ECHO-wide Cohort Data Sharing Policy

I. Introduction and Purpose

A goal of the NIH Environmental influences on Child Health Outcomes (ECHO) Program is to support the ECHO-wide Cohort, composed of multiple, synergistic, longitudinal cohort studies, to investigate the effect of a broad range of early environmental exposures—including physical, chemical, social, behavioral, and biological—on child health and development. In support of that goal, the ***ECHO Data Analysis Center*** (***ECHO DAC***) will create a highly secure environment accessed through a web portal, the ***ECHOPortal***, to support the ECHO Program's needs for a shared collaborative workspace with access to ***ECHO-wide Cohort data***.

The ECHO-wide Cohort Data Sharing Policy balances important goals: to facilitate investigations of the longitudinal impact of prenatal, perinatal, and postnatal environmental exposures on pediatric health; to respect and protect the confidentiality of ECHO Cohort participants contributing data to the ***ECHO-wide Cohort Data Platform***; and to maximize the value of the data generated by this program to the scientific and general community. To access ***ECHO-wide Cohort data***, ***ECHO***, ***ECHO Affiliate***, and ***non-ECHO*** investigators and their institutions must agree to, and abide by, the ECHO-wide Cohort Data Sharing Policy.

The ECHO-wide Cohort Data Sharing Policy aims to promote wide and timely dissemination of data within the ***ECHO Consortium*** and the broader scientific community to ensure their maximum utility and impact. The following principles guided the development of the policy:

- Sharing of data in a timely manner is important for optimizing scientific progress. This policy is consistent with the goals of the NIH data-sharing policy to make public access to digital scientific data the standard for all NIH-funded research. See "Draft NIH Policy for Data Management and Sharing."

- Combining data from multiple cohorts will increase power for addressing important research questions and facilitate addressing the heterogeneity of effects and generalizability.

- Public funds support the data collected by the ECHO Program and, as such, the ECHO-wide Cohort should serve as an accessible resource helping to catalyze the research activities of scientists around the world, including those not supported as a part of the ECHO Program.

- It is important to strike a balance between making data broadly accessible to the scientific community, and the legitimate interest ***ECHO investigators*** have in benefitting from their investment of time and effort to collect and harmonize data.

- The ECHO Program must protect the confidentiality of ECHO Cohort participants contributing data to the ***ECHO-wide Cohort Data Platform***.

II. Scope

This document applies to all ECHO Cohort and Opportunities and Infrastructure Fund (OIF) Awardees contributing data to the ECHO Program and ***ECHO***, ***ECHO Affiliate***, and ***non-ECHO*** investigators who request the use of ***ECHO-wide Cohort data***. It does not apply to data collected for the IDeA States Pediatric Clinical Trials Network as access to those data are governed by that group's data sharing policy.

The ECHO Biospecimen Utilization Policy covers policies related to access to biospecimens collected by ECHO Cohorts. However, this data sharing policy covers all data generated from assays using ECHO biospecimens. Further, NIH's Genomic Data Sharing Policy and this policy cover all data that fall under the NIH Genome Data Sharing Policy (including genomics, epigenomics, and microbiome data) and the NIH policy for Data Management and Sharing (DRAFT available for public comment).

The NIH ECHO Program recognizes that policies of sovereign Native American nations on research and data ownership may be in conflict with this policy, and that stipulations reconciling these policies are necessary. Therefore, language that clarifies sharing of data from Native American tribes to the ECHO-wide Cohort is included in a separate Data Sharing/Use Agreement for the Navajo Birth Cohort Study.

The ECHO Executive Committee, the Steering Committee, and the NIH Program Office will review, revise, and approve this Data Sharing Policy on a regular basis (see section V.).

