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
January 27, 2017

Draft 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 meeting began at 8:15 a.m. on Friday, January 27, 2017, in Building 31, Conference Room 10, on the NIH Campus in Bethesda, Maryland.

Dr. Anderson welcomed members and noted that Drs. Eric Boerwinkle, Molly Carnes, Vivian Lee, Guillermina Lozano, and Keith Reimann were unable to attend the day's meeting. The meeting attendees are identified below. Dr. Anderson also announced that the founding director of the Tribal Health Research Office would be Dr. David Wilson.

Following introductions and announcements from Franziska B. Grieder, D.V.M., Ph.D., Executive Secretary for the NIH Council of Councils, Dr. Anderson 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., Food Allergy Research & Education, Inc., Washington, DC
Sharon Anderson, M.D., Oregon Health & Science University, Portland, OR
Cynthia Barnes-Boyd, Ph.D., R.N., FAAN, University of Illinois Hospital and Health Science System, Chicago, IL
Melissa Brown, M.D., M.N., M.B.A., Thomas Jefferson University, Philadelphia, PA
Jorge L. Contreras, J.D., The University of Utah, Salt Lake City, UT
Jonathan Epstein, M.D., Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA
Hakon Heimer, M.S., Schizophrenia Research Forum, Providence, RI
Patricia D. Hurn, Ph.D., R.N., University of Michigan, Ann Arbor, MI
Terry L. Jernigan, Ph.D., University of California, San Diego, La Jolla, CA
R. Paul Johnson, M.D., Emory University School of Medicine, Atlanta, GA
Sachin Kheterpal, M.D., M.B.A., University of Michigan Medical School, Ann Arbor, MI
Kimberly K. Leslie, M.D., University of Iowa Hospitals and Clinics, Iowa City, IA
Jian-Dong Li, M.D., Ph.D., Georgia State University, Atlanta, GA

Terry Magnuson, Ph.D., The University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC
Edith P. Mitchell, M.D., FACP, Thomas Jefferson University, Philadelphia, PA
Charles P. Mouton, M.D., M.S., Meharry Medical College, Nashville, TN
John Postlethwait, Ph.D., University of Oregon, Eugene, OR
Scout, Ph.D., The Torvus Group, Beverly Hills, CA
J. Leslie Winston, Ph.D., D.D.S., Procter & Gamble Global Oral Care, Mason, OH
Nsedu Obot Witherspoon, M.P.H., Children's Environmental Health Network, Washington, DC
Gail Yokote, M.S., University of California, Davis, Davis, CA

Council Members Absent
Eric Boerwinkle, Ph.D., The University of Texas Health Science Center at Houston, Houston, TX
Molly Carnes, M.D., M.S., University of Wisconsin–Madison, Madison, WI
Vivian S. Lee, M.D., Ph.D., M.B.A., The University of Utah, Salt Lake City, UT
Guillermina Lozano, Ph.D., The University of Texas MD Anderson Cancer Center, Houston, TX
Keith A. Reimann, D.V.M., University of Massachusetts Medical School, Boston, MA

2. Liaisons

Paul M. Coates, Ph.D., Director, Office of Dietary Supplements, DPCPSI
Maureen M. Goodenow, Ph.D., Director, Office of AIDS Research, DPCPSI
William Riley, Ph.D., Director, Office of Behavioral and Social Sciences Research, NIH
Elizabeth Spencer, R.N., representing **Janine Clayton, M.D.,** Director, Office of Research on Women's Health
Elizabeth L. Wilder, Ph.D., Director, Office of Strategic Coordination (OSC), DPCPSI

3. *Ex Officio* Member

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

4. Presenters

Patricia Flatley Brennan, R.N., Ph.D., Director, National Library of Medicine, Interim Associate Director for Data Science, NIH
Stephanie Courchesne-Schlink, Ph.D., Team Leader, OSC, DPCPSI
Michael A. Fischbach, Ph.D., Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco
Michael S. Lauer, M.D., Deputy Director, Office of Extramural Research, NIH
David M. Murray, Ph.D., Director, Office of Disease Prevention, DPCPSI
George M. Santangelo, Ph.D., Director, Office of Portfolio Analysis, DPCPSI
Lawrence A. Tabak, D.D.S., Ph.D., Principal Deputy Director, NIH
Hannah Valantine, M.D., Chief Officer for Scientific Workforce Diversity, NIH

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

Dr. Grieder 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 December 27, 2016.

