CHRONIC KIDNEY DISEASE IN CHILDREN COHORT STUDY (CKiD)

Data Archive

The CKiD Study is a multi-center, cohort study of children aged 6 months to 16 years (at enrollment) with mild to moderate impaired kidney function. Two clinical coordinating centers (CCCs) (at Children's Hospital of Philadelphia and at Children's Mercy Hospital in Kansas City), a central biochemistry laboratory (at the University of Rochester), and a data coordinating center (at Johns Hopkins Bloomberg School of Public Health) have formed a cooperative agreement to conduct a prospective study of chronic kidney disease in children. The study population currently consist of two cohorts. Recruitment of the first cohort of 586 children occurred from January 2005 through August 2009 at 48 pediatric nephrology programs across the United States and two sites in Canada. Recruitment of the second cohort of 305 children occurred from February 2011 to March 2014. Participants complete annual visits and since its inception in 2003, the scientific aims of CKiD have been to determine the risk factors for decline in kidney function and to define how progressive decline in kidney function impacts biomarkers of risk factors for cardiovascular disease; neurocognitive function and behavior; and growth failure and its associated morbidity.

The CKiD data archive contains data collected for Cohort 1 and Cohort 2 at baseline and follow-up up to July 31, 2018. The baseline visit is comprised of two components: V1a and V1b. The first component of the baseline visit (V1a) occurred during the participant's initial visit to the clinical site and included procedures to obtain an iohexol-based GFR measurement. For the purpose of obtaining baseline data on neurocognitive function and growth, the second component of the baseline visit (V1b) occurred within 3 months after the first baseline visit for children between 1 and 3 years of age, and occurred within 6 months for children over the age of 3. The one to six month lag between the two components of visit one was necessary because the procedures needed to measure the GFR did not provide an environment conducive to an unbiased assessment of some of the neurocognitive tests. Markers related to the four specific aims were measured every year (e.g., serum creatinine, Cystatin C, standardized clinical blood pressure obtained with a uniformed centrally-calibrated device, pediatric quality of life, height and weight). GFR was measured annually during the first two years and then every two years thereafter. The cardiovascular component was implemented concurrently with the kidney component after the second year. In turn, the core markers of neurocognitive function and growth was measured in the odd visit years from the third year on.

1. CKiD Documents

The CKiD Documents directory contains:

- **CKiD General Description.** (filename: CKiD General Description_08182020.pdf)
- **CKiD Manual of Procedures (MOP).** As of 2018, the manual was comprised of 56 PDF files including a table of contents and list of version dates for each section. Changes to sections of the MOP are performed throughout the year and are noted by the date in the footer of each section. The directory contains one compiled document of the MOP as of 2018. (filename: CKiD Manual of Procedures_04-01-18.pdf)
- **CKiD Publications.** List of all CKiD study publications as of August 18, 2020. (filename: CKiD Publications_081820.pdf)
- **CKiD Protocol.** Protocol amendments are normally performed annually. The directory contains the 2017 protocol. (filename: CKiD Study Protocol OSMB 06-01-16 to 08-01-17_clean.pdf)

2. CKiD Forms

All CKiD forms contain a footer with the date the form was implemented. Subsequent changes to forms are indicated in the footer in one of two ways:

1. In the case of substantial revision to a form (i.e., a change in the data collected or the addition or deletion of variables), the change will be indicated by assigning a new date to the form. For example, if F01 version 01/01/05 received substantial changes, the version date in the footer would change to the date of implementation of the new version (e.g., from 01/01/**05** to 01/01/**06**).
2. In the case of minor revision to a form (i.e., no change in data collected – correction of typo or change in wording), the change will be indicated by adding a letter after the date in the footer. For example, if F01 version 01/01/05 received minor changes, an "a" would be added to the end of the version date in the footer. If F01 version 01/01/05a were to receive subsequent minor changes, the "a" in the footer would be changed to a "b" (e.g., from 01/01/**05** to 01/01/**05a** to 01/01/**05b**).

The CKiD Forms directory contains baseline forms in PDF format. Multiple versions of some data forms are provided and as previously noted changes to the forms are noted in the footer.