III. Definitions

Item	Description
Analysis Concept	The first notification of a research concept to the ECHO community that includes a brief statement of the study hypothesis or objective and identifies the primary exposures and outcomes. The Coordinating Center will make the concept available to all ECHO awardees for a 2-week period to allow all ***ECHO*** and ***ECHO Affiliate*** investigators an opportunity to review the brief research summary, add to the concept, or request inclusion on the writing team.
Analysis Proposal	A writing team will prepare a detailed ***Analysis Proposal*** (background and motivation, aims/hypotheses, conceptual model, study population, measures, statistical approach, preliminary data, table shells, intended

Item	Description
	journal) for submission to the Publications Committee for review and to the Steering Committee for final approval.
De-identified Data	The term "***de-identified data***" refers to a set of data where ***personally identifiable information (PII)***, specifically geographical locations and dates for ***ECHO-wide Cohort Data***, have been deleted or altered to protect the confidentiality of participants.
ECHOPortal Data Access Request	Writing team submits an attachment to the ***Analysis Proposal*** to request access to individual-level data via the ***ECHOPortal*** for a non-DAC analyst assigned to perform analyses within an approved ***Analysis Proposal***. If the ***DAC*** will be the only entity conducting the analyses, this form is not necessary.
ECHO Data Analysis Center (DAC)	ECHO component responsible for developing the secure environment for accessing ECHO-wide Cohort and OIF project data. Responsible for managing receipt of ***ECHO-wide Cohort data*** from ECHO Cohorts and ***OIF data*** from ***OIF Awardees***, harmonization of ***ECHO-wide Cohort data***, conducting analysis, and providing access to the ***ECHO-wide Cohort data*** and ***OIF data*** for querying, visualization, and analysis.
ECHO Consortium	Includes all ECHO components (IDeA States Pediatric Clinical Trials Network, ECHO Cohorts, Coordinating Center, ***DAC***, Children's Health Exposure Analysis Resource, Patient-reported Outcomes Core, Center for Inherited Disease Research), Principal investigators and key personnel, NIH program staff and project scientists.
ECHO Investigator	Any member of the ***ECHO Consortium*** who is supported by an ECHO grant, subcontract, or consultancy. Support is reflected by the designation as a Principal Investigator (PI), Multiple PI (MPI), or Co-Investigator (Co-I) on the grant or subsequent progress reports.
ECHO Affiliate Investigator	Any investigator sponsored by an ***ECHO investigator***, and located at the sponsoring investigator's institution. ***Affiliate Investigators*** have the same rights and responsibilities for accessing ***ECHO-wide Cohort Data*** and ***OIF Data*** as ***ECHO Investigators***.
Non-ECHO Investigator	Any researcher who is not an ***ECHO Investigator*** or ***ECHO Affiliate Investigator***.
ECHOPortal	A web portal to a High-Performance Computing IT infrastructure developed and maintained by the ***ECHO DAC*** to host, analyze, and control access to data. Behind the ***ECHOPortal*** are tools including an ECHO Cohort metadata catalog, data capture and processing systems,

Item	Description
	and tools for analysis of ***ECHO-wide Cohort data*** (e.g., SAS, R, Winbugs, ArcGIS).
ECHO-wide Cohort Data	Extant or new data the ECHO cohorts will deposit to ***ECHO-wide Cohort data platform*** consistent with the ECHO-wide Cohort Data Collection Protocol.
ECHO-wide Cohort Data Platform	A data platform consisting of essential and recommended common data elements from ECHO cohort awards and their constituent cohorts. This includes specialized data from subsets of cohort awards—at least 2—consistent with the ECHO-wide Cohort Data Collection Protocol. In this policy document, it is also referred to as the ECHO Data Enclave.
FISMA	Federal Information Security Management Act of 2002 that led to the National Institute of Standards and Technology standards for mandatory minimum information security requirements for federal information security systems. RTI International's security and compliance office has certified and accredited the ECHO Data Enclave as having ***FISMA*** moderate security controls.
OIF Awardee	Investigator receiving funding from the NIH Opportunities and Infrastructure Fund (OIF) program. ***OIF awardees*** are considered ***ECHO Investigators***.
OIF Data	Data or technology (e.g., programs, analytical tools) generated by ***OIF Awardees*** as part of their OIF projects.
PII	***Personally Identifiable Information (PII)*** is defined as any information that can be used to uniquely identify a single individual or that can be used with other sources to uniquely identify a single individual. In contrast with PHI, which is federally defined, organizations use a variety of methods to define what constitutes ***PII***. The ECHO Program Manual of Procedures contains the specific data elements that comprise ***PII*** for ***ECHO-wide Cohort data*** (i.e., dates and geographic locations). Access to ***PII*** will be distinct from access to ***de-identified data*** to ensure that reasonable security safeguards protect personal data.