- Minutes from the September 9, 2016, meeting are published on the DPCPSI website. The minutes from this meeting also will be published there.

C. Future Meeting Dates

The next Council meeting will be held on May 26, 2017. The final Council meeting of the year will be held on September 1, 2017.

II. TRACKING UTILITY OF COMMON FUND DATA SETS

Stephanie Courchesne-Schlink, Ph.D., Team Leader for Policy, Planning, Evaluation, and Communications in OSC, explained that the Common Fund is meant to enable a broad range of research in all biomedical disciplines, so it is important to understand if the data produced are having an impact. Utility can be assessed in a number of ways, such as reviewing Web analytics or registrations for sites that require them or tracking the number of times data are downloaded, both of which can be done with minimal time and difficulty. Some programs publish marker papers or suggest citation language, which can be tracked, but this metric relies on correct usage of the citation. Publications likely to cite the data can be reviewed manually, which produces high-quality results but is very labor-intensive.

Common Fund support is time-limited, so utility is put to the test when the support ends and other entities must commit to hosting the data sets. Most data sets have proven sufficiently useful to the biomedical community that they have attained ongoing support through such entities as the National Center for Biotechnology Information (NCBI), other Institutes and Centers, or academic institutions; data also may be uploaded to cloud servers to make it broadly available.

One example of a highly successful Common Fund data set is the Human Microbiome Project (HMP). The data were made available through the Data Analysis and Coordination Center, and the data sets generated by the HMP include more than 3,000 microbial genomes. Such metrics as session length and

number of users suggest that the data are being used frequently and deeply. Additionally, the data from the HMP were integrated into other databases; this happens frequently for Common Fund programs and encourages data accessibility, although tracking becomes more challenging.

Many other data sets that originated in the Common Fund have continued successfully. PubChem, a large repository of small-molecule information, is now hosted by the NCBI. It receives millions of page views per day and a million unique users every month, it is well integrated with many other Web-based scientific resources, and it continues to experience notable annual growth in users despite being long established. The Genotype-Tissue Expression Program, GTEx, generates data about the relationship between genetic variation and gene expression to determine the genetic underpinnings of complex diseases. GTEx data is being widely used and has led to insights into the genetic basis of several diseases. Support will soon end for the Common Fund's Epigenomics program, and robust usage of the Epigenome reference maps has illuminated previously unknown information about epigenetic effects on diseases other than cancer; ENCODE and the International Human Epigenome Consortium are hosting this program's data.

Discussion Highlights

- In response to a question about funding mechanisms that may influence access to data, Dr. Courchesne-Schlink explained that much of the data likely was being used in R01 research. A number of programs work to make the data as accessible as possible and usable for non-experts by providing analysis tools.

- Many challenges remain regarding the storage and tracking of data. For example, a systematic effort to prevent duplicate studies has not been developed, but OSC's working groups maintain an awareness of upcoming publications and have not yet identified significant duplications. The issue of duplicative studies is one that extends far beyond Common Fund datasets.

III. SMALL MOLECULES FROM THE HUMAN MICROBIOTA

Michael A. Fischbach, Ph.D., from the Department of Bioengineering and Therapeutic Sciences at the University of California, San Francisco, noted that the HMP had shaped his career path. He explained that microbes produce chemicals, known as natural products, which frequently are used in human medicine, but scientists do not yet understand why microbes devote significant percentages of their genomes to making these chemicals. Although microbiology focuses on studying microorganisms in laboratory cultures, wild microbes live in complex communities; Dr. Fischbach suggested that natural products may be crucial tools for interspecies interactions in microbe communities.

Natural products traditionally are discovered by scientists traveling to the far corners of the earth in search of new microbes. Dr. Fischbach's team uses microbial genomes to identify genes likely to produce new molecules, with the long-term goal of using raw genome sequences to predict the chemicals the microbes are capable of producing. This is a much more thorough way than laboratory cultures to illuminate all that a microbe can do.

Dr. Fischbach presented several examples of discoveries his laboratory has made using data from the HMP. His team ran a code to identify sequences used to make chemicals through available genomic data on the NCBI database soon after HMP data sets had been made public. The sequences in question were found, as expected, in data from soil bacteria; what was unexpected was their presence in samples from the human microbiome. While continuing to review HMP data, the researchers found large quantities of unknown biosynthetic genes present in abundance; nothing was known about what chemicals these genes might encode or how they could affect humans.