Name	Description
ADVR	Adverse Event Form
F01	Symptoms List
F02	Smoking, Alcohol, Drug Use and Physical Activity
F12	Smoking, Alcohol, Drug Use and Physical Activity Follow-up
F13	General History Follow-up (includes repeating segment on siblings, F13S1)
F13a	Abbreviated General History Follow-up (includes repeating segment on siblings, F13S1)
F14	Medical History Follow-up
F15	Nutritional Assessment and Steroid Use (includes repeating segment on supplements, F15S1)
F16	Sun Exposure
F17	Overall Physical Activity
F19	Hand Grip Test
GH	General History (includes repeating segment on siblings, GHS1)
HLC01	Health Literacy Assessment Coding Sheet
L01	Specimen Collection Form for Visit 1a
L02	Specimen Collection Form for Visit 1b
L03	Local Laboratory – Renal Panel Results
L04	Local Laboratory – CBC Results
L05	Central Laboratory – Renal Panel Tests
L06	Local Laboratory – Urine Assay Results
L07	Central Laboratory – Iohexol Concentration Results
L08	Central Laboratory – Intact Parathyroid Hormone (iPTH) and Wide Range C-Reactive Protein (wrCRP)
L09	Central Laboratory – Lipid Profile
L11	Central Laboratory – Cystatin C Results
L12	Central Laboratory – Iron Tests
L13	Central Laboratory – Vitamin D
L21	Specimen Collection Form for Visit 2
L31	Specimen Collection Form for Visit 3
L41	Specimen Collection Form for Visit 4
L51	Specimen Collection Form for Visit 5 *(form discontinued)*
MEDS	Medication and Supplement Inventory (MEDSUM_FULL)
MH	Medical History (includes repeating segment on nutritional supplements, MHS1)
NRC03A	Cognitive/Development Data Coding Sheet: 12 to < 30 months
NRC03B	Cognitive/Development Data Coding Sheet: 30 months to < 4 years
NRC03C	Cognitive/Development Data Coding Sheet: 4 to < 6 years
NRC03D	Cognitive/Development Data Coding Sheet: 6 to < 17 years and older
NRC03E	Cognitive/Development Data Coding Sheet: 17 years and older
NRC04A	Behavioral Data Coding Sheet: 12 to < 24 months
NRC04B	Behavioral Data Coding Sheet: 2 to < 6 years
NRC04C	Behavioral Data Coding Sheet: 6 to < 8 years
NRC04D	Behavioral Data Coding Sheet: 8 to < 12 years
NRC04E	Behavioral Data Coding Sheet: 12 to < 18years and older
NRC04F	Behavioral Data Coding Sheet: 18 to < 21 years and older
NRC04G	Behavioral Data Coding Sheet: 21 years and older

PE	Physical Exam
PFU01	Phone/In-Person Follow-up Interview Form (includes repeated segment, PFU01S1)
PFU02	Follow-up Site Questionnaire
WFU01	Web-based Follow-up Survey

Prior to 2016, the following forms were not included: F19, HCL01 and WFU01.

3. CKiD Study Data

The CKiD Main Study Data directory contains 52 SAS data files. Each dataset has corresponding codebooks (<Filename>.txt) which are text files and should be opened with a text editor. Of the 52 data files, five (5) of the data files contain repeating segment data: GHS1, F13S1, MHS1, F15S1 and PFU01S1. Repeating segment data are groups of similar data captured within the context of a form. For example, date of birth for each sibling, which is considered repeating segment data, is captured within the general history (GH, F13, F13a (follow-up)) form and the data is stored in GHS1/F13S1. Similarly, information on the types and amounts of nutritional supplements taken is captured within the medical history (MH), nutritional assessment (F15) and Phone/In-Person Follow-up Interview (PFU01) forms, and stored in MHS1/F15S1 and PFU01S1.

