



National Institutes of Health  
*Office of Science Policy*

**NEW!**

# NIH Policy for Data Management & Sharing

Presentation to Johns Hopkins University  
April 15, 2021

**Taunton Paine, MA**

Director, Scientific Data Sharing Policy Division

Office of Science Policy

National Institutes of Health

[sciencepolicy@mail.nih.gov](mailto:sciencepolicy@mail.nih.gov)

# Overview

- **Evolution of NIH Scientific Data Sharing Efforts**
- **Developing the NIH Data Management and Sharing Policy**
- **Overview of the NIH Data Management and Sharing Policy**
- **Implementation Activities**
- **Questions/Discussion**

# Why Share Research Data?

- **Advance rigorous and reproducible research**
  - Enable validation of research results
  - Make high-value datasets accessible
  - Accelerate future research directions
  - Increase opportunities for citation and collaboration



- **Promote public trust in research**
  - Foster transparency and accountability
  - Demonstrate stewardship over taxpayer funds
  - Maximize research participants' contributions
  - Support appropriate protections of research participants' data

# 1980s: Grantee Data & Federal Records

- **1980 SCOTUS: *Forsham v. Harris***<sup>1</sup>

- NIH-funded University Group Diabetes Program found commonly prescribed diabetes drug had 2.5x increased risk of death from heart disease; FDA used the findings to create labeling requirements
- A group of scientists requested raw data (data forms and computer tapes) through the Freedom of Information Act (FOIA); Supreme Court affirmed grantees' data were not agency records and were not subject to FOIA

- **1988 PHS Policy on Distribution of Research Resources**

- Expected unique research resources be made readily available for research purposes to the scientific community after publication

- **1989 NHLBI “L’Enfant Memo”**<sup>2</sup>

- NHLBI Director Claude L’Enfant created policy that grantees and contractors would data available from clinical trials, epidemiological studies, and other large-scale studies within three years of major publications
- Intended to maximize the Federal government’s investment in research

1- <https://www.jameslindlibrary.org/articles/the-trials-and-tribulations-of-the-university-group-diabetes-program-lessons-and-reflections/>

2- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4286227/>

# 1990s: Sharing and Transparency

- **1996 Human Genome Project “Bermuda Principles”**
  - Expected immediate and broad sharing of data
- **1997 FDA Modernization Act and ClinicalTrials.gov**
  - Established ClinicalTrials.gov
- **“Six Cities Study” and 1999 Shelby Amendment**
  - NIH-funded **“Six Cities Study”** found in 1993 that fine particle air pollution (>2.5 microns) reduced lifespans
  - Study findings were used by EPA in 1997 for regulations
  - Findings were challenged by groups who sought access to data, but access was denied on the basis of informed consent and privacy
  - Sen. Shelby created “Shelby Amendment,” later interpreted that federally-funded data underlying regulations must be made available
  - **Beginning of broad consensus that data underlying publicly-funded studies should be made available**



