CHEAR 2016-34

This file contains general information about the CHEAR project data files you are downloading.

Contents:

1. CITING CHEAR DATA

The publication policy, which you agreed to upon registering to the CHEAR data repository, provides guidelines about citations.
https://mychear.chearprogram.org/sites/default/files/CHEAR_Pub_Policy_Final.pdf

CHEAR resources are made possible by NIH funding and must be properly acknowledged in manuscripts (per the journal's requirement), abstracts, posters, and presentations, one or more statements should specify financial acknowledgement as follows:
"Research reported in this publication was supported by the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number [grant(s) (see Table below)]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health."

Investigators who are using *any* CHEAR data that is publicly available should also cite the DOI used to identify the data in the CHEAR Data Repository. This will ensure that the original generators of the data sets will get credit and allow readers to locate and evaluate the provenance of the data.

Grant numbers:

The CHEAR Data Center:

Institution	Principal Investigator	Title	Grant Number
ICAHN School of Medicine at Mount Sinai	Teitelbaum, Susan L.	CHEAR Center for Data Science	U2CES026555
Research Triangle Institute	Fennell, Timothy R	RTI CHEAR Exposure Assessment Hub	U2CES026544

Questions about this policy should be directed to: HHEAR_CC@westat.com.

2. CHEAR GLOSSARY

Participant ID Format

The format of the CHEAR Participant ID is as follows:

Participant ID (PID) = CHEAR assigned participant ID with the structure XXXXXXX. It is a 7 digit numeric code that will uniquely identify a CHEAR participant

Sample ID Format

The format of the CHEAR Sample ID is as follows:

Sample ID (SID) = CHEAR assigned sample ID with the structure **C-XXXXX-SP** (example: C-3XA5P-U)

- C = CHEAR (fixed character prefix)
- Core ID is a 5 digit alpha-numeric code **XXXXX** that will uniquely identify a CHEAR biological sample
- SP=specimen type (study-relevant code listing below)

  BC = Buffy Coat
  BR = Breast Milk
  BU = Buccal Cell
  CB = Cord Blood
  CSF = Cerebrospinal Fluid
  D = DNA
  DBS = Dried Blood Spots
  H = Hair
  ME = Meconium
  N = Nails
  P = Plasma
  PL = Placenta
  R = RNA
  RBC = Red Blood Cells
  S = Serum
  SL = Saliva
  ST = Stool
  SU = Sputum
  T = Teeth
  TI = Tissue
  U = Urine
  WB=Whole Blood
  WBC = White Blood Cells

Targeted Analyte Codes

Chemical Group Code	***Full Chemical Group Name***	***Code***	***Full Chemical Name***	***CAS #***
UDILUTE	Urinary Dilution Factor	CRE	Creatinine	60-27-5
UTE	Trace Elements in urine	Pb	Lead	7439-92-1
UTE	Trace Elements in urine	Mn	Manganese	7439-96-5
UTE	Trace Elements in urine	As	Arsenic	7440-38-2
UTE	Trace Elements in urine	Mo	Molybdenum	7439-98-7
UTE	Trace Elements in urine	Cd	Cadmium	7440-43-9

Targeted Comment Codes

Code	***Definition***
0	Valid measurement
37	Value less than LOD
97	Result repeated with diluted sample

3. DESCRIPTION OF PUBLIC FILES

File Name	*Source*	*File Description*	*DOI**
34_EPI_DATA.csv	Parent study	Epidemiological data submitted from parent study.	10.36043/34_94
34_EPI_DDCB.xlsx	Parent study	Data dictionary and codebook for epidemiological data submitted from parent study.	N/A
34_TARGETED_DATA.csv	CHEAR	Metals and creatinine data in urine samples analyzed by a CHEAR Lab Hub.	10.36043/34_732
34_EPI_META_UM_DATA.csv	CHEAR	Analytic dataset including metabolomics in serum, metals in urine and epidemiologic data for subjects with metabolomics data.	10.36043/CHEAR-2016-34-META-01

*These DOIs do not function as journal article DOIs, which are direct links to an article. Instead, go to https://www.doi.org/ and paste in the HHEAR dataset DOI to access the data. You must have a HHEAR Data Submission and Review Portal (DSRP) Account to access the data.

4. ANALYTIC NOTES

- Please contact the Principal Investigator (PI) of the CHEAR study for questions related to datasets indicated with source "Parent study" on the table in the table above and contact hhearsupport@mssm.edu for all other questions.
    - PI name and contact information at the time of this publication:
      David Christiani, MD, MPH, Harvard School of Public Health (dchris@hsph.harvard.edu)
- Individual lab and epidemiological datasets can be merged on PID variable.
- Dataset *34_EPI_META_UM_DATA.csv* contains epidemiologic and lab data only for subjects with valid metabolomics data.
    - Data structured to be long by subject (PID), method of metabolomics analysis (META_Method), and analysis phase/mode (META_Mode).
- Metadata and source data for metabolomics assays can be found on Metabolomics Workbench using the links provided.