could have utility in clinical trials and replacing animal models. Dr. LaFollette raised the possibility that high levels of validation may not be necessary in every scenario. While validation is important when testing for safety, NAMs could replace animal models in other areas of decision making.

Dr. Baines asked whether 3D cell cultures could be a gold standard to animal models. Dr. Platt said there were MPS that were great models for disease for a specific problem. They could be used to identify signaling pathways for future therapeutic targeting, but Dr. Platt was unsure whether those systems were being used to make clinical decisions. He said he felt there was still much to learn from animal models.

Dr. Baines noted that NIH has been trying to make both clinical trials and cell culture studies more diverse. He asked how researchers could avoid issues with diversity in NAMs. Dr. Platt has found that patients are willing to participate in research, but clinicians are unwilling to approach them. He suggested that clinical training incorporate conversing with all patients and explaining risk in a straightforward way. Dr. LaFollette said cell culture studies could benefit from having evidence of cell sourcing. Scientists can also push back to encourage purposeful variability in their studies, and such principles can be incorporated into training across disciplines. Dr. Tatonetti agreed, encouraging scientists to measure inequities wherever possible. He noted that models are beginning to incorporate validations of equity. Dr. Moore suggested AI and digital twins could contribute to representing the spectrum of humanity. Tests could be performed in parallel with cell lines to augment what is being done in simulation.

Dr. Baines asked how the panelists would engage the public about NAMs and what opportunities they saw for education. Ms. Carder encouraged public workshops and symposiums to be transparent and gather public opinion on NAMs. Dr. Platt agreed, adding that public demonstrations of NAMs could also be useful. Conversations with the public may need to be tailored to specific audiences as well, because different people have different stances on animal research. Dr. Tatonetti suggested that education could also come from the public through advisory boards. Equity could be ensured by having advisory boards of specific demographic groups. Dr. LaFollette encouraged incorporating science communication into scientific training. Scientists also need to be willing to rely on the expertise of others to translate what they do not understand.

Dr. Baines asked the panelists to give their final takeaways of the workshop's first day.

Ms. Carder said she felt that there was great work being done in the development of NAMs but that more discussions were needed to implement suggestions at the ground level. As an example, the workshop brought up trainings, exposure, and education, but it was unclear how those trainings would be put together. She added, however, that today's discussions were moving the needle in the right direction.

Dr. LaFollette reiterated the importance of collaboration. In a response to the WG RFI, she noted the importance of investment, validation, and incentives. Investment in relationship-building and changing scientific minds will be important in advancing NAMs. There should also be greater efforts in connecting with those doing animal research, because those researchers are the target audience for greater NAMs use.

Dr. Moore said the WG needs to be mindful of tensions among different sectors. They also need to think of funding arms at different stages of research, separating innovation and validation.

Dr. Tatonetti encouraged the group to think of the enormous amounts of data generated in the practice of medicine and how they can be used to enhance and create new models.

Dr. Platt suggested implementing training programs at all levels as a way to remove barriers and encourage equitable implementation of NAMs.

Adjournment

Dr. Chang provided preliminary thoughts on the public workshop.

- In Session 1, the speakers emphasized a need for interdisciplinary teams with appropriate knowledge, breaking down silos, creating standard language, and having infrastructure for collaboration. They shared challenges with reproducibility and the potential for NAMs to address this issue.
- In Session 2, the roles of different sectors were highlighted. This discussion had the themes of stakeholder groups thinking about NAMs development and use, science and industry partnerships, regulatory decision making, and consideration of the NAMs lifeycle . There should be engagement with the community early on concerning how technology should be used. This work requires outreach and language to bring together people from siloed areas.
- Session 3 focused on integrated, multisystem models. Speakers brought up the need to build confidence in these models. They need validation, and user needs should be considered from the beginning of development. Models also need long-term investment and interdisciplinary teams in place from the beginning. The speakers felt that NAMs are not able to replace animal models at this point due to various limitations.
- In Session 4, the WG heard from speakers about what they need to address and maximize the use of NAMs. Guiding principles included using different models to address different questions. There also needs to be greater understanding of models, including federated data and transparency about what those models are recapitulating. Researchers and developers need to avoid overselling, which could set the scientific community back.

- Lastly, in Session 5, the speakers discussed how to equitably deploy NAMs. Equity needs to be considered early in the process, and NAMs need to be accessible by all researchers, including those in lower-resource areas. The speakers also shared ideas for workforce development, including the idea of having multiple entry points for scientists with different backgrounds. There is also the potential for NAMs to be introduced earlier in the drug development process. Finally, speakers discussed the potential for virtuous loops in different models, where data from patients trains large-scale models for *in vivo* and *in vitro* NAMs.

Dr. Chang said the WG would consider all ideas and perspectives raised today to inform their report and recommendations on how NIH can advance NAMs most effectively. He reminded attendees of the RFI deadline of September 5, 2023.

The meeting adjourned at 4:54 p.m.

Abbreviations and Acronyms

ACD	Advisory Committee to the Director
AD	Alzheimer's disease
AI	artificial intelligence
AIM-AHEAD	Artificial Intelligence/Machine Learning Consortium to Advance Health Equity and Researcher Diversity
AOP	adverse outcome pathway
AWIC	Animal Welfare Information Center
BETA Center	Center for Biomedical Engineering Technology Acceleration
BMS	Bristol Myers Squibb
BRAIN Initiative	Brain Research Through Advancing Innovative Neurotechnologies®
CPSC	U.S. Consumer Product Safety Commission
DARPA	Defense Advanced Research Projects Agency
DoD	U.S. Department of Defense

EHR	electronic health record
EPA	U.S. Environmental Protection Agency
FASEB	Federation of American Societies for Experimental Biology
FDA	U.S. Food and Drug Administration
ICCVAM	Interagency Coordinating Committee for the Validation of Alternative Methods
ISSCR	International Society for Stem Cell Research
MPS	microphysiological systems
NADPH	National Alliance Against Disparities in Patient Health
NCATS	National Center for Advancing Translational Sciences
NGO	nongovernmental organization
NIBIB	National Institute of Biomedical Imaging and Bioengineering
NICEATM	NTP Interagency Center for the Evaluation of Alternative Toxicological Methods
NIEHS	National Institute of Environmental Health Sciences
NIST	National Institute of Standards and Technology
NTP	National Toxicology Program
NAM	Novel Alternative Methods
ODE	ordinary differential equation
OECD	Organization for Economic Cooperation and Development

OI	organoid intelligence
PGS	polygenic risk score
PXS	polyexposure risk score
RADx	Rapid Acceleration of Diagnostics
RECoDE	Rising Equitable Community Data Ecosystems
RFI	Request for Information
RWJF	Robert Wood Johnson Foundation
SBIR	Small Business Innovation Research
STTR	Small Business Technology Transfer
USDA	U.S. Department of Agriculture
WG	working group