CHEAR Project # 2017-1945

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

Contents:
1. Citing CHEAR data
2. CHEAR Glossary
3. Descriptions of public files
4. Analytic Notes

1. CITING CHEAR DATA

    A. The publication policy, which you agreed to upon registering to the CHEAR data repository, provides guidelines about citations.
       https://hhearprogram.org/sites/default/files/2020-01/HHEAR-ExternalPublicationsPolicy-2019-10-21-508_0.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."

Grant numbers:

Institution	Principal Investigator	Title	Grant Number
Icahn School of Medicine at Mount Sinai	Susan L. Teitelbaum	CHEAR Center for Data Science	U2CES026555

University of Michigan	John D. Meeker	M-CHEAR: Michigan Children's Health Exposure Analysis Resource Laboratory Hub	U2CES026553

B. Publicly available CHEAR data are those data that have been generated by the CHEAR lab(s) combined with the data submitted by CHEAR project investigators. External research investigators who are using *any* publicly available CHEAR data should cite:

    i. The original paper(s) associated with each CHEAR project included in your manuscript. If available, are located [*HHEAR Data Center website study page* hyperlink].

    ii. The DOIs (Digital Object Identifier) associated with the data files used in your manuscript. DOIs are located [*HHEAR Data Center website public study page* hyperlink] as well as section 3 below.

    iii. We also suggest including the following statement in your manuscript:

Data for this project was obtained from the publicly available data in the Human Health Exposure Resource (HHEAR) Data Repository, which has been approved under Icahn School of Medicine at Mount Sinai IRB Protocol # 16-00947.

All of 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.

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

Child Participant ID (CHILD_PID) = When biological specimens are provided by the mother, PID is used to designate who the sample is from (mother or child), and a child PID is also included to designate which child the data is related to if mothers are providing data for more than one child.

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 (1-3 letter code, code list below)

    D = DNA
    U = Urine

Targeted Analyte Codes

Chemical Group Code	Full Chemical Group Name	Code	Full Chemical Name	CAS#
UTE	Urinary trace elements	As	Arsenic	7440-38-2
UTE	Urinary trace elements	Ba	Barium	7440-39-3
UTE	Urinary trace elements	Be	Beryllium	7440-41-7
UTE	Urinary trace elements	Cd	Cadmium	7440-43-9
UTE	Urinary trace elements	Cs	Cesium	7440-46-2
UTE	Urinary trace elements	Cr	Chromium	7440-47-3
UTE	Urinary trace elements	Co	Cobalt	7440-48-4
UTE	Urinary trace elements	Cu	Copper	7440-50-8
UTE	Urinary trace elements	Hg	Mercury	7439-97-6
UTE	Urinary trace elements	Mn	Manganese	7439-96-5
UTE	Urinary trace elements	Mo	Molybdenum	7439-98-7
UTE	Urinary trace elements	Ni	Nickel	7439-98-7
UTE	Urinary trace elements	Pb	Lead	7439-92-1
UTE	Urinary trace elements	Pt	Platinum	7440-06-4
UTE	Urinary trace elements	Sb	Antimony	7440-36-0
UTE	Urinary trace elements	Sn	Tin	7440-31-5
UTE	Urinary trace elements	Tl	Thallium	7440-28-0
UTE	Urinary trace elements	U	Uranium	7440-61-1
UTE	Urinary trace elements	V	Vanadium	7440-62-2
UTE	Urinary trace elements	W	Tungsten	7440-33-7
UTE	Urinary trace elements	Zn	Zinc	7440-66-6

Targeted Comment Codes

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

3. DESCRIPTION OF PUBLIC FILES

Chemical Group Code	Full Chemical Group Name	Code	Full Chemical Name
1945_TARGETED_DATA.csv	CHEAR/ Parent study	Urinary trace elements concentrations in pregnant mothers from 1st and/or 3rd trimesters from CHEAR, and specific gravity concentrations measured by PI.	10.36043/1945_159
1945_UTE_METHODS.docx	CHEAR	Laboratory methods for measurement of urinary trace elements.	
1945_EPI _DATA.csv	Parent study	Epidemiological data including covariates and health outcomes for MADRES cohort.	10.36043/1945_177
1945_EPI_DDCB.xlsx	Parent study	Data dictionary and codebook for epidemiological data.	
MADRES First Trimester Questionnaire_English v. 02.08.17.PDF	Parent study	Questionnaires from Maternal and Developmental Risks from Environmental and Social Stressors (MADRES) cohort data collection.	
MADRES Second Trimester Questionnaire_Spanish_Abridged.PDF			
MADRES Third Trimester Questionnaire_English_Abridged.PDF			
MADRES Third Trimester Questionnaire_Spanish_Abridged.PDF			
1344-methylation_calls.tar.gz	CHEAR	Epigenetics data from infant cord blood: CpG Methylation calls in BedGraph format (files are zipped)	10.36043/1945_148
browser_tracks.tar.gz	CHEAR	Epigenetics data from infant cord blood: Genome browser tracks in BigWig format (files are zipped)	10.36043/1945_149
1945_EPIGEN _MAP.csv	CHEAR	Mapping file to merge epi dataset and epigenetics data.	
1945_EPIGEN_Summary_Report.docx	CHEAR	Laboratory quality control and file summary for epigenetic analysis.	
1945_EPIGEN_METHODS.html	CHEAR	Laboratory methods and quality control for epigenetic analysis.	
SDD-2017-1945_v4.xlsx	CHEAR	A Semantic Data Dictionary (SDD) aims to describe datasets through the use of tables that identify information about data variables' content, description, and format. An SDD uses best practice ontologies to annotate datasets to ensure the document is machine	10.36043/1945_634_2022.2

		readable and conforms to a standard model for scientific studies, in order to maximize opportunities for data harmonization and knowledge sharing [1,2].	

*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.

1. https://tw.rpi.edu/project/SDD

2. The Semantic Data Dictionary - An Approach for Describing and Annotating Data. Rashid S, McCusker J, Pinheiro P, Bax M, Santos H, Das A, Stingone J, McGuinness D. The Semantic Data Dictionary - An Approach for Describing and Annotating Data.Data Intell. Fall 2020;2(4):443-486.

4. ANALYTIC NOTES

- Please contact the Principal Investigator of the CHEAR study for questions related to datasets indicated with source "Parent study" in the table above and contact hhearsupport@mssm.edu for all other questions.
    - The Principal Investigator's name and contact information at the time of this publication is: Dr. Breton at breton@usc.edu.
- When datasets contain a column for SID, merge on SID; otherwise merge on CHILD_PID.
    - Targeted datasets (trace elements) are collected from mothers during pregnancy.
    - Epigenetics data are collected from infant cord blood at birth.
- Specific gravity, a urinary dilution factor, was measured and provided by the PI. However, these specific gravity measurements for each sample are shared with the CHEAR-measured trace elements in the TARGETED dataset.
- Visit our Resources page to download and view helpful tools. (https://hheardatacenter.mssm.edu/resources.asp)
- The HHEAR dataset 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.
    - Request an account at https://hheardatacenter.mssm.edu/.