

Division of Cancer Biology New Grantee Workshop 2025

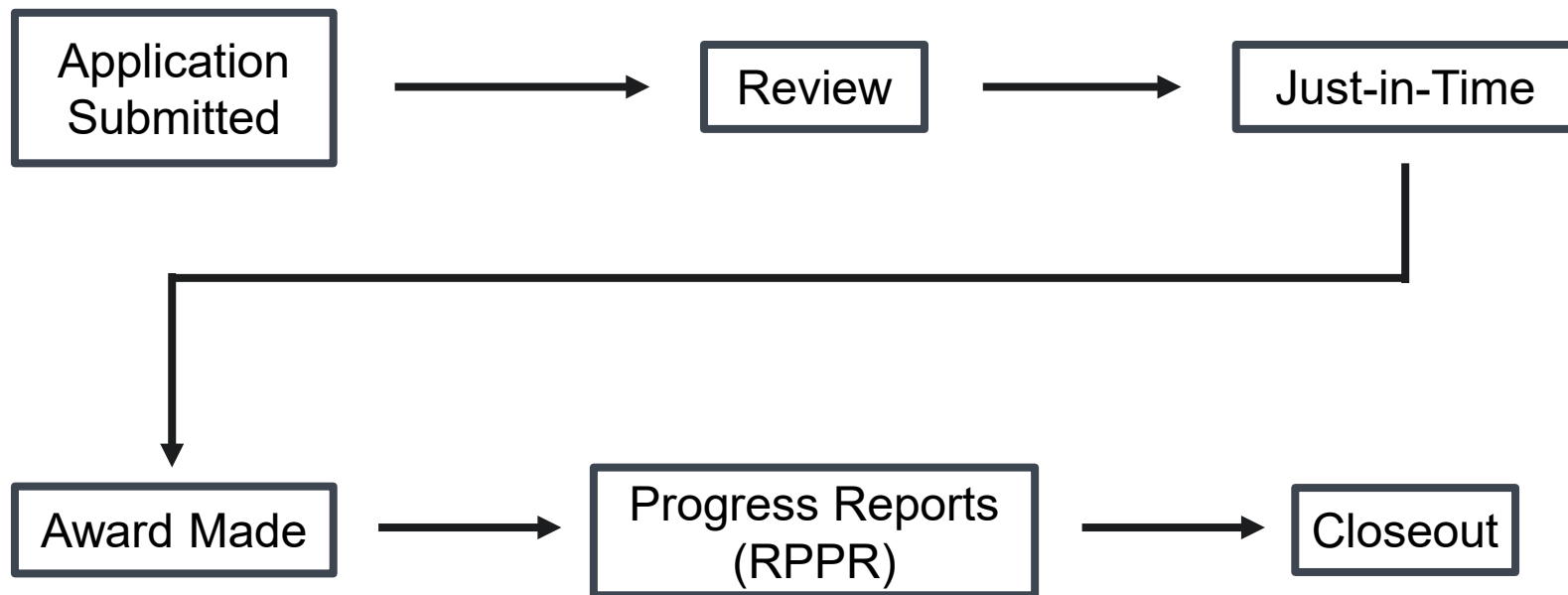
NCI Requirements Regarding Data

Justin Benavidez, PhD

Program Director & Genomic Program Administrator

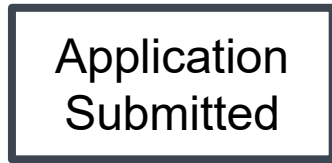
Tumor Metastasis Branch, DCB

A (Very) Simplified View of the NIH Grants Pipeline

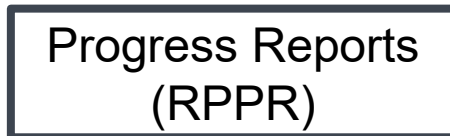


Considerations for Data Sharing are Made Throughout the Grant Pipeline

Write DMS Plan



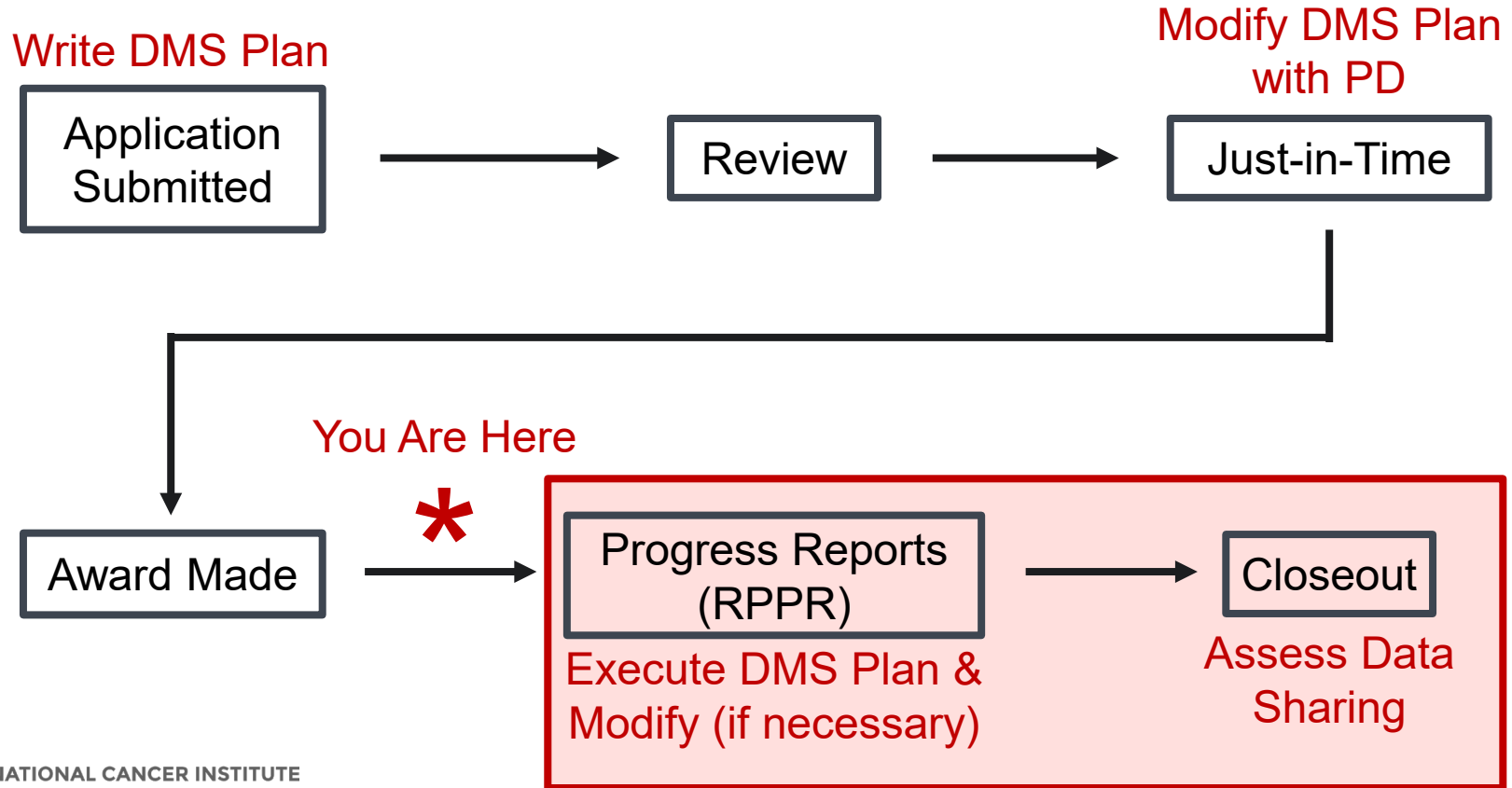
Modify DMS Plan with PD



Execute DMS Plan & Modify (if necessary)

Assess Data Sharing

Data Sharing Considerations Post-Award



NIH Data Sharing Policies: Applicability and Requirements

- **Data Management & Sharing (DMS) Policy (2023)**
 - Applies to all grants generating “scientific data”—all R01s & equivalent
 - PI must write & execute a DMS Plan with input from Program Director
- **Genomic Data Sharing (GDS) Policy (2014)**
 - Applies to grants generating “large-scale” genomic data
 - PI must specify restrictions on data sharing of human genomic data consistent with consent
 - Data must be deposited in an NIH-sponsored repository (dbGaP)

Post-Award DMS Policy Considerations

- **Data Management & Sharing (DMS) Policy (2023)**
 - Applies to all grants generating “scientific data”—all R01s & equivalent
 - PI must write & execute a DMS Plan with input from Program Director
