Vivli

CENTER FOR GLOBAL CLINICAL RESEARCH DATA

Elements of an NIH Data Management and Sharing Plan – Vivli Repository

February 10, 2023

Elements of an NIH Data Management and Sharing Plan

01	Data types	Data to be preserved and shared
02	Related Tools, Software, Code	Tools and software needed to access/manipulate data
03	Common Data Standards	Standards to be applied to scientific data/metadata
04	Data Preservation, Access, Timelines	Repository to be used, persistent unique identifiers, and when/how long data will be available
05	Access, Distribution, Reuse Considerations	Factors for data access, distribution, or reuse
06	Oversight of Data Management	How Plan compliance will be monitored/ managed and by whom

Source: Final NIH Policy for Data Management and Sharing

Vivli

Element 1: Data Types: Vivli Platform

Describe **types of data** and amount of data expected to be generated.

- Vivli is designed for sharing of individual participant level data (IPD) and associated files
- The datasets must be anonymized
- The Vivli team offers support throughout the process

What is included in a shared data package?

Item	Description
Recommended Data Package Set	
Study Protocol	Final protocol with all amendments
Data dictionary	Detailed descriptions of each variable in the dataset, including the definition, source, coding, etc. of the variable
Statistical Analysis Plan	Description of the principal features of the analysis described in the protocol
IPD Dataset	Final cleaned individual participant-level data, anonymized
Anonymization Guidance	Outlines the method used to anonymize the data
Optional	
Analytic code	Software code used to carry out prespecified and additional analyses
Analysis ready IPD data set	Dataset in a format used to carry out a sponsor's analyses
Case report forms	Forms used to collect the data that is described in the protocol for each trial participant

NOTE: This is a subset of the entire full data package and includes the data that underlies the publication findings (tables, figures)

Vivli

Element 2: Related Tools, Software and/or Code

Indicate whether specialized tools are needed to access or manipulate shared scientific data to support replication or reuse, and **name(s) of the needed tool(s) and software**. If applicable, specify how needed tools can be accessed.

- Share the codebook or other explanatory documentation for understanding the data
- Detail any specialized programs or software to open files

Vivli

Element 3: Common Data Standards and Metadata (Vivli repository considerations)

What standards will be applied to the data and metadata?	The plan may indicate if consensus standards exist for scientific data and metadata	Metadata
• Vivli does not currently mandate any data standards • Data is provided in a format that can readily be used for analysis - R, SAS, STATA, Excel, CSV	• CDISC SDTM Standard Data Tabulation Model format is recommended to support for the most efficient data aggregation, re-use, and sharing • CDISC-SDTM is now the standard for those therapeutic areas that have a CDISC standard for data collection https://www.cdisc.org/standards/therapeutic-areas	Metadata is professionally curated by providing the NCT-ID (CT.gov identifier)

Element 4: Data Preservation, Access, and Timelines (Vivli repository considerations)

Provide the name of the repository(ies) where scientific data and metadata will be archived	Describe how the scientific data will be findable and identifiable (i.e., via a persistent unique identifier)	Describe when the scientific data will be made available to other users and for how long data will be available
• A global platform, Vivli offers sustainable platform where metadata are professionally curated. • Data are archived and backed-up to the Microsoft Azure cloud. Data are shared after signing of a DUA.	• Standardized data usage and citation metrics through the Make Data Count standard • ORCID IDs are linkable to datasets using Data CRediT roles • A citable, unique, persistent identifier (DOI) is assigned upon submission minted by DataCite	• The data will be archived and available on the Vivli platform upon request by researchers for a minimum of 10 years after contribution • Data will be made accessible no later than the time of our associated publication or at the end of the grant period (whatever comes first).

Element 5: Access, Distribution and Reuse Considerations (Vivli repository considerations)

Factors affecting subsequent access, distribution or reuse of data related to consent, privacy or confidentiality	State whether access will be controlled and made by the repository after approval	Protections for privacy, rights and confidentiality of human research participants
• Review the informed consent form to ensure sharing is allowable	• Access to the data will be controlled by a managed access process and provided only after approval	• There are multiple levels of protection for managing human participant data in Vivli – all contributors attest that data is anonymized, a managed access process is in place, all requesters sign a DUA and data is securely archived in Microsoft Azure

Element 6: Oversight of Data Management and Sharing (Vivli Repository Example)

Describe how compliance will be monitored and managed	Describe frequency of oversight	Describe the roles who will be conducting the oversight at your institution (include titles and roles)
• The PI will ensure that data are submitted and shared according to the DSMP	• The Data Manager in charge of each program will review if any trials are eligible for sharing each period (quarterly/monthly) as papers are published per their DMSP	• The execution, monitoring and oversight of the DMSP is the responsibility of the trials PI. The plan will be implemented and managed by the project's (PM/ data manager). The Data Manager in charge of each program will be tasked with helping researchers upload data to the appropriate repository monthly.

Vivli

Best practices with human subjects data*

Additional Considerations for Repositories Storing Human Data

Fidelity to Consent: Uses documented procedures to restrict dataset access and use to those that are consistent with participant consent and changes in consent.
Restricted Use Compliant: Uses documented procedures to communicate and enforce data use restrictions, such as preventing reidentification or redistribution to unauthorized users.
Privacy: Implements and provides documentation of measures to protect human subjects' data from inappropriate access.
Plan for Breach: Has security measures that include a response plan for detected data breaches.
Download Control: Controls and audits access to and download of datasets (if download is permitted).
Violations: Has procedures for addressing violations of terms-of-use by users and data mismanagement by the repository.
Request Review: Makes use of an established and transparent process for reviewing data access requests.

* NIH's Desirable Characteristics for Repositories Part II

Fidelity to Consent

Fidelity to consent

From NIH policy: *Employs documented procedures to restrict dataset access and use to those that are consistent with participant consent (such as for use only within the context of research on a specific disease or condition) and changes in consent.*

The consent form must explicitly describe data sharing and/or future use of research data

- Will participants be asked for consent again each time data is shared?
- Scope of use – which diseases?
- Will data will be combined with other data?

Broad consent example

"Information you provide in this study may be used in the future without your additional informed consent by other researchers to advance scientific research and public health. At this time, we do not know the specific details of these future research projects. These projects may involve bringing together information from this study with information from other studies or sources outside typical research settings."

Additional examples

- *NIA Informed consent template*
- *NIAID Informed consent template*
- *OSP* Informed Consent for Secondary Research with Data and Specimens

Vivli

Vivli Search and Access Process

SEARCH

Search Vivli platform for information about available studies.

REQUEST

Request IPD data package.

Each Data Request **reviewed** according to contributor's publicly stated requirements.

ACCESS

Data from approved requests can be **accessed** and **downloaded** after a Data Use Agreement has been signed.

DISSEMINATE

Completed **research results** assigned a DOI.

Researchers may use the Vivli platform to meet their **publication** requirements.

Vivli

What are the costs for sharing, archiving human clinical research data through Vivli?

- $4,000 per study (waived for researchers at Vivli member institutions) (larger than 500 GB are $10,000 per study)
- Anonymization fee (optional) – may be performed at institution - if outsourced allot $10,000 per study through preferred vendors
- Internal team costs for preparing data for sharing, developing supporting documentation

Questions?

- If you have further questions reach out to Vivli by emailing support@vivli.org and we would be happy to help