

# Appendix I: Data Management Plan

## NCEH/ATSDR Data Management Plan Form

**Table 1 – Core DMP Elements** (should be filled out when project approval is sought)

<b>MRID</b> (NCEH/ATSDR metadata repository identifier - for NCEH/ATSDR OD use only.) P1805-0008
<b>*Title</b> Biological and Environmental Sampling of Per- and Polyfluoroalkyl Substances (PFAS)
<b>*Description</b> A statistically based, community sampling design will be used to determine <ul style="list-style-type: none"><li>- PFAS serum concentrations from participants living in communities with exposures to PFAS in drinking water.</li><li>- PFAS urine concentrations from a subset of participants living in communities with exposures to PFAS in drinking water.<ul style="list-style-type: none"><li>o An initial 10% of urine samples will be analyzed for 18 PFAS and creatinine. Additional analytes may be added should methods become available. If the geometric mean PFAS concentrations in this initial subset are elevated compared to the U.S. national reference population, as defined by the 2013-2014 NHANES 95th percentile, all other urine samples from the site will be analyzed.</li></ul></li><li>- PFAS concentrations in indoor dust and tap water samples from a subset (10%) of homes of participants in biological sampling.</li></ul> <p>A questionnaire will be administered to all participants to gather information that can be used to characterize each individual's exposure.</p> <p>The collected data will be analyzed to determine PFAS blood levels for individual participants as well as the community as a whole. Serum and urine concentrations will be compared to reference ranges from nationally representative data, and correlations between environmental and biological PFAS concentrations will be evaluated.</p>
<b>*Last DMP Update</b> May 2, 2018
<b>*Contact Name and Email</b> CDC PI or POC Name (last, first): Worley, Rachel CDC PI or POC e-mail address: idz7@cdc.gov CDC PI or POC phone number: 770-488-1549
<b>Organization</b> ATSDR/DCHI/SSB

## Appendix I: Data Management Plan

**\*Unique Identifier and catalog/database name**

TBD

## Appendix I: Data Management Plan

### **\*Data Access Level(s) – CHECK ALL APPLY**

*(The degree to which the data collected as part of this project could be made publicly available, regardless of whether it has been made available. Projects can have multiple datasets or different data elements within a single dataset that are approved for different levels of public access.)*

#### **PUBLIC Release**

- Public release – Full dataset

*(Dataset can be made available without restrictions; data steward no longer controls data. **This should be the default selection for all datasets unless justified otherwise.**)*

- Public release – Aggregate data

*(Underlying dataset cannot be released or shared, but aggregate/summary data can be made available to public access without restriction)*

Justification (required if selected): In order to protect personally identifiable information, only aggregate summary data will be made available to public access without restriction.

- Public release - Release by ad-hoc request

*(Metadata will be released and the dataset is available by ad-hoc request; data requests CANNOT be denied; no data use agreement or restrictions; data steward no longer controls data.)*

Justification (required if selected):

#### **RESTRICTED Release**

- Restricted use data sharing

*(Dataset is available to particular parties under certain use restrictions or use agreement; data not always under CDC custody. The use restriction/agreement (or template) needs to be attached.)*

Justification (required if selected): Limited data set (including no direct PII) may be shared with approved data requestors for approved scientific purposes via data use agreement.

- Restricted access data sharing

*(Dataset is only available in an RDC; data need to remain under CDC custody.)*

Justification (required if selected):

#### **No Data Release/Sharing**

- No release or data sharing

Justification (required if selected):

## Appendix I: Data Management Plan

### Access Rights/Restrictions

*(Include information regarding access or restrictions based on privacy, security, or other policies of the owner of the data. Include an explanation for the selected "Public Access Level" above.)*

In compliance with federal and state privacy protection laws and regulations, the limited data set may be shared with other federal, state and/or local public health and environmental agencies via data use agreements for research purposes to advance the scientific understanding of human exposures to PFAS.

### License/Other Agreements

*(The license or non-license [i.e., public domain] status with which the dataset will be published. See [Open Licenses](#) for more information. May include DTA, MTA, IAA, MOU or other agreements concerning data use and access.)*

### \*Publisher/Owner

*(The publishing entity and optionally their parent organization(s). This could be the "owner" of the data.)*

CDC/ATSDR

### Access URL(s), If Known

*(URL providing indirect access to the DMP, dataset, data dictionary [variable names and valid values], data collection instrument and other relevant information, including the research protocol if possible.)*

TBD

### Download URL(s), If Known

*(URL providing direct access to a downloadable file of the dataset, summary data, or data tables.)*

TBD

### \*Spatial

*(The range of spatial applicability of a dataset. Could include a geographic region or a named place [city, county, state, region, country].)*

Data will be collected from no less than eight communities across the United States and U.S. Territories.

