**DATA MANAGEMENT AND SHARING SAMPLE TEMPLATE FOR HUMAN SUBJECT CLINICAL RESEARCH DATA (NOT INCLUDING GENOMIC DATA)**

**Element 1: Data Type**

1. **Types and amount of scientific data expected to be generated in the project:**

*(Summarize the types and estimated amount of scientific data expected to be generated in the project).*

**Ex:** This study will collect [data scope] from [instruments, method, survey, experiment, data repository]. [Laboratory results, demographics, clinical observations, surveys, interviews, etc.] will be collected from [number] of research [participants/specimens/experiments] generalizing [number] datasets totaling approximated [amount of data] in size.

**Note:** Include how raw or processed your data will be (i.e. individual, aggregate, summarized, etc.).

1. **Scientific data that will be preserved and shared and the rationale for doing so:**

*(Describe which scientific data from the project will be preserved and shared and provide the rationale for the decision.* ***Researchers with rationale for NOT sharing study data should reference said rationale in this section and section 5a, b, c****).*

**Ex:** Data described in section A will be preserved through institutional storage and retention procedures [reference length of time] and shared through deposition of the data into [repository title].

1. **Metadata, other relevant data, and associated documentation:**

*(Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of scientific data).*

**Ex:** The study protocol, [date collection forms/case report forms, data dictionary, manual of operations, and a glossary of domain-specific terms, etc.] will be submitted.

**Element 2: Related Tools, Software and/or Code**

*(State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed).*

**Ex:** The data should be analyzed with [software, code, versions, packages, etc. that are needed to open, use, and analyze the data as used with the original project].

**Note:** Be specific and complete such that a user can recreate your analysis from the data. Consider important attributes such as the software version and any needed packages or extensions and where they can be found.

**Element 3: Standards**

*(State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources and provide the names of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standard exists).*

**Note:** List which common data standards exist for your scientific data considering standards for terminology, content, and/or data exchange (e.g., ICD-10, CDISC, MINSEQE, DICOM, etc.). Decide which ones you will employ, think about what standards would help someone understand the data you collect and potentially combine it with other related datasets. If the repository you have chosen requires specific standards, list those here. If there are no consensus data standards for your research area, state that here.

**Element 4: Data Preservation, Access, and Associated Timelines**

1. **Repository where scientific data and metadata will be archived:**

*(Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archives; see* [Selecting a Data Repository | Data Sharing (nih.gov)](https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/selecting-a-data-repository). ***Carilion Clinic does not have institutional repository nor the ability to create DOIs (digital object identifiers).****)*

**Ex:** The study will be submitted to [repository title]. [Description of data repository regarding datasets accepted]. [Description as to why this repository was chosen].

**Note:** You also need to indicate when you plan to publish your data. If you are not sharing data, describe how, and for how long your data will be preserved (reference institutional data retention).

1. **How scientific data will be findable and identifiable:**

*(Describe how the scientific data will be finable and identifiable, i.e., via a DOI)*

**Ex:** Scientific data will be findable via a DOI. Metadata associated with the datasets to [repository title] will be submitted. The repository will provide metadata, persistent identifiers, and [if applicable] long-term access for open and controlled access.

**Note:** In order to obtain a DOI for your dataset, Investigators will need to have a DOI generated via the chosen repository or through a collaborator’s institution.

1. **When and how long the scientific data will be made available:**

*(Describe when the scientific data will be made available to other uses (i.e. no later than time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available).*

 **Ex:** Scientific data will be shared as soon as possible. Scientific data included in published manuscripts will be available at the time of publication; all other generated scientific data will be shared no later than the end of the award. The study data will be stored in the repository for at least [length of time].

**Note:** Note that under Section [Section 8.4.2 of the NIH Grants Policy Statement](https://grants.nih.gov/grants/policy/nihgps/html5/section_8/8.4.2_record_retention_and_access.htm), grantee institutions are required to keep the data for 3 years following closeout of a grant or contract agreement. Contracts may specify different time periods. Another note, repositories have differing lengths of data preservation/retention. Familiarize yourself with ***all*** aspects of the chosen repository.

**Element 5: Access, Distribution, or Reuse Considerations**

1. **Factors affecting subsequent access, distribution, or reuse of scientific data:**

*(NIH expects that in drafting plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing. See* [*Frequently Asked Questions (FAQs) | Data Sharing (nih.gov)*](https://sharing.nih.gov/faqs#/data-management-and-sharing-policy.htm) *for examples of justifiable reasons for limiting sharing of data).*

**Ex:** The study dataset will be collected with the following informed **(Carilion Clinic IRB approved)** consent:

 “Your health information could be put in a database and stored there indefinitely (no end date). Your research data will not be labeled with your name or any other information that could be used to identify you. Even though your identifiers will be removed, there may still be a chance that someone could figure out that the information is about you, though we believe it is unlikely that this will happen. Your research data may then be used and shared with other researchers for future unknown research studies without additional informed consent from you. You will not be informed of the details of any future research studies that might be conducted using your health information.”

**OR (If there are justifiable factors affecting subsequent access, distribution, or reuse of scientific data. Justifiable factors affecting subsequent access, distribution, or reuse of scientific data include institutional policies, patent and intellectual property rules, data size limitations, informed consent, privacy and confidentiality protections, revised common rule, federal, state, local rules, tribal rules and genomic data rules. Investigators must then describe applicable restriction/s beyond the below consent language).**

**“**Your research data and personal information collected as a part of the research will not be used or distributed for future research studies, even if your identifiers are removed.”

1. **Whether access to scientific data will be controlled:**

*(State whether access to the scientific data will be controlled access or open access. Controlled access repositories protect data by ensuring that only authorized entities can retrieve data from the repository. For controlled access repositories, describe any control practices (e.g. via a data sharing agreement) and reference its options for limiting access. Open access repositories provide free unrestricted access to scientific and scholarly data).*

**Ex (open access):** De-identified data and documentation will be made available as public use data to the research community via [open access repository title]. The Terms of Use apply and are accepted by those who use or publish to [repository title]. (Discuss in further detail Terms of Use applicable for the repository of choice such as prohibiting attempts to re-identify).

**Ex (controlled access):** To maximize the appropriate sharing of scientific data and protect research participants privacy and confidentiality, access to dataset is made available by [controlled access repository] only when requested by users who meet conditions of use according to [controlled access repository]. (Discuss other control practices listed in repository of choice).

1. **Protections for privacy, rights, and confidentiality of human research participants:**

*(If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g. through de-identification, Certificates of Confidentiality (CoC), and other protective measures)).*

**Ex:** To protect research participants’ privacy and confidentiality, data submitted to the repository will not include personal health information or personally identifiable information such as names, addresses, DOB, etc. Additional protections, such as the approach for managing Health Insurance Portability and Accountability Act identifiers, will be used for de-identification or to provide a limited data set to minimize the risk of participant reidentification.

**Note:** In many cases, only aggregate level data will be shared. Describe as applicable. List if CoC is going to be obtained. Refer to *any* protections utilized. If data will not be shared in order to protect participants’ privacy and confidentiality, explain.

**Element 6: Oversight of Data Management and Sharing**

*(Describe how compliance with this plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).*

**Example language approved by Research Oversight Officials at Carilion Clinic:**

 [Principal Investigator Name/Title] will be responsible for monitoring compliance with the data management (e.g., data capture, documentation, quality review, storage, and backup) aspects of the approved DMS plan at least [insert frequency] and, when appropriate, proposing revisions. Carilion Clinic’s Research and Development, Health Analytics Research Team, Institutional Review Board, and Compliance Office will conduct compliance reviews of the procedures established by each department.