

---

**CTN-0027 CRF Level Data De-Identification Notes**

---

---

**CTN-0027 CRF Level Data De-Identification Notes**

**And**

**Database Descriptions**

**May 17, 2012**

---

---

## CTN-0027 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 and a positive value for days on study refers to an event that occurred after randomization (e.g, 12-month visit date). Age in years at randomization was also calculated based on randomization date and date of birth and entered in the date of birth field.

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 emptied. 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.
- Date field with field type as character were also emptied as they did not represent correct dates.

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:

PATDEID assigns to each study participant a unique value in a variable called PATDEID.

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.

DROPPFILE 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 NIDA Clinical Data Share Website.

### Datasets dropped from the De-identification of the study:

1. **t\_frigen – Genetics dataset**
2. **t\_frplv – Protocol violation log dataset**

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

**ae(Adverse event):** Emptied SUBJECT\_ID (Subject Id), ENTRY\_ID and ENTRY\_DATETIME.

**cidi\_aj (CIDI):** Emptied CIDJ004(other specify), CIDJ005B (other specify), CIDJ018\_7 (other specify)

---

## CTN-0027 CRF Level Data De-Identification Notes

---

**cidi\_I (CIDI):** Emptied CIDL4\_2A (other specify)

**conmeds:** Emptied CRFKEY(Patient id)

**t\_frasf1:** Emptied mcom(Comments), mcom2(Comments), mcom3(Comments), mcom4(Comments), m5b (Other specify), m4b (Other specify), gcom(Comments), gcom2(Comments), gcom3(Comments), gcom4(Comments), g19b, STARTTIM (Start time)

**t\_frasf1a:** Emptied ecom(Comments), ecom2(Comments), ecom3(Comments), ecom4(Comments), e7b

**t\_frasf2a:** Emptied dcom(Comments), dcom2(Comments), dcom3(Comments), dcom4(Comments)

**t\_frasf3:** Emptied lcom(Comments), lcom2(Comments), lcom3(Comments), lcom4(Comments), l16b

**t\_frasf4a:** Emptied fcom(Comments), fcom2(Comments), fcom3(Comments), fcom4(Comments)

**t\_frasf5:** Emptied STOPTIM(time), pcom(Comments), pcom2(Comments), pcom3(Comments), pcom4(Comments)

**t\_frasl1:** Emptied mcom(Comments), mcom2(Comments), mcom3(Comments), mcom4(Comments), gcom4(Comments), m3b, m4b, m5b, gcom2(Comments), gcom3(Comments), g18b, g19b, gcom(Comments)

**t\_frasl1a:** Emptied ecom(Comments), ecom2(Comments), ecom3(Comments), ecom4(Comments)

**t\_frasl2a:** Emptied dcom(Comments), dcom2(Comments), dcom3(Comments), dcom4(Comments)

**t\_frasl3:** Emptied l16b (Other specify)

**t\_frasl3a:** Emptied lcom(Comments), lcom2(Comments), lcom3(Comments), lcom4(Comments)

**t\_frasl4:** Emptied f23c (Other specify)

**t\_frasl4a:** Emptied fcom(Comments), fcom2(Comments), fcom3(Comments), fcom4(Comments)

**t\_frasl5:** Emptied pcom(Comments), pcom2(Comments), pcom3(Comments), pcom4(Comments)

**t\_frABZ (Alcohol Breathalyzer):** Emptied AB002 (comments)

**t\_frcows (Clinical Opiate Withdrawal Scale – Predose):** Emptied cows012a (evaluator's initials)

**t\_frcows2 (Clinical Opiate Withdrawal Scale - Post dose):** Emptied cows012a (evaluator's initials)

**t\_frmenr (Enrollment):** Emptied TXTPINUMBER (Patient ID)

**t\_frDEM (Demographics):** Emptied dem003b5 (other specify), dem004b8 (Comments), dem004d5 (Comments), dem004fs(Comments)