## Appendix I: Data Management Plan

### **\*Temporal**

*(The range of temporal applicability of project)*

Start date of data collection (month/year): October 2018

End date of data collection (month/year): March 2020

## Appendix I: Data Management Plan

**Table 2 – Additional DMP Elements** (should be filled out where possible when project approval is sought; however, many fields can only be filled out later when publication/report is cleared)

<p><b>*Tags/Keywords</b>  <i>(Keywords to help users discover the dataset.)</i>            PFAS, PFOA, PFOS, PFNA, PFHxS</p>										
<p><b>*Intramural or Extramural Project</b>  <input type="checkbox"/> Intramural  <input checked="" type="checkbox"/> Extramural (grant, cooperative agreement, contract, IAA, CDC Foundation, other)            Specify mechanism: Project will be conducted using the ATSDR/DCHI Mission Support contract mechanism.</p>										
<p><b>Project Type – CHECK ALL APPLY</b>  <i>(Multiple selections may apply.)</i></p> <table> <tr> <td><input type="checkbox"/> Research</td> <td><input type="checkbox"/> Emergency</td> </tr> <tr> <td><input checked="" type="checkbox"/> Non-research</td> <td><input checked="" type="checkbox"/> Exposure investigation</td> </tr> <tr> <td><input type="checkbox"/> Surveillance</td> <td><input type="checkbox"/> Ongoing collection</td> </tr> <tr> <td><input type="checkbox"/> Evaluation</td> <td><input type="checkbox"/> Other</td> </tr> </table>	<input type="checkbox"/> Research	<input type="checkbox"/> Emergency	<input checked="" type="checkbox"/> Non-research	<input checked="" type="checkbox"/> Exposure investigation	<input type="checkbox"/> Surveillance	<input type="checkbox"/> Ongoing collection	<input type="checkbox"/> Evaluation	<input type="checkbox"/> Other		
<input type="checkbox"/> Research	<input type="checkbox"/> Emergency									
<input checked="" type="checkbox"/> Non-research	<input checked="" type="checkbox"/> Exposure investigation									
<input type="checkbox"/> Surveillance	<input type="checkbox"/> Ongoing collection									
<input type="checkbox"/> Evaluation	<input type="checkbox"/> Other									
<p><b>Dates</b>            Estimated date of data release/sharing (month/year): December 2020            Preservation expiration date (year that the dataset will be available until): TBD</p>										
<p><b>Data Category</b>  <i>(For explanation of D1 to D10 codes, see Table on page 1)</i></p> <table> <tr> <td><input type="checkbox"/> D1</td> <td><input checked="" type="checkbox"/> D2</td> <td><input type="checkbox"/> D3</td> <td><input type="checkbox"/> D4</td> <td><input type="checkbox"/> D5</td> </tr> <tr> <td><input type="checkbox"/> D6</td> <td><input type="checkbox"/> D7</td> <td><input type="checkbox"/> D8</td> <td><input type="checkbox"/> D9</td> <td><input type="checkbox"/> D10</td> </tr> </table> <p><u>Justification:</u> (provide detailed information about the data category selected above. If <i>D6</i> is selected, provide quantitative estimates of costs in releasing data and expected volume of use. If <i>D7</i> is selected, specify the reason that prevents the owner from releasing/sharing the data.)</p> <p>Data will be collected by a contractor or cooperative agreement partner funded by CDC/ATSDR.</p>	<input type="checkbox"/> D1	<input checked="" type="checkbox"/> D2	<input type="checkbox"/> D3	<input type="checkbox"/> D4	<input type="checkbox"/> D5	<input type="checkbox"/> D6	<input type="checkbox"/> D7	<input type="checkbox"/> D8	<input type="checkbox"/> D9	<input type="checkbox"/> D10
<input type="checkbox"/> D1	<input checked="" type="checkbox"/> D2	<input type="checkbox"/> D3	<input type="checkbox"/> D4	<input type="checkbox"/> D5						
<input type="checkbox"/> D6	<input type="checkbox"/> D7	<input type="checkbox"/> D8	<input type="checkbox"/> D9	<input type="checkbox"/> D10						
<p><b>Population Represented</b>  <i>(e.g., “residents of x,” “inpatients at x,” “users of product x”)</i>            TBD – Residents of no fewer than eight communities with PFAS contaminated drinking water. Exact locations to be determined.</p>										

## Appendix I: Data Management Plan

### **Data Collection Protocol**

*(Brief description with reference to document or website that provides detailed information.)*

Participants in each exposure assessment will complete consent/assent/parental permission forms, provide blood and urine samples and respond to a questionnaire at a centralized location in the community. Participants will be responsible for collecting first morning urine samples in their homes and transporting them to the centralized blood collection location. A subset of participant households will be randomly selected for environmental (i.e., indoor dust and tap water) sampling. Administration of consent forms for environmental sampling and indoor dust and tap water sample collection will take place during a home visit.

