January–February 2023

Journal of Health Care Compliance

DOJ and Sentencing Commission Guidelines: Road Map to Compliance Effectiveness

Breaking Down the Last Barriers to Public Access and Trust: Navigating an Emerging Landscape of Federal Data Management and Sharing Requirements

CIAs and IAs in 2022 – A Year in Review

The HIPAA Right of Access Initiative: A Continued OCR Enforcement Priority

Use of PHI for Non-Patient Purposes

Wolters Kluwer

10041440-0370

Breaking Down the Last Barriers to Public Access and Trust: Navigating an Emerging Landscape of Federal Data Management and Sharing Requirements

Greg Smith / Claire Woodruff / Marisa Zuskar

Greg Smith is a Higher Education Consulting Manager at Huron Consulting Group.

Claire Woodruff is a Higher Education Consulting Analyst at Huron Consulting Group.

Marisa Zuskar is a Higher Education Managing Director at Huron Consulting Group.

According to the *South China Morning Post*, a 55-year-old man who was experiencing shortness of breath and a fever in Wuhan, China was the first confirmed case of COVID-19, tracing back to November 17h 2019; *patient zero*.[1] By March 2020, travelers were taking their last flights for the next months and, for some, years, and the doors of shops, restaurants, and offices were closing for the foreseeable future. Conspiracy theories were abundant, and state and national officials were giving daily, often conflicting, updates. At the time, there was a podcast called *TWiV* ("The Week in Virology"). *TWiV* long predated the pandemic and was hosted by virology, immunology, microbiology, and public health professors from Columbia University Medical Center, University of Florida, and University of Michigan.[2] In early 2020, the podcast was centered around breaking down what we knew about COVID-19; the hosts used episodes to dissect the publicly available scientific data by reviewing studies and outcomes through careful analysis of available literature—most of which is the result of federally funded research. In chaotic circumstances, the best source of control is information.

Access to scientific data and research results turned out to be the lynchpin to our rapid global response to COVID-19. Federal agency public access and data management and sharing policies enabled and fueled a rapid deployment of safety precautions, adoption of personal protective equipment, and the development and expedited roll out of vaccines. And now that the pandemic seems to be managed, there are a plethora of other challenges which the United States (and the

world) can begin and continue to address with free and open access to the results of research—which benefits significantly from taxpayer investment. Federal agency administrators and politicians are now leading the charge to enhance and enforce this ethos via new federal mandates and policies.

The First Frontiers of Federal Public Access Policies

In 2013, the White House Office of Science and Technology Policy ("OSTP") issued an inaugural memo directing federal agencies to develop plans to support increased public access for research results,[3] and since then every agency has developed a policy in accordance with that memo.[4] At the time, the Obama Administration established this policy framework because they were "committed to the proposition that citizens deserve easy access to the results that their tax dollars have paid for,"[5] a direct response to over 65,000 people signing a petition to require free access to scientific journal articles which arose from taxpayer-funded research.[6] But, fast forward to 2019—a Government Accountability Office ("GAO") report found that most agencies subject to the 2013 memoranda had not fully complied with the guidelines and had not made data easily findable or ensured that researchers followed agency public access rules.[7]

This past August, the Biden White House OSTP doubled down by issuing another memo to agency heads, in part driven by the lessons learned from the COVID-19 pandemic, providing even more focused guidance for updating and expanding all federal public access policies to make publications of taxpayer funded research free and openly available.[8] According to the memo, since 2013, 8 million scholarly publications have become available to the public, and 3 million people read these articles every day. The memo cites this impact as the impetus to vastly expand the principles established in 2013. In 2020, OSTP estimated that US taxpayer supported research produced 195,000 to 263,000 articles, or up to 9 percent of 2.9 million papers published worldwide that year.[9]

The National Institutes of Health ("NIH") has been championing public access and is a trailblazer through its own policy framework which is now 25 years in the making, first established through initial iterations of policies and notices dating back to 2003.[10] NIH's newest effort—a comprehensive Data Management and Sharing Policy ("DMSP")—is effective as of January 2023 and it will require researchers and institutions to make data available and accessible at time of publication and that each research proposal and award have a Data Management and Sharing plan ("DMS plan").[11]