**t\_frDIS (Study Medication Change Form):** Emptied DIS005(Comments)

**t\_frINF (Informed consent):** Emptied txtpinumber (patient number)

**t\_frLAB (Labs and urinalysis):** Emptied LAB001b2 (time), LAB002b2 (time), LAB003b2 (time), LAB004b2 (time), LAB005b2 (time), LAB006b2 (time), LAB001b4 (Accession Number), LAB002b4 (Accession Number), LAB003b4 (Accession Number), LAB004b4 (Accession Number), LAB005b4 (Accession Number), LAB006b4 (Accession Number)

**t\_frMDH (Medical history):** Emptied MDH001s (Specify), MDH002s (Specify), MDH003s (Specify), MDH004s (Specify), MDH005s (Specify), MDH006s (Specify), MDH007s (Specify), MDH008s (Specify), MDH009s (Specify),

---

## CTN-0027 CRF Level Data De-Identification Notes

---

MDH010s (Specify), MDH011As (Specify), MDH011Bs (Specify), MDH012s (Specify), MDH013s (Specify), MDH014s (Specify), MDH015s (Specify), MDH016s (Specify), MDH017s (Specify), MDH018s (Specify), MDH019s (Specify), MDH020s (Specify), MDH021s (Specify), MDH018 (Specify), MDH019 (Specify), MDH020 (Specify), MDH021 (Specify), MDH022 (Specify)

**t\_frPEX (Physical examination):** Emptied PEX001b (Comments for abnormal results), PEX002b (Comments for abnormal results), PEX003b (Comments for abnormal results), PEX004b (Comments for abnormal results), PEX005b (Comments for abnormal results), PEX006b (Comments for abnormal results), PEX007b (Comments for abnormal results), PEX008b (Comments for abnormal results), PEX009b (Comments for abnormal results), PEX010b (Comments for abnormal results), PEX012b (Specify), PEX012c (Specify), PEX011b (Comments for abnormal results)

**t\_frPBC (Pregnancy and Birth Control Assessment):** Emptied PBC006 (Comments)

**t\_frPBC\_BL (Pregnancy and Birth Control Assessment):** Emptied PBC006 (Comments)

**t\_fr poc (Pregnancy and Outcome):** Emptied poc007a(specify abnormality), poc007b(specify other), poc009a(indicate reason)

**t\_frPOC1A (Pregnancy and Outcome Contd):** Emptied POC011g1(specify abnormality), POC011g2(specify factors), POC012(comments)

**t\_frRSA (Research session attendance):** Emptied RSA003 (comments)

**t\_fr ran (Randomization):** Emptied inf003 (randomization number)

**t\_frmScr (Screening):** Emptied TXTSCRPATINIT (Node), TCPATNUM (patient number)

**t\_frUDSAB (Urine Drug Screen):** Emptied AB002(Comments), UDS005A (Specify), UDS006A (Specify), UDS007A (Specify), UDS008A (Specify), UDS009A (Specify), UDS010A (Specify), UDS011A (Specify), UDS012A (Specify), UDS013A (Specify)

**t\_frTFB (Timeline follow back):** Emptied fb011b (other specify)

**t\_fr vsf (Vital signs):** Emptied vsf001 (time)

**t\_frSAE2a (Serious Adverse Event):** Emptied SAE013(relevant test including dates), SAE012A (drug lot number)

**t\_frSAE3 (Serious Adverse Event):** Emptied SAE015(relevant history), SAE016 (Comments), SAE017c (email), SAE017d (SAE reporter's name), SAE017f (Monitor's name), SAE017A (Clinician's name).

