ARIC / ARIC-NCS Data Sharing with Investigators, Repositories and Journals

The ARIC Study (including ARIC-NCS) has a history of being very collaborative. Data sharing is with individuals, data repositories, and research consortia. For the latter, in some instances it is just summary statistics that are shared while for others it is individual-level data.

1. Data sharing information for journals

Many journals require authors of manuscripts to make data available for other researchers to try to reproduce results. See the document "ARIC Policy on Datasets for Reproducing Results" at https://aric.cscc.unc.edu/aric9/publications/policies_forms_and_guidelines. For journals that require manuscripts include a statement about data availability, here is an example paragraph that can be modified to suit the purposes of the specific manuscript:

> Data access policies for the ARIC Study specify that research data requests may be submitted to the Steering Committee. These will be promptly reviewed for confidentiality or intellectual property restrictions and will not unreasonably be refused. Individual level participant data may further be restricted by consent, confidentiality or privacy laws/considerations. These policies apply to both clinical and laboratory data. Additional information on how to obtain data, including access to some data through the NHLBI's Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC), which does not need ARIC Study approval, is available at: https://aric.cscc.unc.edu/aric9/researchers/Obtain_Submit_Data.

2. Data sharing with individual investigators

Any scientist who is willing to partner with an ARIC investigator may submit a manuscript proposal or an ancillary study proposal. Once the Publications Committee or Ancillary Studies Committee, respectively, approves the proposal, the investigator and the ARIC Coordinating Center enter into a Data and Materials Distribution Agreement (DMDA). (For an ancillary study requiring use of biospecimens the DMDA also needs to be signed by NHLBI.) After the signing of the DMDA, the Coordinating Center distributes the requested data to the investigator. A nominal fee is charged for the cost of preparing the datasets. The ARIC Publications Committee typically reviews about 200 new manuscript proposals each year. Publications Committee review of manuscript proposals (and completed manuscripts before journal submission), includes those for ARIC-NCS and other ancillary studies.

3. Data sharing with data repositories

A requirement of the ARIC contract is that data be provided to the NHLBI's data repository, BioLINCC. This is mandated not just for data collected under the contract but also for data from ancillary studies, including ARIC-NCS. According to the cohort data sharing agreement with NHLBI, by the end of each contract year (November 14), the Coordinating Center will submit to BioLINCC all visit, annual/semi-annual follow-up call, surveillance, and ancillary study datasets that have been closed in the preceding 12 months. Genotype and relevant phenotype data are also required to be submitted to dbGaP.

For those who have access to the password-protected sections of the ARIC website, detailed information about types of data and associated timing of NIH data repository submission are in memo #63 at https://aric.cscc.unc.edu/aric9/memo.

Once data are submitted to BioLINCC or dbGaP, those repositories make decisions about data sharing without consulting ARIC investigators. Thus the Coordinating Center takes additional steps to de-identify data sent to these repositories and applies the most stringent versions of participants' informed consent choices, such as removing participants who specify that their data may be used only for CVD research. ARIC is also willing to work with other repositories including ones specific to aging and dementia.

4. Summary data sharing with consortia

For some consortia, typically those involved in genomic research, individual studies in the consortia undertake their own analyses (using a specified analysis plan). Summary statistics from these analyses are then shared with a group that meta-analyses the results from the various studies.

5. Individual-level data sharing with consortia

Some consortia request participant-level data from participating studies. Such consortia often undertake pooled analyses rather than meta-analyzing results from the individual studies. An example of this type is the NHLBI Pooled Cohorts Study, coordinated by Columbia University.

6. Hybrid models

The Chronic Kidney Disease Prognosis Consortium, led by ARIC investigators Drs. Coresh and Grams, is a mix of types 3 and 4 above. Some participating studies provide individual data. Those studies unable to send the individual participant level data (de-identified) are sent a standard computer program that is designed to automatically save all output needed for meta- analyses. This consortium has access to all the individual-level ARIC data.

Another type of hybrid is the Global Alzheimer's Association Interactive Network (GAAIN) coordinated by the University of Southern California (USC). In this network a program running on a computer at USC has links to data that are housed locally at each study's site. A GAAIN investigator can query the consortium database to obtain information and summary statistics on data at each "data partner". In order to do any analyses, the investigator then contacts each of the data partners that has relevant data, with a request for the data (after obtaining approval from the data partner's Publications Committee, where relevant). ARIC has signed a Memorandum of Understanding with GAAIN as a data partner and the ARIC Coordinating Center houses a dataset that can be accessed by the USC computer.