U.S. Department of Health and Human Services (HHS)
National Institutes of Health (NIH)
Office of the Director (OD)
Division of Program Coordination, Planning, and Strategic Initiatives (DPCPSI)

Council of Councils Meeting
May 19–20, 2022

Meeting Minutes

Day 1

I. CALL TO ORDER AND INTRODUCTIONS

James M. Anderson, M.D., Ph.D., Director, DPCPSI, welcomed participants, NIH staff members, and members of the public to the meeting of the Council of Councils. Dr. Anderson introduced the new Council of Councils Executive Secretary, Robin I. Kawazoe, Deputy Director, DPCPSI. He announced that Douglas Sheeley, Sc.D., is the new Deputy Director of the Office of Strategic Coordination (OSC). The virtual meeting began at 10:30 a.m. on Thursday, May 19, 2022. Dr. Anderson noted that Andrew P. Feinberg, M.D., M.P.H.; Anna Maria Siega-Riz, Ph.D., M.S.; and Russell N. Van Gelder, M.D., Ph.D., were unable to attend. The meeting attendees are identified below. Dr. Anderson then reviewed the day's agenda.

A. Attendance

1. Council Members

Council Members Present

Chair: James M. Anderson, M.D., Ph.D., Director, DPCPSI
Executive Secretary: Robin I. Kawazoe, Deputy Director, DPCPSI
Maria L. Acebal, J.D., The Aspen Institute, Washington, DC
Maria Rosario G. Araneta, Ph.D., M.P.H., University of California, San Diego, La Jolla, CA
Kristin Ardlie, Ph.D., Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA
Linda Chang, M.D., FAAN, FANA, FISMRM, University of Maryland School of Medicine, Baltimore, MD
Graham A. Colditz, M.D., Dr.P.H., M.P.H., Washington University School of Medicine in St. Louis, St. Louis, MO
Rick Horwitz, Ph.D., Allen Institute for Cell Science, Seattle, WA
Patricia D. Hurn, Ph.D., R.N., University of Michigan, Ann Arbor, MI
Kevin B. Johnson, M.D., M.S., FAAP, FACMI, FIAHSI, FAMIA, Annenberg School for Communication, University of Pennsylvania, Applied Informatics, University of Pennsylvania Health System, Philadelphia, PA
R. Paul Johnson, M.D., Emory University School of Medicine, Atlanta, GA
Karen C. Johnston, M.D., M.Sc., University of Virginia, Charlottesville, VA
Paul J. Kenny, Ph.D., Icahn School of Medicine at Mount Sinai, New York, NY
Sachin Kheterpal, M.D., M.B.A., University of Michigan Medical School, Ann Arbor, MI
Gary A. Koretzky, M.D., Ph.D., Weill Cornell Medical College, New York, NY
Richard D. Krugman, M.D., University of Colorado School of Medicine, Aurora, CO
Jian-Dong Li, M.D., Ph.D., Georgia State University, Atlanta, GA

Kevin C. Kent Lloyd, D.V.M., Ph.D., University of California, Davis, Davis, CA
Edith P. Mitchell, M.D., FACP, FCPP, Thomas Jefferson University, Philadelphia, PA
Charles P. Mouton, M.D., M.S., M.B.A., The University of Texas Medical Branch at Galveston, Galveston, TX
Megan O'Boyle, Phelan-McDermid Syndrome Data Network, Arlington, VA
Rhonda Robinson-Beale, M.D., UnitedHealth Group, Minneapolis, MN
Susan Sanchez, Ph.D., The University of Georgia, Athens, GA
Jean E. Schaffer, M.D., Joslin Diabetes Center, Harvard Medical School, Boston, MA
Scout, Ph.D., National LGBT Cancer Network, Pawtucket, RI

Council Members Absent

Andrew P. Feinberg, M.D., M.P.H., Johns Hopkins University School of Medicine, Baltimore, MD
Anna Maria Siega-Riz, Ph.D., M.S., University of Massachusetts Amherst, Amherst, MA
Russell N. Van Gelder, M.D., Ph.D., University of Washington, Seattle, WA

2. Liaisons

Joseph M. Betz, Ph.D., Acting Director, Office of Dietary Supplements, DPCPSI
Janine A. Clayton, M.D., Director, Office of Research on Women's Health (ORWH), DPCPSI
Maureen M. Goodenow, Ph.D., Director, Office of AIDS Research, DPCPSI
Susan K. Gregurick, Ph.D., Director, Office of Data Science Strategy (ODSS), DPCPSI
Franziska B. Grieder, D.V.M., Ph.D., Director, Office of Research Infrastructure Programs (ORIP), DPCPSI
Christine M. Hunter, Ph.D., ABPP, Acting Director, Office of Behavioral and Social Sciences Research (OBSSR), DPCPSI
Christopher J. Lynch, Ph.D., Acting Director, Office of Nutrition Research, DPCPSI
David M. Murray, Ph.D., Director, Office of Disease Prevention (ODP), DPCPSI
Karen L. Parker, Ph.D., M.S.W., Director, Sexual & Gender Minority Research Office (SGMRO), DPCPSI
George M. Santangelo, Ph.D., Director, Office of Portfolio Analysis, DPCPSI
Marina L. Volkov, Ph.D., Director, Office of Evaluation, Performance, and Reporting, DPCPSI
Elizabeth L. Wilder, Ph.D., Director, OSC, DPCPSI
David R. Wilson, Ph.D., Director, Tribal Health Research Office, DPCPSI

3. *Ex Officio* Member Absent

Tara Schwetz, Ph.D., Acting Principal Deputy Director, NIH

4. Presenters

Jennifer Alvidrez, Ph.D., Health Scientist Administrator, ODP, DPCPSI
Colin Fletcher, Ph.D., Program Director, National Human Genome Research Institute
Susan K. Gregurick, Ph.D., Associate Director for Data Science Strategy and Director, ODSS, DPCPSI
Rick Horwitz, Ph.D., Executive Director, Emeritus and Senior Advisor, Allen Institute for Cell Science, Common Fund Data Ecosystem (CFDE) Working Group Co-Chair, and Council of Councils Member
Christine Hunter, Ph.D., ABPP, Acting Associate Director for Behavioral and Social Sciences Research (BSSR) and Acting Director, OBSSR, DPCPSI
Michael S. Lauer, M.D., Deputy Director for Extramural Research, NIH

Stephanie J. Murphy, V.M.D., Ph.D., DACLAM, Director, Division of Comparative Medicine (DCM), ORIP, DPCPSI
Concepcion (Marie) R. Nierras, Ph.D., Program Leader, OSC, DPCPSI
Andrea T. Norris, M.B.A., Director, Center for Information Technology (CIT), and Chief Information Officer, NIH
Karen L. Parker, Ph.D., M.S.W., Director, SGMRO, DPCPSI
Elizabeth Wilder, Ph.D., Director, OSC, DPCPSI, and CFDE Working Group Co-Chair
Sige Zou, Ph.D., Program Director, DCM, ORIP, DPCPSI

5. NIH Staff and Guests

In addition to Council members, presenters, and Council liaisons, others in attendance included NIH staff and interested members of the public.

B. Announcements and Updates

Robin I. Kawazoe, the Executive Secretary for the NIH Council of Councils, reviewed the following:

- Council members are Special Government Employees during the days of Council meetings and are therefore subject to the rules of conduct governing federal employees.

- Each Council member submitted a financial disclosure form and conflict-of-interest statement in compliance with federal requirements for membership on advisory councils. The financial disclosures are used to assess real and perceived conflicts of interest, and Council members must recuse themselves from the meeting during discussions of any items for which conflicts were identified.

- Time is allotted for discussion between the Council members and presenters, but time for comments from other meeting attendees is limited. The public may submit comments in writing; instructions are available in the *Federal Register* notice for the meeting, which was published on April 18, 2022.

- Minutes from the January 27–28, 2022, meeting are posted on the DPCPSI website. The minutes from this meeting also will be posted there.

C. Future Meeting Dates

The next Council meeting is scheduled to be held virtually September 8–9, 2022.

II. COMMON FUND UPDATE: MOLECULAR TRANSDUCERS OF PHYSICAL ACTIVITY CONSORTIUM (MoTrPAC)

Concepcion (Marie) R. Nierras, Ph.D., Program Leader, OSC, DPCPSI, provided an update on the Molecular Transducers of Physical Activity Consortium (MoTrPAC). The goal of the program is to assemble a comprehensive map of biochemical changes that transmit the health improvement effects of physical activity to all parts of the body. Additionally, the program will investigate the signals that transmit the health-improving effects of physical activity; how these effects are influenced by various factors (e.g., age, sex, body composition, fitness level, mode of exercise training); and how these changes correlate with other known positive effects of exercise (e.g., improved mood, better pain management). MoTrPAC data will be made accessible to other researchers, who are encouraged to develop additional hypotheses and create their own studies based on the program's results.

Dr. Nierras described the design of the MoTrPAC study, which was published in 2020 and comprises preclinical studies in rat models and clinical studies with adult and pediatric participants. The animal studies involve tissue-specific analysis of young adult or aged rats of both sexes that are exposed either to a single bout of exercise or to ongoing training. Up to 19 tissues per animal are evaluated before and after a single bout of exercise and before and after 8 weeks of training. Sex- and age-specific differences also are assessed. In the clinical study, healthy adult participants are categorized as either active (i.e., regular exercisers) or sedentary; the exercise modes of active participants (i.e., endurance exercise or strength training) also are assessed. Following acute exercise, adult participants are subject to blood, muscle, and fat tissue sampling. Nonactive participants are randomized to continue undergoing endurance or strength training or to remain sedentary. The molecular profile of people who exercise regularly is evaluated, in addition to differences before and after 12 weeks of exercise training. The pediatric cohort is assessed as to whether individuals have high or low activity levels. In contrast to the adults, the pediatric participants receive only endurance training and provide only blood samples. These studies will investigate the variability of the human response by age and sex to endurance versus strength training.