One new finding in this data was a set of thiopeptides, which are highly potent agents against gram-positive bacteria. Many other antibiotics bind sites on the ribosome, but these thiopeptides bind a site used by no other antibiotics in clinical use, meaning there would be no interaction between these molecules and any other products currently used. Once this unusual method was identified, researchers began to identify similar sequences in other isolates from the HMP, sometimes in large percentages, which may have implications for individual resistance to infections. Dr. Fischbach emphasized that the discovery of these new molecules is entirely enabled by the HMP—scientists have been searching the world for new molecules for years, but this is the first time they were able to search inside humans.

Dr. Fischbach's team also found a large family of biosynthetic gene clusters in the HMP data; this family is found in high levels in more than 90 percent of HMP subjects. The products of these genes are dipeptide aldehydes; they are chemically similar to a long-known protease inhibitor found in soil. Although the microbes are very common in the human gut, the target was shown to be a site that the researchers had not anticipated.

The average person does not know which chemicals are most abundant in the gut, yet many molecules are made exclusively by gut bacteria, and two thirds of these end up in the bloodstream, sometimes at concentrations comparable to medications. Unknown gut chemicals can have profound effects all over the body. Additionally, the chemicals present vary widely between individuals; two people could eat an identical meal and experience drastically different metabolic effects because the same amino acids can be turned into very different chemical products. For example, Dr. Fischbach's team recently discovered the biosynthetic genes for indole, the precursor of indoxyl sulfate; this chemical, which some microbes produce from tryptophan, is filtered out by healthy kidneys. In those with renal disease, indoxyl sulfate can reach high levels in the blood, which can lead to cardiac events.

The difficulty in isolating a single molecule has made studying these microbes challenging, but when the genome is known, researchers can create experimental situations with a single variable and learn the effects of turning a single chemical on and off by deleting or inserting the genes that encode its production. Data from the HMP make it possible to study these highly abundant microbes in isolation and truly learn their effects.

<u>Discussion Highlights</u>

- Dr. Fischbach confirmed that microbes often are resistant to the chemical they produce, so that they do not kill themselves by producing it; in almost every case, this is the source of resistance genes to a particular antibiotic. Asked to expand upon the role of natural products in interspecies interactions, Dr. Fischbach speculated that bacteria living in complex communities should be resistant to all their neighbors' products; the fact that they are not suggests the microbes prioritize another process, such as growth, over resistance. This leads to adjacent mosaic colonies that affect each other; some microbes can grow only in the presence of another microbe that produces a specific natural product.

- In response to a question about products that affect the host, Dr. Fischbach explained that his team began their studies with products that affect bacteria because antibiotics already are well understood. The peptide aldehyde discussed is more likely to be representative of the majority of molecules—host-facing and with targets that interact with the microbiome.

- The HMP's defined goal of establishing the normal microbiome baseline has allowed researchers to begin comparing the microbiotal changes associated with intestinal and metabolic diseases. The HMP has driven the substantial renewed interest in the field that led to the development of fecal transplants, which completely change the gut bacteria but so far are much safer than

previous interventions. Researchers now are exploring ways to design synthetic communities that mimic human fecal colonies.

- Data infrastructure efforts have been challenged by questions of the appropriate balance of benefiting from shared data and feeding back into the system. Dr. Fischbach supported generous data sharing, explaining that all molecules and natural products made by his team were incorporated into PubChem and other common resources. He noted that intellectual property questions surrounding natural products recently have become more complicated, but his team did protect the peptide aldehyde they discovered. Dr. Fischbach suggested that pharmaceutical companies likely would continue to invest in these projects even if greater protections were in place.

IV. DIVERSIFYING THE PROFESSORIATE: APPROACHES TO RECRUITMENT, RETENTION, AND INCLUSION

Hannah Valantine, M.D., Chief Officer for Scientific Workforce Diversity at the NIH, explained that the Scientific Workforce Diversity Office was created after a 2011 report showed lower rates of R01 funding for African American applicants; the NIH is making an effort to address the question of diversity with scientific rigor and specifically wanted the Office to be headed by a working scientist who understands the issues and culture involved and can communicate those issues to his or her colleagues. Dr. Valantine pointed out that in her own field, organ transplantation, African American patients are at greater risk of organ rejection, so diversity is an important consideration not only in the workforce but also in patient care.

Diversity is important not only for reasons of fairness, but also because a diverse workforce strengthens science. A broad group of people is more likely to widen the scope of inquiry, and data shows that diverse teams lead to more innovation, creative solutions, and outperform less diverse but more expert teams. Additionally, the United States is becoming increasingly diverse and the entire intellectual workforce must be available to discover the best talent.

