National Institutes of Health
National Institute of Diabetes and Digestive and Kidney Diseases

NIDDK Central Repository 20th Anniversary Workshop: Promoting Secondary Research to Accelerate Medical Breakthroughs and Innovation

Virtual Meeting
September 19–20, 2023

EXECUTIVE SUMMARY

Background and Overview

The National Institutes of Health (NIH) sponsored a scientific workshop on September 19–20, 2023, titled "NIDDK Central Repository 20th Anniversary Workshop: Promoting Secondary Research to Accelerate Medical Breakthroughs and Innovation," which was hosted by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The purpose of the workshop was to review the NIDDK Central Repository (NIDDK-CR) program from its early beginnings to its current state and review the vision for its future as an integral component of the NIDDK data ecosystem.

Data and biospecimen repositories were established by the NIDDK in 2003 to enhance the impact of its extensive clinical studies. The NIDDK-CR houses clinical research data and associated biospecimens from NIDDK's extramurally funded multisite clinical studies and has evolved from a traditional catalog of samples and data to a rich state-of-the-art database of samples and associated data, which promises an accelerated rate of research. The NIDDK-CR contains more than 200 data packages from 244 ongoing and concluded studies with millions of associated biospecimens, all generated by clinical research projects across NIDDK's mission areas. These resources are potentially valuable for ancillary studies performed in collaboration with active projects or for secondary research.

Data science at the NIH has been advancing rapidly with the release of the *NIH Strategic Plan for Data Science* under the leadership of the NIH Office of Data Science Strategy (ODSS). This Plan provided a roadmap for modernizing the NIH-funded Biomedical Data Science Ecosystem. NIDDK has invested its resources and, over the past 5 years, has transitioned the NIDDK-CR to the cloud via Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (commonly called STRIDES) and, with ODSS co-funding, has achieved CoreTrustSeal certification and is implementing single sign-on research authentication and authorization services. In parallel with these efforts, the National Library of Medicine (NLM) established a broad 10-year vision for the NIH Data Ecosystem and how it could be harnessed to promote discovery to improve health. In December 2021, the NIDDK released its *Strategic Plan for Research*, which explicitly identified repositories and data science as important opportunities for scientific discovery and priorities for the NIDDK.

The objectives of the NIDDK Central Repository 20th Anniversary Workshop included the following:

- Engage diverse and multidisciplinary experts from the research community to promote the NIDDK-CR resources and the program's broad research portfolio.
- Highlight the valuable resources held under the guardianship of the NIDDK-CR.

- Review how the NIDDK-CR resources can be used in research.

The agenda included keynote speakers, lightning talks, panel discussions, and demonstration workshops. A diverse group of speakers and moderators participated in the sessions, including repository experts from across NIH institutes and centers; investigators who have contributed resources to the NIDDK-CR; secondary analysis researchers who have successfully utilized the NIDDK-CR's resources to advance their fields of science and careers; and data scientists and data science librarians, all with expertise in different areas of the resource life cycle. This workshop directly aligned with NIH and NIDDK's strategic goals to maximize the value and impact of resources generated through NIH-funded research, advance the biomedical data ecosystem, enhance data resources that facilitate research, catalyze scientific discovery, and improve health outcomes. More than 300 participants registered and attended worldwide.

Session 1: Reflecting on 20 Years of the NIDDK Central Repository: Embracing Our Legacy and Shaping Our Future

In May 2001, the NIDDK Advisory Council recommended creating central repositories so that other researchers could access the resources from NIDDK's multisite clinical trials. Subsequently, in 2003, three contracts were awarded for separate data, biosample, and genetics repositories. In 2016, the biosample and genetics repositories were merged into a single entity. The major challenges in 2003 were developing model informed consent and request for application language, establishing governance policies, and persuading study investigators to reposit samples and allow them to be shared.

By 2016, the NIDDK-CR had become productive, growing significantly each year, and it currently includes millions of samples. The major challenges in 2016 were providing custom service to studies (including labels and kits), reviewing consents to ensure that submitted samples were eligible to be shared, and creating a searchable recordkeeping system (for more than 75 studies, some with more than 100 sites). Additionally, too many samples were not of interest. Since that time, the focus has been on making a larger amount of data more searchable and accessible, even to nonspecialists; identifying opportunities to use samples for significant research; and creating virtual catalogues to facilitate searching collections across different repositories.