In addition to datasets for each of the study and laboratory forms, the directory includes 14 summary files: GFRCALIBRATEDSUMMARY (formerly referred to as GFRSUMMARY), KIDHIST, LABMARKERS, GROWTH, CARDIO, ABPM, ECHO, CIMT, MEDSUM_SHORT, NEUROSUMMARY, SOCDEM, NUTRIENTS, GRIPSTRENGTH and VERT_DATEBASE. The **GFRCALIBRATEDSUMMARY** data file contains the glomerular filtration rate measurements based on calibrated iohexol concentrations and IFCC calibrated cystatin C. As of 2019, cystatin C results are standardized to International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) certified reference material; therefore, to convert pre-IFCC results to IFCC-calibrated concentrations, the value is multiplied by 1.17. In GFRCALIBRATEDSUMMARY, new IFCC-calibrated cystatin C concentrations (*CYC_IFCC*) replaced the pre-IFCC cystatin C concentrations (*CYC_DB*). In addition to the calibrated iohexol-based GFR we also provide a bedside estimated GFR using height and serum creatinine and an estimated GFR based on height/serum creatinine and cystatin C and/or BUN (*E2012GFR*). For additional information regarding *E2012GFR* refer to Schwartz et al; Kidney International 2012; 82:445-453. The variables in **KIDHIST** include primary chronic kidney disease (CKD) diagnosis, CKD diagnosis group, date of CKD onset, dates associated with birth, baseline and clinical endpoints of the study (i.e., date of transplant, dialysis, death) and last date free of renal replacement therapy (RRT). For participants who did not experience renal replacement therapy prior to the administrative censoring date (July 31, 2018), a random value between 0.01 and 0.26 was added to the variable duration of time between baseline visit and last date free of renal replacement therapy (*LDATRTFREE* in KIDHIST). The same is done for the variable *LDATALIVE* in KIDHIST, in which participants alive as of July 31, 2018 have a random number between 0.01 and 0.26 added to their time from baseline visit and last date alive. The rationale for doing this is to prohibit the user from determining personal identifiers, namely back-calculating the exact date of study entry (i.e., baseline date) and potentially the exact date of birth. The **LABMARKERS** data file contains variables for laboratory markers such as basic metabolic panel, complete blood count, urine analysis, intact parathyroid, c-reactive protein, lipid panel and iron results as well as calculated proteinuria, acidosis, hypoalbuminemia, abnormal calcium and phosphate (based on KDOQI thresholds), calcium-phosphate product, elevated CRP, anemia and hemoglobin z-scores and percentiles based on age, sex and race per CDC guidelines. The **GROWTH** data file contains height, weight and body mass index (BMI) average, percentile and z-score

variables, and birth history variables. The **CARDIO** data file contains the casual/clinical blood pressure percentile, z-score and index variables. The **ABPM** data file contains ambulatory blood monitoring (ABP) variables. The **ECHO** data file contains echocardiogram measurement data. The **CIMT** data file contains variables for the measurement of carotid intimal medial thickness (cIMT) in a subset of the study population. The **MEDSUM_SHORT** data file contains one record per medication per participant-visit and summarizes whether or not the study participant has been prescribed during the past 30 days any medication that falls into one of several major medication classes including antihypertensives, ESAs, growth hormones, immunosuppressives, anticholinergics, and antidepressants. More detailed information for the medications including dosing amounts, schedules and adherence are provided in the MEDSUM_FULL data file. **NEUROSUMMARY** data file contains key variables from the neurocognitive and behavioral battery, and quality of life data. **SOCDEM** data file contains sociodemographic variables such as race, ethnicity, income and maternal education. **NUTRIENTS** data file contains keys variables from the food frequency questionnaires to assess food intake. Individual level data is summarized into single variables. **GRIPSTRENGTH** data file contains variables from the grip strength assessment. **VERT_DATEBASE** data file contains all the study visits, and status of each visits (i.e., regular, irregular, transitional or disenrolled).

To enhance the existing quality control measure, a supplementary measure was performed. For each variable in each summary file, the first and third quartiles (25th and 75th percentiles), Q1 and Q3 were calculated. Any values that were larger than Q3/Q1 or smaller than 4×Q1 – 3×Q3 were flagged as "far-out values". These limits were based on the methodology of Tukey fences: 3 Tukey fences in the log scale for high values and in the raw scale for low values. The implementation of this quality control methodology yielded subsequent checks for values in the LABMARKERS and ECHO summary files. Flagged values were double-checked to determine whether the measurement entered was a feasible value based on the distribution of the data and/or the participant's previous history for that measurement. The percentages of flagged values were 0.1% in the LABMARKERS and <0.1% in the ECHO summary files. There were no values flagged in the other summary files. After further investigation of the values in the LABMARKERS and ECHO summary files, approximately 20% of the values were corrected. Uncorrected values were recorded as "-90" to indicate that original values was a far-out value.

Prior to 2017, the following summary files were not included in the data upload: LABMARKERS, ABPM and CIMT.

Prior to 2020, the following summary files were not included in the data upload: NUTRIENTS ad GRIPSTRENGH. Also prior to 2020, quality of life data was not included in the NUEROSUMMARY summary files that was uploaded.

