

---

**CTN-0031 CRF Level Data De-Identification Notes**

**And**

**Database Descriptions**

**February 3rd, 2012**

---

---

## CTN-0031 CRF Level Data De-Identification Notes

---

The Case Report Form (CRF)-level de-identified data sets were prepared by reviewing the CRFs and the data dictionary to identify a list of variables that contained, or could potentially contain, identifying information, such as dates and free text fields. The variables were programmatically removed or modified in the raw dataset, resulting in de-identified datasets. All dates were converted to “days on study,” where the date of randomization (BASEDATE) = 0. A negative value for days on study refers to an event that occurred prior to randomization (e.g., date of birth) and a positive value for days on study refers to an event that occurred after randomization (e.g., 12-month visit date). The following free text fields were emptied:

- Any responses derived from “Other. Please specify” or “Specify other reason”. For example, other races, marital status, therapy sessions, SAE categories, reasons study drug not given, abnormalities on ECG and Urine Analysis forms.
- Comment fields
- AE/SAE terms were retained in the database, but any detailed descriptions in a separate field providing additional information the AE or SAE were erased. This includes relevant labs, medical history and medications, drug/alcohol use description, probable cause of death.
- On dosing logs: lot numbers, who administered dose.
- All medical history information.
- Concomitant medication names, purposes, and start year were *retained* in the database, but start dates and months were dropped.
- All protocol violations
- Type of work; job description/parents' type of work; school program name and major; descriptions of living situation, illnesses, concerns, drug/alcohol use, feelings/interests.

Following is a list of all variables in the dataset that were modified or emptied in order to de-identify the data, based on the following commands:

PATIDDEID assigns to each patient a unique value in a variable called PATIDDEID.

BASEDATE names the variable that is used as the base date to calculate days on study. In this study, BASEDATE is date of randomization.

DOS names a variable that is to be converted from a date into days on study.

EMPTY defines a variable to fill with missing values.

DROFFILE defines a data set to drop.

The following documentation presents information on the variable based on the name within the database. To get the corresponding label of the variable, one must look within the data dictionary in Excel that is also provided on the CTN Data Share web site.

**For all the datasets dates were changed to days of study and site was emptied.**

**ran (Randomization):** Removed ran02sp (specify)

**scr (System Screening):** Removed scrpid (Participant ID)

**dtcq8 (Drug taking confidence questionnaire):** Removed dtcq\_pdo and dtcq\_sdo

**infae:** Removed aeviswk

**infsae:** Removed sae03, sae05, sae05b, sae05c, sae06a, sae06b, sae06c, sae06d and sae06e

---

## CTN-0031 CRF Level Data De-Identification Notes

---

**shaq (Stage-12 Monthly Self–Help Activities Questionnaire):** Removed shaq\_1ot (Specify other), shaq\_wh (why), shaq\_32\_sp1 (Specify other), shaq\_32\_sp2 (Specify other), shaq\_32\_sp3 (Specify other), shaq\_32\_sp4 (Specify other), shaq\_32\_sp5 (Specify other) and shaq\_33 (why)

**shaq\_1 (Stage-12 Monthly Self–Help Activities Questionnaire):** Removed shaq\_1ot (Specify other), shaq\_wh (why), shaq\_32\_sp1 (Specify other), shaq\_32\_sp2 (Specify other), shaq\_32\_sp3 (Specify other), shaq\_32\_sp4 (Specify other), shaq\_32\_sp5 (Specify other) and shaq\_33 (why)

**tsee (12-Step Experience and Expectations):** Removed tsee\_ots (Specify other)

**enr (System Enrollment):** Removed pinum (participant number)

**pi (Patient Information):** Removed pinum (participant number)

**dem (Demographics):** Removed dem04\_sp (Specify other)

**dsm1 (DSM-IV Criteria- Substance Related Disorders):** Removed dsm02hsp (Specify other)

**asi1 (ASI-Lite):** Removed asl0g18s (Specify other), asl0g19s (Specify other), asl0g20 (Specify other), asl0gcom (comments), asl0m3sp (Specify other), asl0m4sp (Specify other), asl0m5sp (Specify other) and asl0mcom (comments)

