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

Council of Councils Meeting
September 11, 2020

Meeting Minutes

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. The virtual meeting began at 10:15 a.m. on Friday, September 11, 2020. 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: Franziska B. Grieder, D.V.M., Ph.D., Director, Office of Research Infrastructure Programs (ORIP), 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 MIT and Harvard, Cambridge, MA
Jeffrey R. Botkin, M.D., M.P.H., The University of Utah, Salt Lake City, UT
Linda Chang, M.D., FAAN, FANA, 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
Andrew P. Feinberg, M.D., M.P.H., Johns Hopkins University, Baltimore, MD
Patricia D. Hurn, Ph.D., R.N., University of Michigan, Ann Arbor, MI
Kevin B. Johnson, M.D., M.S., Vanderbilt University Medical Center, Nashville, TN
R. Paul Johnson, M.D., Emory University School of Medicine, Atlanta, GA
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
Michael D. Lairmore, D.V.M., Ph.D., University of California, Davis, Davis, CA
Jian-Dong Li, M.D., Ph.D., Georgia State University, Atlanta, GA
Edith P. Mitchell, M.D., FACP, Thomas Jefferson University, Philadelphia, PA
Charles P. Mouton, M.D., M.S., 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., Blue Cross of Idaho, Meridian, ID
Susan Sanchez, Ph.D., The University of Georgia, Athens, GA

Jean E. Schaffer, M.D., Joslin Diabetes Center, Boston, MA
Scout, Ph.D., National LGBT Cancer Network, Pawtucket, RI
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

Council Members Absent

Rick Horwitz, Ph.D., Allen Institute for Cell Science, 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, DPCPSI
Maureen M. Goodenow, Ph.D., Director, Office of AIDS Research
Susan K. Gregurick, Ph.D., Director, Office of Data Science Strategy (ODSS), DPCPSI
David M. Murray, Ph.D., Director, Office of Disease Prevention, DPCPSI
Karen L. Parker, Ph.D., M.S.W., Director, Sexual & Gender Minority Research Office, DPCPSI
William T. Riley, Ph.D., Director, Office of Behavioral and Social Sciences Research (OBSSR), DPCPSI
Elizabeth L. Wilder, Ph.D., Director, Office of Strategic Coordination, DPCPSI
David R. Wilson, Ph.D., Director, Tribal Health Research Office, DPCPSI

3. *Ex Officio* Members Absent

Lawrence A. Tabak, D.D.S., Ph.D., Principal Deputy Director, NIH

4. Presenters

Francis S. Collins, M.D., Ph.D., Director, NIH
James Coulombe, Ph.D., Chief, Developmental Biology and Structural Variation Branch, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD)
Charles R. Dearolf, Ph.D., Director, Program Development and Support, Office of Intramural Research (OIR), NIH
Susan K. Gregurick, Ph.D., Director, ODSS, DPCPSI
Richard Hodes, M.D., Director, National Institute on Aging (NIA)
Kevin B. Johnson, M.D., M.S., Council of Councils Member
William T. Riley, Ph.D., Director, OBSSR, DPCPSI
Griffin Rodgers, M.D., MACP, Director, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
Norman Sharpless, M.D., Director, National Cancer Institute (NCI)

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

Franziska B. Grieder, D.V.M., Ph.D., 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* notices for the meeting, which were published on August 17 and September 8, 2020.

- Minutes from the May 15, 2020, meeting are posted on the DPCPSI website. The minutes from this meeting also will be posted there.

C. Future Meeting Dates

Future Council meetings are scheduled to be held virtually January 28–29 and either in person or virtually May 20–21 and September 17, 2021. Although these dates are reserved, the duration of each meeting is not yet confirmed.

II. UPDATES FROM THE NIH

Francis S. Collins, M.D., Ph.D., Director of the NIH, expressed gratitude to NIH staff for their hard work in continuing to support the NIH mission during the pandemic. He commented on the recruitment process for five new Institute and Center (IC) directors: Dr. Michael Chiang, who will become Director of the National Eye Institute; Dr. Lindsey Criswell, who will become Director of the National Institute of Arthritis and Musculoskeletal and Skin Diseases; Dr. Rena D'Souza, who will become Director of the National Institute of Dental and Craniofacial Research; Dr. Rick Woychik, who will become Director of the National Institute of Environmental Health Sciences (NIEHS) and National Toxicology Program; and Dr. Shannon Zenk, who will become Director of the National Institute for Nursing Research. He also noted that Dr. Hannah Valantine, the Chief Officer for Scientific Workforce Diversity, is retiring.

Dr. Collins reviewed the trajectory of NIH's budget, which has been increasing since 2015 and received additional funding in response to SARS-CoV-2, and predicted that a continuing resolution would occur after the November election. The additional funding to combat the virus has allowed the NIH to make progress on its research. Dr. Collins outlined the supplements to specific ICs in response to the pandemic and remarked on the unprecedented response by researchers and scientists. He explained that the genetic sequence of the virus was released by China on January 10, 2020. The first vaccine was designed by NIH's Vaccine Research Center in the next 48 hours, based on a technologically advanced and rapid approach using mRNA, and this vaccine now is in phase 3 trials. Dr. Collins commented on the efforts to respond to the pandemic requiring collaboration among many levels of the government, including the Coronavirus Task Force. The NIH now has developed strict public health procedures for its own campus, which have prevented infections from occurring on campus, and NIH staff are working to determine when returning researchers to campus will be safe.