All blood samples will be shipped to the NCEH laboratory for analysis of PFAS concentrations and then shipped to a bio-specimen repository for storage. A subset of urine samples will be shipped to the NCEH laboratory for analysis of PFAS concentrations. The remaining urine samples will be shipped to a bio-specimen repository for storage. Drinking water and indoor dust samples will be shipped to an EPA-accredited laboratory for analysis of PFAS concentrations and then stored.

Detailed data collection procedures are available in the protocol for the assessments.

### **Data Management Protocol**

*(Brief description with reference to physical location(s) or system(s) where data will be housed (e.g., CDC shared network drive, data host system name, SQL database, etc.) and to data formats. Include the locations of dataset both before data release and after data release.)*

All data will be transmitted to ATSDR or ATSDR contractor for incorporation into a centralized data management system. All results will be electronically transmitted in spreadsheet format using a secured and password-protected network.

All documents with personal identifying information (i.e., consent forms, assent forms, collection logs, etc.) will be kept in locked cabinets and all electronic data will be stored on a password-protected network servers behind firewalls, accessible only to those staff working directly with raw data.

De-identified samples will be sent to the laboratories—no individual identifiers will be included. Any reports produced from this information will not identify specific individuals.

### **Process for Omitting Identifying Information**

*(Description of what identifiers are in the database, how they will be removed, and by whom.)*

All samples and data will be identified with a unique code and participant birth date only. Only the project coordinator will have access to the link between the sample identification code and personal identifying information in order to facilitate communication of individual results with participants.

## Appendix I: Data Management Plan

### **Data Quality Protocol (to address issues of privacy protection and statistical stability)**

*(Brief description with reference to document or website that provides detailed information. Describe methods for data validation and error resolution, removal or shielding of any proprietary information, removal or shielding of sensitive information [i.e., data with dual use applicability], removal or shielding of any individually identifying information including indirect identification.)*

All laboratory analyses will be conducted with established procedures for quality assurance and control according to NCEH and EPA protocols. These methods are identified in the protocol for the exposure assessments and are available from NCEH and EPA.

### **Data Retention/Disposal Plan**

*(State when and how the dataset will be archived or destroyed [in accordance with CDC/ATSDR Records Control Schedule: <http://isp-v-maso-apps/RecSched/images/RCS.pdf>].)*

Records will be retained and disposed of in accordance with the CDC Records Control Schedule. Physical copies of assessment materials and reports will be maintained at ATSDR until no longer needed by program officials and will be kept no longer than five years following completion of the exposure assessment in accordance with retention schedules. Computer documents will be disposed of when no longer needed by program officials and will be kept no longer than five years following the study. Personal identifiers will be deleted from records when no longer needed. Disposal methods will include erasing computer files, shredding paper materials, or transferring records to the Federal Records Center when no longer needed for evaluation and analysis. Records are retained for 20 years.

### **Data Analysis Plan**

*(Brief description of planned use of the data. Can include reference to document [e.g., information collection request, research protocol, or other] that provides more detailed information.)*

Data will be analyzed for each individual participant and in aggregate. Detailed information is provided in the protocol.

### **Publication Plan**

*(Brief description of planned CDC-authored and CDC-coauthored publications, including topic, type of publication, and estimated timeline.)*

Aggregate findings from the exposure assessments are expected to be published as ATSDR reports and in the peer reviewed literature. The topic of publications will be human exposure to per- and polyfluoroalkyl substances. Reports and publications are estimated to be released in 2020.

### **Data Release Documentation**

*(List documents provided to users, e.g., variable definitions, codebook, metadata file, guidance on data use.)*

TBD

### **Data Release Format**

*(Recommend to use non-proprietary format when possible, such as CSV, JSON, etc. Also specify data dictionary file format.)*

CSV

### **Data Release Notification**

*(State how potential users will be informed of dataset availability.)*

TBD



## Appendix I: Data Management Plan

Date Form Completed: April 14, 2018

By: Rachel Worley, PhD, Environmental Health Scientist

*Name, Title*

Date Form Last Revised: May 2, 2018

By: Rachel Worley, PhD, Environmental Health Scientist

*Name, Title*