

Practical Approaches for Using the CTN Datashare: Proposing Parameters, Creating Data Tables, and Analyzing Results

Abigail Matthews

The EMMES Corporation

June 14, 2014

amatthews@emmes.com

Outline

- How to Use CTN Datashare
- Demonstration
- Recommendations

No conflicts of interest to report.

Supported by contract HHSN271201200017C/
N01DA-12-2229

datashare.nida.nih.gov

HOW TO USE CTN DATASHARE

How to Use CTN Datashare

- Analysis of only one study
 - Read protocol carefully to identify assessments conducted and any issues with data collection
 - Examine annotated case report forms (CRFs)
 - Used to determine data set name and “domain” for CDISC data
 - Defines data set and variable names for CRF-based data
 - Review de-identification document to determine if any variables of interest have been re-coded or deleted
 - Read primary manuscript which is available from CTN Dissemination Library: ctndisseminationlibrary.org

How to Use Datashare (cont'd)

- Analysis of more than one study
 - Read each individual protocol carefully to:
 - Identify study populations and any subsets that may be of interest for your intended research question
 - Determine how data can be analyzed – e.g., longitudinal data may require using study day not visit numbers
 - Review annotated CRFs and data dictionaries:
 - Data set, domain and variable names
 - Identify common data elements – i.e., data points which are collected similarly across protocols

How to Use Datashare (cont'd)

- Analysis of more than one study (cont'd)
 - Review de-identification document for each study
 - Read all primary outcome papers

- Assessment tab
 - Can identify assessments of interest and associated protocols
 - Obtain listing of assessments for any protocol



DEMONSTRATION



NIDA Data Share

Purpose

The NIDA Data Share web site is an electronic environment that allows data from completed clinical trials to investigators and the public in order to promote new research, encourage further analyses, and disseminate to the scientific community. Secondary analyses produced from data sharing multiply the scientific contribution of the original data and supports the timely release and sharing of final research data from NIH-supported studies for use by others in the translation of research results into knowledge, products and procedures to improve human health. (see <http://grants1.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>).

In order to make the NIDA data available to as wide an audience as possible, the NIDA Data Share web site (<http://datashare.nida.nih.gov>) has been created. As studies are completed and their data become available to that data. The following information will be posted per protocol:

1. Study protocol
2. Reference to study publication of primary outcome
3. Data sets (SAS and ASCII)
4. Annotated case report forms
5. Define file (also known as Data Dictionary)
6. Study-specific de-identification notes

Protection of Human Subjects

Our primary concern in sharing data is the protection of human subjects. The rights and privacy of people who

Assessments Tab

Protocol	FTND
NIDA-CTN-0009	✓
NIDA-CTN-0029	✓
NIDA-CTN-0030	✓
NIDA-CTN-0031	✓
NIDA-CTN-0027	✓
NIDA-CTN-0046	✓
NIDA-CTN-0044	✓
NIDA-CTN-0052	✓

▼ Protocols with similar assessments

Assessment Name	Total	Similar	Matching
Addiction Severity Index Lite (ASI)	20	4	4
Composite International Diagnostic Interview for DSM-IV (CIDI)	11	4	4
<u>DSM IV Criteria Substance Related Disorders (DSMS)</u>	7	4	4
Risk Behavior Survey (RBS)	15	4	4
Time Line Follow Back (TLFB)	9	4	4
Beck Depression Inventory (BDI)	3	3	3
Psychiatric Research Interview for Substance and Mental Disorders (PRISM)	4	3	3
SF-36 Health Status Questionnaire (SF36)	5	3	3
Smoking History Survey (SHS)	3	3	3
Thoughts About Abstinence (TAA)	4	3	3

Assessments Tab

Assessment	NIDA-CTN-0044
BSI	✓
CSS	✓
DSMS	✓
EQ5D	✓
FTND	✓
MC	✓
PHQ	✓
RBS	✓
SAS-SR	✓
TLFB	✓

▼ Protocols with similar assessments

Protocol Title	Protocol Number	Total	Similar	Matching
Starting Treatment with Agonist Replacement Therapies (START)	NIDA-CTN-0027	6	4	4
Smoking Cessation Treatment with Transdermal Nicotine Replacement Therapy in Substance Abuse Rehabilitation Programs	NIDA-CTN-0009	17	3	3
Buprenorphine/Naloxone-Facilitated Rehabilitation for Opioid Dependent Adolescents/Young Adults	NIDA-CTN-0010	10	3	3
Motivational Incentives for Enhanced Drug Abuse Recovery: Drug Free Clinics	NIDA-CTN-0006	5	2	2
Motivational Incentives for Enhanced Drug Abuse Recovery: Methadone Clinics	NIDA-CTN-0007	4	2	2

Data Share Files

Click on Protocol Number of interest to view study-related documentation or download data files.
(DPMC data files are under development.)