# 2000s: Towards Sharing by Default

- **2003 NIH Data Sharing Policy**

- Expects applicants seeking \$500,000 or more to submit a data sharing plan

- **2006 and 2014 NIH GWAS and Genomic Data Sharing policies**

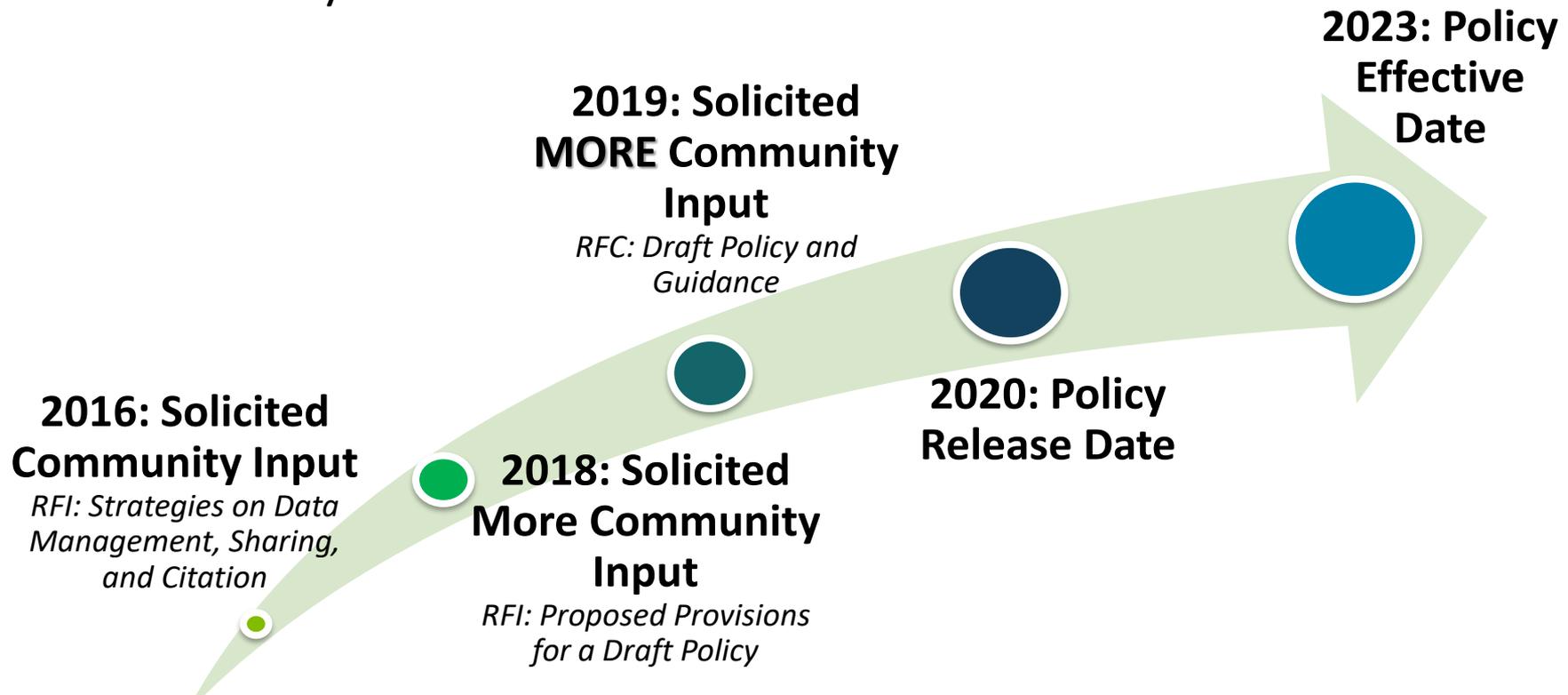
- Established expectation that NIH-funded studies generating “large-scale” genomic data would rapidly submit those data to NIH-designated repositories
- Expects consent for generation, future use, and sharing of human data; institutional certification indicating data use limitations; and oversight by NIH Data Access Committees

- **2013 “Holdren Memo”**

- Required by 2010 America COMPETES Reauthorization Act
- White House Office of Science and Technology Policy directed all Federal Depts. and Agencies with R&D budgets \$100 million+ to work toward requiring data management and sharing plans be submitted from all applicants

# An Iterative Policy Development Process

- Sought public comment at multiple points throughout development
- Tribal Consultation\*  
*\*Details provided in [“NIH Tribal Consultation Report: NIH Draft Policy for Data Management and Sharing”](#)*
- Obtained input from other government agencies & Secretary’s Advisory Committee for Human Research Protections





NATIONAL INSTITUTES OF HEALTH

# NIH Policy for Data Management and Sharing

- **Aims to foster a culture of data stewardship**
- **Submission of Data Management & Sharing Plan for all NIH-funded research** (*how/where/when*)
- **Compliance with the NIH-approved Plan** (*may affect future funding*)
- **Effective January 25, 2023** (*replaces 2003 Data Sharing Policy*)
- **Additional resources available to assist**