In a 2015 PNAS article written with Dr. Francis Collins, Dr. Valantine articulated four areas of focus for strengthening diversity at the NIH: the science of diversity, recruitment retention, sociocultural factors, and sustainability. She described several active diversity programs that address these areas, including the Diversity Consortium Program, which was launched in 2014 to determine the contexts in which various kinds of diversity programs succeed. Experiments are in progress at a variety of institutions; each experiment must work toward given hallmarks of success at three dimensions: the individual student, faculty, and institution/institutional change. The experiments were designed by the institutions and are testing very different hypotheses, such as questions of stereotype threat and student entrepreneurship. Dr. Valantine also discussed the National Research Mentoring Network, which trains scientists to be better mentors and links students with mentor networks that will position them for success. Another approach currently being tested in the NIH Intramural Program, is a recruitment and retention strategy. A search tool was developed to provide search committees with information about highly qualified candidates from underrepresented groups, which refutes myths about a lack of qualified candidates for open positions. Outreach is also critical—Dr. Valantine commented that scientists cannot assume candidates know how it feels to be part of the research community. The Future Research Leaders Conference brings senior Postdocs and junior faculty to NIH during the research festival several days. They network with leaders in the Institutes, and the conference has already produced a number of successfully qualified candidates. These strategies can be adapted quickly by institutions.

Implicit bias is a pervasive problem resulting from the way the brain is wired to deal with large amounts of information quickly by making assumptions. Dr. Valantine shared examples of studies showing that people assume more feminine-looking faculty members are less likely to be scientists, and résumés with female names are rated lower by review committees. The good news is that implicit bias can be overcome with tools that bring awareness.

Dr. Valantine commented that sustaining diversity is a multifactorial problem; for example, the R01 award is a common gateway into the biomedical workforce. However, African Americans submit fewer initial and resubmission applications, the review process gives their applications lower scores, and these applications have a lower likelihood of being funded, reducing the percentage of diverse awardees at each step. Studies of diversity in research indicate that as scientists' progress along the career path, the numbers of individuals from underrepresented groups drop. Data from the past 20 years show a 9-fold increase in scientists from underrepresented groups receiving doctoral degrees, but no corresponding increase in their attaining assistant professor positions; a major gap is in the transition to independent research. The data suggests that if institutions make diverse hiring a focus, these discrepancies can be resolved quickly. Dr. Valantine emphasized that her office will continue to address this problem using data-driven methods with the scientific rigor we use while conducting our science and show that great minds do not think alike—great minds think differently.

Discussion Highlights

- In response to a question about the time investment involved in mentoring, Dr. Valantine commented that there are some, but not enough, mechanisms to address this embedded in the work that the National Research Mentoring Network is doing. She pointed out that the burden of mentoring frequently falls on women and people from underrepresented groups, thereby producing a detrimental effect on a researcher's own career. Recent studies mapping mentor networks show that networks are highly predictive of productivity and success, so it is critical that diversity efforts ensure those from underrepresented groups have access to these networks.

- Experiments are planned to address the critical issue of bias in peer review; data has shown that reviewers make different review comments on applications from women and men. Another opportunity to address diversity is the increasing prevalence of team science, because teams can be structured to include the diversity of participants. Prioritizing research conducted among underserved populations is another area which should be addressed.

- When asked how the research findings discussed will be disseminated, Dr. Valantine noted that the programs supported by the Common Fund and the Intramural Program are both generating publications and dissemination in other formats is in process. The search tool to identify candidates from underrepresented groups is not published yet but will be available soon. Dr. Valantine explained that this tool utilizes field-specific search protocols and can identify candidates at any career stage; she commented that search committees often look first at their own networks, but women and underrepresented groups, frequently, are not part of those networks.

V. NIH UPDATE

Lawrence A. Tabak, D.D.S., Ph.D., Principal Deputy Director of the NIH, described the NIH Innovation Account that was established as part of the implementation of the 21st Century Cures Act, noting that this funding is set aside for the special initiatives but must be appropriated each year, so it requires active engagement from the appropriators. The Innovation Account does not count against budget caps and reauthorizes the NIH for fiscal years 2018 through 2020. These funds are separate from the standard

appropriations process and cannot be moved between programs, but each program will be able to determine the funding timeline most appropriate for its research.