IV. Policy

A. ECHOPortal

ECHOPortal (ECHOPortal.org) is a web portal to secure ***FISMA*** Low and Moderate environments that the ***ECHO DAC*** developed and maintains. The ***FISMA*** Low

environment is used for data management tasks requiring interactive exchange with ECHO cohort sites. User authorization to this environment is role-based to enforce tight control of what users can access and which actions they can perform across the different ***DAC*** systems. The ***FISMA*** Moderate environment contains the ECHO-wide Cohort and ***OIF data*** accessible to qualified researchers. This environment includes a core set of software and analytical tools in addition to data, such that the data stay in the environment and the data analysis is done within the environment. Users are required to use two-factor authentication to access the environment. Their access is granted on an as-needed basis associated with an approved ***analysis proposal*** and is time-limited.

B. Types of Data, Data Submission, and Data Access

The categories below describe the data collected and generated by the ECHO Cohorts and OIF Projects. The ***ECHO investigators*** will share various types of data with the ***ECHO DAC*** according to the policies described after each data type description and they must abide by FAIR data sharing principles, ensuring that data or technology generated from their Cohorts or OIF projects are Findable, Accessible, Interoperable, and Reusable.

1. *Metadata*

Metadata are of four types:

- Information about ECHO Cohorts including key contacts, general study design details, existing data on exposures, outcomes, existing biospecimens, and potential confounders, effect modifiers, and mediators, specimens collected and analyzed;
- Additional documentation on the methods and procedures used by cohorts to collect data prior to ECHO is available for some cohorts (e.g., protocol, data collection forms, codebooks, manual of procedures, visit schedule, data dictionaries);
- Documentation associated with the ECHO-wide Cohort Common Data Model;
- Information about each OIF project, including the name, institution, and contact information of the ***OIF Awardee***, the type and amount of data available or technology developed, and where data are stored or how technology can be accessed (if not stored in the ECHO Enclave and accessible through the ***ECHOPortal***).

a. *Metadata Submission*

ECHO Cohort Awardees will submit their cohort(s)' responses to formal surveys developed by the ***ECHO DAC*** within the timeframe requested by the ***DAC***. ***OIF***

Awardees will submit information about their OIF project to the ***ECHO DAC*** by the end of their 2-year funded OIF project period.

b. *Metadata Access*

The ***ECHOPortal*** Metadata Catalog, accessible only to registered ***ECHO Consortium*** and non-ECHO users, provides information about ECHO Cohorts' study design, existing data on exposures, outcomes, and potential confounders, effect modifiers, and mediators, and biospecimens. For Cohorts whose investigators have granted the ***DAC*** permission to share detailed ECHO Cohort study materials (e.g., protocol, codebook, data dictionary), access is available only to the ***ECHO Consortium*** through the password-protected ***ECHOPortal*** Metadata Catalog.

General information about all OIF projects will also be accessible through the ***ECHOPortal*** Metadata Catalog.

2. *ECHO eaRly dAta subMission Protocol (RAMP)*

RAMP involves cohorts submitting a limited set of existing participant-level data from women and children. The RAMP Manual of Procedures contains detail on variables included in the data submission.

a. *RAMP Data Submission*

ECHO Cohort Awardees participating in RAMP will:

- Share RAMP data for their cohort(s) with the ECHO ***DAC***, with or without personal identifiers depending on decisions of authorized oversight committees/boards (e.g., the parent study's IRB or the ECHO single IRB if the parent study's institutions have ceded review to the ECHO single IRB). ECHO Program does not expect cohorts to obtain new informed consent for RAMP.
- ECHO Cohort Awardees will ensure that their cohort(s) have submitted data elements in RAMP to the ***ECHO DAC*** by January 31, 2018.
- No one will use RAMP data for research purposes.
- Only the ***ECHO DAC*** staff will have access to the RAMP data.