Understanding the New NIH DMSP Policy

Background

The latest NIH DMSP is built on decades of similar policies, such as their foundational 2003 NIH Data Sharing Policy,[12] 2008 Genome-Wide Association Studies Policy,[13] 2014 Genomic Data Sharing ("GDS") Policy,[14] and their 2016 Policy on the Dissemination of NIH-Funded Clinical Trial Information.[15] Planning and design for the newest policy began in 2015 when NIH "initiated the development of a more comprehensive policy" through a series of requests for public comment to seek out feedback from the scientific community via the Federal Register.[16] The NIH requested that the community weigh in on the definition of scientific data, suggestions for what elements specifically should be required in data management and sharing plans, and feedback regarding the timeline and roadmap to a phasic approach toward implementation.[17] NIH's solicitation of public input was compiled and posted publicly,[18] and the feedback was incorporated into the final DMSP.

Purpose

The intent of this new policy is twofold. Free and open public access enables rapid identification and resolution of global public health challenges. NIH's commitment to making research results public is also to foster transparency. Per the new policy, "data sharing enables researchers to rigorously test the validity of research findings." Sharing data sets enables them to be combined and reused in order to build layers and add dimensions to these repositories which equip researchers to "explore new frontiers of discovery."

Definitions

The new policy will apply to "all research, funded or conducted in whole or in part by NIH that results in the generation of scientific data." It defines scientific data as "recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings, regardless of whether the data are used to support scholarly publications." The DMSP is also clear on what NIH does **not** consider scientific data:

- Data *not* necessary for or of sufficient quality to validate and replicate research findings,
- Laboratory notebooks,
- Preliminary analyses,
- Completed case report forms,
- Drafts of scientific papers,
- Plans for future research,
- Peer reviews,
- Communications with colleagues, or
- Physical objects, (*e.g.*, laboratory specimens)

The DMSP's definitions also operationalize data management, data sharing, and metadata. Thoughtfully considering how NIH defines what is and is not scientific data, and what constitutes management and sharing, will inform how institutions and researchers clean, calibrate, and categorize research projects. Understanding these definitions also will help institutions define which research projects and outputs which may be in-scope, and decide how to meet the requirements set forth in the NIH DMSP.

Scope and Requirements

A DMS plan will be required for all applications for funding regardless of funding level. Plans will need to outline how scientific data and subsequent metadata will be managed and shared, and take into account any restrictions or limitations. The NIH Institute, Center or Office must approve compliance with the DMS plan to resolve any issues prior to award. NIH has published a "format page," which includes templatized elements of what constitutes a DMS plan. It can be leveraged by researchers and institutions as a "plug and play" checklist for developing DMS plans.[19]

Considerations and Nuances for NIH DMSP Compliance

NIH has complemented its policy with a compendium of supplemental notices on topics such as the sharing of human research participant data,[20] American Indian/Alaska Native participant data,[21] allowable costs,[22] repository selection,[23] and protecting privacy.[24] In addition, within the policy itself there are implications for institutional reporting and areas which impact multiple functional areas for grantees, such as university libraries, IT offices, and sponsored program offices.

Big-ticket, tactical takeaways from the DMSP which may represent the need for strategic change planning for institutions include:

- ***Use of data repositories:*** NIH is encouraging the use of established repositories[25] and cites the improved adoption of the FAIR principles—Findable, Accessible, Interoperable, Re-usable - as its rationale. Some NIH programs may specify certain data repositories. In circumstances where there is no pre-established data repository, or when NIH has not prescribed the use of an existing repository, NIH's supplemental guidance

establishes characteristics which should be exemplified for repositories which are considered for use, such as specific scenarios which may benefit from certain types of repositories (*e.g.*, cloud-based repositories for large data sets).

- ***Incorporation of new allowable costs:*** The new policy also defines allowable costs associated with implementation. Per NIH, allowable costs include: curating data and developing supporting documentation; preserving and sharing data through selected repository for long-term access and preservation; de-identifying data; local data management considerations; and preparing metadata.[26] Unallowable costs include infrastructure costs included in indirect costs, routine conduct of research costs (including those associated with collecting or gaining access to research data), and costs that are inconsistently charged or double charged as both direct and indirect costs.
- ***New proposal submission, award acceptance, and progress reporting requirements:*** For NIH proposals submitted in January 2023 and after, a compliant DMS plan is required. Sponsored projects offices and authorized organizational officials ("AORs") will need to establish mechanisms for ensuring that they are uniformly reviewing proposals for the existence of such plans and consistently providing feedback plan content. If proposals are awarded, submitted DMS plans will be incorporated by reference into awards, and institutions will be required to treat data in accordance with those plans. Grantees will be required to report the progress of their DMS plan in Research Performance Progress Reports ("RPPR"). Reporting may be more frequent based on direction from NIH Program Officers, who may elect to require updates or revisions to data management and sharing plans. NIH will review for compliance annually. Failure to comply with DMS plans "may result in an enforcement action, including additional special terms and conditions or termination of award, and may affect future funding decisions."