Dr. Collins commented on research efforts at the NIH related to SARS-CoV-2. Accelerating COVID19 Therapeutic Interventions and Vaccines (ACTIV), announced on April 17, is a public–private partnership dedicated to establishing a collaborative framework for prioritizing therapeutic candidates and accelerating vaccine evaluation; accelerating clinical trials of promising agents and leveraging existing clinical trial networks while maintaining rigorous safety standards; and coordinating regulatory processes and leveraging assets among all partners. The partnership includes industry leaders, government leaders, nonprofits, and program management entities. Five ACTIV therapeutics have been prioritized for clinical trials, and ACTIV also works on vaccine development. Operation Warp Speed—the implementation arm to ACTIV's design arm—has invested in six vaccine candidates. Three are in phase 3 trials; two of these are based on mRNA and one is based on a viral vector.

The Rapid Acceleration of Diagnostics (RADx) for COVID-19 program is dedicated to improving testing for SARS-CoV-2, which Dr. Collins stressed is an important supplement to vaccine development in ending the pandemic. RADx has four components. RADx-tech is a competitive, three-phase challenge to identify best candidates for at-home or point-of-care tests; its approaches include funding early innovative diagnostic technologies, advancing late-stage diagnostic technologies to expand testing infrastructure; identifying effective testing implementation strategies for underserved populations; and working closely with other government agencies. RADx phase 2 awards include both point-of-care and laboratory-based tests. RADx Advanced Technology Platforms is a rapid scaleup of advanced technologies to enhance and validate throughput, intended to create ultrahigh-throughput machines and facilities. RADx Underserved Populations is an interlinked community-based demonstration project focused on implementation strategies to enable and enhance testing in vulnerable populations. RADx Radical is intended to develop novel, nontraditional approaches and applications.

Dr. Collins also commented on NCI's involvement with COVID-19 serology, explaining that the NCI runs a serology laboratory for human papillomavirus research that was able to pivot quickly to COVID-19 serology research. The U.S. Food and Drug Administration (FDA) has been able to use this laboratory to conduct performance testing for serology. Dr. Collins pointed out that many experts at the NIH and other health agencies now are being called upon to share their expertise with government leaders and the public, stressing the importance of communication in this public health emergency. He commended Dr. Anthony Fauci, Director of the National Institute for Allergy and Infectious Diseases, for his frank and broad communications with politicians and the public.

Discussion Highlights

- When asked how adequate safety of vaccines will be determined, Dr. Collins explained that time-limited adverse effects—such as a sore arm or a temporary fever—probably will not affect whether a vaccine is approved. The serious adverse event that occurred recently in one of the vaccine trials has not yet been proven to be related to the vaccine, but if it is, such events would be seriously considered when determining whether the vaccine should be approved. Dr. Collins noted that much of the consideration of safety will be within the purview of the Data and Safety Monitoring Board. He added that FDA currently has an efficacy standard of 50 percent, and updated safety guidance may be released soon.
- Dr. Collins commented on the importance of ensuring that enrollment for vaccine clinical trials is diverse, including those from sexual and gender minority populations. He noted that, for many minority populations, the community must be engaged to encourage enrollment.
- Dr. Collins suggested that the urgency of the pandemic has removed some silos between disciplines, and ensuring that they are not reconstructed in the future will be important. The clinical trials network and data sharing systems have become more nationally integrated, and

Dr. Collins hoped that additional challenges to data sharing can be reduced as the coordinated efforts to address the pandemic continue.

- When asked how the NIH will adjudicate the next step in the complicated vaccine development process, Dr. Collins agreed that the outcomes of the trials are uncertain. Operation Warp Speed includes funding for the manufacture of successful vaccines on the idea that all the trials have the potential for success. Distribution issues are likely, and some populations may respond differently to the vaccine. The Centers for Disease Control and Prevention (CDC) is leading the distribution effort, and the National Academy of Sciences developed recommendations for prioritization. The CDC Advisory Committee is working to determine how to apply CDC's framework in a real-world situation, and an expert in supply chains from the U.S. Army is involved in the effort.
- In response to a question about antiretroviral therapy and prophylaxis, particularly in communities resistant to receiving a vaccine, Dr. Collins commented on concerning polls showing a resistance to a vaccine, whether because of a skepticism about the rushed development or other reasons. He emphasized the importance of preparing prophylaxis. Trials are in progress with monoclonal antibodies for high-risk situations, such as nursing homes. Dr. Collins hoped another generation of antivirals specific to COVID-19 would be developed now that remdesivir has been proven to be successful for people already in the hospital. He emphasized the need to continue working on many types of treatments, particularly because this will not be the last coronavirus pandemic and better preparation is necessary.

III. REVIEW, DISCUSSION AND VOTE—REVISED COUNCIL OF COUNCILS OPERATING PROCEDURES

Dr. Anderson explained the need to update the Council procedures to include concept clearance for the Lasker Clinical Research Scholars program. The procedures also must be updated to clarify that the sum of the administrative supplements for projects addressing coronavirus research, previously discussed at the May meeting, is allowed by the NIH to exceed the cost of the parent award, if appropriately justified.

Discussion Highlights

- Dr. Anderson clarified that some ICs may have their own policies against supplements' exceeding parent awards; most supplements likely would not reach this threshold, but Dr. Anderson stressed the necessity of this change for rapid investment in infrastructure to respond to the pandemic.
- Council members suggested updates at future meetings on how often administrative supplements have exceeded the amounts of the parent awards.