**asi3 (ASI-Lite cont.):** Removed asl0ecom (comments)

**asi7 (ASI-Lite:- Drug/Alcohol use):** Removed asl0dcom (comments)

**asi8 (ASI-Lite:- Legal Status) :** Removed asl12com (comments) and asl0l16s (Specify other)

**asi9 (ASI-Lite:- Legal Status (cont.)):** Removed asl0lcom (comments)

**asi12 (ASI-Lite:- Family/ Social Relationships (cont.)):** Removed asl0fcom (comments)

**asi13 (ASI Lite:- Psychiatric Status):** Removed asl0pcom (comments)

**mdhx (Medical AND Psychiatric History):** Removed mdhx1sp (Specify other), mdhx2sp (Specify other), mdhx3sp (Specify other), mdhx4sp (Specify other), mdhx5sp (Specify other), mdhx6sp (Specify other), mdhx7sp (Specify other), mdhx8sp (Specify other), mdhx9sp (Specify other), mdhx10sp (Specify other), mdhx11sp (Specify other), mdhx12sp (Specify other), MDH013SP (Specify other), mdhx14sp (Specify other), mdhx15sp (Specify other), mdhx16sp (Specify other), mdhx17sp (Specify other), mdhx18sp (Specify other), mdhx19sp (Specify other) and mdhx20sp (Specify other)

**sue (SUBSTANCE USE EVENTS):** Removed sue03gsp (Specify other), sue04asp (Specify other) and sue05 (Comments)

**asf1 (ASI Lite Follow-Up):** Removed asf0g19s (Specify other), asf0gcom (comments), asf0m4sp (Specify other), asf0m5sp (Specify other) and asf0mcom (comments)

**asf2 (ASI-Lite Follow-Up: Employment/Support Status):** Removed asf0e7sp (Comments)

**asf3 (ASI-Lite Follow-Up: Employment/Support Status (cont.)):** Removed asf0ecom (Comments)

**asf7 (ASI-Lite Follow-Up: Alcohol/Drugs):** Removed asf0dcom (Comments)

**asf8 (ASI-Lite Follow-Up: Legal Status):** Removed asf12com (Comments)

**asf9 (ASI-Lite Follow-Up: Legal Status (cont.)):** Removed asf01com (Comments)

**asf11 (ASI-Lite Follow-up: Family/Social Relationships (cont.)):** Removed asf0f23s (Specify other)

---

## CTN-0031 CRF Level Data De-Identification Notes

---

**asf12 (ASI-Lite Follow-up: General Information):** Removed asf0fcom (Comments)

**asf13 (ASI-Lite Follow-up: Psychiatric Status):** Removed asf0pcom (Comments)

**sae (Serious Adverse Event):** Removed sae05(Other Specify) , sae06a (Comments), sae06b (Comments) and sae06c (Comments)

**term (Study Termination):** Removed term02gs (Other reason specify)

---

## CTN-0031 Database Descriptions

---

Databases for this protocol are available for download in both SAS and ASCII/comma separated value (CSV) format. When you retrieve the datasets from the data download page on the CTN Data Share website, a ZIP file will be provided containing the datasets.

Note that when you download the SAS datasets, the files will be SAS transport files. The transport.xpt contains all the SAS datasets available. The following code can be executed in SAS to retrieve the individual SAS databases.

```
libname newlib "[Name of Path where you want to store the SAS datasets. E.g.: c:\temp\CTN0031] ";
filename tranfile "[Name of path where you store transport.xpt. E.g.: C:\temp\ transport.xpt";
proc cimport library=newlib infile=tranfile; run;
```

Also in the SAS ZIP file are SAS transport files corresponding to each individual database. These files are provided to allow import into SPSS and differ slightly from the SAS datasets in the Transport.xpt and from the data dictionaries provided. We describe below specific modifications to the SAS datasets to accommodate SPSS import.