MoTrPAC consists of multiple investigators recruiting participants at up to 11 clinical sites, as well as six preclinical animal study sites. Human and animal samples are dispatched to chemical analysis sites, the resulting data are evaluated and organized within a bioinformatics center, and the whole enterprise is managed via a coordinating center and a data monitoring center. The animal studies have been performed, and the samples and data presently are being analyzed. Analysis of the young adult rats that underwent progressive training is complete, and a series of manuscripts is scheduled to be submitted for publication this year. Dr. Nierras highlighted results from the young adult rat study and explained that every analyzed tissue exhibited responses to exercise—including unexpected tissues, such as kidney and adrenal tissues. More than 40,000 unique analytes were regulated at multiple levels over the course of training. Several major molecular pathways related to metabolism, inflammation, extracellular matrix remodeling, and nutrient absorption were identified as dominant responses. Strong sex-specific responses also were observed. Next steps include comparing responses between young adult and older animals to assess the effects of aging on the physiological response to exercise.

Recruitment for MoTrPAC clinical studies began in October 2019 before being suspended in March 2020 because of the COVID-19 pandemic. Samples from the pre-COVID participant cohort—which included sedentary adults, highly active adults, and pediatric participants—presently are being analyzed and will be used for comparison with the preclinical data sets. After discussions on how to resume recruitment and exercise during an airborne pandemic, all clinical sites have reopened with newly developed safety protocols, which have been posted publicly to serve as a reference for other researchers. Dr. Nierras detailed the multi-omics analysis of human samples, which will assess gene expression, protein levels, post-translational modifications, metabolites, and more to provide a detailed molecular profile of participants to integrate with their clinical, physiological, and psychosocial information. Additionally, ancillary studies already have leveraged MoTrPAC protocols, samples, and data. For example, other groups have used MoTrPAC protocols to study populations with diabetes or with inflammatory diseases. MoTrPAC's ancillary studies policy is publicly available on the MoTrPAC website. Dr. Nierras thanked MoTrPAC leadership, management, investigators, and participants for their contributions.

Discussion Highlights

- In response to a question about the ethnic and geographical diversity of MoTrPAC study participants, Dr. Nierras commented that when the MoTrPAC study is completed, it will be the largest study examining the link between exercise and human health improvements. She added that close to 40 percent of participants come from underrepresented ethnic populations.

- When asked about bridging studies to compare exercise responses between species, Dr. Nierras explained that although certain molecular responses identified in rats could not be assayed in humans, surrogate markers could be used in their stead. She provided the example of an adrenal tissue response in rats that could be evaluated using blood biomarkers or hormone levels in human participants. Dr. Nierras pointed out that the uniform genetic background of the rat model was an advantage in the study. She added that genetic variation was a hallmark of the clinical studies, and that further experiments were being initiated in other species (e.g., mice).

- In response to a question about genomic characterization of the human study participants, Dr. Nierras answered that all human participants are undergoing whole-genome sequencing. Additionally, all human samples are being subject to RNA sequencing, epigenetic assays, targeted and untargeted metabolomics, lipidomics, and exosome assays. Because of the limited availability of human tissue samples, assays are prioritized according to the information that they can provide.

- When asked whether MoTrPAC participants could be enrolled in the *All of Us* Research Program, Dr. Nierras affirmed that MoTrPAC participants receive information about enrolling in the program but added that she did not know the proportion of participants who ultimately enrolled in *All of Us*.

- In response to a question about incorporating social determinants of health into the study, Dr. Nierras responded that these data were not being collected directly but that psychosocial effects of exercise training (e.g., pain management, quality of sleep) were being evaluated.

- In response to a question about recruiting participants with particular medical needs, Dr. Nierras pointed out that MoTrPAC participants were intended to be generally healthy individuals. The program is not evaluating special populations but is encouraging other investigators to use MoTrPAC protocols and reference data sets for their own focused studies.

III. ORIP CONCEPT CLEARANCE: REISSUANCE—ANIMAL MODELS AND RELATED BIOLOGICAL MATERIALS FOR RESEARCH (R21) (VOTE)

Stephanie J. Murphy, V.M.D., Ph.D., DACLAM, Director, DCM, ORIP, DPCPSI, introduced for concept clearance the reissue of the Development of Animal Models and Related Biological Materials for Research program, which uses the R21 mechanism. The objective of this program is to encourage innovative research to develop, characterize, and improve animal models, biological materials, and novel technologies to better understand human health and disease, as well as seek projects aimed at improving the diagnosis and control of diseases that interfere with animal use for biomedical research. The funds available and the anticipated number of awards for this program are contingent on NIH appropriations and the submission of meritorious applications. The award project period is 2 years, and the Council action is a vote for approval of reissuance of the concept.

The Animal Models R21 program was established in 2007 by the National Center for Research Resources and has continued to evolve under the administration of ORIP since 2012. To align with ORIP's mission of awarding grants to support research resources—such as animal models of human disease—this R21 program meets the demand for animal models that are more predictable and accessible for biomedical research and addresses the need for technological advancement for developing animal models. ORIP has issued four funding opportunity announcements (FOAs) for the Animal Models R21 program since 2013. Although the award rate is higher under the most recently completed FOA than in the previous two completed FOAs, the number of applications and awards listed for this FOA are lower. This is in part because this FOA was active for a 2-year period, compared with the 3-year periods of the other FOAs,

and partly because of decreases in application submissions due to COVID-19 impacts and restrictions. This FOA also was transitioned from a program announcement (PA) to a PA with special receipt, referral, or review considerations (PAR), which uses specific language for what types of projects would be considered responsive or nonresponsive. Consequently, although fewer applications were submitted under this PAR, they were more in alignment with ORIP's program priorities, resulting in a higher award rate.

A major theme of the *ORIP Strategic Plan for 2021–2025* is to facilitate the development and ensure the availability of the highest quality and most useful animal models and related resources for the advancement of research on human disease. As part of ORIP's NIH-wide emphasis, ORIP seeks to improve and disseminate the best animal models that are of interest to multiple NIH Institutes and Centers (ICs). To align with ORIP's NIH-wide mission, proposed R21 projects must have broad application to multiple NIH ICs and must explore multiple body systems or evaluate diseases that impact multiple body systems.

ORIP's Animal Models R21 program has made significant progress and impacts since 2013. For the past three completed FOAs, most R21 applications and awards were focused on animal model and technology development, with the primary model being the mouse, followed by the fly, and then zebrafish. The award rate averaged across the past three completed FOAs was approximately 20 percent. Awards made under the first two completed FOAs resulted in 114 and 87 publications, respectively, with approximately 80 percent of these awards having one or more publications. Although the most recently completed FOA ended only last year, three publications have been reported under this FOA so far.

Publications associated with the first completed FOA were cited 2,165 times between 2014 and March 2022, and publications associated with the second completed FOA were cited 796 times between 2018 and March 2022. A number of these high-risk, high-reward studies have led to the development of novel techniques, methodologies, and applications that have impacted or are impacting biomedical research. One example is an R21 award to the University of Utah for an application titled Gene Targeting in Zebrafish: Building Models to Assay Disease Genes, which led to two publications, with one publication focused on precise gene editing of the zebrafish genome and efficient recovery of recessive and phenotypically silent conditional mutations. A paper published in 2016 in the journal *Developmental Cell* was cited 107 times. Dr. Murphy requested that the Council vote to approve the reissue of the concept for Animal Models and Related Biological Materials for Research.

Discussion Highlights

- The discussants, Drs. Susan Sanchez and Patricia Hurn, provided their comments. Dr. Sanchez expressed enthusiasm for the concept and its past work supporting this critical area of biomedical science. Dr. Hurn also supported the reissue, noting the success of the previous iteration, but asked Dr. Murphy about the concept's broad focus. Dr. Murphy explained that the most recent ORIP strategic plan emphasizes NIH-wide impacts, so language was added to encourage investigators to consider the needs of multiple ICs.

- Dr. Murphy confirmed that no major changes to the language or intent of the reissuance are planned. ORIP will encourage those who participate to deposit the technologies, tools, and models they develop into appropriate biorepositories to share across the community.

Vote

A motion to approve the reissue of the Animal Models and Related Biological Materials for Research concept was forwarded and seconded. The motion passed with no abstentions.

IV. ORIP CONCEPT CLEARANCE: DEVELOPMENT OF RESOURCES AND TECHNOLOGIES FOR ENHANCING RIGOR, REPRODUCIBILITY, AND TRANSLATABILITY OF ANIMAL MODELS IN BIOMEDICAL RESEARCH (VOTE)

Sige Zou, Ph.D., Program Director, DCM, ORIP, DPCPSI, introduced a concept clearance for the development of resources and technologies for enhancing rigor, reproducibility, and translatability of animal models in biomedical research. He explained that the number of awards will be contingent on NIH appropriations and the submission of meritorious applications. The project period of each award will range from 2 to 5 years, depending on funding mechanisms.

Animal models play essential roles in the discovery of basic biological mechanisms, understanding of etiology of human diseases, and development of treatments. Preclinical research using animal models, however, often involves a high level of irreproducibility, as well as a high attrition rate for drug development. The *ORIP Strategic Plan 2021–2025* lists goals related to facilitating the development and ensuring the availability of best animal models; improving the dissemination of models; and advancing applications of new technologies to improve the generation, care, preservation, and distribution of animal models.

ORIP organized a workshop on Validation of Animal Models and Tools for Biomedical Research in collaboration with multiple NIH ICs, with help from research communities. This workshop was held as 10 virtual sessions between late 2020 and early 2021. Each session featured five to seven presentations with 200 to 500 attendees, including researchers outside the United States, as well as NIH program officials. Participants discussed the status of and the needs related to the validation of animal models and the rigor and reproducibility of animal research. Recommendations addressed genetic technologies, systematic phenotyping, screening technologies, artificial intelligence (AI) strategies, standardization and reporting, imaging facilities, and stock centers.