No. of Records	Dataset
1582	abpm.sas7bdat
112	advr.sas7bdat
6612	cardio.sas7bdat
477	cimt.sas7bdat
1709	echo.sas7bdat
4519	f01.sas7bdat
443	f02.sas7bdat
2136	f12.sas7bdat
1389	f13a.sas7bdat
2255	f13.sas7bdat

2574	f13s1.sas7bdat
3638	f14.sas7bdat
4359	f15.sas7bdat
192	f15s1.sas7bdat
1665	f16.sas7bdat
1928	f17.sas7bdat
884	f19.sas7bdat
6074	gfrcalibratedsummary.sas7bdat
894	gh.sas7bdat
1082	ghs1.sas7bdat
729	gripstrength.sas7bdat
6679	growth.sas7bdat
624	hlc01.sas7bdat
891	kidhist.sas7bdat
891	l01.sas7bdat
827	l02.sas7bdat
4840	l03.sas7bdat
4947	l04.sas7bdat
4946	l05.sas7bdat
3631	l06.sas7bdat
2946	l07.sas7bdat
2616	l08.sas7bdat
2233	l09.sas7bdat
4761	l11.sas7bdat
2545	l12.sas7bdat
124	l13.sas7bdat
807	l21.sas7bdat
1635	l31.sas7bdat
1428	l41.sas7bdat
179	l51.sas7bdat
5776	labmarkers.sas7bdat
21458	medsum_full.sas7bdat
5172	medsum_short.sas7bdat
891	mh.sas7bdat
14	mhs1.sas7bdat
4649	neurosummary.sas7bdat
29	nrc03a.sas7bdat
55	nrc03b.sas7bdat
126	nrc03c.sas7bdat
1875	nrc03d.sas7bdat
356	nrc03e.sas7bdat
10	nrc04a.sas7bdat
200	nrc04b.sas7bdat
176	nrc04c.sas7bdat
617	nrc04d.sas7bdat
1290	nrc04e.sas7bdat
219	nrc04f.sas7bdat
93	nrc04g.sas7bdat
3219	nutrients.sas7bdat
5776	pe.sas7bdat
792	pfu01.sas7bdat
824	pfu01s1.sas7bdat
835	pfu02.sas7bdat
5775	socdem.sas7bdat
7562	vert_datebase.sas7bdat
103	wfu01.sas7bdat

Data Issues
Baseline and follow-up data for participants (586 Cohort 1 and 305 Cohort 2) through July 31, 2018, is included in the archived data.

All dates, except as noted, in the files have been converted to duration in years from baseline enrollment visit to protect the identity of participants. Dates of birth and baseline enrollment in KIDHIST are provided as year only.

All variables which contain information which might be used to identify the participant (e.g., unusual physical characteristics, geographic information) have been set to missing. A list of these variables and their description is below:

Table	Variable	Description
ADVR	AEINITL	Initials of person completing form
ADVR	AEIRSPOT	Specify other iohexol reaction
F01	SYRVINIT	Initials of person reviewing form
F02	ADSMKAGE	Age at which first smoked whole cigarette
F02	ADSMQTYR	Age quit smoking - years
F02	ADSMKDAV	Average number cigarettes smoked per week before quitting
F02	ADSMAVRG	Average number cigarettes smoked per week now
F12	ADINITL	Initials of person completing form
F13	GHINITL	Initials of person completing form
F13/F13a	GHRVINIT	Initials of coordinator reviewing form
F13/F13a	GHCURZIP	Current zip code of child's primary household
F13/F13a	GHPRVZIP	Previous zip code of child's primary household
F13/F13a	GHPGCZIP	Current zip of parent/guardian
F13/F13a	GHSPOAPH	Specify other adult living in primary household
F14	MHINITL	Initials of person completing form
F14	MHRVINIT	Initials of person reviewing form
F14	MHOTINSP	Specify other type of insurance
F14	MHRCOTSP	Specify other health care location
F15	NSINITL	Initials of person completing form
F16	SEINITL	Initials of person completing form
F17	PAINITL	Initials of person completing form
F19	HGINITL	Initials of person completing form
GH	GHINITL	Initials of person completing form
GH	GHRVINIT	Initials of coordinator reviewing form
GH	GHCURZIP	Current zip code of child's primary household
GH	GHPRVZIP	Previous zip code of child's primary household
GH	GHSPOAPH	Specify other adult living in primary household
GH	GHYRMOVD	Year participant moved to U.S. or Canada
GH	GHYRTOUS	Year participant moved to U.S.
GHS1	GHSIBDOB	Sibling’s date of birth
HLC01	HLINITL	Initials of person completing form