1. Renaming of SAS dataset to conform to 8 character limits:

- a. t\_frasf10 was renamed to t\_fras15
- b. t\_frasf11 was renamed to t\_fras16
- c. t\_frasf12 was renamed to t\_fras17
- d. t\_frasf13 was renamed to t\_fras18
- e. t\_frasl10 was renamed to t\_fras28
- f. t\_frasl11 was renamed to t\_fras29
- g. t\_frasl12 was renamed to t\_fras30
- h. t\_frasl13 was renamed to t\_fras31
- i. t\_frstd1 was renamed to t\_frstd71
- j. t\_frfrmcas was renamed to t\_frfr48

2. Renaming of variables to conform to naming of variables to be no longer than 8 characters.

Old Variable Name	New Variable Name	Dataset
PATIENTID	PATI0764	SAS.AB
PATIENTNUMBER	PATI0765	SAS.AB
ENTRY_DATETIME	ENTR0659	SAS.AE
FORMINDEX	FORM0662	SAS.AE
ITEMSETINDEX	ITEM0714	SAS.AE
MERGE_DATETIME	MERG0758	SAS.AE
PATIENTID	PATI0764	SAS.AE
PATIENTNUMBER	PATI0765	SAS.AE
SUBJECT_ID	SUBJ0957	SAS.AE
VISITINDEX	VISI1164	SAS.AE
PATIENTID	PATI0764	SAS.ASF1
PATIENTNUMBER	PATI0765	SAS.ASF1
PATIENTID	PATI0764	SAS.ASF10
PATIENTNUMBER	PATI0765	SAS.ASF10
PATIENTID	PATI0764	SAS.ASF11

Old Variable Name	New Variable Name	Dataset
PATIENTNUMBER	PATI0765	SAS.ASF11
PATIENTID	PATI0764	SAS.ASF12
PATIENTNUMBER	PATI0765	SAS.ASF12
PATIENTID	PATI0764	SAS.ASF13
PATIENTNUMBER	PATI0765	SAS.ASF13
PATIENTID	PATI0764	SAS.ASF2
PATIENTNUMBER	PATI0765	SAS.ASF2
PATIENTID	PATI0764	SAS.ASF3
PATIENTNUMBER	PATI0765	SAS.ASF3
PATIENTID	PATI0764	SAS.ASF4
PATIENTNUMBER	PATI0765	SAS.ASF4
PATIENTID	PATI0764	SAS.ASF5
PATIENTNUMBER	PATI0765	SAS.ASF5
PATIENTID	PATI0764	SAS.ASF6
PATIENTNUMBER	PATI0765	SAS.ASF6

## CTN-0031 Database Descriptions

Old Variable Name	New Variable Name	Dataset
PATIENTID	PATI0764	SAS.ASF7
PATIENTNUMBER	PATI0765	SAS.ASF7
PATIENTID	PATI0764	SAS.ASF8
PATIENTNUMBER	PATI0765	SAS.ASF8
PATIENTID	PATI0764	SAS.ASF9
PATIENTNUMBER	PATI0765	SAS.ASF9
PATIENTID	PATI0764	SAS.ASI1
PATIENTNUMBER	PATI0765	SAS.ASI1
PATIENTID	PATI0764	SAS.ASI10
PATIENTNUMBER	PATI0765	SAS.ASI10
PATIENTID	PATI0764	SAS.ASI11
PATIENTNUMBER	PATI0765	SAS.ASI11
PATIENTID	PATI0764	SAS.ASI12
PATIENTNUMBER	PATI0765	SAS.ASI12
PATIENTID	PATI0764	SAS.ASI13
PATIENTNUMBER	PATI0765	SAS.ASI13
PATIENTID	PATI0764	SAS.ASI2
PATIENTNUMBER	PATI0765	SAS.ASI2
PATIENTID	PATI0764	SAS.ASI3
PATIENTNUMBER	PATI0765	SAS.ASI3
PATIENTID	PATI0764	SAS.ASI4
PATIENTNUMBER	PATI0765	SAS.ASI4
PATIENTID	PATI0764	SAS.ASI5
PATIENTNUMBER	PATI0765	SAS.ASI5
PATIENTID	PATI0764	SAS.ASI6
PATIENTNUMBER	PATI0765	SAS.ASI6
PATIENTID	PATI0764	SAS.ASI7
PATIENTNUMBER	PATI0765	SAS.ASI7
PATIENTID	PATI0764	SAS.ASI8
PATIENTNUMBER	PATI0765	SAS.ASI8
PATIENTID	PATI0764	SAS.ASI9
PATIENTNUMBER	PATI0765	SAS.ASI9
PATIENTID	PATI0764	SAS.DEM
PATIENTNUMBER	PATI0765	SAS.DEM
PATIENTID	PATI0764	SAS.DSM1
PATIENTNUMBER	PATI0765	SAS.DSM1
PATIENTID	PATI0764	SAS.DSM2
PATIENTNUMBER	PATI0765	SAS.DSM2
PATIENTID	PATI0764	SAS.DSM3
PATIENTNUMBER	PATI0765	SAS.DSM3
PATIENTID	PATI0764	SAS.DSM4
PATIENTNUMBER	PATI0765	SAS.DSM4
PATIENTID	PATI0764	SAS.DSM5