Based on the recommendations of workshop participants, ORIP proposed an initiative to support research- and resource-related research projects aimed at developing broadly applicable technologies, tools, and resources for validating animal models and enhancing rigor, reproducibility, and translatability of animal research. The initiative will align with ORIP's NIH-wide mission and strategic plan. In addition, ORIP will seek additional input from NIH ICs to identify resource gaps and needs, including those in AI strategies, and deep integration approaches.

Suitable projects for the initiative include resources or technologies to facilitate phenotyping at multiple scales; technologies for integrating multi-omics, biochemical, physiological, morphological, and behavioral data; strategies that allow user-friendly informatic searches and integrative mining of data for comparative human–animal biology; high-throughput imaging technologies for integrative analysis of cells and cellular networks across animal species; and resources that facilitate collaborations between basic scientists and clinical researchers in the use of multiple animal models for studying human diseases. Potential mechanisms for this initiative include R01, R21, R24, U24, and small business grant mechanisms. Dr. Zou requested that the Council vote for approval of the concept for Development of Resources and Technologies for Enhancing Rigor, Reproducibility, and Translatability of Animal Models in Biomedical Research.

<u>Discussion Highlights</u>

- The discussants, Drs. Kevin C. Kent Lloyd and Jian-Dong Li, provided their comments. Dr. Lloyd expressed support for the concept. He noted that progress has been made in the area of scientific reproducibility, but more work is needed, particularly related to understanding

experimental variables. Additionally, he noted that a working group of the Advisory Committee to the Director filed a report in June 2021 on this topic. He added that new technologies provide further opportunities in this area. Dr. Li also expressed support for the concept. He underscored the importance of animal models for research and commented on the need for new technologies, tools, and resources for validating animal models for successful translational applications.

Vote

A motion to approve the Development of Resources and Technologies for Enhancing Rigor, Reproducibility, and Translatability of Animal Models in Biomedical Research concept was forwarded and seconded. The motion passed with no abstentions.

V. REPORT OF THE COUNCIL OF COUNCILS WORKING GROUP ON BSSR INTEGRATION (VOTE)

Christine Hunter, Ph.D., ABPP, Acting Associate Director for BSSR, NIH, and Acting Director, OBSSR, DPCPSI, introduced the NIH Council of Councils Working Group report on the integration of BSSR at the NIH. The working group, a special advisory panel of behavioral scientists and community experts, was charged by Congress with assessing BSSR integration and recommending ways to increase integration and realize the benefits to the overall health of BSSR at the NIH. The working group members were selected from IC Advisory Councils with an emphasis on diversity of perspective. The working group assessed the current status of BSSR in NIH-supported research and training and identified existing processes that should continue or be enhanced, as well as new opportunities for enhancing processes, coordination, and integration of BSSR into research conducted across the Institutes, Centers, and Offices (ICOs). They also were charged with seeking input from experts in BSSR on the function and structure of the NIH research enterprise and with preparing a report that includes recommendations on ways to encourage greater BSSR integration and relevance to the research supported across the NIH. The working group explicitly focused on both social and behavioral science and on extramural, rather than intramural, aspects; they also undertook a broad examination of how the NIH operates to support integration rather than a detailed look at integration in specific disciplines. Dr. Hunter encouraged Council members to read the full report.

The group reviewed 24 IC strategic plans, four topic-specific NIH-wide strategic plans, and the *NIH-Wide Strategic Plan for Fiscal Years 2021–2025*. Eight strategic plans, or 28.5 percent, were identified as having significant BSSR integration, and another eight had moderate BSSR integration. The remaining 12 strategic plans, or 42.9 percent, had only nominal BSSR integration. Dr. Hunter pointed out that in some plans BSSR was only listed among a long list of disciplines or outcomes. When funding was analyzed using Research, Condition, and Disease Categorization (RCDC) codes, 10 ICs, or 41.7 percent, were identified with significant BSSR portfolios, defined as more than 20 percent of overall grant funding. Three (12.5%) had moderate levels of BSSR, defined as 10 to 20 percent of overall funding, and 11 (48.5%) had nominal BSSR funding, or less than 10 percent of overall grant funding. Among resource and center grants (P50, P30, and R24 grants), BSSR counted for 10 percent of the overall NIH portfolio. Training grants had relatively low numbers compared with the overall NIH portfolio, and the range over 10 years was 248 to 425 grants annually.

A survey also was conducted that focused on BSSR activities in the past three fiscal years (FYs), including research or training workshops and FOAs, NIH-wide research or training initiatives, internal and external IC collaboration on BSSR initiatives, BSSR-relevant scientific communication, and IC BSSR staff, including number of staff and models of staffing. The group also assessed IC advisory council membership as of December 2021 with two levels of BSSR expert review. The working group relied on expert input from the NIH, staff members, and qualitative assessments. Dr. Hunter emphasized

that the issue of BSSR integration is an NIH-wide cultural consideration and a multilayered issue, so looking at patterns is critical.

Dr. Hunter summarized the working group recommendations, pointing out that they focus on opportunities for increased BSSR integration, representation, and resources at multiple levels. First, the working group recommends that as strategic plans are revised or new strategic plans are developed, the NIH should ensure BSSR is more consistently included and linked to the IC mission and priorities. The absence of BSSR sends a message that it may be of lesser importance, and lack of inclusion limits the internal and external focus on BSSR initiatives and funding. Strategic plans are designed to define and operationalize how an IC's mission is advanced. Although inclusion in a strategic plan does not always equate to funding or other dimensions of integration, strategic plans are an important indicator of an IC's priorities, they communicate research priorities to both internal and external audiences, and drive initiative development and spending. So, if BSSR is not represented, an IC is less likely to prioritize BSSR. Twelve strategic plans—about 43 percent of those assessed—were identified as having nominal BSSR integration. The working group recommends that NIH leadership support and encourage the ICs and NIH-wide groups to work with the OBSSR and BSSR staff within their ICs when they develop their next strategic plans to help identify important BSSR goals that are relevant to each IC mission.

The working group recommends that the NIH evaluate and monitor the distribution of BSSR staff in the agency and identify strategies to address gaps in the number and diversity of BSSR staff. BSSR cannot be well integrated and maximally contribute to the broader NIH research mission unless BSSR expertise is represented consistently, and several ICs have limited or no program staff with BSSR as their primary expertise. ICs with dedicated work units or staff had more BSSR integration, and ICs using a distributed portfolio model had less BSSR integration.

Another recommendation was that the NIH bring IC Advisory Council representation into alignment with the policy requiring a minimum of two members on each Council with behavioral or public health expertise. Five IC Advisory Councils, or 21 percent of those assessed, had only one member with primary expertise in behavioral science or public health. A lack of high-level advisory representation hampers an IC's ability to develop, consider, and advance BSSR-relevant initiatives and grant funding. Although some Advisory Councils may be technically in compliance if they did not require primary expertise in BSSR, the working group felt that broader high-level advisory representation would reflect the intent of the policy and was important to BSSR integration.

The working group also recommends that the NIH continue to evaluate and monitor the composition of scientific review panels to ensure they adequately reflect BSSR knowledge and expertise and rapidly address any systematic gaps and biases. The Center for Scientific Review is conducting an ongoing quality improvement process and engaging in efforts to enhance fairness and reduce bias. Because of the central role of peer review in the NIH research enterprise, the working group noted that without intentional and ongoing vigilance to BSSR representation, some of the most innovative and integrated BSSR likely would not fare well in review Currently, BSSR applications are reviewed predominantly in BSSR-focused study sections. For BSSR to be well integrated into the broader biomedical research enterprise, the NIH also needs to ensure adequate BSSR expertise on study sections whose primary focus might not be BSSR but where BSSR factors, outcomes, and methods are included or should be included.

For grant funding and resources, the working group recommends that the NIH direct ICs with nominal BSSR in their portfolios to work with the OBSSR to identify opportunities to increase the application of BSSR in their research and training initiatives. BSSR funding at ICs varies from more than 50 percent of their portfolios to less than 5 percent. In 11 ICs—or 12 counting OD—BSSR represents less than 10 percent of total funding. Although there is no right percentage of funding, more BSSR should be encouraged for the ICs with more nominal levels of funding. The NIH also should identify gaps and

address opportunities to increase centers, resource grants, and trial networks that include BSSR capacity and focus. BSSR accounts for only 10 percent of all center resource and center grant funding, which represent significant long-term investments and commitments by the IC to provide additional support to enhance research development, capability, efficiency, and productivity in an area of science. Without dedicated resources and support, BSSR is at a disadvantage.

The NIH also should increase resources allocated to the OBSSR for staff and initiatives. The congressional language is highly aligned and consistent with OBSSR's mission, priorities, and activities, and the OBSSR is well positioned to support BSSR integration across the NIH. The Office has had good success addressing crosscutting scientific, training, and methodology gaps. Increased resources would build on its success and accelerate the pace of integration. The OBSSR could lead or facilitate the development of high-priority and crosscutting initiatives and resource initiatives. The recommendation includes seeding important new areas of science and providing subject-matter expert support to the ICs.

Finally, the working group recommends that the NIH engage BSSR expertise early and throughout the development and implementation of new research policies and practices. This recommendation was based on discussion among the working group members and their knowledge and experience of senior investigators with long track records of funding at the NIH rather than on a quantitative assessment. The NIH should involve BSSR experts and consider BSSR methods, measures, and practices as part of research policy development and implementation. This will ensure the policies are appropriate to BSSR and that these researchers can adhere to these policies without disrupting scientific progress.

The working group also identified four crosscutting considerations that are not unique to BSSR or specific to the working group charge, but the working group considered them important to share with NIH leadership. First, the NIH should enhance approaches to measurement of NIH funding. The current tracking tools, such as RCDC, are good but do not provide quantitative metrics of the extent to which any given category is represented in a grant. A grant may count as BSSR even when BSSR is a relatively small component of the research, which makes it difficult to look at nuanced issues and degrees of integration. Second, the NIH should use existing and new BSSR evidence to support the NIH efforts to enhance the diversity of the NIH and extramural research workforce. Although the working group was impressed with NIH's efforts to increase diversity to date, members wanted to draw attention to the need to use evidence-based approaches to address workforce diversity. Third, the NIH should foster team science and multidisciplinary integration, which is increasingly important as scientific research continues to grow in scale and complexity, yet many funding opportunities do not encourage or directly require team science. Academic incentives also do not always reward team science, so the working group encourages the NIH to consider ways to encourage the kind of science likely to answer the most important questions. Finally, the NIH should enhance the engagement of BSSR to inform improved conduct of science. Behavioral and social factors play an important role in how science is conducted, reported, and referenced, and these factors also are critical in understanding what builds trust in science. The working group recommends intentional inclusion of BSSR in advancing equitable and rigorous conduct of science.