NIDDK's vision for the NIDDK-CR was for investigators not conducting the initial studies to realize the full potential for existing resources to generate new findings and accelerate medical breakthroughs that impact health and health care outcomes. The mission was and still is to provide access to the wider research community, thus expanding the usefulness of the resources generated from NIDDK extramurally funded multicenter clinical studies beyond the end of the studies. The NIDDK-CR has emerged as an international resource supporting large NIDDK-funded active and concluded clinical studies consisting of resources from more than 160 collections across more than 100 diseases, disorders, and topics (e.g., diabetes and other endocrine, metabolic, and digestive diseases) extending across 6 major research areas within NIDDK divisions. The NIDDK-CR is 1 of 27 NIH-supported control access repositories that provides materials to secondary investigators around the globe.

The website platform was first launched in 2004 with only 6 collections and was relaunched in 2021 as Resources for Research (R4R), with 160 collections. The R4R contains data and resources from more than 240 interventional and observational clinical studies, 60 percent of which are available to the external research community upon request. Over the years, the number of collections has grown and the demand for resources has increased. The NIDDK, with the support of the ODSS, has implemented system enhancements and process improvements and performed the necessary metrics to assess the utilization and the utility of R4R, as well as employed the relevant parts of the repositories to align with FAIR

(Findable, Accessible, Interoperable, and Reusable) and TRUST (Transparency, Responsibility, User, Sustainability, Technology) principles. The NIDDK-CR has 5,570 registered users and more than 800 approved investigators who have produced more than 500 publications or abstracts.

Plans for the NIDDK-CR in the Upcoming Years

- Continue modernizing the repository, establishing partnerships within and outside the NIH, and leveraging existing opportunities that align NIH and ODSS objectives.
- Actively participate in efforts to increase the role of repositories in advancing the biomedical ecosystem.
- Improve data harmonization across projects and enhance artificial intelligence (AI) readiness, engaging with different parts of the community and potentially increasing the pipeline of secondary researchers.

Session II: Modernizing the Data Repository Ecosystem to Accelerate the Pace of Biomedical Research

The NIH NLM, which has been serving science and society since 1836, is a knowledge hub that powers and is powered by AI and comprises health standards, research literature, biomedical data and tools, library resources, and direct-to-consumer health information. In February 2023, the NIH and NLM launched a new initiative—Digital NIH: Innovation, Technology, and Computation for the Future of NIH—as a framework to define high-priority capabilities and manage NIH technology investments across functional areas common to all Institutes and Centers to achieve NIH outcomes. The Digital NIH framework applies across extramural research management; intramural clinical and basic research; administration and management; and crosscutting capabilities (e.g., security and workforce). Implementing Digital NIH is a multiyear journey that will iteratively explore solutions to prioritize capabilities over the next several years. Digital NIH provides the infrastructure to modernize the biomedical data resource ecosystem, and the *NIH Strategic Plan for Data Science* provides the data content, making it accessible.

The *NIH 2023–2028 Strategic Plan for Data Science* has five goals: Boost capabilities to sustain the NIH Data Management and Sharing (DMS) Policy; establish programs to enhance human-derived data for research; create new opportunities in methods and artificial intelligence; increase support for a federated biomedical research data infrastructure; and strengthen a broader community of data science.

New initiatives include developing common data elements (CDEs) to build quality into data acquisition; considering the data life cycle to glean the most out of every research project; and developing a data catalog for making data findable, which will be a curated catalog of biomedical data sets from selected publicly available repositories. Data science experts further elaborated on CDEs. The readiness of the NIH and its investigators to move into a data CDE strategy that will enhance data reuse is exciting. In fact, in research, investigators should think of CDEs as a team sport and not define them in a vacuum. They should make use of professional societies and gatherings of researchers and share websites, which are critical tools for moving the conversation forward. Investigators are encouraged to think about the concepts they are trying to capture and then look for equivalent measures of the concept, rather than jump right to the variable and the question and answer that they want to use, which is too specific. At the concept level, CDEs are a tool that requires iterations of testing and exploring.