Old Variable Name	New Variable Name	Dataset
PATIENTNUMBER	PATI0765	SAS.DSM5
PATIENTNUMBER	PATI0765	SAS.DTCQ8
PATIENTID	PATI0764	SAS.ENR
PATIENTNUMBER	PATI0765	SAS.ENR
PATIENTNUMBER	PATI0765	SAS.HAQ
PATIENTID	PATI0764	SAS.IEX
PATIENTNUMBER	PATI0765	SAS.IEX
PATIENTID	PATI0764	SAS.INFAE
PATIENTNUMBER	PATI0765	SAS.INFAE
PATIENTID	PATI0764	SAS.INFSAE
PATIENTNUMBER	PATI0765	SAS.INFSAE
PATIENTID	PATI0764	SAS.MDHX
PATIENTNUMBER	PATI0765	SAS.MDHX
PATIENTID	PATI0764	SAS.PI
PATIENTNUMBER	PATI0765	SAS.PI
PATIENTNUMBER	PATI0765	SAS.PSS
PATIENTID	PATI0764	SAS.RAN
PATIENTNUMBER	PATI0765	SAS.RAN
ENTRY_DATETIME	ENTR0659	SAS.SAE
FORMINDEX	FORM0662	SAS.SAE
ITEMSETINDEX	ITEM0714	SAS.SAE
MERGE_DATETIME	MERG0758	SAS.SAE
PATIENTID	PATI0764	SAS.SAE
PATIENTNUMBER	PATI0765	SAS.SAE
SUBJECT_ID	SUBJ0957	SAS.SAE
VISITINDEX	VISI1164	SAS.SAE
PATIENTID	PATI0764	SAS.SAE2
PATIENTNUMBER	PATI0765	SAS.SAE2
PATIENTID	PATI0764	SAS.SCR
PATIENTNUMBER	PATI0765	SAS.SCR
PATIENTNUMBER	PATI0765	SAS.SEQP
PATIENTNUMBER	PATI0765	SAS.SHAQ
SHAQ_29_1	SHAQ0859	SAS.SHAQ
SHAQ_29_10	SHAQ0860	SAS.SHAQ
SHAQ_29_11	SHAQ0861	SAS.SHAQ
SHAQ_29_12	SHAQ0862	SAS.SHAQ
SHAQ_29_2	SHAQ0863	SAS.SHAQ
SHAQ_29_3	SHAQ0864	SAS.SHAQ
SHAQ_29_4	SHAQ0865	SAS.SHAQ
SHAQ_29_5	SHAQ0866	SAS.SHAQ
SHAQ_29_6	SHAQ0867	SAS.SHAQ
SHAQ_29_7	SHAQ0868	SAS.SHAQ
SHAQ_29_8	SHAQ0869	SAS.SHAQ