- **Genomic Data Sharing (GDS) Policy (2014)**
 - Applies to grants generating “large-scale” genomic data
 - PI must specify restrictions on data sharing of human genomic data consistent with consent
 - Data must be deposited in an NIH-sponsored repository (dbGaP)

Making Changes to the DMS Plan Post-Award

- The DMS Policy **allows modifications** to DMS plans post-award when plans change:
 - New data types being generated
 - New samples with specific data sharing restrictions
- Send modifications to your Program Director ***proactively*** to avoid delays or other issues
- RPPRs will also include a question about changes to your DMS Plan
- Your PD will assess whether changes are appropriate and may ask clarifying questions

Reporting Data Sharing in Progress Reports (RPPRs)

- Annual progress reports now include a tabular section to report data generation, deposition, and release
- Your Program Director will review & determine whether your DMS plan is being followed and may reach out for more details

Data Type ^	Has data been generated to date? ▾	Has the data been shared? ▾	Status of Data Sharing ▾	Repository ▾	Unique Identifiers/Digital Object Identifier ▾
array-derived genotype data ...	Y	Y	Shared	Data Hub	ABC1234
Demographic Data ...	Y	N	Data not yet published will share in future	Not Applicable	Not Applicable
Phenotypic and clinical data ...	N	N	Not yet expected to be shared	Not Applicable	Not Applicable