3. *Data Submission for the ECHO-wide Cohort Data Collection Protocol*

The ECHO-wide Cohort Data Collection Protocol specifies data elements a cohort may collect across the life course from preconception through adolescence. This protocol facilitates the creation of the ***ECHO-wide Cohort Data Platform***. The NIH ECHO Program recognizes that it is critical to maximize the benefits of sharing while protecting the interests of the ***ECHO investigators***, who have invested time and effort to collect the data. The NIH requires ECHO Cohort Awardees and their

constituent cohorts to deposit with the ***ECHO DAC*** all existing and newly collected essential and recommended data elements included in the ECHO-wide Cohort Data Collection Protocol. This requirement is in alignment with the original expectations of the ECHO RFA to leverage existing and new data.

The NIH ECHO Program does not require the ECHO Cohort Awardees to deposit data that the ECHO-wide Cohort Protocol does not specify. However, awardees may choose to deposit this type of data for future integration with the ***ECHO-wide Cohort Data Platform***, if the ***Data Analysis Center*** approves.

ECHO Cohort Awardees will ensure that:

- Their cohort(s) submit existing essential and recommended data elements pertaining to the ECHO-wide Cohort Data Collection Protocol on participants enrolled in ECHO Cohorts for which the cohorts have obtained IRB approval by March 31, 2020.
- Their cohort(s) submit newly collected data to the ***DAC*** at least every 6 months. For data collected with ***ECHO DAC***-developed data collection tools, cohorts will, by default, submit those data to the ***ECHO DAC*** in real time.
- The ***DAC*** will sequester existing and newly collected data from the Cohorts in a restricted area accessible for its review and harmonization efforts before making any data from a cohort available for analysis. The ***ECHO DAC*** and each Cohort's Principal Investigators will make an integrated decision over when the ***DAC*** adds a Cohort's data elements to the ***ECHO-wide Cohort Data platform***. The lag time will be guided by the following considerations:
    - Pre-processing of data at the Cohort level if applicable,
    - The ***DAC*** ensuring data quality and completing data harmonization processes,
    - Any evidence of selection bias from analyzing data from cohort participants seen earlier vs. later in a data collection cycle.

4. *Data Submission by OIF Awardees*

OIF awardees must ensure that they make data or technology generated from their OIF projects available to ***ECHO***, ***ECHO Affiliate***, and ***non-ECHO*** investigators. Accordingly, ***OIF Awardees*** must develop a data sharing plan, which outlines plans for sharing ***OIF data*** or technology generated by the OIF project. In consultation with the ***ECHO DAC***, ***OIF Awardees*** will determine the best mechanism for sharing data resulting from their OIF project, whether through the ***DAC***'s ***ECHOPortal*** or some other facility or repository. Regardless of the sharing mechanism specified in the data sharing plan, ***OIF Awardees*** must ensure that their ***OIF data*** or technology be made available for sharing by the end of their 2-year funded OIF project period.

As described in Section 3 for ***ECHO-wide Cohort Data***, the ***DAC*** will initially sequester all ***OIF data*** it receives until making a joint decision with the ***OIF Awardee*** that the ***OIF data*** are ready for broader sharing.

C. ECHO-wide Cohort Data Collection Protocol: Data Access, Approval Requirements, and Terms of Access

1. *ECHO-wide Cohort Data Collection Protocol Data Access within ECHO Data Enclave*

All investigators follow a similar process to use ECHO-wide Cohort and ***OIF Data*** within ***ECHOPortal*** for an approved ***Analysis Proposal***. For more details, refer to the ECHO Publications Policy. ***OIF Awardees*** who require access to ***ECHO-wide Cohort data*** for completion of the OIF project are granted an exception to this rule as the OIF application review process is deemed equivalent to submitting and receiving approval for an ***Analysis Proposal***.