## CTN-0031 Database Descriptions

Old Variable Name	New Variable Name	Dataset
SHAQ_29_9	SHAQ0870	SAS.SHAQ
SHAQ_29_ST1	SHAQ0871	SAS.SHAQ
SHAQ_29_ST10	SHAQ0872	SAS.SHAQ
SHAQ_29_ST11	SHAQ0873	SAS.SHAQ
SHAQ_29_ST12	SHAQ0874	SAS.SHAQ
SHAQ_29_ST2	SHAQ0875	SAS.SHAQ
SHAQ_29_ST3	SHAQ0876	SAS.SHAQ
SHAQ_29_ST4	SHAQ0877	SAS.SHAQ
SHAQ_29_ST5	SHAQ0878	SAS.SHAQ
SHAQ_29_ST6	SHAQ0879	SAS.SHAQ
SHAQ_29_ST7	SHAQ0880	SAS.SHAQ
SHAQ_29_ST8	SHAQ0881	SAS.SHAQ
SHAQ_29_ST9	SHAQ0882	SAS.SHAQ
SHAQ_32_DY1	SHAQ0888	SAS.SHAQ
SHAQ_32_DY2	SHAQ0889	SAS.SHAQ
SHAQ_32_DY3	SHAQ0890	SAS.SHAQ
SHAQ_32_DY4	SHAQ0891	SAS.SHAQ
SHAQ_32_DY5	SHAQ0892	SAS.SHAQ
SHAQ_32_SP1	SHAQ0893	SAS.SHAQ
SHAQ_32_SP2	SHAQ0894	SAS.SHAQ
SHAQ_32_SP3	SHAQ0895	SAS.SHAQ
SHAQ_32_SP4	SHAQ0896	SAS.SHAQ
SHAQ_32_SP5	SHAQ0897	SAS.SHAQ
PATIENTNUMBER	PATI0765	SAS.SHAQ_1
SHAQ_29_1	SHAQ0859	SAS.SHAQ_1
SHAQ_29_10	SHAQ0860	SAS.SHAQ_1
SHAQ_29_11	SHAQ0861	SAS.SHAQ_1
SHAQ_29_12	SHAQ0862	SAS.SHAQ_1
SHAQ_29_2	SHAQ0863	SAS.SHAQ_1
SHAQ_29_3	SHAQ0864	SAS.SHAQ_1
SHAQ_29_4	SHAQ0865	SAS.SHAQ_1
SHAQ_29_5	SHAQ0866	SAS.SHAQ_1
SHAQ_29_6	SHAQ0867	SAS.SHAQ_1
SHAQ_29_7	SHAQ0868	SAS.SHAQ_1
SHAQ_29_8	SHAQ0869	SAS.SHAQ_1
SHAQ_29_9	SHAQ0870	SAS.SHAQ_1
SHAQ_29_ST1	SHAQ0871	SAS.SHAQ_1
SHAQ_29_ST10	SHAQ0872	SAS.SHAQ_1
SHAQ_29_ST11	SHAQ0873	SAS.SHAQ_1
SHAQ_29_ST12	SHAQ0874	SAS.SHAQ_1
SHAQ_29_ST2	SHAQ0875	SAS.SHAQ_1
SHAQ_29_ST3	SHAQ0876	SAS.SHAQ_1
SHAQ_29_ST4	SHAQ0877	SAS.SHAQ_1