Session III: Evolution of the Traditional Repositories—NIDDK-CR and Others

Cutting-edge technologies have advanced over the past 20 years, and similarly, NIH repositories also have evolved during this time period, beginning in 2000 to 2003, with the launch of PubMed Central and the NIDDK-CR. A detailed list of NIH repositories can be found on the Repositories for Sharing Scientific Data webpage. In 2022–2023, the NIH launched the Generalist Repository Ecosystem Initiative (commonly called GREI) and released the new DMS Policy. The NIDDK-CR and the repositories across the NIH encourage technology influenced projects but also are driven by the needs of the research community and by the requests of the program officers across the NIH. These policies and scientific research needs are what compels technology as it evolves. The research community has been able to ask new and exciting questions by using the available data and the biospecimens that are offered for reuse.

Other NIH repositories and resources include the Kidney Precision Medicine Project (KPMP), the GenitoUrinary Development Molecular Anatomy Project (GUDMAP), ReBuilding a Kidney (RBK), and BioData Catalyst. The KPMP aims to establish a human kidney disease atlas and data repository. GUDMAP provides data and tools that facilitate research on the genitourinary tract for the scientific and medical community. The RBK is an NIDDK-funded consortium coordinating research to generate or repair nephrons that can function within the kidney. The National Heart, Lung, and Blood Institute (NHLBI) BioData Catalyst, a cloud-based ecosystem of secure workspaces, increases data availability and accessibility; ensures data governance and security; and develops a culture of data sharing and reuse for heart, lung, blood, and sleep data.

NIH experts in the field discussed common challenges working with disparate data sets and opportunities for using technology to improve and integrate legacy data across studies. They described, from their perspective, opportunities for addressing some of these challenges through new technologies, as well as where and how these technologies can be used to bring disparate collections together for new and innovative work. Examples of the challenges and opportunities are summarized below.

Research Gaps and Challenges

- No single repository can host all collections. A major challenge is interoperability, and it is more of an issue for non-omics data types (e.g., NIDDK-specific).
- Data standards and ontologies are lacking, as are software tool best practices.
- NIDDK data challenges include heterogeneity of data, data quality, and consistency.
- Data privacy and security—protecting privacy while maximizing data reuse—remains a concern.
- Large language models (LLMs) are revolutionizing the standardization of scientific data in terms of format, annotation, data quality, and data description and documentation, but they are underutilized.

Research Opportunities

- Support open collaboration to improve reproducibility.
- Make knowledge machine actionable by integrating knowledge and data.
- Establish platforms that link data, tools, and knowledge (e.g., the Human Islet Research Network Pancreas Knowledgebase Program).

- Support data integration platforms, machine learning, and AI.
- Establish data standards and ontologies, as well as secure cloud-based analytic platforms.
- Consider open data standards and application programming interfaces (APIs).
- Understand ways that technology can be used to address some of the current challenges regarding data standards, architecture, and common APIs.

Session IV: FAIRization of NIH-generated Resources: Meeting the Intent of the 2023 Data Management and Sharing Policy

The panelists, who were experts in data standards, biospecimen data standards, data curation, and data science research, discussed practical applications of FAIR principles for data and biospecimen deposition and sharing for secondary uses, with examples of what it means to make resources "FAIR" in meeting the intent of the 2023 DMS Policy. When researchers use CDEs for addressing FAIRness, data are more interoperable, reusable, and reproducible. CDEs also save time and resources by streamlining setup and facilitating harmonization across systems. National Cancer Institute (NCI) data requirements for biobanking in support of cancer epidemiological research focus on quality and reuse of standard operating procedures (SOPs); conformity of practices and use of those SOPs; reproducibility, especially for rare cancers; and comparability. The Data Curation Network (DCN) ensures that data sets meet the FAIR guiding principles by implementing the CURATED (Check, Understand, Request, Augment, Transform, Evaluate, and Document) model, which begins with checking files to ensure they open and conclude with documenting all curation activities throughout the process. To begin a large data repository or data project, such as GUDMAP, researchers first generate the appropriate metadata. FAIR was readily adopted and is a guiding principle of the NIDDK-sponsored Analysis, Technology, Leadership, Administration, and Science-Data to Knowledge (ATLAS-D2K) Center.

Panelists highlighted new applications of FAIR principles. Data generators who are collecting their information for publication using the resources of the ATLAS-D2K Center are encouraged to create a collection that encompasses data, not just biospecimens. NCI soon will launch a publication data repository that will allow curated access to the data sets that are within publications and will be linked to the study. The DCN hosts workshops to educate information professionals in data curation. Research gaps and challenges were discussed, and examples are summarized below.