Vote

A motion to approve the proposed modifications to approve the changes to include the Lasker Award and to Section IV, Part C of the Council of Councils Operating Procedures was forwarded and seconded. The motion passed with no abstentions.

IV. ODSS CONCEPT CLEARANCE: SMART HEALTH AND BIOMEDICAL RESEARCH IN THE ERA OF ARTIFICIAL INTELLIGENCE AND ADVANCED DATA SCIENCE

William Riley, Ph.D., the Director of OBSSR, explained that the reissue of the Smart Health and Biomedical Research in the Era of Artificial Intelligence and Advanced Data Science initiative, a joint initiative between the NIH and the National Science Foundation (NSF), aims to accelerate the development and integration of innovative computer and information science and engineering approaches

to support the transformation of health and medicine. Dr. Riley outlined the background of the first two rounds of this initiative, in 2013 and 2018. Previously, sensor technologies at NSF were developed but not matured to a point where they could be used in clinical health research, so this project was developed to bridge the translational gap. Projects must be integrative, making contributions to at least two disciplines, and must address a key health issue. Each project is expected to include several trainees, and the budget is $300,000 in total costs per year for 4 years. Initially, sensor technologies were most transformative for social and behavioral scientists because they allowed for the assessment and intervention of behaviors and the social context in which those behaviors occurred. ODSS did not exist at that time, but now it can assume the lead from OBSSR.

NSF takes the lead in reviewing applications, and the NIH addresses and modifies those reviews to ensure that the applications will be usable in clinical research. Currently, 10 NIH ICs participate, with 35 NSF applications and 32 NIH applications funded. Almost 100 NIH principal investigators have been funded, 65 percent of whom are early-stage or new investigators. Dr. Riley reviewed several example studies in the program, including a smartphone app using sonar to track chest and abdomen movements for people with sleep apnea, the modification of this system to detect opioid overdose, and the use of similar ultrasound technology to detect childhood ear infections.

Susan Gregurick, Ph.D., the Director of the ODSS, explained that artificial intelligence (AI) can be useful in many ways, but many of these AI methods require new and innovative crosscutting technologies. Key areas of the Smart and Connected Health Program align with goals in NIH's strategic plan for data science and include supporting tools for interoperable and federated digital platforms, funding new approaches in artificial intelligence to transform data science, investing in continued success of multimodal sensors, developing new approaches to support individuals in participating in their own health, and providing methods to improve the interpretation of complex medical images. This program focuses on increasing collaboration among computer scientists, engineers, and biomedical researchers to provide interdisciplinary research and training. Dr. Gregurick anticipated that this program would expand the network of researchers engaging with the NIH ICs, almost all of which have signed on to participate in the reissue, providing significant outreach to the broader NIH community.

Discussion Highlights

- Dr. Anderson commented that this is an example of the kind of metrics the Council requested to demonstrate proof of success for concept renewals. Many other ICs are interested in collecting research examples from their programs following this model, and other metrics will be identified when appropriate.
- The discussants, Drs. Sachin Kheterpal and Paul Kenny, provided their comments. Dr. Kheterpal expressed his strong support but encouraged transparency about members of review panels to ensure that engineers and clinicians are equally represented, and the projects are fully integrated between NSF and the NIH. Dr. Kheterpal also recommended clear encouragement of types of unstructured data beyond medical images. Dr. Gregurick planned to take Dr. Kheterpal's suggestion of a higher budget under advisement.
- In response to a question about directing project support to Small Business Innovation Research (SBIR) or Small Business Technology Transfer (STTR) programs, Dr. Riley explained that some of the program's projects already have been funded by SBIR grants and that ICs will continue to use specific funds set aside for SBIR and STTR programs.
- When asked whether interoperability could be a requirement, Dr. Gregurick explained that interoperability is an active area of research and difficult to achieve, so requiring it would be difficult, but the program tries to encourage interoperability as much as possible.

- Dr. Gregurick pointed out that NIH will publish a NOTICE and each participating IC will include language with its own area of interest, to direct applicants specifically to an IC's research mission.

Vote

A motion to approve the Smart Health and Biomedical Research in the Era of Artificial Intelligence and Advanced Data Science concept with the consideration of suggestions made during the discussion was forwarded and seconded. The motion passed with no abstentions.

V. COMMON FUND CONCEPT CLEARANCE: NUTRITION FOR PRECISION HEALTH, POWERED BY THE *ALL OF US* RESEARCH PROGRAM

Griffin Rodgers, M.D., the Director of the NIDDK, presented a new concept for a nutrition study that would leverage the *All of Us* network to provide an evidence base for individualized nutrition recommendations. The proposal was developed by staff across 17 ICs and will be co-chaired by the Nutrition Research Task Force co-chairs and managed by program leaders at the Common Fund, NIDDK, and *All of Us* program. Nutrition is integral to many processes within the missions of most ICs, and diseases linked to poor diet are the most frequent preventable cause of death, accounting for one in five deaths. The current dietary recommendations are one-size-fits-all, despite evidence that nutritional needs vary widely among individuals. Nutritional status is a complex interplay of dietary intake, microbiome ecology, biology of metabolism, genetics, and environmental exposures. More targeted, precision interventions are the future of nutrition, but an evidence base must be developed.

