Once you download xpt file (named “transport.xpt), use the sas program (appended at the end of this document) to generate sas data sets.

Demographic information is recorded in a SAS data set (or an equivalent CSV data set) named DEM. The variables you are interested in are DEM01DT(age=|DEM01DT|/365.25),DEM02(gender), DEM03(ethnicity), DEM04 and DEM04A-DEM04G (race categories). The corresponding meaning of the values for each variable can be found in an annotated pdf file named “CTN31A\_CDD\_annotated” you have already had.

If you also want education information of the participants, you may need to download CTN0031 (parent study). Education information (years of education) is recorded in a data set named “ASI2” . The variable name is ASL0E1A (years of education). However, I find there is a problem to link the parent and child studies. I will report to the company (Emmes) that generated these deidentified data sets to see whether there is a link key to provide to you or publish on the NIDA CTN data share.

/\*\*\*\*

Programmer: Lian Hu

Description:

1. use transport.xpt in CTN0031A to generate sas datasets

2. link education variables in parent study (CTN031, ASI2) to child study (CTN031A DEM)\_\_\_failed. Will report.

\*\*\*\*/

/\*\*\*CTN31A downloaded from CTN data share on 12/8/2016\*\*\*/

option nofmterr;

libname CTN031A "E:\CTN\CCTN DATASHARE DOWNLOAD\CTN0031A\Sasdata"; \*\*\*change path to actual path\*\*;

filename importin "E:\CTN\CCTN DATASHARE DOWNLOAD\CTN0031A\Xptdata\transport.xpt"; \*\*\*change path to transport.xpt to actual path\*\*;

**proc** **cimport** infile=importin library=CTN031A memtype=data;

**run**;

\*\*\*\*a total of 28 sas data sets. Use DEM to locate demographics including age, gender, ethnicity and race\*\*\*\*;

**data** dem;

set CTN031A.dem;

keep patdeid DEM01DT DEM02 DEM03 DEM04 DEM04A DEM04B

DEM04C DEM04D DEM04E DEM04F DEM04G;

**run**; \*\*\*obs=244\*\*\*;

libname CTN031 "E:\CTN\CCTN DATASHARE DOWNLOAD\CTN0031\SASdata\_newdownload";\*\*\*change to actual path\*\*;

filename importin "E:\CTN\CCTN DATASHARE DOWNLOAD\CTN0031\SASdata\_newdownload\sas-transport-crf-files\_nida-ctn-0031\transport.xpt"; \*\*change to actual path\*\*\*;

**proc** **cimport** infile=importin library=CTN031 memtype=data;

**run**;

**data** edu\_ctn31;

set CTN031.asi2;

keep ASL0E1A patdeid;

**run**; \*\*\*ASL0E1A is “years of education” in parent study\*\*\*;

\*\*\*\*Currently, there is no appropriate link on patdeid between CTN0031 and CTN0031A. We will report this issue for you at a later time\*\*\*\*

HI Gwendolyn,

For now, I give you a description on how to identify demographics information in ctn31A data sets. As you stated that you do not have sas program on your computer, I am copying the code along with the descriptions in a word document. You can discuss with your statistician if needed.

You also said you want to have education level of the participants. I find that this may be only recorded in the parent study. I have a description in the attached file of which data set and variable name can be used to retrieve this information, including the sas code.  However, I was unable to link the two studies by either patdeid (patient unique identifier) or other demographics. I will report this issue to the company (Emmes) that created the data sets to see whether they can provide the link key file.

Hope this helps

Lian