Study Number ▲	Study Title	Investigator	Release Date
NIDA-CTN-0001	Buprenorphine/Naloxone versus Clonidine for Inpatient Opiate Detoxification	Walter Ling, M.D.	05/08/2006
NIDA-CTN-0002	Buprenorphine/Naloxone versus Clonidine for Outpatient Opiate Detoxification	Walter Ling, M.D.	07/26/2006
NIDA-CTN-0003	Suboxone (Buprenorphine/Naloxone) Taper: A Comparison of Two Schedules	Walter Ling, M.D.	05/29/2008
NIDA-CTN-0004	MET to Improve Treatment Engagement and Outcome in Subjects Seeking Treatment for Substance Abuse	Kathleen Carroll, Ph.D.	10/03/2007
NIDA-CTN-0005	MI (Motivational Interviewing) to Improve Treatment Engagement and Outcome in Subjects Seeking Treatment for Substance Abuse	Kathleen Carroll, Ph.D.	09/08/2006
NIDA-CTN-0006	Motivational Incentives for Enhanced Drug Abuse Recovery: Drug Free Clinics	Maxine Stitzer, Ph.D.	10/31/2006
NIDA-CTN-0007	Motivational Incentives for Enhanced Drug Abuse Recovery: Methadone Clinics	Maxine Stitzer, Ph.D.	12/13/2006
NIDA-CTN-0008	Assessment of the National Drug Abuse Clinical Trials Network: A Baseline for Investigating Diffusion of Innovation	Dennis McCarty, Ph.D.	01/25/2007
NIDA-CTN-0009	Smoking Cessation Treatment with Transdermal Nicotine Replacement Therapy in Substance Abuse Rehabilitation Programs	Malcolm Reid, Ph.D.	12/03/2007

NIDA-CTN-0006

[<< Previous Protocol](#) [Next Protocol >>](#)

Last Modified: Jul 01, 2011

Study Number: NIDA-CTN-0006

Investigator: [Maxine Stitzer, Ph.D.](#)

Study Title: Motivational Incentives for Enhanced Drug Abuse Recovery: Drug Free Clinics

Short Description: To study treatment incentives, interventions that can motivate clients to attend treatment sessions and initiate and sustain abstinence, compared to treatment outcomes in Community Treatment Programs.

Release Date: 10/31/2006

Study Description:

To improve treatment outcomes, it is necessary to use interventions that can motivate clients to attend treatment and initiate and sustain abstinence. Extensive research has consistently shown that client motivation for positive behavioral change can be influenced by the use of novel interventions based on established behavioral principles. In numerous controlled experimental studies in both outpatient cocaine treatment and methadone maintenance programs, procedures involving the provision of tangible incentives to clients for drug-free urines resulted in sustained periods of abstinence and retention in treatment. The initial study in Community Treatment Programs will utilize incentive values considerably lower than those typically used in research clinics.

Keywords: [Motivational Incentives \(2\)](#)

▶ [Assessment Summary](#)

▶ [Protocols using similar assessments](#)

General Protocol Information

 [CTN-0006 Primary Manuscript](#) 

 [CTN-0006 Study Information](#)

 [CTN-0006 Study Protocol](#) 

CDISC Data

 [CTN-0006 Data Files](#)

 [CTN-0006 Annotated CRF](#)

 [CTN-0006 Study-Specific De-identification Notes](#)

NIDA-CTN-0044

Last Modified: May 15, 2014

Study Number: NIDA-CTN-0044

Investigator: [Edward V. Nunes, MD](#)

Study Title: Web-delivery of Evidence-Based, Psychosocial Treatment for Substance Use Disorders

Short Description: A randomized study to evaluate the effectiveness of including an interactive, web-based version of the Community Reinforcement Approach (CRA) intervention plus incentives targeting drug abstinence and treatment participation as part of community-based, outpatient substance abuse treatment.