Because nutrition is a crosscutting issue, Dr. Rodgers suggested that this program fits well within the Common Fund. He outlined a study conducted in Israel that shows the differences among individuals on the same diet, noting that 90 percent of the predictive model came from microbiome compositional data. Measurements using -omics technology are very costly, leading investigators to study only selective components of their models rather than perform a comprehensive assessment. Additionally, models are not predictive across populations and diets—results of the study in Israel were not predictive for studies in populations in the midwestern United States. *All of Us* already has a large cohort with a wealth of data, and the program is committed to diversity and inclusion. The program has built extensive infrastructure, can leverage genomic and electronic health record data, and uses data sharing policies that ensure any qualified researcher can access the data. The Nutrition for Precision Health program expects to be able to return additional value to *All of Us* by introducing new types of data not previously available to collect, such as microbiome or detailed dietary intake data. Information returned to participants could include dietary recommendations tailored to an individual or a particular subgroup.

Dr. Rodgers outlined the proposed structure of the program. During the first 5 years, researchers will recruit 10,000 participants for a modular, nested discovery science study, with continuous glucose monitoring, biological measure collection, and mixed-meal challenge tests for each participant. Researchers also will examine other social, community, and contextual factors that could be sources of individual variability in dietary responses, as well as conduct microbiome ecology and behavioral, physiological, proteomic, and metabolomic assessments. The program also will leverage *All of Us* data on genomics, electronic health records, and surveys. In Module 1, researchers will examine all participants' responses to their usual diet and obtain blood, urine, and stool samples for -omics analysis; in a smaller subset, researchers will examine responses to challenge diets in controlled feeding studies for either 1,500 free-living participants in Module 2 or 500 domiciled participants in Module 3. The extensive data collection and monitoring of these patients will lead to dietary recommendations tailored to population subgroups or individuals, and the program can involve tests of new smart device tools, data models, and artificial intelligence approaches to advance the field of precision nutrition. The primary deliverable of the first phase is the development of algorithms that will predict individual responses to diets. In the

second phase of the program, further studies will be conducted to validate these emerging predictive algorithms.

Five initiatives are proposed for this research, managed by a collaborative NIH work group with expertise in many relevant areas. The first initiative will establish data and study coordination through a new research coordinating center supplemented by two existing *All of Us* coordination awards. Clinical centers—in collaboration with *All of Us* health care participant organizations (HCPOs), which will recruit and enroll *All of Us* participants into the nutrition study—will perform the feeding studies and sample collections. Three data generation centers will perform metabolic phenotyping, microbiome analysis, and dietary assessment. An artificial intelligence, bioinformatics, and data modeling center will integrate data-driven and mechanistic approaches with mathematic and computational modeling to develop the intended algorithms that can predict biological responses. The existing *All of Us* biobank is proposed for receiving, processing, recording, and storing the bio samples and metadata collected at the clinical centers. Dr. Rodgers emphasized that this will be the first study to collaborate with *All of Us*, so a post-award planning group will be formed. The total budget for this project is $155,900, and the project will result in a landmark study that would showcase the initial investments in *All of Us* with many key accomplishments for moving nutrition science forward and fueling discovery science for many years.

Discussion Highlights

- The discussants, Drs. Anna Maria Siega-Riz and Maria Rosario Araneta, provided their comments. Dr. Siega-Riz emphasized the necessity for more nuanced nutritional understanding for individuals. Dr. Araneta supported the concept enthusiastically and commented on the potential to revolutionize diabetes management and prevention. She also suggested enrolling nonobese Asians at high risk for diabetes.
- When asked for additional information on the diversity of participants in *All of Us*, Dr. Rodgers explained that more than half of enrollees are part of racial and ethnic minority groups, and more than 80 percent of participants are part of groups traditionally underrepresented in biomedical research, including sexual and gender minorities and people often excluded because of income status, educational status, geography, access to care, and disabilities. Enrollment also is planned for individuals with chronic disease. Dr. Rodgers added that religious diversity also is included in that group.
- In response to a question about how long-term outcomes will be assessed when the concept proposals are for only 5 years, Dr. Rodgers clarified that *All of Us* plans to track participants for many years, so their electronic health records will be accessible for assessing long-term changes to their health status. This approach also will allow participants' data to be analyzed with artificial intelligence, which may identify unexpected connections.
- When asked about the plan to use newer nutritional assessments, Dr. Chris Lynch of the Office of Nutrition Research explained that, although new tools have not been validated, one goal of the project will be to improve dietary assessment measures in conjunction with some controlled feeding studies.
- Dr. Rodgers commented that meal timing for the first study will be made as consistent as possible and added that a 1-year planning period will optimize the protocols, site selection, and other factors. Dr. Lynch added that the organizing committee includes researchers from the NIEHS, so the exposome and food exposures will be considered, and samples of food wrappings will be collected.
- Council members suggested extending the studies into mouse research to gather data on factors difficult to study in humans. Dr. Rodgers responded that the stakeholders for this study want to

focus on human experimentation, but he co-chairs Molecular Transducers of Physical Activity Consortium (MoTrPAC), which has a component to study animal models and less accessible tissues.

- In response to a question about socioeconomic effects on diet, Dr. Rodgers clarified that the components of the project focusing on the exposome will include socioeconomic factors.
- When asked about outreach and educational studies, Dr. Rodgers commented that education is not part of this effort, but it is an important aspect of 2020-2030 Strategic Plan for NIH Nutrition Research.
- Dr. Josh Denny of *All of Us* explained that ways to enroll participants not currently in a healthcare provider organization.

Vote

A motion to approve the Nutrition for Precision Health Powered by *All of Us* Research Program concept was forwarded and seconded. The motion passed with no abstentions.

VI. 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 155 ORIP applications with requested first-year direct costs of $670,925,846.