**If data has not been generated and/or shared as outlined in the approved Plan, describe why, and identify any corrective actions that have or will be taken to comply with the approved plan.*

Post-Award GDS Policy Considerations

- **Data Management & Sharing (DMS) Policy (2023)**
 - Applies to all grants generating “scientific data”—all R01s & equivalent
 - PI must write & execute a DMS Plan with input from Program Director
- **Genomic Data Sharing (GDS) Policy (2014)**
 - Applies to grants generating “large-scale” genomic data
 - PI must specify restrictions on data sharing of human genomic data consistent with consent
 - Data must be deposited in an NIH-sponsored repository (dbGaP)

Transition to GDS Applicability: Genomic Data Thresholds

- While reviewing an RPPR, your Program Director may reach out for details to determine whether a threshold for “large-scale” genomic data has been surpassed—***even if your award was not initially subject to the GDS Policy***
- A “specimen” in this context depends on the technology used—***for single-cell techniques, each cell constitutes a specimen***

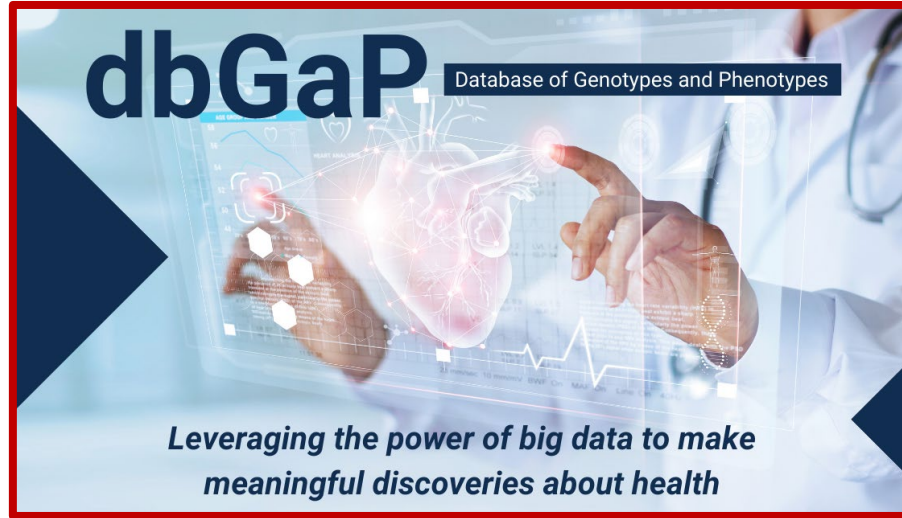
Data Type	# of Specimens (Human)	# of Specimens (Nonhuman)
GWAS	1000	500
<100 genes	1000	500
≥100 genes	100	50
Genome-wide	100	50
ChIP-seq	100	50
Metagenome	100	50

[More details on cancer.gov](https://www.cancer.gov)

GDS-Applicable Awardees Generating Human Data Must Provide an Institutional Certification (IC) Form

- Blank IC Forms can be found on the [NIH Data Sharing website](#) and come in three variants:
 - Samples collected **before** Jan 25, 2015, with **and** without consent
 - Samples collected **after** Jan 25, 2015
- IC Forms specify restrictions on sharing of **human** genomic data consistent with consent, and must be signed by an institutional signing official
 - This applies even if proposed research is **not** human subjects research (i.e., human cell lines, PDXs, etc.)

dbGaP: An NIH Repository for Controlled-Access Data



- The **D**atabase of **G**enotypes and **P**henotypes (dbGaP): NIH-managed repository for genomic data
- Common controlled-access repository for human genomic data—required for human data subject to the GDS policy

Initiating Data Deposition into dbGaP

1. Contact your **Program Director** and indicate that you'd like to deposit data to dbGaP
2. Your PD will put you in contact with the appropriate **Genomic Program Administrator** (for DCB, that's me!)
3. The GPA will register your study and send a link to begin data submission
4. After depositing study metadata (in collaboration with dbGaP staff), you'll receive an accession number for your study

DCB Resources for Data Sharing Questions



Soumya Korrapati, PhD

DMS Policy Expert

soumya.korrapati@nih.gov



Justin Benavidez, PhD

GPA/GDS Expert

justin.benavidez@nih.gov

2024 NIH Public Access Policy

- Applies to manuscripts accepted for publication ***on or after December 31, 2025***
- All NIH-funded publications ***must be submitted to PubMed Central*** for public availability ***without embargo*** upon the official date of publication
- This does ***not*** require payment of an article-processing charge (APC)
- PIs/institutions may upload the accepted manuscript directly to PubMed Central

Thank you!



**NATIONAL
CANCER
INSTITUTE**

cancer.gov

cancer.gov/espanol