

## CHRONIC RENAL INSUFFICIENCY COHORT (CRIC) STUDY

### CRIC Data Archive: Recruitment, Follow-up Phase 2 Data, and CRIC Phase 3

The CRIC data archive contains data collected during the baseline and phase 2 follow-up visits for all core CRIC and Hispanic CRIC (HCRIC) participants. In total, the CRIC Study (CRIC and HCRIC combined) pre-screened 9,370 individuals, screened 6,277 potential enrollees, and recruited 3,939 participants from April 8, 2003 through September 3, 2008. This archive includes the study documentation, data collection forms, study data, and the dataset integrity check for the recruitment and follow-up phase 2 data up to 2018.

In addition, this data package includes data from CRIC phase 3 which includes data from new subjects separate from Phase 1 through 2018. The variable “ENROL\_PHS” in the “personlevel.sas7bdat” derived dataset notes whether the subject is from Phase 1 (1) or from Phase 3 (3).

The files in the archive are organized into the following directories:

1. Documentation
2. Forms
3. Data
4. DSICs

### 1. Documentation Directory

The Documentation Directory contains five subdirectories:

- The CRIC **Manual of Operations** Subdirectory houses all versions of the CRIC MOP (i.e., Versions 1.0 and 2.0), all amendments (i.e., Addendum 1 and 2), and a list of revisions from Version 1.0 to 2.0 corresponding to the recruitment period.
- The CRIC **Protocol** Subdirectory contains all versions (i.e., Versions 1.0 through 3.0) of the CRIC Study protocol and all amendments (i.e., Amendments 1-4) that governed the recruitment period of the study from April 8, 2003 through September 3, 2008.
- The CRIC **Publications** Subdirectory includes two files:
  - A list of CRIC publications through November 1, 2010 with links to PubMed abstracts can be found in “CRIC Publications\_20101101.pdf.”
  - A list of CRIC presentations to-date is provided in “CRIC Presentations\_20101220.xls.”
- The CRIC Phase 1 Subdirectory includes:
  - The CRIC **Contents** Subdirectory includes a contents of the datasets including:
    - The folder “Dataset contents\_derived”, containing:
      - The spreadsheet “CRIC MAY2020\_NIDDK\_Derived\_Variables\_Dictionary.xlsx”, which is the high-level data dictionary for the variables in the derived personlevel , sa\_allc, sa\_cvd, sa\_cvd\_cnsr, visitlevel, all\_cvd\_events, anc\_pwv, anc\_vitd, ecg

- A document (contents\_final\_mdata.pdf) containing the contents of the nine derived datasets.
  - A subfolder folder called “Derived variables\_documentation, which includes documents providing definitions for the derived variables.
  - all\_cvd\_events\_contents.pdf
  - anc\_pwv\_contents.pdf\*
  - anc\_vitd\_contents.pdf\*
  - CRIC SEP2019\_NIDDK\_Derived\_Variables\_Dictionary.xlsx\*
  - ecg\_contents.pdf\*
  - personlevel\_contents.pdf\*
  - sa\_allc\_contents.pdf\*
  - sa\_cvd\_cnsr\_contents.pdf\*
  - sa\_cvd\_contents.pdf\*
  - visitlevel\_contents.pdf\*
- The folder “Dataset contents\_raw”, containing:
    - SAS PROC CONTENTS files for each of the twenty-one CRF and Non-CRF raw SAS data files.
    - A folder called “dictionaries and documents”, containing:
      - Two high-level data dictionaries-- CRF Data Dictionary.xls and Non-CRF Data Dictionary.xls.
      - A listing of the raw and intermediate data files, “Raw and Intermediate Data File List.xls”.
      - Documentation for the Diet History Questionnaire calculations.
      - Documentation about the physical activity “PHYACT” data preparation.
- The CRIC Phase 3 Subdirectory includes:
    - The CRIC **Contents** Subdirectory includes a contents of the datasets including:
      - The folder “Dataset contents\_derived”, containing:
        - The spreadsheet “CRIC 112015\_NIDDK\_Analysis Data\_Dictionary.xlsx”, which is the high-level data dictionary for the variables in the derived personlevel , sa\_allc, sa\_cvd, sa\_cvd\_cnsr, visitlevel.
        - A document (contents\_final\_mdata.pdf) containing the contents of the six derived datasets.
        - A subfolder folder called “Derived variables\_documentation, which includes documents providing definitions for the derived variables.