# The Devil is in the Details...

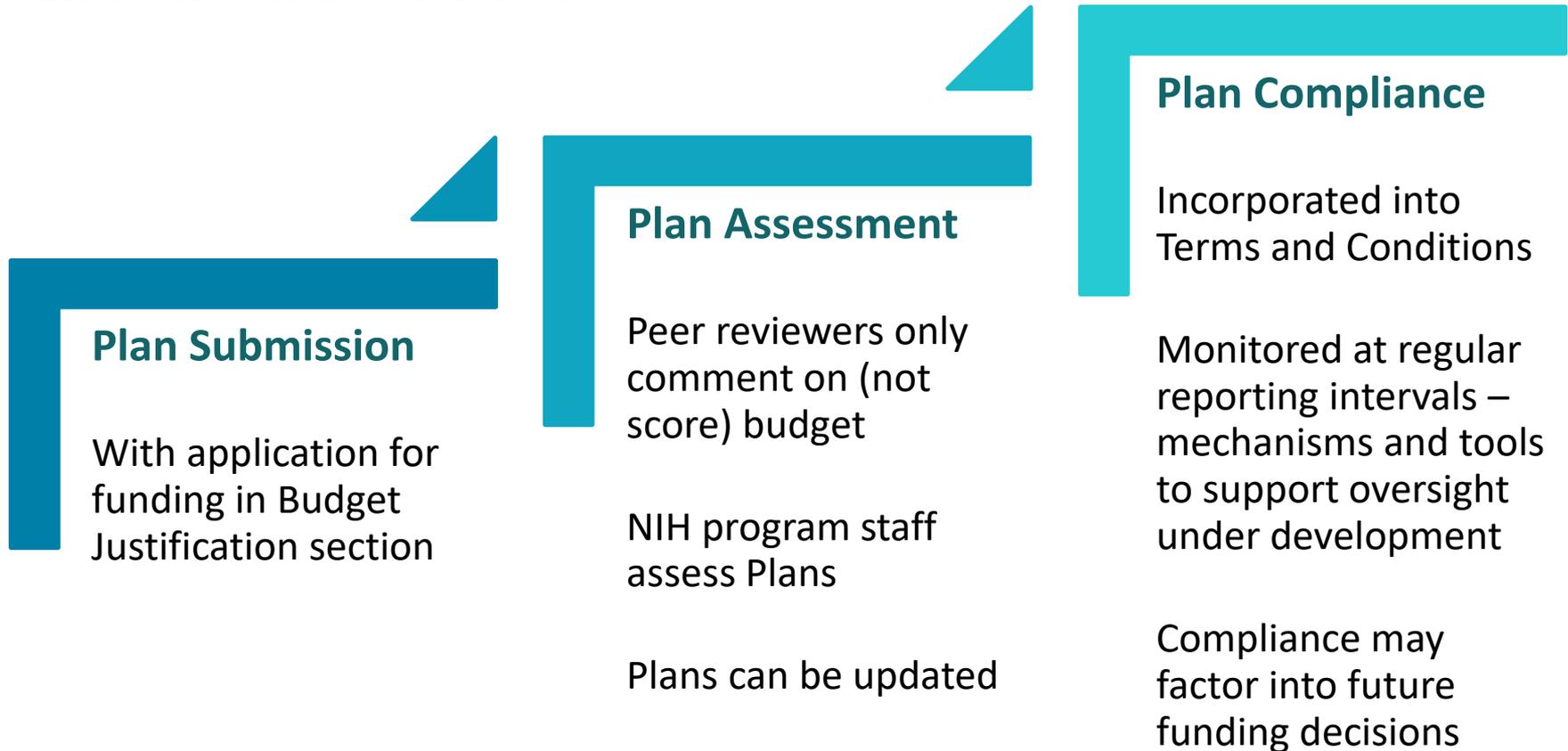
- **Scope:** All NIH-supported research generating *scientific data*, i.e., those data commonly accepted in scientific community “as of sufficient quality to validate and replicate research findings”
- **SHARING SHOULD BE ...**
  - **The default practice**
    - Maximize appropriate data sharing; Plans may justify exceptions (i.e., ethical, legal, technical factors)
  - **Responsibly implemented**
    - Plans should outline protection of privacy, rights, and confidentiality; existing laws, regulations, and policies continue to apply
  - **Timely**
    - **When to share data?** No later than publication or end of award (if unpublished); other relevant requirements and expectations (e.g., repository policies, retention requirements, journal policies) for minimum time frames



'Which brings us to my next point.'

# Plan Submission and Review

## Extramural Grant Awards\*



*\*Analogous requirements for contracts, OTAs, IRP*

# Repository Selection for Managing and Sharing Data

- **Use of established repositories are encouraged**
  - e.g., NIH [BMIC list of repositories](#)
- **Supplemental information provided to help investigators identify appropriate data repositories**
  - e.g., use of persistent unique identifiers, attached metadata, facilitates quality assurance
- **NIH ICs may designate specific data repository(ies)**





## Supplemental Info to the Policy: **Allowable Costs**

- **Reasonable costs allowed in budget requests**
  - Curating data/developing supporting documentation
  - Preserving/sharing data through repositories
  - Local data management considerations
- **NOT considered data sharing costs**
  - Infrastructure costs typically included in indirect costs
  - Costs associated with the routine conduct of research (e.g., costs of gaining access to research data)

# Resources for Implementation

## What's Next?

- **Outreach activities:** engaging community to make aware of the policy and understand implementation challenges (e.g., [April 28-29 2021 NAS workshop on the culture of data management & sharing](#))
- **Enhance compliance mechanisms:** building tools to streamline compliance and facilitate tracking, public posting of Plans
- **Harmonize data sharing expectations** and reduce redundancy
- **Develop resources:** develop FAQs, guidance, additional supp. information (i.e., for AI/AN data), cost estimating tools
- **Strengthen incentives** for data sharing, e.g., through data citation

*~2-year implementation window!*

# Thank You!

- [OSP Data Management and Sharing Website](#)
- [NOT-OD-21-013](#) – Final NIH Policy for Data Management and Sharing
- [NOT-OD-21-014](#) – Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan
- [NOT-OD-21-015](#) – Supplemental Information to the NIH Policy for Data Management and Sharing: Allowable Costs for Data Management and Sharing
- [NOT-OD-21-016](#) – Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research

**Questions?** [sciencepolicy@mail.nih.gov](mailto:sciencepolicy@mail.nih.gov)