Old Variable Name	New Variable Name	Dataset
SHAQ_29_ST5	SHAQ0878	SAS.SHAQ_1
SHAQ_29_ST6	SHAQ0879	SAS.SHAQ_1
SHAQ_29_ST7	SHAQ0880	SAS.SHAQ_1
SHAQ_29_ST8	SHAQ0881	SAS.SHAQ_1
SHAQ_29_ST9	SHAQ0882	SAS.SHAQ_1
SHAQ_32_DY1	SHAQ0888	SAS.SHAQ_1
SHAQ_32_DY2	SHAQ0889	SAS.SHAQ_1
SHAQ_32_DY3	SHAQ0890	SAS.SHAQ_1
SHAQ_32_DY4	SHAQ0891	SAS.SHAQ_1
SHAQ_32_DY5	SHAQ0892	SAS.SHAQ_1
SHAQ_32_SP1	SHAQ0893	SAS.SHAQ_1
SHAQ_32_SP2	SHAQ0894	SAS.SHAQ_1
SHAQ_32_SP3	SHAQ0895	SAS.SHAQ_1
SHAQ_32_SP4	SHAQ0896	SAS.SHAQ_1
SHAQ_32_SP5	SHAQ0897	SAS.SHAQ_1
PATIENTNUMBER	PATI0765	SAS.SIBSR
PATIENTNUMBER	PATI0765	SAS.SOCRATES
SOCRATES_1	SOCR0936	SAS.SOCRATES
SOCRATES_10	SOCR0937	SAS.SOCRATES
SOCRATES_11	SOCR0938	SAS.SOCRATES
SOCRATES_12	SOCR0939	SAS.SOCRATES
SOCRATES_13	SOCR0940	SAS.SOCRATES
SOCRATES_14	SOCR0941	SAS.SOCRATES
SOCRATES_15	SOCR0942	SAS.SOCRATES
SOCRATES_16	SOCR0943	SAS.SOCRATES
SOCRATES_17	SOCR0944	SAS.SOCRATES
SOCRATES_18	SOCR0945	SAS.SOCRATES
SOCRATES_19	SOCR0946	SAS.SOCRATES
SOCRATES_2	SOCR0947	SAS.SOCRATES
SOCRATES_3	SOCR0948	SAS.SOCRATES
SOCRATES_4	SOCR0949	SAS.SOCRATES
SOCRATES_5	SOCR0950	SAS.SOCRATES
SOCRATES_6	SOCR0951	SAS.SOCRATES
SOCRATES_7	SOCR0952	SAS.SOCRATES
SOCRATES_8	SOCR0953	SAS.SOCRATES
SOCRATES_9	SOCR0954	SAS.SOCRATES
SOCRATES_DT	SOCR0955	SAS.SOCRATES
PATIENTID	PATI0764	SAS.SUC
PATIENTNUMBER	PATI0765	SAS.SUC
PATIENTID	PATI0764	SAS.SUCLEAD
PATIENTNUMBER	PATI0765	SAS.SUCLEAD
PATIENTID	PATI0764	SAS.SUE
PATIENTNUMBER	PATI0765	SAS.SUE

---

## CTN-0031 Database Descriptions

---

Old Variable Name	New Variable Name	Dataset
SUE01ATT_NORM	SUE00984	SAS.SUE
SUE01ATT_UNIT	SUE00985	SAS.SUE
PATIENTNUMBER	PATI0765	SAS.SUSS
PATIENTNUMBER	PATI0765	SAS.SYRAAP
SYRAAP_10	SYRA1029	SAS.SYRAAP
SYRAAP_11	SYRA1030	SAS.SYRAAP
SYRAAP_12	SYRA1031	SAS.SYRAAP
SYRAAP_13	SYRA1032	SAS.SYRAAP
SYRAAP_14	SYRA1033	SAS.SYRAAP
SYRAAP_15	SYRA1034	SAS.SYRAAP
SYRAAP_DT	SYRA1043	SAS.SYRAAP
PATIENTID	PATI0764	SAS.TERM
PATIENTNUMBER	PATI0765	SAS.TERM
PATIENTNUMBER	PATI0765	SAS.TSEE
TSEE_CMAT	TSEE1071	SAS.TSEE
TSEE_SOST	TSEE1083	SAS.TSEE
PATIENTID	PATI0764	SAS.TSR1
PATIENTNUMBER	PATI0765	SAS.TSR1
PATIENTID	PATI0764	SAS.TSR2
PATIENTNUMBER	PATI0765	SAS.TSR2
PATIENTID	PATI0764	SAS.TSR11
PATIENTNUMBER	PATI0765	SAS.TSR11
PATIENTID	PATI0764	SAS.TSR12
PATIENTNUMBER	PATI0765	SAS.TSR12
PATIENTID	PATI0764	SAS.TSRLEAD
PATIENTNUMBER	PATI0765	SAS.TSRLEAD
PATIENTID	PATI0764	SAS.TSRO1
PATIENTNUMBER	PATI0765	SAS.TSRO1
PATIENTID	PATI0764	SAS.TSRO2
PATIENTNUMBER	PATI0765	SAS.TSRO2
PATIENTID	PATI0764	SAS.UDS
PATIENTNUMBER	PATI0765	SAS.UDS
PATIENTID	PATI0764	SAS.VISR
PATIENTNUMBER	PATI0765	SAS.VISR
PATIENTID	PATI0764	SAS.VIST
PATIENTNUMBER	PATI0765	SAS.VIST

---

## **CTN-0031 Database Descriptions**

---

THIS PAGE INTENTIONALLY LEFT BLANK.