

Instructions for the Derived Dataset Submission Form

I. General Instructions

This fillable PDF form is completed by ARIC analysts/investigators who submit derived variable dataset(s) to the ARIC CC for general distribution to the formal list of ARIC data recipients. The dataset(s) are created and verified with no CC involvement except for the distribution. The investigator creates all accompanying documentation (labels, codes, dictionaries) to be distributed with the data.

This form describes the documentation and datasets to be distributed. Submit the completed, signed form to arichelp@unc.edu. The CC will process the form and send the investigator a secure invitation to upload materials.

Any data files that require manipulation by the CC (e.g., create derived variables) do NOT need this form and should follow the standard CC procedures for SC processing and distribution. Contact arichelp@unc.edu if you have question about whether or not to use this form.

Please complete one form per unique dataset. Datasets containing the same observations and variables, but saved as different file types require only 1 completed form, for example the same dataset in SAS and Stata. Datasets that contain different content should have separate submission forms.

II. Detailed Instructions for Each Item

Designated contact:

- **Name of designated contact:** Enter the first and last name of the individual who will be listed as the contact for questions related to the data described in this submission form. This designated contact must have an ARIC website account so their contact information is available via the ARIC directory. If necessary, request an ARIC website account using the 'Create new account' link on the ARIC website.

Dataset Information:

- **Dataset title:** Enter a descriptive title for the dataset. This could be the dataset label found in the SAS dataset (ex: Visit 6 Physical Activity Derived Dataset created on 13JUN2018).
- **Dataset name:** This is the specific dataset name (ex: v6paderived).
- **Dataset file type(s):** Check all file types included in the dataset submission. A single dataset can be submitted as multiple file types (ex: if SAS (.sas7bdat) and Stata (.dta) datasets are submitted, then check both "SAS" and "Stata").
- **If other, please specify:** Enter the dataset file type for those where other file type was selected (ex: using if "Other" was checked above for a CSV file, then enter "CSV" in this field).
- **Related ARIC component(s):** Check all ARIC components that were used to create the variables included in the dataset.
- **If visit or ancillary, please specify:** If visit or ancillary was checked above, then provide the visit number(s) and/or the ancillary study name(s).
- **Number of records:** Enter the number of observations or rows, excluding the header or variable names, in the dataset.

- **Number of variables:** Enter the number of variables, including the ID or SubjectID variable, in the dataset.
- **Multiple records per Subject ID:** If there are multiple rows or observation per ID, then select "Yes". If only one row or observation per ID, then select "No".
- **Labels provided for all variables:** Variables labels may be provided as part of the dataset itself and/or as part of the data documentation.
- **Units included:** If units are provided in the data documentation, labels, or variable name, then select "Yes". If no units are provided, then select "No". If none of the included variables have units, then select "N/A".
- **Formats/coded values included:** If variable formats or coded values are provided in the data documentation or labels, then select "Yes". If no formats or coded values are provided, then select "No". If none of the included variables have formats or coded values (i.e. all variables are continuous numbers or descriptive text strings), then select "N/A".
- **Dataset description:** Provide a brief description of the dataset. If separate documentation provides a description, then you may refer to that documentation.
- **Derived variable algorithm(s):** Provide the algorithm or definition(s) for any derived variables. If provided in separate documentation, then refer to the document with the details. This may look like a derived variable dictionary.
- **Source dataset(s) and variables:** List any datasets and/or variables used in the creation of the dataset and its derived variables. If provided in separate documentation, then you may refer to that documentation.
- **Analytic Recommendations:** Provide any information that might be helpful for other investigators to know when using the derived dataset. If provided in separate documentation, then you may refer to that documentation.

Approving Distribution of Dataset

- **Signature of designated contact:** The designated contact may sign using a digital signature or may print, sign by hand, and scan. By signing this form, the designated contact approves the distribution of this dataset and its documentation to all ARIC data recipients.
- **Date:** Enter the date the form is signed by the designated contact.

Coordinating Center Use Only – CSCC Staff ONLY

- **Dataset approved:** The assigned CC reviewer checks this field once they confirm the submitted dataset is approved.
- **Documentation approved:** The assigned CC reviewer checks this field once the submitted form and data documentation are approved. Before checking this box, the reviewer needs to confirm the designated contact has an account on the ARIC website and that the email listed in the ARIC directory is current.
- **Reviewer initials:** Once the assigned CC reviewer completes their review of the dataset and documentation, the CC reviewer enters their first, middle (if applicable), and last initials in this field.
- **Raw dataset location in SOURCE:** After the data are saved to the appropriate location in Source, the assigned CC reviewer enters the folder location here for reference.

- **Date approved:** The assigned CC reviewer enters the date that the CC approved and distributed the derived dataset and documentation. Data are distributed by posting a single zipped file to the Collaborator Derived section of the ARIC Secure Download Datasets webpage.