<u>Discussion Highlights</u>

- Dr. Hunter clarified that BSSR's lower success rates may be related to its increased likelihood of using human or clinical research, which can slow publication. Such differences may be related more to the nature of the research than to discrimination against BSSR. Dr. Hunter agreed that publicizing NIH's support for BSSR applications is something OBSSR can do along with broader NIH efforts. Dr. Hunter also pointed out that the most recent NIH-wide strategic plan more explicitly supports behavioral and social sciences.

- When asked about qualitative feedback for IC leaders and how BSSR should be prioritized, Dr. Hunter commented that the working group assessed multiple levels of integration to show the cumulative effect of a lack of BSSR at each IC. Some opportunities to encourage more BSSR reflect the change in traditional ways of thinking, particularly in the wake of the COVID-19 pandemic, which highlighted that people do not always behave in predictable ways.

- Council members emphasized the importance of communicating to the community and undoing a long history of behavioral and social sciences' not being included.

- Dr. Hunter clarified that the assessment of measures or quality of life surveys was not within the purview of this working group but agreed it is an important area of focus for the NIH and OBSSR.

- Dr. Anderson plans to include a letter when he submits the report to the NIH Acting Director, noting the Council's comments that communication about the importance of BSSR to the whole enterprise of improving human health and the importance of measuring progress.

Vote

A motion to concur with the report on the integration of BSSR at the NIH was forwarded and seconded. The motion passed with two abstentions.

VI. ODP CONCEPT CLEARANCE: ADVANCING PREVENTION RESEARCH FOR HEALTH EQUITY (ADVANCE) INITIATIVES (VOTE)

Jennifer Alvidrez, Ph.D., Health Scientist Administrator, ODP, DPCPSI, presented on the ADVANCE Initiatives concept, which will solicit research projects to develop and test new preventive interventions or new implementation strategies for existing interventions that address leading risk factors in populations that experience health disparities. Funding will be contingent on NIH appropriations and availability of funds from ICOs. The ODP will prioritize ADVANCE FOAs for its own co-funding support. The awards will last from 2 to 5 years.

The mission of the ODP is to improve public health by increasing the scope, quality, dissemination, and impact of prevention research supported by the NIH. It also aims to provide leadership for the development, coordination, and implementation of prevention research in collaboration with NIH ICs and other partners. The ODP elevated its focus on health disparities from considering disparities a crosscutting theme to addressing them as a new strategic priority in 2022, and ADVANCE is one activity related to this change. In the Healthy People 2020 end-of-decade progress report, health disparities related to race, ethnicity, educational attainment, family income, and geographic location had not been significantly reduced over the last decade, suggesting a need for new strategies.

ODP also conducted portfolio analyses to identify NIH's efforts in this area between FY 2012 and FY 2019 based on the leading risk factors for death and disability, which have documented disparities in incidence or prevalence in one or more NIH health disparities populations. As knowledge about the epidemiology of and risk factors for chronic disease has increased over time, it might be expected that prevention research would also increase. However, support for new prevention research did not change over these 8 years. Fewer than one in three prevention projects measured one of the leading risk factors for death or disability, and fewer than one in nine included a randomized intervention to address a leading risk factor. Furthermore, fewer than one in 19 used an intervention to address a leading risk factor in health disparity populations.

ODP discussed these findings with the ICOs to assess their interest in a new initiative to promote preventive intervention research with populations that experience health disparities. The initiative, intended to be driven by ICOs and supported by the ODP, was presented to ICO Directors in February 2021. Twenty-four ICOs agreed to participate in initial planning discussions. When surveyed about ICO-specific interests in the leading risk factors, related preventive services, causes of death, and populations that experience health disparities, four clusters were identified reflecting shared interests, and four ADVANCE workgroups were created.

The topics are (1) cardiometabolic risk factors; (2) alcohol, tobacco, and other drug use as a risk factor; (3) mental health as a risk factor; and (4) cancer screening and preventive services. Workgroups are co-chaired by ODP and ICO representatives, and ICO members decide which groups to join. Members are encouraged to think broadly about prevention research priorities, gaps, and opportunities in their area, as well as target populations and age groups. Workgroups establish their own timelines and can decide which activities to pursue in developing FOAs, and they can determine the content of their FOAs, which may be specific to one workgroup or involve multiple workgroups.

Several focus areas apply across ADVANCE FOAs. One is the emphasis on NIH-designated populations experiencing health disparities, which include racial and ethnic minorities, sexual and gender minority (SGM) populations, socioeconomically disadvantaged populations, and underserved rural populations. Other priority populations—such as people with disabilities, veterans, or people with HIV—may be included in combination with health disparity populations. Another focus area is prospective preventive interventions, including primary, secondary, or universal prevention interventions, which may be evaluated using randomized controlled trials or rigorous quasi-experimental designs. Dr. Alvidrez pointed out that the majority of the prevention portfolio is observational, so more interventional projects are needed. Additional focus areas across ADVANCE include multilevel interventions addressing social determinants of health—particularly those that operate beyond addressing individual knowledge, attitudes, and behavior—and collaboration with community partners and service providers to enhance intervention acceptability, feasibility, and sustainability.

The workgroups currently are working to identify ICO gaps and priorities, which may differ across workgroup topic areas. ICOs may identify gaps and priorities related to specific health disparity populations or subgroups, specific risk factors of interest, or outcomes of interest. -Outcomes are not limited to leading causes of death but may include changes in risk factors, screening or early detection, or health status that can be changed within the typical 5-year NIH project period. Anticipated FOAs across ADVANCE workgroups may include notices of special interest to solicit applications relevant to IC priorities, R01 PARs to solicit full-scale preventive intervention projects, or network PARs to support preventive intervention projects that share common features and coordinated data collection. FOAs also could support training, professional development, or capacity-building FOAs if the workgroups determine that resource or personnel development is needed to successfully conduct research in this area. Each workgroup is anticipated to generate one to three FOAs beginning in summer 2022 for funding beginning in FY 2023.

Discussion Highlights

- Dr. Anderson highlighted the impact of ODP's recent initiative to manually curate and score characteristics of years of grants, revealing that a large portion of the current disease prevention portfolio is observational or relies on reuse of existing data and does not focus on populations with disproportionately poor health outcomes.

- The discussants, Drs. Maria Rosario Araneta and Graham Colditz, provided their comments. Dr. Araneta enthusiastically supported the concept but asked for additional considerations:

inclusion of a fifth cluster encompassing infectious diseases, inflammation, and immunodeficiency given the frequency of COVID-19 and diabetes; inclusion of vaccination as a structural intervention for the prevention of a number of conditions; and inclusion of pediatric conditions and structural interventions for schools. Dr. Araneta also suggested that ADVANCE prioritize structural interventions that are more enduring, require that planning and implementation of interventions be conducted in partnership with communities affected by health inequities, and implement a longitudinal component to show long-term results. Dr. Alvidrez explained that as an NIH-wide effort, participating ICOs contribute their own interests, resulting in flexibility about what can be included. She confirmed that vaccinations have been discussed but noted that longitudinal interventions will be determined by the success of the first cycle, although she agreed that the ability to extend beyond the 5-year project period is desired.

- Dr. Colditz expressed enthusiasm for the potential of multilevel interventions to reach populations that experience health disparities, noting that the cross-ICO workgroups reflect the way that people live their lives without being defined by a single disease.

- When asked about the potential for interagency collaboration, Dr. Alvidrez explained that ODP has some existing partnerships with HHS and other federal agencies, which could be accessed after initial FOAs are implemented.

- In response to a question about the inclusion of intersectionality, Dr. Alvidrez pointed out that although the FOAs may not focus on intersectionality, the cross-ICO nature of the workgroups allows the identification of intersectional populations, and FOAs could solicit intervention projects that involve such populations.

- When asked about the use of known implementation science strategies, Dr. Alvidrez explained that some disparities could be caused by an inability of evidence-based interventions to reach that community, which is an implementation science issue. In other cases, no interventions have been proven to work for a particular issue, so new interventions must be developed with the cooperation of the community. ODP strongly supports determining the appropriate methodology for each project. Council members suggested that past studies could be revisited to ensure that results can be translated to populations in need.

- In response to a question about implementation to affect change and continue follow-up, Dr. Alvidrez pointed out that although the resulting projects will vary in their longitudinal trajectories, in the first cycle the project intends to avoid research that will rapidly depart a community. The projects will need to be developed with the structure to sustain the programs, reflecting a need in the community and providing a pathway to continue if the intervention is effective. Renewal or continuation of awards likely will be possible, but sustainability should be considered from the beginning.

Vote

A motion to approve the ADVANCE Initiatives concept was forwarded and seconded. The motion passed with no abstentions.

VII. SGMRO UPDATE: THE NATIONAL ACADEMIES OF SCIENCE, ENGINEERING, AND MEDICINE (NASEM) CONSENSUS STUDY REPORT, *MEASURING SEX, GENDER IDENTITY, AND SEXUAL ORIENTATION*

Karen L. Parker, Ph.D., M.S.W., Director, SGMRO, DPCPSI, explained that the SGMRO helps people navigate issues so that they can be more inclusive in their work. Some SGMRO activities include coordinating SGM health research activities across the NIH, representing the NIH at conferences and events focused on SGM research, serving as a resource for the extramural and NIH communities about SGM-related research activities, connecting extramural researchers with key NIH contacts, convening conferences and workshops to inform priority-setting and research activities, collaborating with NIH ICs on the development of SGM health research reports, leading implementation of the *NIH Strategic Plan to Advance Research on the Health and Well-being of Sexual & Gender Minorities,* and leveraging resources and developing initiatives to support SGM health research. SGMRO also works closely with extramural researchers focused on SGM health research and ensures that trainees and early-stage investigators learn how to navigate the NIH. One of the major goals of the strategic plan is to increase data collection in SGM populations.