**t\_frterm (Study Termination):** Emptied ter003gs(Other specify), ter003fs(Other specify), ter004(dose date known)

---

## CTN-0027 CRF Level Data De-Identification Notes

---

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 NIDA Clinical 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\CTN0027] ";  
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\_FRASF1A was renamed to T\_FRAS12
- b. T\_FRASF2A was renamed to T\_FRAS14
- c. T\_FRASF4A was renamed to T\_FRAS17
- d. T\_FRASL1A was renamed to T\_FRAS20
- e. T\_FRASL2A was renamed to T\_FRAS22
- f. T\_FRASL3A was renamed to T\_FRAS24
- g. T\_FRASL4A was renamed to T\_FRAS26
- h. T\_FRCOWS2 was renamed to T\_FRCO29
- i. T\_FRPBC\_BL was renamed to T\_FRPB43
- j. T\_FRPOC1A was renamed to T\_FRPO47
- k. T\_FRRBS0A was renamed to T\_FRRB49
- l. T\_FRSAE2A was renamed to T\_FRSA54
- m. T\_FRSAE2B was renamed to T\_FRSA55
- n. T\_FRUDSAB was renamed to T\_FRUD59

2. Truncation of fields to make sure that the field length is less than 200 characters.

Variable	Len	Member
CTS_REASON	2000	AE
CTS_REASON	2000	CONMED
ASSESSMENT	500	FIG
CTS_REASON	2000	SAE
SAE003	500	SAE
GCOM	250	T_FRASF1
GCOM2	250	T_FRASF1

Variable	Len	Member
GCOM3	250	T_FRASF1
GCOM4	250	T_FRASF1
M7	255	T_FRASF1
M8	255	T_FRASF1
MCOM	250	T_FRASF1
MCOM2	250	T_FRASF1
MCOM3	250	T_FRASF1

---

**CTN-0027 Database Descriptions**

---

Variable	Len	Member
MCOM4	250	T_FRASF1
E7	255	T_FRASF1A
E20	255	T_FRASF1A
E21	255	T_FRASF1A
ECOM	250	T_FRASF1A
ECOM2	250	T_FRASF1A
ECOM3	250	T_FRASF1A
ECOM4	250	T_FRASF1A
D14	255	T_FRASF2
D10C	255	T_FRASF2
D11C	255	T_FRASF2
D12C	255	T_FRASF2
D3C	255	T_FRASF2
D4AC	255	T_FRASF2
D4C	255	T_FRASF2
D5C	255	T_FRASF2
D6C	255	T_FRASF2
D7C	255	T_FRASF2
D8C	255	T_FRASF2
D9C	255	T_FRASF2
D28	255	T_FRASF2A
D29	255	T_FRASF2A
D30	255	T_FRASF2A
D31	255	T_FRASF2A
DCOM	250	T_FRASF2A
DCOM2	250	T_FRASF2A
DCOM3	250	T_FRASF2A
DCOM4	250	T_FRASF2A
L2	255	T_FRASF3
L25	255	T_FRASF3
L28	255	T_FRASF3
L29	255	T_FRASF3
LCOM	250	T_FRASF3
LCOM2	250	T_FRASF3
LCOM3	250	T_FRASF3
LCOM4	250	T_FRASF3
F1	255	T_FRASF4

Variable	Len	Member
F4	255	T_FRASF4
F32	255	T_FRASF4A
F33	255	T_FRASF4A
F34	255	T_FRASF4A
F35	255	T_FRASF4A
FCOM	250	T_FRASF4A
FCOM2	250	T_FRASF4A
FCOM3	250	T_FRASF4A
FCOM4	250	T_FRASF4A
P13	255	T_FRASF5
P14	255	T_FRASF5
PCOM	250	T_FRASF5
PCOM2	250	T_FRASF5
PCOM3	250	T_FRASF5
PCOM4	250	T_FRASF5
G17	255	T_FRASL1
GCOM	250	T_FRASL1
GCOM2	250	T_FRASL1
GCOM3	250	T_FRASL1
GCOM4	250	T_FRASL1
M7	255	T_FRASL1
M8	255	T_FRASL1
MCOM	250	T_FRASL1
MCOM2	250	T_FRASL1
MCOM3	250	T_FRASL1
MCOM4	250	T_FRASL1
E7	255	T_FRASL1A
E10	255	T_FRASL1A
E20	255	T_FRASL1A
E21	255	T_FRASL1A
ECOM	250	T_FRASL1A
ECOM2	250	T_FRASL1A
ECOM3	250	T_FRASL1A
ECOM4	250	T_FRASL1A
D10C	255	T_FRASL2
D11C	255	T_FRASL2
D12C	255	T_FRASL2