Release Date: 05/14/2014

Study Description:

This was a randomized, controlled trial in which 507 participants in outpatient treatment for substance use disorders were recruited from ten sites to participate in a study comparing 1) treatment as usual (TAU) to 2) a modification of TAU, which included access to the Therapeutic Education System (TES), a computerized psychosocial intervention which combines skill building modules based on the Community Reinforcement Approach (CRA) with incentives contingent upon abstinence from drugs and completion of TES modules. The purpose of this study was to evaluate the relative effectiveness of the interventions on the primary outcomes of a) drug abstinence during active treatment and b) treatment retention. The study also evaluated a number of secondary outcome measures.

Keywords: [web-based therapy \(1\)](#), [telemedicine \(1\)](#), [computer assisted therapy \(1\)](#)

▶ [Assessment Summary](#)

▶ [Protocols using similar assessments](#)

[<< Previous Protocol](#) [Next Protocol >>](#)

General Protocol Information

 [CTN-0044 Study Information](#)

 [CTN-0044 Study Protocol](#) 

CRF Level Data

 [CTN-0044 Data Files](#)

 [CTN-0044 Data Dictionary](#)

 [CTN-0044 CRFs](#)

 [CTN-0044 De-identification Notes](#)

DM:

- Make SITEID blank.
- Recode several ('AMERICAN INDIAN OR ALASKAN NATIVE', 'ASIAN', 'MULTIPLE', 'OTHER (SPECIFY)') low count RACEs to "OTHER".

DS: Clean.

SUPPDS: Months and days were removed from discontinuation reasons.

IE: Clean.

QS:

- Some QSORRES and QSSTRESC = MM/DD/YY so made to be MM/DD/YY.
- Some QSORRES contained unique race categories which have been recoded to OTHER.

SC:

- Dropped NODE.
- Dropped all RACE based SCTESTs that could identify someone based on a specific race entry. RACE should be gotten from DM.RACE.

SUPPSC:

All records are potentially identifying info. This dataset should not be given as de-identified data.

SE: Clean.

SU: Clean.

SV: Clean.

TA, TE, TI, TS, and TV are all trial metadata files and have no identifying information

TU: Clean.

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 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.

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.

Note: BFSID on BFS dataset was masked with a random number.

For all the datasets dates were changed to days of study, site and randomization date (RANDDT) was emptied.

enra (Screening enrollment): Removed sccnstdt and sccomm (comments)

enrb (Randomization enrollment): Removed RAASMTDT, RAOTHRSP (Specify) and RAJUDGSP (Specify)

enrz (Screen Failures): Removed RAASMTDT, RAOTHRSP (Specify) and RAJUDGSP (Specify)

ad1 (Adverse Event): Removed (Specify), a1aebsp, a1comm, a1hospad, a1hospc

ad2 (Serious Adverse Event Summary): Removed a2summ, a2medhx, a2follup and a2addinf

ad3 (Serious Adverse Event Medical Reviewer): Removed a3addinf and a3comm

Dem (Demographics): Removed deothrsp (Specify) and demcomm (Comments)

alb (Alcohol Breathalyze): Removed ab2othsp (Specify), ab1othsp (Specify) and albcomm (Comments)

Ready to Download Data

1. Choose a data structure (if available)
 - CDISC (FDA format – long and narrow)
 - One row per question
 - Variable name/question described in one or more columns
 - All values/answers in one column
 - Minimizes amount of merging
 - CRF-based (wide)
 - Each CRF is a separate data set
 - Each question/variable is a separate column
 - May requires merging of a substantial number of data sets

DEMOGRAPHICS (DEM)

DOMAIN: DM and SC

NIDA-CTN-0009		Smoking Cessation Treatment With Transdermal Nicotine Replacement Therapy In Substance Abuse Rehabilitation Programs					Seq. Num.: <u>1</u>									
NODE:	<input type="checkbox"/> <input type="checkbox"/>	STUDY WEEK: <i>(circle one)</i>	EPOCH		Screening	Baseline	VISITNUM									
CTP-SITE ID:	<input type="checkbox"/> <input type="checkbox"/> - <input type="checkbox"/> <input type="checkbox"/>		-1	1	2	3	4	5	6	7	8	9	10	11	12	13
PARTICIPANT ID:	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> USUBJID		14	15	16	17	18	19	20	21	22	23	24	25	26	
Assessment Date: ___ / ___ / ___ (mm / dd / yyyy)						STAFF ID: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>										
<input type="checkbox"/> FORM COMPLETION STATUS		1=Form completed as required 2=Participant refused 3=Responsible person did not complete			4=Not enough time at the visit 5=Participant did not attend visit 6=Other (specify: _____)											