Dr. Parker outlined some of the current challenges related to measurement of SGM populations. No official standards for measuring SGM populations have been defined, meaning that even when questions are asked, pooling data and understanding the status of SGM populations is difficult. Labels and definitions are fluid and rapidly changing, particularly among younger populations. A recent poll showed that one in five Gen Z adults identified in some way as a member of an SGM population; as these SGM populations age, the need to capture the populations within research and surveillance systems will continue to be critical. Many terms used within the SGM community are unfamiliar to sexual majorities and cisgender populations, which can hinder collection and validation of data. Official statistics require time series (i.e., repeated measures over time), so any measures defined will need to be capable of such use. Production survey vehicles to investigate how these questions persist over time and how they could be modified are lacking. Translations into languages other than English are needed, and any survey questions must work for both interviewer- and self-administered response modes, as well as when asking proxies to respond for other household members. Dr. Parker also noted the particular dearth of research on measuring intersex traits or persons with differences of sex development (DSD).

NIH commissioned from NASEM to convene a panel of experts to review the existing knowledge base related to SGM-related measurement, make recommendations for specific measures, and provide guidance for their use. The effort was led by SGMRO and co-funded by 18 other NIH components. NASEM was tasked with (1) reviewing current measures and the methodological issues related to measuring sex as a nonbinary construct, gender identity, and sexual orientation in surveys and research studies, in administrative settings, and in clinical settings and (2) producing a consensus report with conclusions and recommendations on guiding principles for collecting data on sex, gender identity, and sexual orientation, as well as recommended measures for these constructs in different settings.

Although SGMRO recognizes that some questions would not be appropriate at all times with all people, the Office asked NASEM to focus on measures that can be used in the U.S. English-speaking adult population. More detailed response options may be necessary for measures used within LGBTQI+ populations, and modifications to the recommendations may be needed for use within younger populations. The measures also should prioritize representation of indigenous SGM populations, a space where the NIH has a particular dearth of research. The scope also focuses solely on measures of identity for counting and identifying members of sexual minority populations—more methodological research already has been conducted in this dimension, and identity is more relevant for measuring disparities in treatment and outcomes compared with such dimensions as attraction and behavior.

The following guiding principles were identified: (1) inclusiveness, or the principle that people deserve to count and be counted; (2) precision, or the use of precise terminology that reflects the constructs of interest, particularly constructs of sex and gender; (3) autonomy, or the need to respect individual identity and autonomy; (4) parsimony, or the collection of only necessary data; and (5) privacy, or the use of data in a manner that benefits respondents and respects their privacy and confidentiality.

The report recommended that the NIH collect and report data on gender by default, but collection of data on sex as a biological variable should be limited to circumstances in which information about sex traits is relevant. Collection of data on sex as a biological variable should be accompanied by collection of data on gender, and data collection efforts should not conflate sex as a biological variable with gender or otherwise treat the concepts as interchangeable. Such response options as "I don't know" and "prefer not to answer" should be used only when a response is required. Reporting of the use of write-in categories in published tabulations of responses is strongly encouraged, but continued testing and use of write-in response options is needed. NIH's Sex as a Biological Variable Policy (SABV) can work in collaboration with collection of gender identity and nonbinary sex data. Collecting gender identity data in human studies does not subvert or replace the need and requirement for SABV to continue in basic research. The SGMRO is working with the ORWH to develop specific information related to ensuring that SABV aligns with SGM research and data.

Dr. Parker noted that a Two-Spirit response category for American Indian and Alaska Native (AI/AN) respondents should be included in both the sexual orientation and gender identity groups only when Indigenous populations can identify themselves. Because Two-Spirit is a term by and for Indigenous peoples and is culturally anchored with meaning and, potentially, social status, it is not appropriate for use by non-Indigenous populations.

Dr. Parker showed the recommended measure for collecting data on sexual orientation, which asks respondents to select one option that best represents how they think of themselves, including lesbian or gay, straight (i.e., not gay or lesbian), bisexual, Two-Spirit (if the respondent is AI/AN), or a different term provided in free text. Recommended topics for future research include alternate wording for the "straight" response option, preferred ordering of response categories, guidelines for measures that capture other dimensions of sexual orientation (i.e., attraction, behavior), addition of sexual orientation response options that may be more prevalent in subsets of the LGBTQI+ population (e.g., queer, questioning), evaluation of existing measures and identification of best practices for collecting data among sexual minority adolescents, and impact of proxy reporting of sexual orientation identity.

The recommended measure for gender identity is a two-step question that asks respondents what sex was assigned at birth on their original birth certificate and their current gender identity, which includes a Two-Spirit option for AI/AN respondents and free text option. Recommended topics for future research include testing current gender-specific response categories and optimal response ordering, developing alternative two-step gender measures that offer an inclusive count of gender minorities without asking sex assigned at birth, assessing the inclusion of "nonbinary" in gender identity response categories, measuring utility of a nonbinary response when asking about sex assigned at birth to capture the intersex population, and determining how to count responses in terms of gender. Additionally, the two-step measure should be tested beyond general population assessments of English-speaking adults, and a "select all that apply" option should be evaluated for the current gender question.

The recommended measure for gathering information on nonbinary sex asks respondents whether they have ever been diagnosed by a medical doctor or other health professional with an intersex condition or DSD, or whether they were born with or developed naturally in puberty genitals, reproductive organs, or chromosomal patterns that do not fit standard definitions of male and female. Dr. Parker noted that the committee recognizes the wordiness of this question, but she explained that very little research has been

conducted on any measures to capture this population. The committee considered three measures, and this was the most complex. Recommended topics for future research in this area include additional testing of single-item intersex and DSD status questions; assessing the efficacy of the three intersex and DSD measures highlighted in the report to determine which measure most effectively identifies the intersex and DSD populations in a range of settings; determining the effects of including definitions and examples of terms used in intersex status questions, such as "intersex," "DSD," and specific intersex variations; and identifying the effects of proxy reporting of intersex or DSD status, particularly of caregivers reporting children's status.

In the future, the Office plans to collaborate with the SGM Research Coordinating Committee to map opportunities at ICOs to add or update recommended questions as data systems and surveys are aligned. It also plans to update and refresh the SGMRO website and related measurement webpages to feature and highlight the recommendations from the NASEM report. The SGMRO will collaborate and provide technical assistance across NIH, HHS, and interagency workgroups to add the NASEM recommendations where appropriate or develop testing of new measures outlined in the future research recommendations. Dr. Parker reiterated that these questions, although imperfect, can serve as a starting point, and the advocacy community has expressed broad support and urged federal agencies and other groups to adopt the recommendations.

Discussion Highlights

- In response to a question about the next steps to increase researchers' comfort with these questions, Dr. Parker commented that the SGMRO believes that research that does not assess SGM identities is not rigorous and could lead to misinterpretation of data. People in medical schools have expressed strong interest in more education about these aspects. The SGMRO is working with colleagues across the NIH to determine how to enact funding opportunities focused specifically on education and on using these measures in specific settings. The Office also is working with the NIH to develop a resource that will address issues of cultural competency related to gender, which extends beyond SGM populations. The Clinical Center, in particular, has focused on such initiatives as having pronoun information available. She suggested that visibility and education will help; some is the responsibility of the SGMRO and NIH, and other efforts will require integrating education more broadly.

- Dr. Scout commented on the LGBTQI+ advocacy community's excitement about this report and urged attendees to adopt the recommendations quickly. Training is available, as are administrative supplements for data collection.

- When asked whether other NIH groups might incentivize or strongly encourage adoption of these measures, Dr. Parker responded that the SGMRO is meeting with representatives from the PhenX Toolkit, a place many researchers look for measures, and discussing with *All of Us* representatives how to monitor their data collection to be more aligned with the NASEM report. Dr. Parker also noted that SGMRO representatives participate in many committees. She added that the President's budget for 2023 includes a potential SOGI Research Center which would support research opportunities across government to implement and study SOGI-related questions.

- When asked about provider pushback, Dr. Parker pointed out that studies have shown that patients were comfortable answering the questions, but the majority of providers were nervous about asking the questions, and privacy concerns can be overcome. She emphasized that both these areas offer opportunities to address concerns, and this report serves as a jumping-off point.

VIII. ADJOURNMENT FOR THE DAY

Dr. Anderson adjourned the meeting at 2:31 p.m. on May 19, 2022.

Day 2

I. REVIEW OF GRANT APPLICATIONS

This portion of the meeting was closed to the public, in accordance with the provisions set forth in Sections 552(b)(c)(4) and 552(b)(c)(6), Title 5, U.S. Code and Section 10(d) of the Federal Advisory Committee Act, as amended (5 U.S.C. Appendix).[1] Members were instructed to exit the meeting if they deemed that their participation in the deliberation of any matter before the Council would represent a real or perceived conflict of interest. Members were asked to sign a conflict-of-interest/confidentiality certification to this effect. The *en bloc* vote for concurrence with the initial review recommendations was affirmed by all Council members present. During the closed session, the Council concurred with the review of 122 ORIP applications with requested first-year direct costs of $42,023,538, and 939 Common Fund applications with requested first-year direct costs of $939,350,917.

II. CALL TO ORDER AND INTRODUCTION TO DATA PRESENTATIONS

Dr. Anderson welcomed participants, NIH staff members, and members of the public to the second day of the meeting and reviewed the day's agenda. The open session of the virtual meeting began at 11:50 a.m. on Friday, May 20, 2022.