- The folder “Dataset contents\_raw”, containing:
  - SAS PROC CONTENTS files for each of the twenty-one CRF and Non-CRF raw SAS data files.
  - A folder called “dictionaries and documents”, containing:
    - Two high-level data dictionaries-- CRF Data Dictionary.xls and Non-CRF Data Dictionary.xls.
    - A listing of the raw and intermediate data files, “Raw and Intermediate Data File List.xls”.
    - Documentation for the Diet History Questionnaire calculations.
    - Documentation about the physical activity “PHYACT” data preparation.
- The CRIC **Formats** Subdirectory includes the following formats file:
  - ‘CRIC\_2019\_formats.pdf’ – decodes for each of the formats, the SAS format data saved at ‘P:\niddk\dataset\_files\CRIC\_V10\Phase 1 and Phase 3\formats.sas7bcat’

## 2. Forms Directory

The Forms Directory contains the following:

- A spreadsheet named “CRIC CRF\_Data File\_Linkage.xls” that links each of the Case Report Forms (CRFs) to its corresponding dataset. Multiple versions of some data forms are provided. It is noted that variables were added as the study proceeded; consequently a large number of missing values may be found for some variables (see Table 1 below).
- The set of 26 data entry CRFs (in pdf format) corresponding to data transferred to the Repository from the screening and baseline visits. All versions of a given CRF active during the recruitment period are provided.
- A pdf file named “CRIC Copyrighted CRFs” provides the names of copyrighted forms used in CRIC recruitment. Users should note due to copyright issues, these forms are not released by the repository.

## 3. Data Directory

The Data Directory contains the following:

- The **CRIC Ancillary Data** Subdirectory includes folders for any of 10 manuscripts that have been published as of 11/1/2010 and included data analysis. Each manuscript folder (see Table 2) includes at a minimum:
  - A “List of files sent” document
  - A SAS analytic data file (.sas7bdat)
  - A SAS Proc Contents output file
  - A formats file

Table 1. Manuscript title, journal and date of publication, and corresponding file name.

File Name	Manuscript Title	Journal	Year of publication
MS004	Chronic Renal Insufficiency Cohort (CRIC) Study: baseline characteristics and associations with kidney function	Clin J Am Soc Nephrol	2009
MS005	Hypertension awareness, treatment, and control in adults with CKD: results from the Chronic Renal Insufficiency Cohort (CRIC) Study	Am J Kid Dis	2010
MS008	Variability of creatinine measurements in clinical laboratories: results from the CRIC study	Am J Nephrol	2010
MS016	Chronic kidney disease and cognitive function in older adults: findings from the chronic renal insufficiency cohort cognitive study	J Am Geriat Soc	2010
MS018	Aortic pulse pressure is associated with carotid IMT in chronic kidney disease: report from Chronic Renal Insufficiency Cohort.	Am J Hyper	2009
MS022	Aortic PMV in chronic kidney disease: a CRIC ancillary study	Am J Hyper	2010
M0022a	Correlation between pulse wave velocity and other measures of arterial stiffness in chronic kidney disease	Clinical Nephrology	2007
MS0034	Prevalence of ocular fundus pathology in patients with chronic kidney disease	Clin J Am Soc Nephro	2010
MS037	Chronic kidney disease and prevalent atrial fibrillation: the Chronic Renal Insufficiency Cohort (CRIC)	Am Heart J	2010
MS065	Central pulse pressure in chronic kidney disease: A Chronic Renal Insufficiency Cohort Ancillary Study	Hypertension	2010

A pdf document, “CRIC Manuscript\_Analytical Sections.pdf,” provides complete citations and describes the statistical methods used for all published manuscripts.

- The **CRIC Study Data** Subdirectory provides:
  - Phase 1 and Phase 3:
    - A CRIC **Derived Data** Subdirectory containing nine SAS derived data files: “all\_cvd\_events.sas7bdat”, “anc\_pwv.sas7bdat”, “anc\_vitd.sas7bdat”, “ecg.sas7bdat”, “personlevel.sas7bdat”, “sa\_allc.sas7bdat”, “sa\_cvd\_cnsr.sas7bdat”, “sa\_cvd.sas7bdat”, and “visitlevel.sas7bdat” with variables copied directly from the raw data files, as well as calculated and categorical variables based on variables from the raw data files.
    - A CRIC **Raw Data** Subdirectory containing thirty-four SAS CRF and Non-CRF raw and intermediate data files pertaining to the CRIC recruitment and follow-up time frame.
    - formats.sas7bcats: containing the CRIC SAS format library.

#### 4. DSICs

The CRIC Dataset Integrity Checks (DSICs) directory contains four reports based on an examination of the data

in the repository by statisticians and quality control specialists at the Repository: Atrial\_Fibrillation\_DSIC, Baseline\_DSIC, NGAL\_DSIC, Followup\_Phase\_2\_DSIC, and Followup\_Phase\_3\_DSIC. Datasets were checked for completeness, consistency, and usability. Published data were compared to values recalculated from the CRIC data in the NIDDK Repository.