Dr. Tabak reviewed specific initiatives, beginning with the goals of the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative: to understand how individual cells and neural circuits interact to enable the spectrum of human behavior. The long-term goal of the BRAIN Initiative is to be able to monitor and regulate brain circuits to diagnose and treat neurological and mental health disorders. Dr. Tabak commented that the BRAIN Initiative was off to a strong start, having released almost two dozen Funding Opportunity Announcements (FOAs) totaling up to $100 million in new awards this fiscal year.

The Precision Medicine Initiative (PMI) aims to create a biomedical data resource that reflects the diverse population of the United States, including people of every age, race, ethnicity, socioeconomic status, geographic location, and health status. The PMI requires researchers to build a series of tools that enable ease of data collection and foster active partnerships with community groups and health care providers. Dr. Tabak commented that the first version protocol is near completion and there are plans for community engagement and outreach, though more work is required to build public confidence.

Dr. Tabak recounted the goals of the Cancer Moonshot Initiative: to be able to manipulate tumor pathways for treatment, understand tumor genes, create better models to test hypotheses, and encourage greater collaboration. The Moonshot is designed to build on longstanding National Cancer Institute investments but remain a unique and separate program to enhance cancer detection. He noted that the release of many FOAs is anticipated, which will require a large review effort.

The Regenerative Medicine Initiative works with the U.S. Food and Drug Administration to support stem cell research to further the field. This initiative includes a novel aspect requiring matching funds from applicants; operationalization of this component has not yet been determined. The Next Generation Researchers Initiative requires that the NIH Director coordinate all policies and programs aimed at new researchers and adjusts the NIH loan repayment program to match trainees' current burdens. The Eureka Prize Competition supports identifying areas of biomedical science that could advance through a prize competition, particularly related to conditions that have a disparity between research investments and treatment investments. The 21st Century Cures Act also requires an NIH-wide strategic plan that is updated every 6 years and the establishment of a working group to enhance rigor and reproducibility. Dr. Tabak relayed some of the steps the NIH already has taken regarding reproducibility, including revised applications and reviews and increased investigator stability. The Cures Act also encourages sharing of data generated from NIH-funded research and requires that research related to sexual and gender minority populations develop "valid and reliable methods," and the address "methodological challenges."

Dr. Tabak provided a brief update about data science efforts, which are a cross-agency concern. The role of the Associate Director for Data Science is to set the vision for the trans-NIH data science effort and oversee the Big Data to Knowledge (BD2K) Program; Dr. Patricia Flatley Brennan recently has taken on the position in an interim capacity, in addition to her duties as the Director of the National Library of Medicine. The BD2K Program will be managed as a Common Fund program through the DPCPSI OSC.

Dr. Tabak then commented on the presidential transition, noting that Dr. Collins will continue in his position as the NIH Director, although the length of his tenure has not been defined. Dr. Tabak complimented the preparation of the transition teams, which contributed to a smooth transition. He noted that the current hiring freeze is not uncommon in a new administration; he expected an exceptions process for positions key to public safety. Dr. Tabak suggested that the 21st Century Cures Act and the programs it encourages, such as research related to the health of sexual and gender minority populations and cross-

departmental initiatives, are unlikely to be modified because of the bipartisan support for the creation of the act.

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 room 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 574 ORIP applications with requested first-year direct costs of $325,900,786.

VII. NIH PERSPECTIVES ON ENHANCING SCIENTIFIC RIGOR AND REPRODUCIBILITY

Michael S. Lauer, M.D., Director of the Office of Extramural Research, presented several examples to demonstrate the challenge of rigor and reproducibility efforts. Small samples are more likely to produce extreme findings that seem significant, but attempts to reproduce these studies will fail to repeat the extreme result. Humans instinctively develop narratives and meaning; even experienced statisticians can have difficulty reviewing scattered data without attempting to impose a story on the results. Many of the reproducibility problems in science can be ascribed to similar common cognitive biases.

The vast quantities of information now accessible to science have made a difficult field even more challenging by providing more opportunities for error and reducing the ability to eliminate all potential variables. Dr. Lauer demonstrated the ease with which data can be manipulated to show a significant result by selectively including or excluding variables. Most science done today involves so many variables and potential outcomes that finding a seemingly significant result for any given variable is almost certain.