VII. REPORT FROM THE SEQUENCE READ ARCHIVE DATA WORKING GROUP OF THE COUNCIL OF COUNCILS

Kevin Johnson, M.D., the Co-Chair of the Sequence Read Archive Data Working Group of the Council of Councils, explained that the Sequence Read Archive (SRA) is one of NIH's largest and most diverse data sets and represents the diversity of the genome throughout the tree of life. The SRA links diseases with genetic and epigenetic variations, bioinformatics, and evolutionary biology. The public part of the SRA, which is 8.8 petabytes, is composed of approximately 50 percent genomic information, 40 percent RNAseq information, and 10 percent prokaryotic and metagenomic information. The data and metadata also are used for bioinformatics methods development. The total size of the SRA currently is about 13.4 petabytes; because copies are on both the Google Cloud Platform and Amazon Web Services, the size of total data under consideration is 26.8 petabytes. The large size of the database provided the opportunity to test the NIH Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES) initiative. The SRA currently has more than 10 million records, each growing exponentially; more than 1.2 million visitors downloaded large pieces of the data set in 2019, with 20 percent of visits from cloud IP addresses. Data currently exist both on and off the cloud and are provided in a variety of file formats, which are standardized to the Extract, Transform, and Load (ETL) format; the ETL format currently is the only format in which researchers can access the data.

[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.

The SRA is expected to grow exponentially with the amount of data expected from both projects already funded and new data types; the existence of both original and ETL data formats increases the amount of data stored in the SRA. The largest single component of the ETL data are base quality score (BQS) data, which are representations of the probability of error at each base call. Many file types have one BQS per letter of sequence, resulting in a large amount of data that are difficult to compress. The current growth rate of the SRA is projected to quickly exceed the NIH budget for storage and maintenance, and the SRA Data Working Group was charged with providing recommendations to the Council on key factors for storing and managing the SRA data, as well as evaluating and identifying solutions. The Working Group also was charged with evaluating the use of BQS data and various format compression strategies.

The Working Group delivered its final report in September 2020 and recommended developing a new model for SRA data storage and retrieval in the cloud, working with the National Center for Biotechnology Information (NCBI) and the NIH to communicate that model, and continuing research to inform changes to the model over time. The chosen model allows the combination of hot storage (i.e., data immediately available upon request) and cold storage (i.e., data that could take 48 hours to make available for analysis). The most active ETL data will be available in hot storage, and the half of SRA data less frequently accessed will be placed in cold storage. All data in their original formats with BQS data will be retained in cold storage, and two versions of ETL data—one with BQS and one without—will be maintained. The Working Group recommends that the NCBI monitor data usage and determine the appropriate cloud storage location for each data set depending on usage data. Users should be informed as they are retrieving data where the data are stored and the costs associated with retrieving the data. The Working Group also recommends that the NCBI limit the amount of data users can request "thawed" from cold to hot storage without approval, which would prevent the accidental overuse of NIH resources. These limits should be defined by a sliding time interval window to allow users to access data in a timeframe that fits their research needs. These recommendations require algorithmic management of the data set, but the NCBI already is engaged in this task and will be able to monitor usage as needed.

Dr. Johnson reviewed an analysis of how data currently are accessed. Between May and October 2019, 50 percent of unique data records were accessed, implying that keeping less than 50 percent of data in hot storage could potentially increase the costs because of the high degree of thaw required. The Working Group focused much discussion on communication of the new storage model, because many users remain unfamiliar with aspects of cloud computing and storage. The Working Group recommended developing communication materials focused on non-cloud-based sources of data; if the NIH makes BQS data available only in the cloud, associated equity challenges must be considered and addressed. Cost models should be clearly communicated, including specific information on storage and computing costs and how responsibility for the costs is distributed between the user and the NIH. Potential users should be provided with education on cloud usage, hot and cold data access, and compute time monitoring.

Dr. Johnson emphasized that the SRA remains in its early stages, so continued research to inform changes to the model is necessary. The Working Group recommends that the NIH monitor the costs of the current model to adjust over time based on actual costs and determine whether different strategies are needed for different cloud service providers, which currently have different costs. The Working Group also recommends that the NIH monitor use and adjust policies to ensure that no subset of researchers bears an undue burden as a consequence of data format. The NIH should consider intramural and extramural support for efforts that explore new compression strategies, optimize code, and increase efficiencies in the cloud to reduce computing costs—academic, industry, and fee-for-service communities should be engaged to address software optimization challenges.

Dr. Gregurick reported that, in line with the SRA Data Working Group recommendations, the NIH issued a request for information (RFI) to solicit community input and received more than 70 responses from U.S. and international participants. The majority of responders indicated that BQS is important and needs

additional pipelines, but approximately half of participants indicated that file formats without BQS could be useful. A small percentage of respondents compute in the cloud, but the majority download data for computing. A majority of participants valued access to original-format data. Dr. Gregurick noted that further analysis of the RFI responses is ongoing.

Discussion Highlights

- Dr. Anderson reminded Council members that they would vote on whether to accept the Working Group's report; recommendations would not be edited, but additional comments would be conveyed to Dr. Collins in a letter for his consideration when evaluating the recommendations.
- When asked why BQS remains relevant, Dr. Johnson noted that more than 50 percent of RFI responses preferred to retain BQS data, and the Working Group also strongly preferred this. He explained that the size of BQS could be addressed with newer techniques, and the preference for maintaining BQS could change over time, but the community's continued preference for BQS is clear.
- In response to a question about the affordability of the SRA for smaller laboratories, Dr. Johnson explained that the NCBI maintains some ability to provide data for free at a slightly slower pace.
- Dr. Johnson acknowledged concerns about data security and commented that the Working Group hopes the industry is able to provide appropriate security. After discussion of other models, this short-term strategy was determined to be the best option for the NIH. Dr. Gregurick noted that the NIH and ODSS pay close attention to the equity of cloud computing.

