WIHS Public Data Set (PDS) FAQs

Where can I find more information about the WIHS?

- [WIHS public website](#) offers an overview of the WIHS.
- [Data collection forms](#) (also available in the public data set) are provided by visit.
- [Manual of Operations (MOO)](#) includes information on the baseline interview and follow-up interviews.

How do I read in the data?

To read in an ASCII file into SAS:

data f22hx;
    INFILE "d:\data\f22hx.dat" LRECL= 5500 TRUNCOVER;
    %INCLUDE "d:\Sasinput\f22hx.inp";
run;

The LRECL statement ensures that SAS reads beyond the 256 character in the file. An LRECL of 5500 should be wide enough to read any of the WIHS tables.

The TRUNCOVER statement tells SAS that if it comes to the end of the line, to stop reading in data from that line of data (and leaves any "missed" variables as missing).

This is an old paper, but it does talk about TRUNCOVER and MISSOVER, another option used with reading ASCII files: http://www2.sas.com/proceedings/sugi26/p009-26.pdf

The %INCLUDE statement reads in the SAS INPUT statement that has been saved in f22hx.inp (e.g., input pubid visit, etc.).

How do the forms combine together (there is no mapping of the tables)?

Most forms may be combined using ID/VISIT. There are some forms that may have >1 record per visit, such as the medication forms (DRUG1, DRUG2, DRUG3). Within some forms, there are repeating segments (>=1 record per visit) which are provided as separate tables. Examples of repeating segments include pregnancy outcomes (F23s1) and ARV usage (F22s3, F22MEDs3). Information on the interview is available in the Investigators section of the [WIHS public website](#).

We recommend you review the [Question by Question Specifications](#) (QxQs) available on the WIHS public website for any outcomes or areas of interest included in your analysis. They are available under each visit link and will help you to understand the additional nuances of our questions.

We've included an excel sheet that shows a list of all variables that have ever been collected at a WIHS visit and how many responses we have for each of these variables.

New for Version 16:

\csv directory (beta version)
 -- CSV data files are new as of WIHS P16. We recommend using SAS to read the *.dat files with the SAS INPUT statements. The CSV files are provided as an alternative method but may not always read properly into a program. Please notify DACC of any issues (mwccs@jhu.edu)
 -- Data files in CSV format. Variable names are included as the first line of each file
 -- Please do not use MS Excel to view the CSV data as it may cause truncated data.