An underpowered study that shows a positive relationship between variables is likely to be false because it lacks the power to detect an association even if one actually exists; conversely, an underpowered study that gives a negative result is likely to be dismissed precisely because it was underpowered. Thus, underpowered studies waste resources and are unethical, and they should not be conducted in any case. Transparent reporting—reporting sample size estimation, randomization, blindness, and data handling—could help optimize the predictive value of clinical or preclinical research, but many current papers do not address these components. Dr. Lauer suggested that funders are beginning to recognize these considerations and demand stronger statistics.

The NIH's recent updates to application and peer review policies require four key components: the scientific premise forming the basis of the proposed research; rigorous experimental design for robust and unbiased results; consideration of relevant biological variables, such as potential confounders and sex; and authentication of the data studied. Many resources are available on the NIH's website to assist

[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.

applicants, including articles, references, and detailed explanations of the information required for each component.

Clinical trials are a particular area of interest to the rigor and reproducibility effort. Trials often either do not get published or are irreproducible, and frequently they are reviewed by individuals who are not familiar with clinical trial methodology. The NIH now requires the clinical trial application to be under a dedicated FOA rather than a parent R01 announcement, which allows individuals with the appropriate clinical trial knowledge to be incorporated in the review panel. The application also must include the intervention and primary endpoint, and reporting is required.

Many stakeholders have united to make science more robust, and interest is spreading to fields outside of biomedical science. Acknowledgement is an important step toward addressing problems; outlying data points and complex variables cannot be ignored or removed. Best practices to prevent these problems could include publishing protocols, sharing data and analysis coding, defining exploratory studies, and attaining independent confirmation. Addressing universal cognitive bias is more important than ever, because science is more successful than ever.

Discussion Highlights

- Although the new requirements apply to all NIH applications, much interest has been focused on preclinical research. The clinical trial applications have not yet changed; Dr. Lauer anticipates the changes will be applied within several years. He expressed the hope that posing questions about rigor will encourage applicants to consider these questions and result in more rigorous studies.

- When asked whether more stringent requirements would reduce innovation, Dr. Lauer commented that small exploratory studies posed no threat to rigor if the exploratory nature was acknowledged. He added that poorly designed studies would not promote innovation. Although larger studies are more expensive, Dr. Lauer proposed that fewer studies with better designs would be more productive than using limited resources inefficiently.

- In response to a suggestion that a greater focus should be placed on defining determinants of success, Dr. Lauer pointed out that in situations with a small effect size, a very large trial is needed to demonstrate a positive effect, in which case researchers must determine whether the larger study is worth the resource investment needed to find a small effect.

- Regarding the possibility of overstating science's flaws in addition to overstating its successes, Dr. Lauer acknowledged the need for balance and emphasized that recognizing cognitive bias and approaching data thoughtfully are the most important strategies to prevent data issues.

VIII. OFFICE OF DISEASE PREVENTION'S FISCAL YEAR 2014–2018 STRATEGIC PLAN: A MID-COURSE REVIEW

David M. Murray, Ph.D., Director of the Office of Disease Prevention (ODP), reviewed ODP's progress at the midpoint of its strategic plan. This strategic plan has six priorities, each of which has a team dedicated toward its goals. He explained that ODP serves as a prevention science methods resource to assist any groups in incorporating the strongest methods possible. Under Strategic Priority I, Dr. Murray's team developed machine learning tools to better categorize prevention research. In assessing the accuracy of the previous classification system, Dr. Murray's team noticed that most Type 1 R01 prevention studies included observational designs, which are less likely to be reproducible, and a much lower percentage included clinical trial designs, which generate reproducible results more often. If this statistic remains consistent in the assessment of the remaining mechanisms, the office will push for more clinical trials.

Under Strategic Priority II, the team worked with a number of other groups and offices to identify evidence gaps and areas for additional investment by collecting data and offering workshops on prevention topics. One of their important activities is an annual survey of the ICs to identify current activities that may address Insufficient Evidence Statements reported by the US Preventive Services Task Force. ODP has created 5 new Scientific Interest Groups to develop workshops and FOAs to address some of these areas.

Strategic Priority III addresses methods improvement. Dr. Murray's team has worked to identify training opportunities and prevention research methods already in existence and made them easy to find on the ODP website. They also have created a database of prevention methods experts and a tool that helps scientific review officers find methods experts for their panels. The team has organized webinars, talks, and workshops addressing relevant prevention science methods issues and has established an Early Career Investigators Award. Additionally, an online course was created to familiarize investigators with methods specific to group randomized trials. The Team is also proposing language related to these issues also for incorporation into FOAs to encourage investigators to consider these issues early in the process.