III. UPDATE ON THE NIH SCIENCE AND TECHNOLOGY RESEARCH INFRASTRUCTURE FOR DISCOVERY, EXPERIMENTATION, AND SUSTAINABILITY (STRIDES) INITIATIVE

Andrea T. Norris, M.B.A., Director, CIT, and Chief Information Officer, NIH, provided an update on the STRIDES Initiative. New capabilities for computation and data storage have enhanced opportunities for science and scientific research. The large amount of data generated, however, requires cloud storage and access. To move researchers and research assets into cloud platforms, ensure the researchers have the skills to work in these new environments, and support the national research ecosystem, the NIH negotiated with major cloud providers to provide cloud capabilities, support, and training at a discount to any institution that the NIH funds. This allowed partnership in such new areas as artificial intelligence (AI) and machine learning (ML). In 3 years, the STRIDES program has moved more than 170 petabytes of scientific data into the cloud across Google Cloud Platform and Amazon Web Services; a new partnership with Microsoft Azure, the third major commercial platform in the United States, has been added. Cost savings range from 10 to 25 percent, compared with typical pricing. NIH-funded researchers have used more than 250 million compute hours, more than 700 research programs have been enrolled, and more than 4,200 people have been trained in general cloud use. The NIH also develops training programs for using the cloud for specific biomedical research activities. Through the STRIDES initiative, the NIH also has achieved more than $31 million in cost savings in the form of discounts and cloud credits. As program usage increases, the commercial providers grow, allowing them to provide further advantages to the NIH.

[1] For the record, it is noted that members absented themselves from the meeting when the Council discussed applications (a) from their respective institutions or (b) in which a conflict of interest may have occurred. This procedure applied only to applications that were discussed individually, not to *en bloc* actions.

When a research institution enrolls in the STRIDES program, the vendor works with the institution to provide discounts to any NIH-funded researcher. STRIDES is supporting many important research programs, including the Common Fund, and the ODSS is a strong partner with this program given its data science themes. Adoption and use have increased quickly, yielding substantial amounts of computationally oriented science, analytics, and data work across all areas of science. The STRIDES agreement allows free storage of and access to certain types of data repositories and data sets, including the Sequence Read Archive, a widely used data resource, and SARS-CoV-2 analytic data. Another key benefit of STRIDES is the speed and agility provided by this service, which accelerates the pace at which science can be conducted. The initiative also has been critical in allowing the NIH to perform AI/ML research in the cloud.

Training has been critical for the initiative, and the NIH has partnered with both vendors and other training providers. Before the COVID-19 pandemic, training was conducted regionally. Other methods include online training, instructor-led training, and codeathons. This approach has affected the program's strategic focus on building capacity with minority-serving institutions, Institutional Development Award (IDeA) states (i.e., states with low levels of NIH funding), and other institutions that may not typically have as much involvement in the computational space for biomedical research. Previously, many such institutions could not engage in this type of research because significant campus computational infrastructure was necessary, but cloud computing eliminates that requirement. Efforts in this area have included collaborative research and development projects, targeted engagement and training efforts, and special research credits from cloud providers. After several pilot programs, new award supplements have been announced to help institutions in IDeA states develop and optimize pipelines, algorithms, and data analysis. The ODSS also has supported creation of the NIH Cloud Lab in the STRIDES environment to allow researchers to experiment with data and strengthen their skills without making any irreversible changes. Primary use cases include exploring the cloud consoles, supplementing cloud training, experimenting with simple cloud solutions, and benchmarking costs.

Ms. Norris pointed out that another essential component of STRIDES was the need to develop the NIH cloud infrastructure within the commercial platforms. The Enterprise Cloud Platforms have been developed—through each of the three commercial providers—for NIH intramural researchers at a lower cost. These new platforms provide secure, dedicated network connectivity; 80 to 90 percent of necessary applied and inheritable cybersecurity controls; optimized cloud environments; and extended federated login. Challenges tend to be the same across ICs and institutions, but the research community has been able to take advantage of the program and the capabilities offered by the cloud. Ms. Norris pointed out that not all work can be conducted in the cloud—she anticipated that the percentage of work in the cloud will continue to increase, but some on-premise capabilities remain necessary. Researchers also tend to need help and guidance in using the system effectively, showing the importance of advisory and technological support.

Ms. Norris concluded by listing the following near-term initiatives: Continue to increase the adoption and use of STRIDES; explore expansion of partnerships to include widely used biomedical software and programs; scale adoption of enterprise cloud platforms and support within the NIH; and develop more scalable, sustainable cloud training strategies. Ms. Norris and her team are pushing for the adoption of STRIDES wherever appropriate, including encouraging applicants for computationally oriented awards to use the program. She emphasized that use would not be required. The program also will continue to increase the capacity of minority-serving and historically underrepresented institutions. Governance work regarding how the NIH uses STRIDES in the cloud, as well as how the NIH works with the extramural community, is in progress. Partnerships may be expanded, and additional computational tools and applications may be added. Finally, best practices for scalability, as well as training and support, will continue to be evaluated.

Discussion Highlights

- When asked about the funding model as the program continues to grow, Ms. Norris commented that partnerships with vendors likely will continue, providing ongoing benefits for the NIH in terms of discounts and for the vendors in terms of increased business. The savings provided by the discounts offset the cost of the STRIDES support team, and any extra funding can support more NIH research. Participating institutions use a variety of approaches to balance the research costs and the savings from the discounts.

- Although Ms. Norris agreed that an interoperable ecosystem regardless of vendor was ultimately desirable, she acknowledged that the program is not yet close to reaching this goal. Currently, they are working to make data more findable, accessible, interoperable, and reusable (FAIR), as well as sustainable, and ensuring that NIH's big data programs can interoperate. Another challenge is ontology across unique scientific domains. Ms. Norris added that working on a single platform in the future was unlikely.

- When asked if the NIH should become a cloud provider, Ms. Norris replied that this may not have been considered. The NIH generally focuses on advancing the science and discovery, rather than supporting the mechanics. Although her team works collaboratively with the vendors to gather information as the program proceeds, they prefer not to manipulate the market, which is why the NIH partnered with the top three cloud platform providers that the community already used.

- Council members urged Ms. Norris to encourage commercial vendors to implement a usage cap to ensure researchers remain under budget. Ms. Norris agreed that STRIDES may have this leverage as a major user of these platforms.

- In response to a question about the challenges of cybersecurity approvals, Ms. Norris reiterated that STRIDES has 80 to 90 percent of controls in place on each platform; although they do not yet have full approval of a standard control, the current standards are high and the program is working to improve.

- Council members suggested a presentation by a STRIDES user at a future Council meeting.

IV. REPORT OF THE CFDE WORKING GROUP (VOTE)

Rick Horwitz, Ph.D., Executive Director, Emeritus and Senior Advisor, Allen Institute for Cell Science, CFDE Working Group Co-Chair, and Council of Councils Member and Elizabeth Wilder, Ph.D., Director, OSC, DPCPSI, and CFDE Working Group Co-Chair presented on the CFDE Working Group. Common Fund programs catalyze biomedical discovery and innovation, and many programs develop large data sets intended to be shared with the biomedical research community to accelerate science and promote innovation. However, the data sets were siloed—each program delivered large data sets to its own data coordinating center and had its own metadata and tools. Although the Common Fund focuses on collaborating across the NIH, it was not successfully collaborating across Common Fund programs, creating redundancies and missing opportunities for discovery.

The CFDE was created in 2019 for an initial 3-year period with the intent of fostering new science by creating an ecosystem that enables greater access to and use of Common Fund data resources. The program asked how researchers would combine data sets from multiple Common Fund programs more effectively, what happens to Common Fund data after the program's funding ends, and how to train users to work with data in the cloud. The CFDE aims to enable users to query across and use multiple data sets,

sustain Common Fund data and tools, and train users to work with Common Fund data. The CFDE is intended to support and enhance the value of many Common Fund programs as core infrastructure.

Today the CFDE has one coordinating center and 11 Common Fund program data coordinating centers. The CFDE Portal allows users to find relevant data sets by searching across metadata, and the Data Coordinating Center joint projects explore scientific questions through interrogation of multiple data sets. The program has 22 research awards to enhance the utility and use of Common Fund data sets and is developing an assessment plan to measure the progress and impacts of the program.

Dr. Horwitz explained that a Council of Councils working group was charged with reviewing the current scope and goals of the CFDE, as well as its progress to date, and making recommendations for future scope and goals in the following areas: findability and accessibility of data, data harmonization and interoperability, cloud workspaces, the CFDE scope and strategy in the context of related NIH activities, sustaining access to data and tools after Common Fund programs end, and training and outreach to enhance access to and use of the data. The working group made several key observations. First, the challenges that CFDE faces with respect to managing data resources are shared by data managers throughout the biomedical research enterprise. The CFDE will not solve these challenges alone, and some are beyond the scope of the CFDE. The pilot phase of the CFDE has been impressive and accomplished much in 3 years, including development of a discovery portal, establishment of a metadata model, piloting a cloud-based workspace, training users, supporting exploration across data sets, and establishing a community of investigators who are collaborating to build the ecosystem. The chief metric of success for the CFDE is discovery—if the CFDE is successful, in 5 years investigators will be using Common Fund data for new discoveries and purposes.

The working group defined several recommendations. First, the group recommends that the NIH support queries across data sets through use of community-based standards, development of robust use cases and workflows for diverse data users, development of innovative methods for data search, expanded use of knowledge graphs that provide relationships across data sets, and citation of data and tool contributors by CFDE users. User experiences should shape further development of CFDE. Additionally, CFDE should actively evolve through development of tools for new uses or users and making data sets available in the cloud. Continued attention to NIH-wide and other data communities is essential.

To support sustainability and accessibility, the working group also noted that public repositories offer an increasingly attractive option for sustaining data beyond Common Fund support. Repositories developed by Common Fund programs should continue to transition to support from NIH ICs or other entities if they prove useful to the community. CFDE should not be a long-term repository for Common Fund data sets. Repository and data management are NIH-wide concerns beyond the scope of the CFDE, although the CFDE should participate in these efforts as an important client of these repositories. Cloud storage may be an option, but barriers to democratization need to be considered.