An ***ECHO DAC*** analyst will conduct all analyses within an approved ***Analysis Proposal*** if an investigator reports a Conflict of Interest (COI) and for all analyses requiring use of ***PII*** (i.e., dates, geographic locations). For more details on identification and reporting of COIs, refer to the ECHO Conflict of Interest Policy.

For ***Analysis Proposals*** that include data from the Navajo Birth Cohort Study (NBCS), an ***ECHO DAC*** or a NBCS analyst must conduct all analyses.

The ***ECHO DAC*** will grant data access to analysts designated by an investigator who is a writing team lead upon (1) the ECHO Steering Committee approving the ***Analysis Proposal*** or award of an OIF project requiring access to ***ECHO-wide Cohort data***, and (2) the ***DAC*** approval of the ***ECHOPortal data access request*** for any non-***DAC*** analyst.

For operational details, refer to ECHO Manual of Operations.

- ***Non-ECHO Investigator*** Access:

Non-ECHO investigators will be able to submit Analysis Concepts based on the first release of ECHO-wide Cohort Data Collection Protocol data 12 months after ***ECHO*** and ***ECHO Affiliate*** investigators have access to these data and 6 months after for subsequent releases. This embargo period is consistent with the goals stated in the "NIH Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research." *For operational details, refer to the ECHO Manual of Operations.*

2. *ECHO De-identified Data for the General Scientific Community:*

In addition to the above stipulations for performing analyses of data with personal identifiers within the ECHO-wide Cohort data enclave, the ***DAC***, the NIH ECHO Program Office, and the Steering Committee will also develop policies

and procedures for access to a public-use version of these data in the future to the general scientific community without the need for an ***Analysis Proposal***. The NIH ECHO Program Office, the Steering Committee, and the ***DAC*** will specify the process for producing and accessing these data in a future update to this policy.

3. *Data Access Approval Requirements*

Investigators seeking access to data via the ***ECHOPortal*** to conduct analysis on individual-level data must first meet the following data access approval requirements. They are applicable to ***ECHO***, ***ECHO Affiliate***, and ***non-ECHO*** investigators:

a. *Investigator Access to Data via ECHOPortal*

- Investigators with an ECHO Steering Committee-approved ***Analysis Proposal*** or an active OIF award will have access only to those data needed for the approved ***Analysis Proposal*** or OIF project.

  The investigator can delegate the responsibility for conducting the analysis to a ***DAC*** analyst(s) or a non-***DAC*** analyst(s). Non-***DAC*** analysts will only have access to ***de-identified data***. An ***ECHO DAC*** analyst must conduct all analyses that require access to ***PII***.

  To delegate to a non-***DAC*** analyst, the investigator must identify the non-***DAC*** analyst on their ***Analysis Proposal*** or OIF application, and the investigator must submit an ***ECHOPortal Data Access Request*** Form as an attachment. The form must include a "Delegation of Authority" signed by the responsible investigator from the institution where the analyst is employed that acknowledges accountability for all activities the analyst conducts behind the ***ECHOPortal***. The non-***DAC*** analyst must complete training in the protection of human subjects and IT security awareness and attest to adhere to the terms and conditions of data access. *For operational details, refer to ECHO Manual of Operations.*

- All investigators, whether ***ECHO***, ***ECHO Affiliate***, or ***non-ECHO***, need IRB approval to begin analyses. For ECHO investigators and Affiliates covered locally for the ECHO-wide Cohort Data Collection Protocol, this may only require notification of the local IRB about the proposed analysis.

4. *Terms of Data Access*

- For *all* ***analysis proposals*** and OIF projects, a ***DAC*** analyst will conduct analyses that require the use of ***PII*** (i.e., dates, geographic locations). No investigator (***ECHO***, ***ECHO Affiliate***, or ***non-ECHO***) will have access to ***PII*** data through ***ECHOPortal***.

- Only analysts specified in the ***Analysis Proposal*** or OIF application will receive access to de-identified individual-level data via the ***ECHOPortal***.