Not Just NIH: OSTP's Latest Directive to Other Federal R&D Sponsors

OSTP's latest policy guidance instructs all agencies to follow NIH's lead as soon as possible and without delay; the drop dead date for agency compliance being December 31, 2025. President Biden has long been lambasting the shortcomings of U.S. public access policy. In his remarks to the American Association for Cancer Research in 2016, which were highlighted in the announcement from the White House which accompanied the August OSTP memo, Biden said:[27]

> *"For anyone to get access to that publication, they have to pay hundreds, or even thousands, of dollars to subscribe to a single journal. And here's the kicker — the journal owns the data for a year. The taxpayers fund $5 billion a year in cancer research every year, but once it's published, nearly all of that taxpayer-funded research sits behind walls. Tell me how this is moving the process along more rapidly."*

And while it is clear from OSTP's latest memo and the Biden comments that this is a priority, NIH's long, calculated road to its latest DSMP and our experience with the federal government's rollout of other sweeping policies and initiatives for R&D (*e.g.*, topics as research security, or uniform federal grant proposal application forms) might mean that institutions and grantees are in a "wait and see" mode for some time. But timeline details contained within the OSTP memo might predict that other large grant-making agencies are next on the shortlist for implementation of DMSP's similar to the NIH.

Timeline Details and Impacted Agencies

Agencies who spend more than $100 million on research and development annually are expected to submit their draft policy no later than 180 days after the memo issue date, while agencies with less than $100 million in annual R&D expenditures have 360 days. This means that in early 2023 we should see new DMSPs from Department of Defense, NASA, Department of Energy, National Science Foundation, USDA, Departments of Interior, Transportation, Commerce, the EPA, Homeland Security, and Veterans Affairs, to name a few. Some, such as NSF, already have long standing public access policies similar to NIH.[28] But one thing is clear: More prescriptive policies are coming. Researchers and institutions who rely on federal funding for scientific research will see the impact of these policies in the coming years.

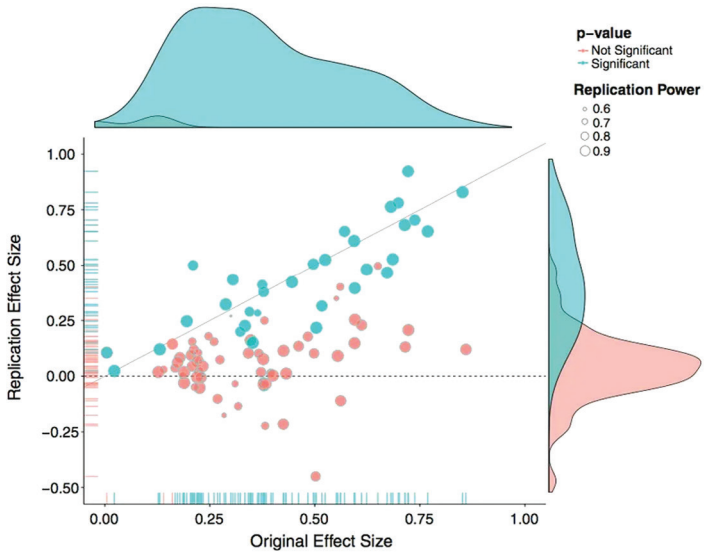
Overview of Key Dates

- **February 21, 2023:** Federal agencies with more than $100 million in annual R&D expenditures will update or develop draft DMSPs.
- **August 20, 2023:** Federal agencies with less than $100 million in annual R&D expenditures will update or develop draft DMSPs.
- **December 31, 2024:** Agencies should have finalized and published full DMSPs for public comment. They should have an effective date no later than one year after their plan was published.
- **December 31, 2025:** Latest date for all federal agencies to have updated public access policies in effect.