Vote

A motion to approve the report of the SRA Data Working Group was forwarded and seconded. The motion passed with no abstentions and one nay.

Dr. Gregurick requested that the SRA Working Group continue work with a new charge to focus on the evaluation of the SRA as a resource and related issues, which could include the analysis and evaluation of strategies for or changes to SRA data storage, management, and access; impact for the biomedical research community; recommendations for data retention, or data usage, data models that maintain sustainable costs for the NIH and preserve community access; and vision for future needs or opportunities, including sustaining the SRA as a community resource. Priorities for 2021 will be to examine the data related to the scientific impact of the SRA, value to the community, access, cost, usage, and other areas that may inform this evaluation. Council members offered no comments or questions on this charge.

Vote

A motion to approve the revised charge to the SRA Data Working Group was forwarded and seconded. The motion passed with no abstentions.

VIII. NIH LASKER CLINICAL RESEARCH SCHOLARS PROGRAM

Charles Dearolf, Ph.D., the Director of Program Development and Support at the OIR, explained that the NIH Lasker Clinical Research Scholars Program provides strong support to a limited number of early-stage researchers. The NIH wants to encourage clinically trained investigators to pursue careers in biomedical research. This unique partnership involves the intramural and extramural research programs at the NIH and the Albert and Mary Lasker Foundation. The participating ICs provide funding and support, and the Lasker Foundation provides other benefits and networking opportunities. Candidates apply by submitting a proposal for an R01-like project and letters of recommendation. During the first phase, the

Intramural Research Program provides full support for the scholars, who are given independent tenure-track investigator positions for 5 years at the NIH. After this phase, investigators can either continue on this track or receive an R00 component with $500,000 in direct costs per year that can be used at an outside institution. Candidates are early-stage clinical researchers who are not already tenured and who have a clinical license to practice in the United States.

Benefits to scholars include protected time to conduct research as their primary goal with full funding from the intramural programs. Scholars have no service obligations or formal teaching responsibilities; some may be able to obtain adjunct appointments if they previously were assistant professors at outside medical centers. Benefits for the extramural component include significant funding for their research and documented success in the NIH peer-reviewed application process. Benefits provided by the Lasker Foundation include the opportunity to attend the annual Lasker Awards luncheon, as well as a breakfast before the luncheon with the opportunity to interact with the prior year's winners. The Foundation also sponsors several Lasker Lessons in Leadership events, and the scholars are highlighted in reports and on the website.

The program began in 2011 and has grown to include 33 past and current scholars selected from 121 applicants, including investigators from underrepresented minorities and slightly more women than men. Currently, 11 ICs have at least one scholar. Scholars are reviewed by the sponsoring ICs and Boards of Scientific Counselors, and all scholars are in good standing. Long-term outcomes show that the first scholars have received early tenure, taken on high-level administrative leadership roles, and become productive and well-respected researchers. Most of the past scholars have chosen to remain at the NIH, but one will be requesting the R00 extramural funds and will be taking a leadership position at an external institution.

Discussion Highlights

- Dr. Anderson asked for clarification on whether the program encourages participants to connect with the extramural community or remain on the intramural research path. Dr. Dearolf responded that both outcomes are considered successful and the program's main focus is ensuring that scholars become leaders in academic medicine.
- The discussants, Drs. Russell Van Gelder and Jean Schaffer, provided their comments. Dr. Van Gelder commented on the uniqueness of the program and expressed his support for the principle. He raised concerns about the new DPCPSI oversight, requesting additional information on outcomes, metrics of success, and formalized mentorship. He also questioned the balance of benefits provided by the Lasker Foundation compared to those provided by the NIH and requested more information on the budget and past progress. Dr. Dearolf explained that each IC provides funding for scholars as appropriate for its own budget and outlined the application and selection process. Oversight is provided by each IC's Board of Scientific Counselors. Each scholar has a mentoring committee, and a staff member at the OIR serves as a mentor to tenure-track investigators.
- Dr. Schaffer commented on the innovative nature and relative newness of the program and encouraged overall continuation. She asked whether the spectrum of clinical care at the Clinical Center affects the research chosen by the investigators. Dr. Dearolf explained that a new funding announcement is released each year, featuring potential research areas of interest suggested by participating ICs. Applicants are encouraged to contact the intramural program at the IC of interest to ensure that programmatic interest in their research is likely before applying. Dr. Dearolf clarified that, although the Clinical Center does not have an emergency room or walk-in service, the Clinical Center can provide support in recruiting patients for the protocols developed by the scholars.

- When asked how longer-term career awardees compare to those in other career mechanisms, Dr. Dearolf acknowledged that such an evaluation has not been conducted, but some scholars on other awards, such as K or R awards, have chosen to participate in the Lasker Program rather than keep those awards.
- Council members agreed that additional evaluations and metrics for success are needed. Dr. Anderson suggested that Council members vote to renew the R00 component for the current scholars at this meeting and schedule a presentation with additional metrics for a future Council meeting.