To meet the goals of Strategic Priority IV, the team develops initiatives to address the research gaps and opportunities identified under Strategic Priority II. They have created new trans-NIH scientific interest groups to address issues of child and adult screening, genetics of prevention, policy and legislative evaluation and intervention, and comorbid disease prevention; these groups will make recommendations for workshops, meetings, or FOAs in these areas. They also have provided input on strategic plans for Institutes, Centers, and Offices, as well as for the NIH-wide strategic plan.

Strategic Priority V relates to ODP's charge to disseminate effective, preventive interventions. One of the largest sections on the SPV portion of the ODP website is a database of intervention critiques collected by groups across the country. ODP also participates in dissemination and implementation research throughout the NIH.

Under Strategic Priority VI, the team has worked to increase the visibility of prevention research, largely through expanding their website and social media presence, creating an email list, and presenting the resources they have created at scientific meetings around the country. One of the important sections of the website is called Resources for Researchers, a centralized location for information related to NIH-funded research in prevention.

Strategic planning for the next period has begun among the ODP staff; conversations with stakeholders will occur in the spring and early summer, and a Request for Information will be released in the late summer or early fall to acquire public input on the second strategic plan.

Discussion Highlights

- In response to a question about disseminating epidemiologic research to the public, Dr. Murray recommended that the NIH could fund more follow-up studies, which are necessary to provide the public with a clear statement of benefit.

- Dr. Murray explained that recipients of the mailing list were identified through the Type 1 R01 assessment and the database of prevention research experts; there also is a standard mailing list that people can join through the website. Their social media presence is modest because most outreach efforts are directed at the prevention research community, rather than the general public.

IX. DATA SCIENCE AT THE NIH: PIVOTING TO THE FUTURE

Patricia Flatley Brennan, R.N., Ph.D., Director of the National Library of Medicine (NLM) and interim Associate Director for Data Science, presented on her vision and direction for the NLM and trans-NIH data issues in this rapidly evolving period of data science. She explained that biomedicine was first studied through experience and observation, then through the refinement of those observations into focused experiments. Technological advances later allowed the addition of computation, and now science and scientific discovery are driven by data. Dr. Brennan commented that the systematic exploration at the heart of biomedicine had not changed, but the substrate had evolved to enable new types of discovery at an accelerated pace.

Data must be findable and they must be accessible, meaning that scientists must know where the data are and there must be pathways to acquire them. Data also must be interoperable, or possess the ability to link one data element to other observations or other types of data, and data must be reusable. Dr. Brennan emphasized that in addition to increasing data storage and access, new methods for structuring, managing, and analyzing data must be developed.

Dr. Brennan commented on the vision for the NLM, noting that it is older than the NIH and over time the library can best be considered a dynamic interplay of medicine and information. Today the NLM hosts more than 4 million visitors a day to its electronic resources, and the NCBI moves massive amounts of data—about 50 terabytes a day—in support of discovery. The NLM must have the ability to store information and make it accessible, which is an updated extension of the historical role of librarians.

It is critical to develop ways to assess the value of data and forecast whether that value will change over the lifecycle of the data, because the value of data will guide how to determine its preservation and worth the investment; the NLM will engage economists and futurists to create probabilistic models for data preservation. Dr. Brennan noted that preservation strategies for the most fragile media will be required to allow future access and retention. A key component of preservation is creating and promoting standards, and NLM staff will work with communities around the world to ensure that data can be interoperable across users and computational environments. Dr. Brennan commented that the NLM is known for building tools that are accessible from platforms all over the world, and continuing this legacy requires constant updating.

A data-sophisticated workforce is required to create and maintain the NLM's data efforts; the NLM serves data scientists, research scientists, and clinicians, and each of these groups interacts with data at a different level of the data continuum. Additionally, patients have become part of the clinical workforce by actively participating in their own care and collecting their own health information data. Open science allows engagement with diverse users and collaboration across disciplines, and resolution to data science challenges in other fields, such as mathematics or astronomy, can be applied to health science.

Dr. Brennan commented on the plan to integrate the NLM with the BD2K Initiative and build on its successes, including supporting research activities and resource services within and outside the NIH through collaboration with the 12 Centers of Excellence around the country and testing such new models as cloud storage. Dr. Collins and the NIH leadership have embraced the data science efforts as a discovery pathway for the future, and the NLM is ready to drive the future of data science on the NIH campus to build on what is known, engage those already invested, and accelerate the effort to have safe, findable, accessible, and interoperable data.