HLC01	HLFCINIT	Initials of person completing form
L01	SCINITL	Initials of person completing form
L01	SCFAXNUM	Fax number for results
L01	SCRCNAME	Recipient's name
L02	SPINITL	Initials of person completing form
L02	SPWBINIT	Whole blood initials
L02	SPSPOTFN	Specify other reason 10 fingernails not collected
L02	SPSPOTHR	Specify other reason 20 hair fibers not collected
L02	SPSPOTTN	Specify other reason 10 toenails not collected
L03	BMINITL	Initials of person completing form
L04	CBINITL	Initials of person completing form
L06	UAINITL	Initials of person completing form
L06	UARNAVAL	Specify reason urine assay test results not available
L07	GCRNAVAL	Specify reason Iohexol concentration results not available
L11	CCRNAVAL	Specify reason Cystatin C test results not available
L21	SCINITL	Initials of person completing form
L21	SCFAXNUM	Fax number for results
L21	SCRCNAME	Recipient's name
L21	SPCSPOTTN	Specify other reason 10 toenails not collected
L23	VTINITL	Initials of person completing form
L31	SCINITL	Initials of person completing form
L31	SMWBINIT	Whole blood initials
L41	SCINITL	Initials of person completing form
L41	SCSPOTTN	Specify other reason 10 toenails not collected
L51	SMINITL	Initials of person completing form
L51	SMWBINIT	Whole blood, initials
MH	MHINITL	Initials of person completing form
MH	MHRVINIT	Initials of person reviewing form
MH	MHOTINSP	Specify other type of insurance
MH	MHRCOTSP	Specify other health care location
NRC03A	C1PSINIT	Initials of psychologist completing form
NRC03B	C2PSINIT	Initials of psychologist completing form
NRC03C	C3PSINIT	Initials of psychologist completing form
NRC03D	C4PSINIT	Initials of psychologist completing form
NRC03E	C5PSINIT	Initials of psychologist completing form
NRC04A	B1INIT	Initials of coordinator reviewing form
NRC04B	B2INIT	Initials of coordinator reviewing form
NRC04C	B3INIT	Initials of coordinator reviewing form
NRC04D	B4INIT	Initials of coordinator reviewing form
NRC04E	B5INIT	Initials of coordinator reviewing form
NRC04F	B6INIT	Initials of coordinator reviewing form
NRC04G	B7INIT	Initials of coordinator reviewing form
PE	PEINITLS	Initials of examiner
PE	PEBPINIT	Initials of blood pressure reader
PE	PESPTANR	Specify problem completing Tanner Staging

PE	PESPVITL	Problem completing – vital signs
PE	PESPBLPM	Problem completing – blood pressure
PFU01	PFINITL	Initials of person completing form
PFU01	PFCURZIP	Current zip code of child's primary household
PFU01	PFPROMSP	Specify other method reporting vital status
PFU01	PFMMRKT	Date most recent transplant – month
PFU01	PFDMRKT	Date most recent transplant – day
PFU01	PFYMRKT	Date most recent transplant – year
PFU01	PFMWLIST	Date listed for deceased donor transplant – month
PFU01	PFDWLIST	Date listed for deceased donor transplant – day
PFU01	PFYWLIST	Date listed for deceased donor transplant – year
PFU01	PFMRRDLY	Date most recent regular dialysis started – month
PFU01	PFDRRDLY	Date most recent regular dialysis started – day
PFU01	PFYRRDLY	Date most recent regular dialysis started – year
PFU01	PFMMRDLY	Date most recent dialysis started – month
PFU01	PFDMRDLY	Date most recent dialysis started – day
PFU01	PFYMRDLY	Date most recent dialysis started – year
PFU02	P2INITL	Initials of person completing form
PFU02	P2SCLAB	Lab used to measure serum creatinine
PFU02	P2IRNAVL	Specify reason iron tests results not available
WFU01	PFMMRKT	Date most recent transplant – month
WFU01	PFDMRKT	Date most recent transplant – day
WFU01	PFYMRKT	Date most recent transplant – year
WFU01	PFMWLIST	Date listed for deceased donor transplant – month
WFU01	PFDWLIST	Date listed for deceased donor transplant – day
WFU01	PFYWLIST	Date listed for deceased donor transplant – year
WFU01	PFMMRDLY	Date most recent dialysis started – month
WFU01	PFDMRDLY	Date most recent dialysis started – day
WFU01	PFYMRDLY	Date most recent dialysis started – year