

NIH Data Management and Sharing (DMS) Policy Overview

Soumya Korrapati, PhD

Office of the Director

Division of Cancer Biology/NCI

Overview of the NIH Data Management and Sharing (DMS) plan

Consists of 6 elements - **Findable, Accessible, Interoperable, Reusable data**

1. Sample type, amounts, data types; data preservation and sharing; metadata and associate documentation
2. Documentation of software tools
3. Documentation of data standards
4. Repositories, data reuse, timelines
5. Data access, distribution and reuse
6. Data management and oversight

DMS plan- Element 1

Element 1A: Describe ALL data types, not only omics data

- ✓ Source of data (i.e., species; human, mouse, zebra fish, human cell lines, mouse cell lines, PDX, primary tumor etc.; to determine access/distribution/reuse considerations)
- ✓ Sample number (to determine if Genomic Data Sharing (GDS) policy applies)**
- ✓ Amounts (volume) of data (to determine what the DMS budget will look like, if any!)

** There is NO need for a separate GDS plan/table.
GDS plan is harmonized into the DMS plan.

DMS plan- Element 1 (cont'd)

Element **1B**: Data types that will be **preserved and shared**.

- ✓ All data need not be shared
- ✓ Ethical, legal, and technical factors guide extent of preserving and sharing

Element **1C**: Share all **metadata** and other associated relevant documentation.

- ✓ Data about how a dataset or resource came about, but also how it is internally structured. e.g., the unit of analysis, collection method, sampling procedure, sample size, categories, variables, etc.
- ✓ Metadata have to be gathered by the researchers according to best practice in their research community and will be published together with the data.

DMS plan- Element 2

Element 2: Share **specialized tools, software and or code.**

- ✓ Information about accessing those tools/software
- ✓ Use of open-source code and tools

DMS plan- Element 3

Element **3**: Indicate what **data standards** will be applied to scientific data.

- ✓ data formats
- ✓ data dictionaries
- ✓ data identifiers, definitions etc.

DMS plan- Element 4

Element **4A: Repository(s)** where scientific data and metadata will be archived.

- ✓ Data preserved and shared through established repositories
- ✓ If no known established repositories, then use [Generalist Repositories](#)
- ✓ For plans with human data under *GDS policy*, data HAVE to be shared in [NIH-supported repositories](#)
- ✓ For plans with non-human data under *GDS policy*, data can be shared in any established data repository, whether NIH-supported or not.

DMS plan- Element 4 (cont'd)

Element **4B**: How archived data will be **findable and identifiable**.

- ✓ Accession numbers
- ✓ Digital object identifiers (doi)
- ✓ Unique persistent identifiers

Element **4C**: **When** and for **how long** data will be archived.

- ✓ Data should be made available at the time of publication or before the end of the award, whichever comes first
- ✓ For plans with data under *GDS policy*, data must be shared within 9 months after the full dataset is collected and data have been cleaned and Qced
- ✓ Repositories generally set time limits on data availability

DMS plan- Element 5

Element **5A**: Factors affecting **subsequent access, distribution or reuse of data**.

- ✓ Broad data sharing is highly encouraged
- ✓ Mention and justify if broad data sharing is not possible

Element **5B**: Indicate data will be **open access or controlled access**.

- ✓ Mention and justify if data will be controlled access
 - Data generated from human subjects (PDX, primary tumors, organoids, primary cell lines etc.) be deposited with controlled access

Element **5C**: Mention what protections for privacy, rights and confidentiality of human research participants will be followed.

DMS plan- Element 6

Element 6: Data **sharing and management** oversight at parent institution.

Summary

- NIH's goal is to promote a culture in which data management and sharing are recognized as an integral component of a biomedical research project, rather than an administrative or additive one.
- Division of Cancer Biology recognizes the initial challenges but believes data sharing will greatly benefit the cancer research community by reducing unnecessary data replication and wastage of precious resources, while facilitating transparency, reproducibility, discovery and innovation.
- Feel free to contact our DMS policy expert with any queries regarding the NIH Data Management and Sharing Policy.

[MORE
INFORMATION](#)



[CONTACT US](#)



Thank You



**NATIONAL
CANCER
INSTITUTE**

[cancer.gov](https://www.cancer.gov)

[cancer.gov/espanol](https://www.cancer.gov/espanol)