Vote

A motion to renew the Limited Competition Lasker Clinical Research Scholars Transition Award was forwarded and seconded. The motion passed with two abstentions.

IX. COMMON FUND CONCEPT CLEARANCE: CELLULAR SENESCENCE NETWORK

Norman Sharpless, M.D., the Director of the NCI, and Richard Hodes, M.D., the Director of the NIA, presented on the concept for the Cellular Senescence Network, which would address an emerging scientific area of trans-NIH interest. Dr. Sharpless explained that senescent cells form in response to cellular damage, serving as an important anticancer mechanism but simultaneously inducing cellular aging in a variety of tissues and organs. The phenotype of cellular senescence has been recognized since the 1950s, but in recent years the wide variety of kinds of cellular damage that can induce this phenotype has become clear. The major regulators of senescence are tumor-suppressor genes, and senescence contributes to health through a number of mechanisms, including suppressing cancer, healing wounds, and limiting atherosclerotic plaques. However, emerging data suggest that senescent cells also induce pathology in a variety of tissues and contribute negatively to aging and pathogenesis in such ways as secreting pro-inflammatory cytokines and accumulating in places where they can cause anatomic problems. Dr. Sharpless suggested that many scientists believe senescent cells can be removed from a tissue when they are not playing a vital function, a concept known as senolysis, advanced as a therapeutic approach in a variety of human age-associated disease conditions. He outlined several studies providing evidence that clearing senescent cells diminishes age-associated pathologies in a number of tissues. Trans-NIH interest in this topic is driven by the potential importance in many human phenotypes. Research challenges include the heterogeneity of the senescent phenotype and the wide variety of inducers of senescence and impact on tissues. Dr. Sharpless reiterated the need for this Common Fund program to address critical research gaps in understanding the biology and physiology of senescent cells and aging.

Dr. Hodes pointed out that although cellular senescence is an extremely active area of research, the manifestations are so different that infrastructure and an organizing principle are needed to engage the growing body of critical experimental data. To develop this concept, a RFI was released, and three think tanks identified several important areas of consensus. A multidimensional atlas should be created to categorize and characterize senescent cells and identify their heterogeneity by induction setting and cell type. A set of "gold standard" biomarkers should be identified, although Dr. Hodes pointed out the possibility that no single such standard exists for all senescent cells. Experimental and computational predictive models should be established to analyze the causal effects of various perturbations on senescence, and the imaging and visualization tools necessary to trace senescence at the whole-body and cellular levels should be developed. Dr. Hodes emphasized the importance of deploying perturbation tools and demonstrating the consequences of those perturbations on senescent cells *in vivo*, particularly when working toward translation.

The Cellular Senescence Program focuses on the formation of the Atlas of Cellular Senescence in four dimensions, including time. The atlas will consist of a searchable database to capture multiomic data from multiple tissues during both normal and pathologic processes, as well as capture the relationships among tissues and the relationship of senescent cells to the microenvironment of non-senescent cells. A taxonomy also will be developed to classify cellular senescence. This program will be highly transformative and catalytic because it will affect many areas of basic and translational science, and many ICs were involved in the generation of the initiative. The initiative includes six tissue-mapping centers, each with an administrative core and three functional research areas—a biospecimen collection unit; a data analysis and computational modeling core; and a molecular, cellular, and tissue analysis unit. Technology development projects also will be proposed. The initiative proposes requests for applications (RFAs) for both of the first 2 years to build on the studies established from the first RFA. At the core of the initiative is a Consortium Organization and Data Coordinating Center, which will serve as a hub for organizing the complex and extensive data and making these data interoperable, sustainable, transparent, and available to the research community. The total cost of the proposed initiative is $144 million. Dr. Hodes emphasized the breadth of the input used to develop this program and the breadth of research the program could inspire.

Discussion Highlights

- The discussants, Drs. Schaffer and Van Gelder, provided their comments. Both expressed support for the program but noted concerns about whether technologies should be further developed and whether the full spectrum of disease states would be represented in the atlas.
- When asked about the balance between studying normal and diseased senescence, Dr. Sharpless reiterated that the phenotype is very heterogeneous and emphasized that a detailed understanding of what the process means in different tissues and different types of cells in the same tissues is needed. He pointed out that although many studies have been conducted in individual tissues and organs, the proposed atlas is necessary to develop a thorough enough understanding to predict human biology. He added that the conflicting roles of senescence—its contribution to aging and its cancer prevention function—mean that goal-oriented research on the process is lacking, making the NIH a good fit for conducting basic research.
- Dr. Hodes commented that the technology to develop the atlas already exists, and the initial studies will lead to further technological development and research discovery.
- In response to a question about animal models, Dr. Hodes agreed that although human tissues are the main emphasis, the program should be flexible to developments that could be explored better with animal models. Dr. Van Gelder suggested consideration of a parallel, smaller-scope atlas for mouse or other small-animal models.
- When asked about the diversity of the samples within the atlas, Dr. Hodes emphasized that funding decisions will include consideration of the diversity of tissue collections. He clarified that the approach to ensuring adequate breadth and diversity will be to solicit initial applications that show expertise in at least two tissue areas. The balance of these areas and additional applications will need to be considered rigorously and with advice from investigators and external advisory groups. Dr. Sharpless added that the coordinating center will ensure that redundancies are minimized and the diversity of tissues and disease states studied is maximized. Leadership across the NIH also strongly supports this initiative.
- Dr. Hodes emphasized that he anticipates and encourages multi-investigator, multisite applications to increase diversity and reduce silos; communication will be a critical aspect of these studies. Dr. Sharpless added that the increase in virtual communication in response to the pandemic could increase collaboration among institutions.