Research Gaps and Challenges

- Using the vocabulary is challenging for data scientists and researchers because no dedicated disciplinary data repository has been identified.
- With the large data sets, the amount of shareable data is not well understood.
- Data sets need to be able to stand alone as published objects and not rely on research articles for context.
- Biohoarding (building but not sharing a quantifiable collection of tissue samples) is a major issue and reduces the sustainability of repositories and the use of approximately 10 percent of specimens in repositories worldwide.
- Although CDEs are considered a good approach to data standards, the field is slow to implement the necessary changes.

Session V: Secondary Research to Accelerate Medical Breakthroughs and Innovation

Trends indicate that approximately 30 percent of the world's data volume is being generated by the health care industry. The projection is that by 2025, the compounded annual growth rate of data for health care will reach 36 percent; however, these data are siloed across systems, and harmonization remains a challenge. Three components for secondary data analysis success were outlined: (1) data that are standardized, high quality, curated (context), and harmonized; (2) aggregation and processing that incorporates semantic interoperability, de-duplication, and de-identification; and (3) tools and presentation that include scalable statistical methods and pipelines and advanced analytics platforms.

Innovation, Opportunities, and Breakthroughs

Innovations regarding secondary data analysis include addressing standards and interoperability. Recent strides in the acceptability of AI and machine learning have influenced tools and presentation. Much of health care data comes from electronic health records (EHRs). The Health Information Technology for Economic and Clinical Health Act of 2009 succeeded in promoting the widespread adoption of EHRs, but over time, it has resulted in data traps with varying local standards at each individual organizational level. Example innovations and breakthroughs are summarized below.

- Research-based data marts and collaborations, including the Observational Medical Outcomes Partnership Common Data Model (CDM), European Health Data & Evidence Network, and National Patient-Centered Clinical Research Network CDM.
- 21st Century Cures Act prohibition of information-blocking and the Trusted Exchange Framework Common Agreement.
- Cosmos, a cross-institution single EHR vendor–facilitated data aggregation tool.
- Unlocking of unstructured data using AI and machine learning.
- LLMs for data de-identification and semantic interoperability.
- AI to advance predictive heterogeneity of treatment-effect studies and reinterpretation of negative randomized control trials.

Session VI: Supporting NIDDK-Funded Research and Expanding NIDDK's Data Ecosystem

Panelists provided examples of ways NIDDK-funded research is strengthening the data ecosystem and accelerating medical breakthroughs through secondary research.

The NIDDK Information Network (dkNET) provides a single point of access to information about diverse research resources that advance the mission of the NIDDK. The dkNET platform provides tools and resources to support the FAIR principles, includes a Hypothesis Center for bioinformatics training, assists researchers in meeting the requirements of the DMS Policy, and helps to increase the impact of NIDDK research and data.

The Type 2 Diabetes Knowledge Portal (T2DKP) was established in 2015 to make genetic association data available to the wider research community. The T2DKP began with 25 traits and 9 data sets and uniquely shared summary statistics for human genetic data. Other research communities, including cardiovascular disease (CVD) and type 1 diabetes (T1D), have since shared such summary statistics and are supporting the open-access resource Common Metabolic Disease Knowledge Portal, which focuses on

six metabolic diseases. Some ways to sustain this NIDDK resource in the future are to capitalize on the efforts of data repositories and disease and domain-specific resources; create cross-talk among these resources; and create the structures for sharing within these resources.

The Vivli data-sharing platform includes trials from any disease, country, sponsor, funder, or investigator. To implement FAIR principles, Vivli has three search modes to enable scientifically precise queries; uses standardized metadata and terms from the Cochrane vocabulary to describe trial data; provides a single process to access data from multiple data contributors and platforms; and provides a secure research environment prepopulated with common analytic tools. Ninety-three studies listed on Vivli redirect to the NIDDK-CR/R4R; some are archived on Vivli. Requesters are redirected to the NIDDK for processing and approval and to access data. The NIDDK–Vivli collaboration enables integration with another 7,000 member data sets in Vivli.

Panelists highlighted suggestions for improving data sharing across data sets and platforms:

- Establish a coalition of interested parties to share their results and receive guidance on publishing their data in a rapid time frame.
- Understand the mission of the different stakeholders, facilitate developing a trust relationship, and ensure that incentives are aligned with their missions.
- Emulate the Pharmaceutical Research and Manufacturers of America and the European Federation of Pharmaceutical Industries and Associations, which established the Principles for Clinical Trial Data Sharing used by trade associations.
- Increase awareness of the existing resources to inform researchers' choice of a repository in which to deposit data and samples.
- Provide feedback to the research community on the progress of writing data sharing plans to meet the requirements of the NIH DMS Policy.

