

This document serves as an information sheet for the final data files from CTN-0047. The following is a list of all files contained in this data transfer:

Documentation:

0047_A_AnnotatedCRFs.pdf
0047_B_AnnotatedCRFs.pdf
0047_Z_AnnotatedCRFs.pdf
CTN0047-dictionary.xls
CTN0047-nulled-values.csv
Readme-Revised.doc

Data Files:

Listed in CTN0047-dictionary.xls - Tables tab

De-Identification

- For a detailed list of nulled columns please see file CTN0047-nulled-values.csv
- All dates have been converted to days post randomization
- All columns with names, initials and phone numbers have been nulled
- Columns for comments and specification are nulled as free text may contain identifying information

Identifiers

- The unique identifier for each participant is contained in the variable called PATID.
- The visit information is in variable VISNO which takes the following values.

VISNO	VISIT NAME
00	Baseline
01	Booster 1 (Day 3)
02	Booster 2 (Day 7)
03	1 Month
04	3 Month
05	6 Month
06	12 Month

- Each dataset (except HSL- hair sample lab results) contains PROT (Protocol Number), PATID and RANDDT(Randomization Date) which are not listed in the Annotated CRF.

Protocol Segments

- Segment A: Contains data from screening and baseline assessments (as documented in 0047_A_AnnotatedCRFs.pdf).
- Segment B: Completion of enrollment form (ENR-B) triggers randomization, thus this protocol segment contains all post-randomization assessments (as documented in 0047_B_AnnotatedCRFs.pdf). Note that only the data in this segment that are collected pre-randomization are on the ENR-B form.

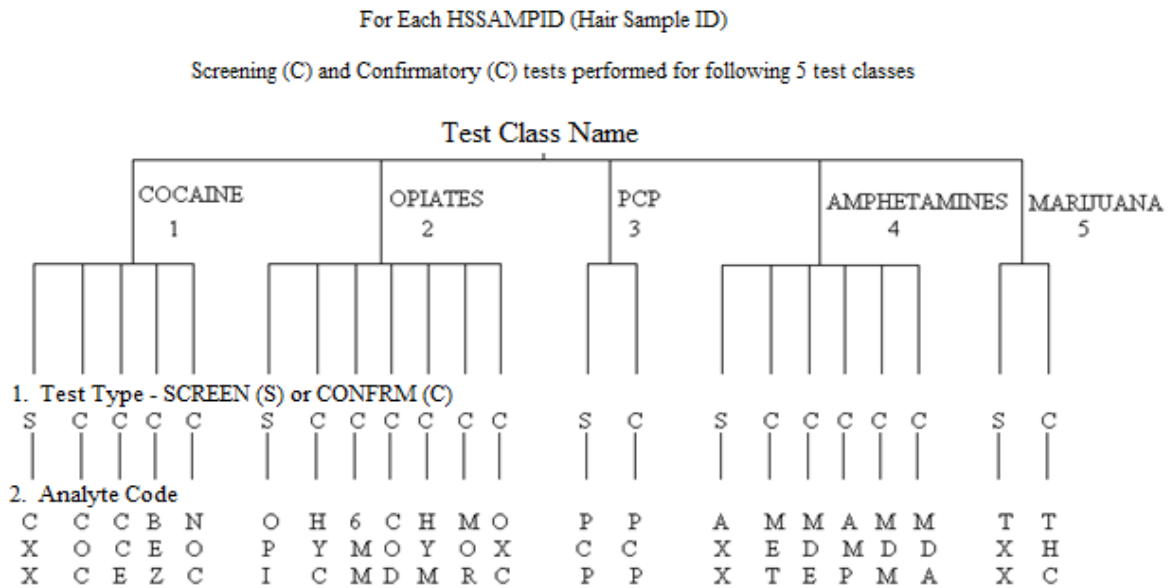
- Segment Z: Contains information on people who were consented, but not randomized (as documented in 0047_Z_AnnotatedCRFs.pdf).

Randomization Information

- The randomized treatment arm information is in ENROLL dataset. The variables TRTRUE and TRTSHOWN contain the treatment arm information.
- TREATDT is the randomization/consent date in the ENROLL dataset.

Hair Sample LAB results (hsl.sas7bdat)

- This data set contains hair sample results from the lab.
- For each hair sample ID there are 22 records.
- First the screening test is performed for each of the five test classes (Cocaine, Opiates, PCP, Amphetamines and Marijuana).
- Then the confirmatory tests are performed within each of these test classes. The following schema shows the test classes, type of test and analyte code for each drug class. The analyte name and the data dictionary for hsl dataset are in 0047_HSL_DataDictionary.doc.
- The hair sample results are in variable testresult. {TNP – Test Not Performed, DET – Detected (Positive), NOT – Not Detected (Negative)}



Summary

- The summary of the database names listed by protocol segment, data collected by, and field names is in Summary.xls