- The ***ECHO DAC*** will grant access only to data sets or data views required to implement the analyses specified in the approved ***Analysis Proposal*** or OIF application. The ***ECHO DAC*** staff will provide consultation to determine the data needed for analysis.
- The analyst can remove summary-level results from the ***ECHOPortal*** after review and approval at the ***DAC***, but not individual-level data. Dissemination of those summary results in the form of publication is subject to further approval via the Publications Committee Review process.
- In alignment with the Publications Policy, time limitations on data access may apply. Generally, progress from ***Analysis Proposal*** approval to submission of the resulting manuscript should be one year or less, unless otherwise designated by the Publications Committee. The ECHO Steering Committee will determine the consequences of failure to adhere to this timeline (e.g., suspension of access to data for a determined period of time).

D. Data Privacy Breach and Abuse and Penalties for Non-Compliance with Data Sharing Policy

The ***DAC*** may monitor activities conducted on the computer system accessing ***ECHOPortal*** to facilitate protection against unauthorized access, and to verify security procedures, survivability, and operational security. During monitoring, the ***DAC*** may examine, record, copy, and use information for authorized purposes. Unauthorized use or access may subject the user to criminal prosecution. If the ***DAC*** or an individual outside the ***DAC*** becomes aware of a suspected or actual privacy or information security incident, or a violation of the terms and conditions of data access, they must report the event to the ECHO ***DAC*** and CC immediately as it may impact confidentiality, integrity, or availability of ***ECHOPortal*** data and resources. Whoever discovers such an event must report to appropriate institutional authorities, including Privacy Officers, in accordance with their institutional policies. The ECHO CC must report the incident to the ECHO single IRB, and the ECHO NIH Program Office within 72 hours after receiving the report and may notify the ECHO Steering Committee. The report will include the following information as available:

- Type of data affected (e.g., ***PII***, other)
- Approximate date and time of the incident
- Detailed description of the incident (e.g., amount of data involved, specific systems, servers, and IP addresses involved)
- Incident response taken to remediate the incident (e.g., notification of the data originator when ***PII*** is involved)
- Identify a mechanism to prevent or minimize future risk

The ***ECHO DAC*** will work with the ECHO Steering Committee and the NIH Program Office to remedy any ongoing concerns, or revise processes to reduce the risk of a similar event recurring.

V. Review and Revision

The ECHO Executive Committee, the Steering Committee, and the NIH Program Office will review and approve updates to all ECHO Program related documents (e.g., Manuals of Procedures, ECHO-wide Cohort Data Sharing Policy) at least every 12 months. The relevant parties will be notified of any changes through standard ECHO Program channels.

VI. Supporting documents

A. ECHOPortal Data Access Request Form

B. Delegation of Authority

VII. References

A. ECHO Publications Policy

B. ECHO Biospecimen Utilization Policy

C. NIH Genome Data Sharing Policy (Paltoo et al, Nature Genetics 2014 46:934–8). See at https://osp.od.nih.gov/scientific-sharing/genomic-data-sharing/

D. NIH Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research http://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf

E. RAMP Manual of Procedures

F. ECHO Manual of Operations

G. ECHO Conflict of Interest Policy

VIII. History of Change

Section Affected	Version Date	Changes Made
Introduction, Scope, Definitions, and Sections A, B, and C under Part IV	10/Jan/2020	Revised Policy to include reference to ***OIF Awardees*** and applicability of Data Sharing Policy to OIF projects and OIF-generated data. Also included ***ECHO Affiliate Investigator*** definition and applicability of the Policy to this investigator type. Added requirement that DAC analyst must conduct all analyses requiring access to PII, and DAC analyst or NBCS analyst must conduct all analyses when data from the NBCS is involved. Other minor editorial changes throughout the document for consistency in wording and refinement of description of **DAC** processes.
N/A	08/Nov/2017	New Policy

Approval Page

Name/Title	Signature/Date
Approved by Steering Committee via post-meeting survey	January 31, 2020