---

**CTN-0027 Database Descriptions**

---

Variable	Len	Member
D36C	255	T_FRASL2
D3C	255	T_FRASL2
D4AC	255	T_FRASL2
D4C	255	T_FRASL2
D5C	255	T_FRASL2
D6C	255	T_FRASL2
D7C	255	T_FRASL2
D8C	255	T_FRASL2
D9C	255	T_FRASL2
D28	255	T_FRASL2A
D29	255	T_FRASL2A
D30	255	T_FRASL2A
D31	255	T_FRASL2A
DCOM	250	T_FRASL2A
DCOM2	250	T_FRASL2A
DCOM3	250	T_FRASL2A
DCOM4	250	T_FRASL2A
L2	255	T_FRASL3
L25	255	T_FRASL3A
L28	255	T_FRASL3A
L29	255	T_FRASL3A
LCOM	250	T_FRASL3A
LCOM2	250	T_FRASL3A
LCOM3	250	T_FRASL3A
LCOM4	250	T_FRASL3A
F1	255	T_FRASL4
F4	255	T_FRASL4
F32	255	T_FRASL4A
F33	255	T_FRASL4A
F34	255	T_FRASL4A
F35	255	T_FRASL4A
FCOM	250	T_FRASL4A
FCOM2	250	T_FRASL4A
FCOM3	250	T_FRASL4A
FCOM4	250	T_FRASL4A
P13	255	T_FRASL5
P14	255	T_FRASL5

Variable	Len	Member
PCOM	250	T_FRASL5
PCOM2	250	T_FRASL5
PCOM3	250	T_FRASL5
PCOM4	250	T_FRASL5
COWS001	255	T_FRCOWS
COWS002	255	T_FRCOWS
COWS003	255	T_FRCOWS
COWS004	255	T_FRCOWS
COWS005	255	T_FRCOWS
COWS006	255	T_FRCOWS
COWS007	255	T_FRCOWS
COWS008	255	T_FRCOWS
COWS009	255	T_FRCOWS
COWS010	255	T_FRCOWS
COWS011	255	T_FRCOWS
COWS012	255	T_FRCOWS
COWS001	255	T_FRCOWS2
COWS002	255	T_FRCOWS2
COWS003	255	T_FRCOWS2
COWS004	255	T_FRCOWS2
COWS005	255	T_FRCOWS2
COWS006	255	T_FRCOWS2
COWS007	255	T_FRCOWS2
COWS008	255	T_FRCOWS2
COWS009	255	T_FRCOWS2
COWS010	255	T_FRCOWS2
COWS011	255	T_FRCOWS2
COWS012	255	T_FRCOWS2
DEM003B1	255	T_FRDEM
DEM003B2	255	T_FRDEM
DEM003B3	255	T_FRDEM
DEM003B4	255	T_FRDEM
DEM004A	255	T_FRDEM
DEM004B	255	T_FRDEM
DEM004B1	255	T_FRDEM
DEM004B2	255	T_FRDEM
DEM004B3	255	T_FRDEM

