CHEAR Project # 2017-1729

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

Contents:

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	HHEAR Center for Data Science	U2CES026555
Emory University	Miller, Gary W.	National Exposure Assessment Laboratory at Emory	U2CES026560

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 [https://hheardatacenter.mssm.edu/Search/Study]
    ii. The DOIs (Digital Object Identifier) associated with the data files used in your manuscript. DOIs are located *https://hheardatacenter.mssm.edu/PublicFile/ViewPublicFile?projectId=69* as well as section 3 below.
    iii. We also suggest including the following statement in your manuscript:

Data for this analysis were obtained from the publicly available 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
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 (relevant specimen code below):

BC = Buffy Coat	**BR** = Breast Milk	**BU** = Buccal Cell
CB = Cord Blood	**CSF** = Cerebrospinal Fluid	**DNA** = 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	Analyte Code	Full Analyte Name	CAS #
UPAH	Polycyclic aromatic hydrocarbons	FLUO2_FLUO3	2-hydroxyfluorene + 3-hydroxyfluorene	2443-58-5_6344-67-8
UPAH	Polycyclic aromatic hydrocarbons	NAP1	1-hydroxynaphthalene	90-15-3
UPAH	Polycyclic aromatic hydrocarbons	NAP2	2-hydroxynaphthalene	135-19-3
UPAH	Polycyclic aromatic hydrocarbons	PHEN1	1-hydroxyphenanthrene	2433-56-9
UPAH	Polycyclic aromatic hydrocarbons	PHEN2	2-hydroxyphenanthrene	605-55-0
UPAH	Polycyclic aromatic hydrocarbons	PHEN3	3-hydroxyphenanthrene	605-87-8
UPAH	Polycyclic aromatic hydrocarbons	PHEN4	4-hydroxyphenanthrene	7651-86-7
UPAH	Polycyclic aromatic hydrocarbons	PYR1	1-hydroxypyrene	5315-79-7
INFLAM	Cytokines and Other Inflammatory Markers	IL10	Interleukin 10	130068-27-8
INFLAM	Cytokines and Other Inflammatory Markers	IL1B	Interleukin1B	
INFLAM	Cytokines and Other Inflammatory Markers	IL6	Interleukin6	308067-66-5
INFLAM	Cytokines and Other Inflammatory Markers	TNFA	Tumor necrosis factor A	308079-78-9
OXID	Oxidative Stress Markers	OS8OHDG	8-OH-deoxyguanosine	88847-89-6
OXID	Oxidative Stress Markers	OSMDA	Malonodialdehyde	542-78-9

Targeted Comment Codes

Code	*Definition*
0	Valid measurement
15	Interfering substances present (no concentration reported)
37	Value less than LOD

3. DESCRIPTION OF PUBLIC FILES (example below – change to your specific files)

File Name	*Source*	*File Description*	*DOI**
1729 _EPI_DATA.csv	Parent study	Epidemiological data for pregnant mothers in this study including demographic and birth related outcome data.	10.36043/1729_705
1729_EPI_DDCB.xlsx	Parent study	Data dictionary and codebook corresponding to the 1729_EPI_DATA.csv epidemiological data set.	NA
1729_TARGETED_DATA.csv	CHEAR	Urinary PAH, oxidative stress and inflammatory marker concentrations from pregnant mothers.	10.36043/1729_707
1729_META_MAP.csv	CHEAR	Metabolomics data with appended SIDs are hosted on Metabolomics Workbench; the SID-PID map can be used to link these data to other data files for this project	NA
1729_META_LINK	CHEAR	Hyperlink to Metabolomics Workbench for annotated and full feature metabolomic data	10.21228/M8T112
SDD-2017-1729.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 readable and conforms to a standard model for scientific studies, in order to maximize opportunities for data harmonization and knowledge sharing [1,2].	TBD

*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.

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:

      Beate Ritz, MD, PhD
      University of California Los Angeles
      michelle.ritz@mountsinai.org

- *Data sets may be merged on PID (CHEAR particpant ID)*
- *Biospecimens (urine and/or serum) obtained from mother during pregnancy at 1st, 2nd or 3rd trimesters or at day of delivery.*
- 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/.

Field Code Changed