Vote

A motion to approve the concept for the Cellular Senescence Network was forwarded and seconded. The motion passed with no abstentions.

X. COMMON FUND CONCEPT CLEARANCE: GABRIELLA MILLER KIDS FIRST PEDIATRIC RESEARCH PROGRAM: PLANS FOR FISCAL YEARS (FYs) 2022–2024

James Coulombe, Ph.D., the Chief of the Developmental Biology and Structural Variation Branch at the NICHD, presented on the Gabriella Miller Kids First Research Program for FYs 2022–2024, a Common Fund program that appropriates $12.6 million per year for pediatric research. Although the program was signed into law in 2014 and authorized for 10 years, Congress still must approve those funds yearly. The program uses collaborative research and data sharing with the aim of accelerating pediatric research and improving diagnostics and therapeutics for patients with childhood cancer and structural birth defects. Although these defects and cancers are leading causes of death in children, individual conditions are rare, which has been a challenge for collecting sufficient patient populations for genomic studies. A child with a birth defect has a higher risk of developing cancer, so studying these conditions together is important.

The program began with two main initiatives. The first was to identify a cohort of children with pediatric cancer and/or structural birth defects and collect high-quality whole-genome sequence data from them and their families. The second was to build a data resource to facilitate studies of these conditions. Forty cohorts representing a diversity of childhood cancers and structural birth defects have been or currently are being sequenced. In addition to the investigators who submitted these samples, more than 150 data access requests from the external research community have been submitted, and more requests are expected as additional data sets are released. The Kids First data resource is a cloud-based platform designed to empower and accelerate collaborative research, and it offers a data resource portal to help users find data to address their scientific questions and a cloud-based workspace where researchers can co-analyze multiple data sets across conditions or multiple NIH data efforts. To date, 18 publications have resulted from this research, and this number is expected to grow considerably over time. Kids First also has many collaborations, most recently aiming to enable researchers to combine and cross-analyze data generated by a variety of programs and pioneer interoperability across programs.

Three initiatives are proposed for the final 3 years of the program to build on the infrastructure, data sets, and expertise of the community to strengthen data sharing and collaborative discovery. The first initiative will continue adding data to the resource and expanding the data types included. The second initiative is to continue developing and improving the data resource, and the third initiative is to engage the expertise of the research community to improve the utility of the data. This could include increasing the depth of phenotypic and clinical data available in the data resource. Ongoing activities will continue through other sources of funding, including pursuing interoperability and looking for new avenues of interoperability, as well as pursuing other NIH-wide collaborations. The data resource already is of high value to the pediatric research community and plays a central role in establishing interoperability across NIH data resources.

Discussion Highlights

- The discussants, Drs. Richard Klugman and Jeffrey Botkin, provided their comments. Both supported the program strongly and looked forward to the future of the program.
- When asked whether the 150 data access requests were considered a significant number, Dr. Coulombe explained that the data access requests are for the relatively few data sets available for public access. The number of available data sets is the result of the long time required to

prepare and release the data; he reiterated that additional cohorts will be released for access soon and the number of requests is expected to grow rapidly.

- In response to a question about data richness, Dr. Coulombe explained that the sequencing is funded through an unusual mechanism that provides the investigator supplying the samples with access to the sequencing center services but no funding. The third proposed initiative would supply small grants to enable investigators to add richer data than might be available without funding, including phenotypic, clinical, and environmental data that are key to making the genomic data thoroughly useful.
- When asked about collaboration with CDC, Dr. Coulombe explained that although CDC has collected many samples, the consents for these samples do not allow genomic data sharing.
- When asked about potential concurrence with work conducted in *All of Us*, Dr. Coulombe clarified that the consent processes for *All of Us* cause difficulties in sharing genomic data. He emphasized that the ultimate goal is to build a data platform tailored to pediatric research as part of a federation of data platforms enabling access and cross-analysis of NIH-supported data sets. The Gabriella Miller Kids First program is making considerable efforts toward this kind of interoperability and it is hoped that *All of Us* data eventually could be incorporated. He added that the program is open to collaboration with other programs that could provide additional data with appropriate consents.
- Council members commented on the importance of this program and the value of the resource. Dr. Coulombe commented on the lengthy process of preparing the data for sharing and hoped that the NIH could improve those rates eventually, pointing out that Kids First has taken several initiatives to speed up various parts of the process.
- In response to a question about community engagement strategies, Dr. Coulombe explained that a public outreach component is part of the Data Resource Center and has been active in contacting patient advocacy groups. He noted that the birth defects community is fragmented, and because applications for Kids First sequencing services must be submitted through research investigator-initiated applications rather than by community organizations, these challenges have limited the ability to engage with communities.

Vote

A motion to approve the concept clearance for the Gabriella Miller Kids First Pediatric Research Program plans for FYs 2022–2024 was forwarded and seconded. The motion passed with no abstentions.

XI. 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 January 2021 and also will be virtual. An additional meeting will be scheduled to gather Council members' input on the NIH Strategic Plan for FYs 2021–2025.

XII. ADJOURNMENT

Dr. Anderson adjourned the meeting at 4:46 p.m. on September 11, 2020.

XIII. CERTIFICATION

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

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

Franziska B. Grieder, D.V.M., Ph.D. Executive Secretary, NIH Council of Councils Director, ORIP, DPCPSI, OD, NIH	Date