The 'Reproducibility Crisis' and Public Distrust

In addition to the public health and broader technology innovation benefits of a free and open environment for sharing the results of publicly funded research, another aspect of public access policies is targeting some of this public distrust, allowing for a greater reuse of data and supporting continued inquiry. This shift involves increasing responsibility for uniformly managing and sharing data for both researchers and their institutions.

Data management and sharing in the scientific research space has been an important topic in recent years due to what has been labeled the "reproducibility crisis." In a study published in August 2015 in *Science*, researchers attempted to replicate results from 100 psychology studies from top journals and "overwhelmingly" failed.[29] In an interview with Vox, the lead author on the study, Professor Brian Nosek from University of Virginia said that "pre-registration [putting the study design on an open database before running the study, so you can't change the methods if you get results you don't like] is an important feature of doing confirmatory analysis in research." Professor Nosek's own study was only able to powerfully replicate the results of studies 36 percent of the time.[30] In the figure from the article, each dot represents a study where you can see the replication effect size relative to the original effect size; the diagonal line



represents replication effect size equal to the original and points below the dotted line showed effects in the opposite direction of the original study. By mandating that federally sponsored research projects publicly share research data, specifically with the goal of validation and replication (as is the case in the NIH purpose statement), the government is taking direct aim at the crisis which Professor Nosek's study highlights, and which underpins public distrust.

Prior policies and frameworks from sponsors like NIH seemed to frame such responsibilities as expectation, rather than a mandate. The newest policies clearly define the requirements as a facet of compliantly stewarding federally funded research projects. By making clear that agencies will be involved in reviewing proposed plans as well as monitoring progress for executing against those plans during the course of projects, and stipulating that non-compliance may impact future funding decisions, the Biden OSTP is clearly layering teeth into the Obama administration's initial 2013 framework.

Preparing for the New Culture of Data Management and Sharing

Institutions must be prepared to implement policies and practices emerging from federal requirements and to adapt to the evolving data management and sharing culture as agencies adopt new, stronger policies. Institutions can consider an "assess and implement" approach to tailor implementation to their organizational structure and programmatic focus.

- **Establish data management and sharing project governance**
  - ❏ Thoughtfully consider institutional stakeholders which may include existing data governance committees or councils, researchers and programmatic leaders and laboratory staff, as well as business and operational functions such as risk, IT, library, and legal groups.
- **Leverage existing investments**
  - ❏ Engage existing data librarians to help provide information about repositories and general data management and sharing practice recommendations.
  - ❏ Inventory institutionally housed research data repositories or subscriptions to external data or code repositories, and employ electronic lab notebook technologies already; cataloguing those resources and ensuring research data management and sharing practices are well integrated with those resources will enable success.
- **Recognize and address key compliance risk areas**
  - ❏ *Data security:* Discover and understand the various mechanisms by which data is stored, managed, and shared and determine alignment with institutional information security protocols to ensure that project specific DMS plans incorporate best practice cyber hygiene.
  - ❏ *Privacy:* Develop parameters and central standards for how personally identifiable information ("PII") or protected health information ("PHI") might be de-identified or otherwise unable to be attributed to a specific individual in datasets uploaded into repositories. Some of the foundational frameworks for privacy compliance include:
    - Health Insurance Portability and Accountability Act (HIPAA) 45 CFR Parts 160, 162, and 164
    - Health Information for Technology and Economic and Clinical Health Act (HITECH) Title 13 U.S.C. §
    - Family Education Rights and Privacy Act (FERPA) 20 U.S.C. § 1232g; 34 CFR Part 99
    - General Data Protection Regulation (GDPR) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016

  - The Federal Trade Commission's Health Breach Notification—16 CFR §318
  - Various state laws
- **Embark on a project toward establishing institutional resources for complying with federal data management and sharing policies**
  - ❑ *Initiate discovery:* Hold focused meetings with key stakeholders to understand the current state of your institution's readiness to comply with new and anticipated federal mandates.
  - ❑ *Formalize and validate observations:* Document the outcome of your discovery activities and validate those findings as a basis for future action.
  - ❑ *Develop recommendations:* Armed with what you have observed, develop targeted actions toward implementing an institutional data management and sharing program.
  - ❑ *Build an implementation roadmap:* Leveraging a phased and team-based approach, define the workstreams which may be required for execution, and set a timeline with key milestones and routine statusing to ensure progress.