---

**CTN-0027 Database Descriptions**

---

Variable	Len	Member
DEM004B4	255	T_FRDEM
DEM004B5	255	T_FRDEM
DEM004B6	255	T_FRDEM
DEM004B7	255	T_FRDEM
DEM004C	255	T_FRDEM
DEM004D	255	T_FRDEM
DEM004D1	255	T_FRDEM
DEM004D2	255	T_FRDEM
DEM004D3	255	T_FRDEM
DEM004D4	255	T_FRDEM
DEM004E	255	T_FRDEM
DEM004F	255	T_FRDEM
DEM004G	255	T_FRDEM
DEM004H	255	T_FRDEM
DIS004B	255	T_FRDIS
DIS004C	255	T_FRDIS
DIS004C1	255	T_FRDIS
DIS004C2	255	T_FRDIS
DIS004C3	255	T_FRDIS
DIS004C4	255	T_FRDIS
PEX001A	255	T_FRPEX
PEX002A	255	T_FRPEX
PEX003A	255	T_FRPEX
PEX004A	255	T_FRPEX
PEX005A	255	T_FRPEX
PEX006A	255	T_FRPEX
PEX007A	255	T_FRPEX
PEX008A	255	T_FRPEX
PEX009A	255	T_FRPEX
PEX010A	255	T_FRPEX
PEX011A	255	T_FRPEX
PEX012A	255	T_FRPEX
POC002A	255	T_FRPOC
POC002B	255	T_FRPOC
POC002C	255	T_FRPOC
RBS0D1B	255	T_FRRBSF
RBS0D1C	255	T_FRRBSF

Variable	Len	Member
RBS0D2B	255	T_FRRBSF
RBS0D2C	255	T_FRRBSF
RBS0G1B	255	T_FRRBSF
RBS0G1C	255	T_FRRBSF
RBS0G2B	255	T_FRRBSF
RBS0G2C	255	T_FRRBSF
RBS0H1B	255	T_FRRBSF
RBS0H1C	255	T_FRRBSF
RBS0I1B	255	T_FRRBSF
RBS0I1C	255	T_FRRBSF
RBS0D1B	255	T_FRRBSM
RBS0D1C	255	T_FRRBSM
RBS0D2B	255	T_FRRBSM
RBS0D2C	255	T_FRRBSM
RBS0E1B	255	T_FRRBSM
RBS0E1C	255	T_FRRBSM
RBS0E2B	255	T_FRRBSM
RBS0E2C	255	T_FRRBSM
RBS0F1B	255	T_FRRBSM
RBS0F1C	255	T_FRRBSM
RBS0G1B	255	T_FRRBSM
RBS0G1C	255	T_FRRBSM
RBS0G2B	255	T_FRRBSM
RBS0G2C	255	T_FRRBSM
RBS0I1B	255	T_FRRBSM
RBS0I1C	255	T_FRRBSM
TER003A	255	T_FRTERM
TER003AA	255	T_FRTERM
TER003AB	255	T_FRTERM
TER003AC	255	T_FRTERM
TER003B	255	T_FRTERM
TER003C	255	T_FRTERM
TER003D	255	T_FRTERM
TER003E	255	T_FRTERM
TER003F	255	T_FRTERM
TER003FA	255	T_FRTERM
TER003FB	255	T_FRTERM



---

**CTN-0027 Database Descriptions**

---

Variable	Len	Member
TER003FD	255	T_FRTERM
TER003FE	255	T_FRTERM
TER003FF	255	T_FRTERM
TER003FG	255	T_FRTERM
TER003FH	255	T_FRTERM
TER003FI	255	T_FRTERM
TER003G	255	T_FRTERM
TER003NA	255	T_FRTERM
TER03FC1	255	T_FRTERM
TER03FC2	255	T_FRTERM
TER03FC3	255	T_FRTERM
TER03FC4	255	T_FRTERM
TFB001C	255	T_FRTFB
TFB002B	255	T_FRTFB
TFB003B	255	T_FRTFB
TFB004B	255	T_FRTFB
TFB005B	255	T_FRTFB
TFB006B	255	T_FRTFB
TFB007B	255	T_FRTFB
TFB008B	255	T_FRTFB
TFB009B	255	T_FRTFB
TFB010B	255	T_FRTFB
TFB011B	255	T_FRTFB
TFB011C	255	T_FRTFB

---

## **CTN-0027 Database Descriptions**

---

THIS PAGE INTENTIONALLY LEFT BLANK.