Lightning Talk Sessions: NIDDK-CR Secondary Research—Scientific Discoveries

Research has been featured in numerous peer-reviewed journals, editorials, and commentaries and has stimulated other projects and further scientific discoveries. To promote additional secondary analysis, graphical user interfaces are publicly shared, programing codes are uploaded in GitHub, and tools are made public after clearing regulatory reviews. Invited speakers presented their research in two lightning talk sessions.

Methods for Biomarker Discovery in Diabetes, Kidney, and Urologic Diseases

Resource Access Program (X01), Research Project Grant (R01), and Administrative Review Supplement awardees discussed their scientific discoveries in disease mechanisms, pathogenic processes, disease progression, and clinical responses achieved or made possible with resources available through the NIDDK-CR. Examples of key findings are highlighted below.

- A general increase in small peptide abundance in interstitial cystitis/bladder pain syndrome (IC/BPS)–specific peptides that may act as putative biomarkers was identified in human urine samples from female patients with IC/BPS who participated in the NIDDK-sponsored Multidisciplinary Approach to the Study of Chronic Pelvic Pain Research Network.

Additionally, the well-known biomarker antiproliferative factor peptide was not consistently found in IC/BPS patient urine, potentially limiting its utility as a biomarker.

- A new pathway for CVD complications in T1D was discovered using longitudinal serum samples from the Diabetes Control and Complications Trial and Epidemiology of Diabetes Interventions and Complications Cohort.
- A panel of metabolites and clinical phenotype variables that offer potential as a screening tool for metabolic dysfunction–associated steatotic liver disease was determined using Nonalcoholic Steatohepatitis Clinical Research Network highly phenotyped cohorts.
- Higher baseline levels of tumor necrosis factor receptor 1 (TNFR1), TNFR2, TNF-alpha, and interleukin 10 were associated with increased risks of end-stage kidney disease; the association was identified using stored serum samples from the African American Study of Kidney Disease and Hypertension.

The speakers noted several helpful approaches when working with the NIDDK-CR: (1) Request the minimum amount of sample needed for a study because minimal impact requests may be routed through a more streamlined process compared to a full proposal; (2) Interact with and leverage the expertise of NIDDK-CR personnel; and (3) Contact NIDDK-CR support staff as early as possible to assess sample availability.

NIDDK-CR Data Secondary Analyses, Qualitative and Mixed Methods Studies

Data science awardees presented their discoveries in disease mechanisms, pathogenic processes, progression, and clinical responses achieved or made possible with resources available through the NIDDK-CR. Research methods included statistical analysis, meta-analysis, unsupervised machine learning, and mathematical models. Examples of research questions and findings are summarized below.

- The variability in adiposity associated with increased risk of morbidity and mortality in type 2 diabetes (T2D) using data from the Action for Health in Diabetes (commonly called Look AHEAD) Trial.
- Data from the HALT-Polycystic Kidney Disease (PKD) A and B and Focal Segmental Glomerulosclerosis (FSGS)-FONT studies were used to validate two surrogate endpoints—glomerular filtration rate slope and albuminuria—for chronic kidney disease (CKD) Epidemiology Collaboration Clinical Trials (CKD-EPI CT).
- Six distinct clusters of islet antibody development—with respect to timing, type, and titer—were identified in individuals with T1D ages 1 to 12 with genetic susceptibility, all using data from The Environmental Determinants of Diabetes in the Young (TEDDY) study.
- C-peptide preservation was proportional with HbA1c improvements over the time course investigated using data from the T1D TrialNet, suggesting that C-peptide is a reasonably likely surrogate endpoint in T1D. These studies have informed regulatory submissions and approvals.

The speakers noted several helpful approaches when working with the NIDDK-CR: (1) Check all the necessary requirements from the Institutional Review Board; (2) Ensure proper data management and security processes; (3) Keep the NIDDK updated on the manuscript; (4) Strictly follow the acknowledgement guidelines; and (4) Work closely with grants and contracts offices to explain the nature of data-use agreements.