Conclusion

The writing is on the wall: academic medical centers, hospitals, universities and other institutions who receive federal support for research and development will need to be ready to comply with the new NIH DMSP. Organizations who pursue and administer federal R&D awards and their researchers will also need the infrastructure to support compliance with all anticipated federal agency policies following OSTP's most recent guidance with further reaching impacts. As in the early days of the COVID-19 pandemic, when access to information was one of the few areas of control we had, your institutional leaders will need to be informed in order to chart their own path as the ecosphere of taxpayer-funded research results stabilizes toward one which prioritizes both public access and public trust.

Endnotes

1. Coronavirus: China's first confirmed Covid-19 case traced back to November 17, South China Morning Post (March 13, 2020), *https://www.scmp.com*.
2. This Week in Virology, *https://www.microbe.tv*.
3. Office *of Science and Technology Policy, "Memorandum for the Heads of Executive Departments and Agencies" (Feb. 22, 2013), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf*.
4. OSTP *Congressional Report, https://www.whitehouse.gov/wp-content/uploads/2022/02/2021-Public-Access-Congressional-Report_OSTP.pdf*.
5. Expanding Public Access to the Results of Federally Funded Research, *https://www.whitehouse.gov*.
6. We the People: Your Voice in Our Government, Increasing Public Access to the Results of Scientific Research, *https://petitions.obamawhitehouse.archives.gov/petition/require-free-access-over-internet-scientific-journal-articles-arising-taxpayer-funded/*.
7. Federal Research: Additional Actions Needed to Improve Public Access to Research Results, GAO-*20-18 (Nov. 21, 2019), https://www.gao.gov/products/gao-20-81*.
8. OSTP Issues Guidance to Make Federally Funded Research Freely Available Without Delay, (Aug. *25, 2022), https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/*.
9. White House requires immediate public access to all U.S.-funded research papers by 2025, *https://www.science.org/content/article/white-house-requires-immediate-public-access-all-u-s--funded-research-papers-2025*.
10. NIH, *Statement on Final NIH Policy for Data Management and Sharing (Oct. 29, 2020), https://www.nih.gov/about-nih/who-we-are/nih-director/statements/statement-final-nih-policy-data-management-sharing*.
11. *https://www.nih.gov/about-nih/who-we-are/nih-director/statements/statement-final-nih-policy-data-management-sharing*.
12. *https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html*.
13. *https://grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html*.
14. *https://grants.nih.gov/grants/guide/notice-files/not-od-14-124.html*.
15. *https://grants.nih.gov/grants/guide/notice-files/not-od-16-149.html*.
16. *https://grants.nih.gov/grants/nih-public-access-plan.pdf*.

17. NOT-OD-19-014: Request for Information (RFI) on Proposed Provisions for a Draft Data Management and Sharing Policy for NIH Funded or Supported Research, *https://grants.nih.gov/grants/guide/notice-files/NOT-OD-19-014.html*.
18. *https://osp.od.nih.gov/wp-content/uploads/RFI_Final_Report_Feb2020.pdf*.
19. OMB No. 0925-0001 and 0925-0002, DMS Plan Format Page, *https://www.nih.gov/*.
20. *https://grants.nih.gov/grants/guide/notice-files/NOT-OD-22-213.html*.
21. *https://grants.nih.gov/grants/guide/notice-files/NOT-OD-22-214.html*.
22. *https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-015.html*.
23. *https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html*.
24. *https://grants.nih.gov/grants/guide/notice-files/NOT-OD-22-131.html*.
25. NIH, Repositories for Sharing Scientific Data, *https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data*.
26. NIH, Budgeting for Data Management & Sharing, *https://sharing.nih.gov/data-management-and-sharing-policy/planning-and-budgeting-for-data-management-and-sharing/budgeting-for-data-management-sharing*.
27. OSTP Issues Guidance to Make Federally Funded Research Freely Available Without Delay (Aug. 25, 2022), *https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/*.
28. National Science Foundation, Dissemination and Sharing of Research Results, *https://www.nsf.gov/bfa/dias/policy/dmp.jsp*.
29. Vox, "Scientists often fail when they try to replicate studies. This psychologist explains why." (Aug. *27, 2015), https://www.vox.com/2015/8/27/9212161/psychology-replication*.
30. *https://www.science.org/doi/10.1126/science.aac4716*.