Finally, the working group noted that training needs to be an enhanced priority for the CFDE, and the focus of training should be on enabling the use of Common Fund data sets. The CFDE should consider diverse training approaches, including emphasizing scalable training. Individualized training efforts also are valuable and should emphasize early-stage investigators and investigators from lower-resourced institutions.

Discussion Highlights

- In response to a question about the dynamic nature of data, Dr. Wilder explained that the working group concluded that the CFDE should be an active participant in conversations about the NIH-wide problem of what data upkeep entails. However, the CFDE cannot solve the issue, which is a

larger concern that affects long-term management of data sets everywhere. Partnerships with entities working to address this issue could be an option.

- Dr. Horwitz reiterated the working group's recommendation for increased emphasis on training; Council members urged Dr. Anderson to convey the need for increased training in his letter to the Acting Director.

- In response to a question about the lack of research involving human trials reflected in the report, Dr. Wilder clarified that the CFDE focuses on data generated by current Common Fund programs, which include significant amounts of basic science. Some programs—such as Kids First and MoTrPAC—do include human data, but the extent to which the CFDE will include human data will depend on future Common Fund programs that collect that type of data.

Vote

A motion for concurrence with the Working Group Report on the CFDE was forwarded and seconded. The motion passed with one abstention.

V. NIH DATA MANAGEMENT AND SHARING POLICY

Michael S. Lauer, M.D., Deputy Director for Extramural Research, NIH, provided an update on the implementation of the NIH Policy for Data Management and Sharing. Data sharing is expected and often required, and data sharing policies are not new—the first NIH Data Sharing Policy was established in 2003, followed by the 2014 NIH Genomic Data Sharing Policy, which included human and nonhuman genetic data, and the 2016 NIH Policy on Dissemination of NIH-Funded Clinical Trial Information. Several ICs and programs have their own specific policies and guidelines. Development of the NIH Policy for Data Management and Sharing began in 2016 with the 21st Century Cures Act, was developed into a draft policy and guidance in 2019, and will become effective in January 2023, replacing the 2003 Data Sharing Policy.

The NIH Policy for Data Management and Sharing has two basic requirements: submission of a data management and sharing plan for all NIH-funded research and compliance with the ICO-approved plan. Potential grant recipients are asked to develop a plan because the NIH funds a diversity of science types and cannot prescribe what data sharing looks like for each project. The policy expects sharing to be the default practice. Data sharing should be maximized, with justifiable limits for technical, ethical, and legal factors. The plan should be implemented responsibly and outline the protection of privacy, rights, and confidentiality. It also should abide by existing laws, regulations, and policies. Data sharing also should be prospectively planned for at all stages of the research process. The NIH expects that the data sharing plan may change over time, but any changes would be made in close coordination with NIH staff.

When this policy takes effect, proposed data sharing plans will be submitted with each application, including a brief description in the budget justification and the full plan as a separate attachment. The plan will be assessed during the peer review process—the reviewers will comment on, but not score, the budget. NIH program staff then will assess the plan, which can be revised at this time. The plan then will be incorporated into the terms and conditions of the grant, and compliance will be monitored at regular intervals and may affect future funding decisions.

Key parameters of a flexible policy are key. All data should be managed, but not all data need to be shared. Data from all NIH-supported research generating scientific data—defined as recorded factual material of sufficient quality to validate and replicate research findings—should be shared, including both published and unpublished data. Data that do not need to be shared include laboratory notebooks, preliminary analyses, case report forms, and physical objects. The shared data should be accessible as

soon as possible, but no later than publication or the end of an award. Journal or repository policies may affect how long the data should be shared.

Webinars, FAQs, and draft guidance for researchers working with AI/AN participants are available now; before the plan is launched, additional information—such as cost considerations and information on protecting privacy—will be available. Sample plans also will be developed. The Genomic Data Sharing Policy will be harmonized to ensure that applicants have to submit only one plan. After implementation, the policy will be assessed regularly for short- and long-term goals, and incentives for data sharing will be developed. Dr. Lauer emphasized the need to generate a culture in which data sharing is positive and researchers are incentivized to share, similar to current incentives to publish.

Dr. Lauer demonstrated the new sharing.nih.gov site, which allows users to explore the areas in which the NIH has sharing policies. Key resources on the site include policy overview infographics, a decision tool to help users determine which policies apply to their research, step-by-step instructions for some tasks, information on planning and budgeting for a data management sharing policy, and best practices. The site also will separate projects submitted before and after the new policy.

Dr. Lauer outlined some questions raised to date. Because the policy is flexible, it also has some ambiguity, so some have asked how to interpret the definition of scientific data, where data should be shared, and how the policy relates to other ICO data sharing policies. Questions about funding have been raised, such as whether the NIH will raise budget caps, what happens after the award ends, and how program officers will assess plans. Dr. Lauer confirmed that the plans will be public and compliance will be enforced similarly to other NIH compliance processes. Dr. Lauer encouraged attendees to pay attention to outreach activities and the sharing.nih.gov website.

Discussion Highlights

- Dr. Lauer explained that they are still considering how to apply the policy to programs in which data sharing is already a component, but it could involve a very short sharing plan that refers readers to the elements addressed in the program application.

- Council members suggested defining specific parameters for preliminary data.

- In response to a question about sharing data back to communities of patients or organizations that provided the data, Dr. Lauer responded that some programs have addressed this, but the concept needs to be considered more generally, and particularly when related to data from human participants. He emphasized that the policy represents a minimum of what would be expected, so projects may have their own requirements about sharing data back to participants. Council members encouraged a culture shift beyond the bare minimum, noting that research is intended to help people, rather than solely serve as a basis for publication.

- When asked about examples of success in this area, Dr. Lauer pointed to the National Science Foundation, which conducts different kinds of research but has engaged in sharing for a long time. In other fields—such as economics and library science—data sharing has become routine. He noted that human data in this area will present some new challenges.

- Dr. Lauer confirmed that the data sharing policy will not be considered in the evaluation of grants and that this would be emphasized to investigators and study sections.

VI. PARTNERSHIPS CONTRIBUTING TO DATA MANAGEMENT AND SHARING POLICY IMPLEMENTATION THROUGH FAIR DATA SHARING

Susan K. Gregurick, Ph.D., Associate Director for Data Science, NIH, and Director, ODSS, DPCPSI, presented an overview of NIH's data management and sharing capabilities. ODSS's mission is to catalyze new capabilities in biomedical data science by providing NIH-wide leadership and coordination for modernization of the NIH data resource ecosystem, developing a diverse and talented data science workforce, and building strategic partnerships to develop and disseminate advanced technologies and methods. The NIH Strategic Plan for Data Science articulates goals and objectives to create data infrastructure. Dr. Gregurick emphasized that the NIH recognizes the challenges ahead.

The NIH has supported the generation of a wealth of biomedical data that are available and reusable to research communities; not all of these data, however, can be used efficiently and effectively by AI/ML applications. During the past year, the NIH has provided supplement awards to NIH investigators to support collaborations to make data FAIR and AI/ML-ready. The most common biomedical focus areas include Alzheimer's disease, cardiovascular disease, and aging. The most common data types include imaging, electronic health records, -omics, and speech.

Dr. Gregurick emphasized that the NIH strongly encourages open-access data sharing repositories as a first choice for data sharing. She noted that researchers across the biomedical enterprise have different expectations and experiences, and the need for data sharing differs across disciplines. To help researchers locate appropriate resources for sharing their data, and to promote awareness of resources where data sets can be located for reuse, the Trans-NIH BioMedical Informatics Coordinating Committee maintains lists of several types of domain-specific data sharing resources. Additionally, PubMed Central article data sets now are available on the cloud through the Amazon Web Services Open Data Sponsorship Program. This effort was made possible through the STRIDES Initiative partnership.

The NIH supports a variety of data repositories and knowledge bases of differing sizes and complexity and at different levels of maturity, and each has the potential to bring value to a given research area. During the past year, the ODSS has supported existing NIH repositories to leverage supplemental funding to enhance their ability to align with FAIR and TRUST (transparency, responsibility, user focus, sustainability, and technology) principles. To date, 17 awards have been granted, including three to intramural research repositories. Topics include traumatic brain injury, obesity, nutrition, mental health, and immune response.

Data resources are important research tools, and the NIH also supports funding opportunities for new and existing data repositories and knowledge bases to be awarded as NIH biomedical resources. The Biomedical Data Repositories and Knowledgebases Program supports investigator-initiated, sustainable data resource development driven by critical research needs. To date, seven awards have been granted. The projects must fill a scientific need or gap, encourage adoption of good data management practices, engage the research community to contribute and use data, and govern data life cycle and preservation.

The Generalist Repository Ecosystem Initiative is soliciting applications from generalist repositories that are working together to implement consistent capabilities, create better access to and discovery of NIH-funded data, conduct outreach and training on FAIR data practices, and engage the research community. This initiative will enhance abilities to align with the desired characteristics of data repositories, implement browsing and searching functions for data across supported generalist repositories, develop consistent metadata models across the generalist repositories, conduct quality control for NIH-funded data, enable the connectivity of digital objects, support researcher use cases, develop educational materials on how to create FAIR data and manage data, and develop implementation metrics.

The Federation of American Societies for Experimental Biology (FASEB) DataWorks! Prize fuels the vision and an annual challenge to showcase the benefits of research data management, while also recognizing and rewarding teams whose research demonstrates the power of data sharing by seeking new and innovative approaches to data sharing in the field of biological and biomedical research. A total of $500,000 is available for as many as 12 awards. The FASEB also is seeking researchers to align with the NIH 2023 Data Management and Sharing Policy for additional awards.

The ODSS has partnered with the National Library of Medicine to support training through the Data Curation Network, which is composed of professional data curators, data management experts, data repository administrators, and scientists and scholars from academic institutions and nonprofits. Learning sessions provide approaches and methods and best practices for how to better manage, curate, share, and promote transparency, reproducibility, and reuse of research data. ODSS also is hosting a monthly seminar series to highlight exemplars of data sharing and reuse.