Demonstration Workshops

In the "Accessing NIDDK Central Repository Resources Workshop," participants learned from the NIDDK-CR support team at Booz Allen Hamilton how to navigate and access resources in R4R. They also discussed challenges with and opportunities for using secondary data and identifying the appropriate data sets and biospecimens for research.

During the "NIDDK-CR Analytics Workspace Demonstration Workshop," participants learned how to navigate and access NIDDK-CR resources and learned of the launch of the first in a series of NIDDK-CR Data Challenges. The aim of this initial Challenge is to enhance data shared through the NIDDK-CR for AI-driven secondary research. The NIDDK is soliciting innovative approaches to improve the utility of longitudinal studies focused on T1D, which include TEDDY and TrialNet. Further details can be accessed from the challenge.gov website. Participants also observed a demonstration of the new analytics workspace (Amazon Web Services Service Workbench) that will be piloted in the Data-Centric Challenge.

Summary and Conclusion

Data and biospecimen repositories, whether stewarded by the NIH or supported by the NIH in the extramural community, are a significant research investment. This 2-day workshop provided participants the opportunity to learn about the foundation of the NIDDK-CR and its influence on scientific discoveries as an international resource, efforts to modernize the NIDDK-CR, and NIH's and NIDDK's commitment to driving innovation and enhancing utilization of the NIDDK-funded resources. Experts in the field highlighted that data-driven research and clinical care are experiencing exciting developments. A roadmap for ensuring high-quality data collection, sharing, and reuse of data collected through biomedical research and in the course of clinical care can be addressed by a life cycle approach. The NIDDK-CR Data-Centric Challenge will help inform expanding NIDDK's data ecosystem.

Planning Committee

- Beena Akolkar, Ph.D., NIDDK, NIH
- Dee Cooper, M.H.A., NIDDK, NIH
- Sharon Lawlor, M.B.A., NIDDK, NIH
- Chris Mullins, Ph.D., NIDDK, NIH
- Ludmilla Pawlikowska, Ph.D., NIDDK, NIH
- Rebecca Rodriguez, Ph.D., NIDDK, NIH

Moderators

- Beena Akolkar, Ph.D., NIDDK, NIH
- Dee Cooper, M.H.A., NIDDK, NIH
- Gregory Germino, M.D., NIDDK, NIH
- Mike Keller, Ph.D., Booz Allen Hamilton
- Ludmilla Pawlikowska, Ph.D., NIDDK, NIH
- Erik Schultes, Ph.D., Leiden University
- Lisa Spain, Ph.D., NIDDK, NIH

Speakers

- Patricia Flatley Brennan, Ph.D., RN, NLM, NIH
- Eric Brunskill, Ph.D., NIDDK, NIH
- Regina Bures, Ph.D., NHLBI, NIH
- Noël Burtt, Broad Institute
- Juhi Chaudhari, M.S., Tufts University School of Medicine
- James Checco, Ph.D., University of Nebraska-Lincoln
- Theresa Chen, M.D., M.H.S., University of California San Francisco
- Anya Dabic, Booz Allen Hamilton
- Jen Darragh, M.L.I.S., Duke University
- Alissa Fujimoto, M.S., Kylie Snow, M.P.H., Booz Allen Hamilton
- Gregory Germino, M.D., NIDDK, NIH
- Jeffrey Grethe, Ph.D., University of California, San Diego
- Joseph Hedrick, Ph.D., Critical Path Institute
- Marianne Henderson, M.S., CPC, PMP, NCI, NIH
- Sashawn Lawrence, M.P.H., Booz Allen Hamilton
- Rebecca Li, Ph.D., Vivli
- Myra Lipes, M.D., Harvard Medical School
- Sejal Mistry, Ph.D., University of Utah Health
- Griffin P. Rodgers, M.D., MACP, NIDDK, NIH
- Rebekah Rasooly, Ph.D., NICHD, NIH
- Veerasamy Ravichandran, Ph.D., NIDDK, NIH
- Rebecca Rodriguez, Ph.D., NIDDK, NIH
- Prasanna Santhanam, M.D., M.B.B.S., Johns Hopkins Medicine
- Yasir Tarabichi, M.D., M.S.C.R., MetroHealth and Case Western Reserve University
- Robin Taylor, M.L.L.S., NLM, NIH
- M. Todd Valerius, Ph.D., Brigham and Women's Hospital
- Miriam Vos, M.D., M.S.P.H., Emory University School of Medicine
- Xujing Wang, Ph.D., NIDDK, NIH