<u>Discussion Highlights</u>

- Regarding the involvement of the National Science Foundation (NSF), Dr. Brennan explained two joint initiatives within the BD2K program to address methodological challenges and archive

sustainability. She suggested additional opportunities to work with the NSF to establish data structures for very large data sets and develop additional tools for the increasingly connected health environment.

- In response to a question about private-public partnerships, Dr. Brennan emphasized that the solution to data management and data science challenges cannot be exclusively a government solution—the structure of data repositories requires an enormous investment from the private sector. The larger question revolves around the economic model for data storage; Dr. Brennan expressed her hope that the acceleration of discovery also would translate into new economic products but noted this has not always been true in the history of knowledge building. Several possible economic models can be considered, depending on whether the costs fall to the researchers, the public, or the user, and Dr. Brennan theorized that the eventual solution would be a hybrid of these models.

- Data monetization and security are critical ethical questions; there are plans to explore these issues within the BD2K Initiative and beyond. Data storage models of the future likely will be distributed, so data stored in multiple locations will need to be accessed, which adds questions about data transport. Dr. Brennan commented that the NLM is the best host for such conversations, because of its familiarity with public funding and public access issues.

- Large portions of the data in the world currently are stored in private servers, so intellectual property will be a large component of future conversations. A data-sophisticated workforce can help predict upcoming questions and develop solutions for diverse stakeholders.

X. NEW METRICS AND TOOLS FOR EVALUATING THE IMPACT OF NIH-FUNDED RESEARCH

George M. Santangelo, Ph.D., Director of the Office of Portfolio Analysis (OPA), DPCPSI, explained that OPA divides its efforts between coordination activities and development of the science of portfolio analysis, which helps scientists understand how best to accelerate scientific progress. OPA frequently addresses questions of gaps, opportunities, and overlap in the NIH portfolio, and it has developed many tools that are available across the NIH to gauge productivity using diverse metrics.

Dr. Santangelo demonstrated the use of the Relative Citation Ratio (RCR), a metric that assesses influence at the level of individual articles, and provides a validated, mathematically-sound alternative to journal impact factor. The *iTrans* tool can track articles cited by clinical trials or guidelines since their publication, providing a method of visualizing the translational productivity of funding awards. Dr. Santangelo demonstrated the difference between the journal impact factor and RCR for papers published by the recipients of awards in two distinct areas of biomedical research, explaining that because the RCR normalizes citations at the article level, it provides a more nuanced picture of the success of NIH-funded research.

Both the *iCite* and *iTrans* tools are accessible through the *iSearch* tool, which is a broad portfolio analysis platform developed by Dr. Santangelo's team to more efficiently respond to frequent questions about trans-NIH analyses. *iSearch* provides easy-to-use access to a carefully curated, extensively linked group of datasets, including grants, publications, clinical trials, patents, and approved drugs. *iSearch* is fast, comprehensive, and amenable to free text queries. It also allows real-time data exploration, and the data are updated frequently.

Dr. Santangelo demonstrated the *iSearch* tool, pointing out the various modules available. He explained that the interactive aspect of this tool utilizes facets, which allow users to filter the data by such criteria as

Institute, fiscal year, or award status. When asked when this tool would be available to the public, Dr. Santangelo explained that some components require further streamlining to correspond to other public-facing NIH resources. Regarding potential pushback from other aggregating services, Dr. Santangelo noted that his team has partnerships with outside research companies that are supportive of openly curated data sets.

Dr. Santangelo demonstrated *iCite* and *iTrans*, noting that results from *iSearch* can be transferred easily to *iCite* and *iTrans*. He suggested that the ability for users to upload text to find awards and publications could be a useful future addition.

XI. CLOSING REMARKS

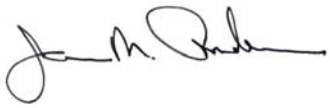
Dr. Anderson thanked the Council members and speakers for their contributions at this meeting. He reminded the members that the next Council meeting will be held on May 26, 2017.

XII. ADJOURNMENT

Dr. Anderson adjourned the meeting at 3:51 p.m. on January 27, 2017.

XIII. CERTIFICATION

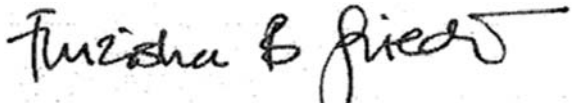
I hereby certify that, to the best of my knowledge, the foregoing summary minutes are accurate and complete.



03/14/2017

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

Date



3/14/2017

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

Date