HCHS/SOL Study Data Sharing Policy and Plan

In accordance with the NHLBI plan for implementing the [2023 NIH Data Management and Sharing (DMS) Policy](#), in its cohort studies, the HCHS/SOL Coordinating Center (CC) will implement an **annual data submission process to NIH data repositories including BioLINCC, the database of Genotypes and Phenotypes (dbGaP) and BioData Catalyst (BDC)**. This data sharing plan applies to all frozen[1] HCHS/SOL datasets, including those from ancillary studies. Interim data releases of data not intended for peer-reviewed publication are not subject to the policy. Submissions to NIH repositories will occur annually two months after the end of each HCHS/SOL CC contract year (i.e., November 29th contract year ending, January 31th submission deadline).

HCHS/SOL Data Recipients include HCHS/SOL field centers (Bronx, Chicago, Miami and San Diego), Central Lab (Minnesota), baseline Reading Centers, and Affiliated Institutions.

Timeline of Data Sharing

Type of Data	Timing of Data Closeout / Freeze and Recipients	Timing of NIH Repository Submission and Recipients
Follow-Up	Freeze annually by contact year AND share with the **HCHS/SOL Data Recipients.**	Annual January submission to **BioLINCC, or BDC/dbGaP (per NIH).**
Incident Events	Freeze by **event year that has over 90% events reviewed**, with a minimum one-year lag due to abstraction, review, and adjudication processes for CHD, HF, stroke and pulmonary events, AND share with the **HCHS/SOL Data Recipients**.	Annual January submission to **BioLINCC, or BDC/dbGaP (per NIH),** provided there are frozen incident events data and shared with the **HCHS/SOL Data Recipients.**
Clinic Visit (Raw)	The CC will freeze clinic visit data in raw form and datasets from reading center AND share with **the HCHS/SOL Data Recipients** within six months of completing CDART database lock[2].	During Annual January submission to **BioLINCC, or BDC/dbGaP (per NIH)** when applicable**.**
Derived (Visit data including non-genomic biospecimen analytes, and Follow-up)	The CC will freeze derived datasets AND share with the **HCHS/SOL Data Recipients** within six months of freezing.	Annual January submission to **BioLINCC, or BDC/dbGaP (per NIH).**

---

[1] A frozen dataset refers to a dataset that has been finalized at a specific point in time, meaning no further changes will be made to its contents. It has the masked ID for investigator's use instead of the original participant ID.

[2] Database lock is a restriction applied to the data management system at the conclusion of data collection and cleaning. Its purpose is to prevent any further modifications, ensuring that the dataset remains stable and unaltered for final analysis, reporting, and archiving.

Type of Data	Timing of Data Closeout / Freeze and Recipients	Timing of NIH Repository Submission and Recipients
Ancillary Study with data collection	For non-omics data, the CC will freeze ancillary study datasets AND share with the **Ancillary Study PI within six months** of completing CDART database lock or receiving final external data files, whichever is later. For omics data, timing is based on the 2023 NIH Data Management and Sharing Policy	During Annual January submission **simultaneously to HCHS/SOL Data Recipients, BioLINCC, or BDC/dbGaP (per NIH)** for any non-omics data shared with Ancillary Study PI by July 31st of the previous year. Otherwise, submission will occur the following year. For omics data, only submitted to dbGaP by the Ancillary Study Genetics Lab.
Ancillary study with non-genomic biospecimen processing ONLY	After the laboratory completes its internal quality control (QC) procedures, the assay data will be submitted to the CC. The CC will perform final QC verification, after which the dataset will be frozen and provided to the **Ancillary Study PI**.	During the annual January submission **simultaneously with HCHS/SOL Data Recipients, BioLINCC, or BDC/dbGaP (per NIH)** for any data shared with Ancillary Study PI by July 31st of the previous year. Otherwise, submission will occur the following year.
Ancillary Study with Secondary Data Analysis ONLY	Not Applicable	Not Applicable
Multi-omic Datasets Subject to NIH DMS Policy[3] (e.g., genomics, DNA methylation, RNA sequencing)	Timing is based on the 2023 NIH Data Management and Sharing Policy	Only submitted to dbGaP by the Ancillary Study Genetics Lab.
Raw Reading Center Data (e.g., imaging files and signals data not otherwise shared through summary data)	Completed by the Reading Center at the end of ancillary study project period.	Only submitted directly to BDC by the Reading Centers.
Geocoded data	Not applicable Geocoded data cannot be shared. It is available at UNC Secure Research Workspace (SRW).	Not applicable Geocoded data cannot be shared.

[3] The NIH DMS policy refers to the broad sharing of multi-omics which includes, but is not limited to, the following: genomics and epigenetics (e.g., WGS, WES, EPIC arrays, targeted panels), transcriptomics (e.g., RNA-seq), microbiomics (e.g., bacteria, virus), proteomics (e.g., mass spec, RPPA) and metabolomics (e.g., mass spec).

Historical Context and Planning

- Previously, data collected from main study visits and follow-up were shared with BioLINCC two years after distribution to HCHS/SOL Data Recipients; Ancillary Study data were shared with BioLINCC three years after sharing with the Ancillary Study PI.
- Section 5.8 of the HCHS/SOL Ancillary Studies Policy was updated to reflect the revised data sharing policy and timelines.
- Item 11 in section C of the Ancillary Study proposal form was updated to reflect the revised data sharing policy and timelines.
- Investigators are encouraged to align their research activities and manuscript submissions with the updated timelines outlined in this Data Sharing Policy.

For questions, please contact the HCHS/SOL Coordinating Center (HCHSAdministration@unc.edu).