Discussion Highlights

- When asked about hidden biases in data. Dr. Gregurick stated that the ODSS posted a Notice of Special Interest to develop and advance ethical frameworks for the use of AI/ML in biomedical and behavioral sciences, focusing on data and methods. She emphasized the importance of considering best practices during data generation and noted that the Bridge to Artificial Intelligence program is also addressing this topic.

- Dr. Gregurick explained that plans for assessing and addressing noncompliance will be overseen by the Office of Extramural Research but will be implemented through each Institute and program. She noted the importance of community engagement in developing expectations for data management and sharing. An effort in Europe, the ELIXIR program, has funded an effort to help researchers develop their research data management plans called RDMKit.

VII. COMMON FUND FINAL REPORT: KNOCKOUT MOUSE PHENOTYPING PROGRAM (KOMP2)

Colin Fletcher, Ph.D., KOMP2 Program Director, NHGRI, presented the final KOMP2 report. He explained that the vision and goal of a knockout mouse program (KOMP) began in 2003 against the backdrop of the release of the draft mouse genome, consisting of 20,000 well-defined genes, and growing frustration in the wider biomedical research community about the lack of access to mouse knockout technology. The NIH and the genomics community convened for a mouse genome-wide targeted mutagenesis meeting in 2003 to discuss the possibility of creating a resource for the research community. This resource would need to address the major challenges in developing mouse knockout technologies. The high cost and technical challenges slowed production of gene knockouts—even if 500 per year were produced, 40 years of production would be required to span the entire genome. Additionally, availability was low, with few strains developed by specialists in only certain laboratories deposited in public repositories. Another challenge was the low level of coverage because phenotyping was not comprehensive at the time. The low level of resource sharing further impeded comprehensive phenotyping.

The NIH and the genomics community recognized that a coordinated project to systematically knock out all mouse genes would have a tremendous benefit to the research community. They agreed that such a project would include a comprehensive genome-wide resource of mutant embryonic stem (ES) cell lines, in which most known mouse genes could be knocked out within 5 years. The community also recommended that ES cells be converted into mice at a rate consistent with project funding and the ability of the worldwide scientific community to analyze and phenotype them by a limited set of broad and cost-

effective screens. Additionally, all ES cell clones and mice, as frozen embryos or sperm, should be available to any researcher at minimal cost. Finally, all mouse phenotyping and reporter expression data should be deposited into a public database.

Dr. Fletcher detailed the overall implementation of this vision. In 2006, the NIH launched KOMP as an NIH-wide 5-year program, with a budget of $56.6 million from participating NIH ICs. KOMP created 8,500 ES cell lines in C57BL/6N mice, with null or conditional alleles and containing a reporter to evaluate gene expression. After successful completion of KOMP, the NIH launched KOMP2 in 2011 as an NIH-wide and Common Fund 10-year program, with a budget of $225 million. KOMP2 has completed its 10-year cycle and has closed as a Common Fund project. Since its inception, KOMP2 created 5,500 mouse lines (2,500 ES/3,000 CRISPR) and established broad phenotyping across many domains. KOMP reagents are distributed by the ORIP-supported Mutant Mouse Resource and Research Center network, and KOMP has a web portal that is operated by the European Bioinformatics Institute.

Dr. Fletcher noted that KOMP was a collaborative effort with international partners, which allowed the program to double production. In the first phase of KOMP (2006–2011), both the NIH-funded research laboratories and the European Conditional Mouse Mutagenesis Programme generated 8,500 knockouts simultaneously, totaling 17 ES cell clones and allowing KOMP to cover the mouse genome completely within a 5-year period. From 2011 to the present, KOMP2 worked with the International Mouse Phenotyping Consortium (IMPC) and strengthened relationships with other international phenotyping laboratories. Twenty-two active partners—located in Australia, Asia, Africa, Europe, and North America—comprise the IMPC. KOMP2 and international partners generated more than 10,000 mouse knockout strains, and nearly 9,000 have been phenotyped.

Dr. Fletcher described KOMP2's implementation process. Each center operated independently, and in parallel, generated and phenotyped its own mice using approved standard operating procedures. The initial step was to check for viability at weaning. Because the breeding plan uses heterozygous intercrosses of the knockouts, the expectation is that 25 percent of the offspring would be homozygous nulls. Among breeding pairs, an inability to produce such offspring likely reflected embryonic lethality. The animals then were shifted into the embryo phenotyping assays and imaged at the various embryonic stages to determine the point of lethality. Results showed that 25 percent of genes have embryonic lethality. Two types of lethality were observed. Cell autonomous lethals occur early, prior to gastrulation, and overlap cell culture lethals. Developmental lethals present later in gestation and are unique genes that would not be identified as lethal in cell culture. Ten percent of genes were subviable, with animals surviving birth before presenting lethality sometime before weaning. Subviability presents as a variable phenotype, and sometimes only a tiny percentage of these animals survive. Next, fertility was checked. In a comprehensive pipeline, assays across many biophysical domains were completed between 9 and 16 weeks. At 16 weeks, mice were euthanized, necropsies and clinical chemistry were completed, and data were uploaded into the database. Quality-checked data were entered into the core data archive, analyzed for outliers, annotated with phenotype ontology terms, and visualization was completed.

Results can be accessed via the IMPC portal and are presented as a combination of easy-to-scan tables and graphs, allowing the user to then find more detailed information on the gene page, including homozygous subviability. An ideogram of the test results on the front page provides a summary overview of the phenotypes by using colored images that represent various biological domains. An interactive Manhattan plot provides an overview of quantitative data and links to each assay result. In addition, application programming interfaces are available for programmatic data access.

Dr. Fletcher highlighted the benefits of the program. KOMP provides the research community a resource to explore the dark protein coding genes. Of the 475 protein kinase genes in humans, only 30 to 40 have been thoroughly investigated. A structural impediment to studying poorly annotated and poorly studied

genes is that without preliminary data or a hypothesis, securing R01 funding to investigate this area of the genome is a challenge. KOMP2, in collaboration with the Common Fund Illuminating the Druggable Genome project, has knockout coverage that extends across the entire protein kinase family. Importantly, KOMP2 found that dark genes are a rich area of phenotypes. For example, in the deafness assays, the team evaluated 2,000 cell lines and found 65 knockouts that were deaf; 51 were novel associations. Even among the 14 known human deafness genes, two previously had no mouse knockout. KOMP2 is providing necessary reagents and insights into novel genotype/phenotype correlations to the research community.

Another benefit of KOMP2 is that it provides insight into biology in terms of sex as a biological variable and pleiotropy (i.e., pleiotropic gene effects). Cohorts of knockouts included seven male and seven female animals. The results showed that 17 percent of knockout phenotypes show differences between male and female animals. In addition, pleiotropy was commonly observed, in fact, 77% of genes are associated with 2 or more phenotypes.

Overall, KOMP2 has identified multiple phenotypes in nearly all knockouts and this finding presents a major shift in thinking about gene function that can be observed only through this comprehensive and broad examination of knockout mice.

As a further benefit, Dr. Fletcher highlighted KOMP2's translational aspects in relation to human disease. Although the KOMP2 lines are not generated as specific human disease alleles, good correlation to some human diseases was observed. Examples include a candidate gene for congenital diaphragmatic hernia, Cdc42bpbin, investigated in collaboration with the Common Fund Gabriella Miller Kids First Pediatric Research Program.

Finally, Dr. Fletcher explained that the KOMP Working Group tracks impact and uptake by the community of the KOMP2 resources and resulting publications. The research community is purchasing and using KOMP/IMPC mice; since inception, 4,442 publications have been generated using IMPC reagents, with 500 to 600 published annually by R01 researchers.

KOMP2/IMPC is aiming to phenotype 13,000 genes by completing 3,000 additional knockouts in the next 5 years. Approximately 1,200 will be NIH-wide funded. The expectation is to issue awards in summer 2022 for a final 5-year project, which will further expand the comprehensive, genome-wide collection of knockouts generated by this program.

Discussion Highlights

- When asked about strain differences and phenotype comparisons to previously published strains, Dr. Fletcher first noted KOMP2's goal of comprehensively phenotyping 13,000 mice. This effort complements the 4,000 already contributed by the broader community, which—when combined—now provides coverage of mouse–human one-to-one homolog. He explained that a vast amount of community phenotyping has been performed on mixed-strain backgrounds. KOMP2 investigators conducted analyses to determine whether any extra phenotypes were being added. He anticipates focusing on side-by-side comparisons over the next 5 years of the project. Dr. Anderson pointed out that KOMP2 has established a baseline and background for rigor and reproducibility for studies in mouse models.

- In response to a question about integrating data in KOMP2 with other databases, Dr. Fletcher pointed out that KOMP2 annotation terms, not raw data, are exported to Mouse Genome Informatics, then imported to the gene page. Information on the alleles can be viewed and compared in that database. The Mouse Genome Database data have not been incorporated into

the KOMP database, but the publications containing phenotyping using the KOMP2 resources have been uploaded.

VIII. CLOSING REMARKS

Dr. Anderson thanked the Council members and speakers for their contributions at this meeting. He reminded the members that the next Council meeting is scheduled for September 8–9, 2022, and also will be virtual.

IX. ADJOURNMENT

Dr. Anderson adjourned the meeting at 2:56 p.m. on May 20, 2022.

X. CERTIFICATION

I hereby certify that, to the best of my knowledge, the foregoing summary minutes are accurate and complete.

James M. Anderson -S Digitally signed by James M. Anderson -S Date: 2022.06.28 16:50:40 -04'00' — 6/28/22

James M. Anderson, M.D., Ph.D.
Chair, NIH Council of Councils
Director, DPCPSI, OD, NIH

Date

Robin I. Kawazoe -S Digitally signed by Robin I. Kawazoe -S Date: 2022.06.28 16:38:04 -04'00' — 6/28/22

Robin I. Kawazoe
Executive Secretary, NIH Council of Councils
Deputy Director, DPCPSI, OD, NIH

Date