DEMOGRAPHICS

1. Sex **DMSEX**

(1 = Male, 2=female)

2. Date of Birth

___ / ___ / ___
m m d d y y y y

DMBRTHDTC

3. **ETHNICITY/RACE:** For each of the following, please circle '0' for 'NO' or '1' for 'YES'. You have the option of not answering by circling '1' for 'Participant chooses not to answer'. For those categories with further specification, please select at least one sub-category.

SCTEST

NO	YES	DMRACE (If multiple, then = 'MULTIPLE')
0	1	White
0	1	Black, African American or Negro SCTEST
0	1	American Indian or Alaska Native
0	1	Spanish, Hispanic or Latino (enter "1" below for all that apply)
		___ Mexican, Mexican-American, or Chicano
		___ Puerto Rican
		___ Cuban
		___ Other (specify) _____

**QNAM=SOTHERS
QLABEL=SPANISH, HISPANIC, OR LATINO: OTHER TEXT
IDVAR=SCSEQ**

CDISC SAS File (CTN-0009)

DOMAIN	USUBJID	EPOCH	VSTESTCD	VSTEST	VSORRES	VSORRESU	VISIT	VISITNUM	VSDY
VS	09_000748	SCREENING	DBP	DIASTOLI	76	MMHG	SCREENING	-3	-23
VS	09_000748	BASELINE	DBP	DIASTOLI	86	MMHG	BASELINE	-2	-14
VS	09_000748	PRETDQ	DBP	DIASTOLI	77	MMHG	PRE-TDQ	-1	6
VS	09_000748	TREATMENT	DBP	DIASTOLI	71	MMHG	TREATMENT	1	14
VS	09_000748	TREATMENT	DBP	DIASTOLI	79	MMHG	TREATMENT	2	21
VS	09_000748	TREATMENT	DBP	DIASTOLI	83	MMHG	TREATMENT	3	27
VS	09_000748	TREATMENT	DBP	DIASTOLI	74	MMHG	TREATMENT	4	34
VS	09_000748	TREATMENT	DBP	DIASTOLI	71	MMHG	TREATMENT	5	42
VS	09_000748	TREATMENT	DBP	DIASTOLI	79	MMHG	TREATMENT	6	49
VS	09_000748	TREATMENT	DBP	DIASTOLI	74	MMHG	TREATMENT	7	57
VS	09_000748	TREATMENT	DBP	DIASTOLI	84	MMHG	TREATMENT	8	63
VS	09_000748	FOLLOW UP	DBP	DIASTOLI	84	MMHG	FOLLOW UP	9	69
VS	09_000748	FOLLOW UP	DBP	DIASTOLI	83	MMHG	FOLLOW UP	13	105
VS	09_000748	SCREENING	PULSE	PULSE	93	BEATS/MIN	SCREENING	-3	-23
VS	09_000748	BASELINE	PULSE	PULSE	88	BEATS/MIN	BASELINE	-2	-14
VS	09_000748	PRETDQ	PULSE	PULSE	87	BEATS/MIN	PRE-TDQ	-1	6
VS	09_000748	TREATMENT	PULSE	PULSE	83	BEATS/MIN	TREATMENT	1	14
VS	09_000748	TREATMENT	PULSE	PULSE	90	BEATS/MIN	TREATMENT	2	21
VS	09_000748	TREATMENT	PULSE	PULSE	82	BEATS/MIN	TREATMENT	3	27
VS	09_000748	TREATMENT	PULSE	PULSE	72	BEATS/MIN	TREATMENT	4	34
VS	09_000748	TREATMENT	PULSE	PULSE	86	BEATS/MIN	TREATMENT	5	42
VS	09_000748	TREATMENT	PULSE	PULSE	79	BEATS/MIN	TREATMENT	6	49
VS	09_000748	TREATMENT	PULSE	PULSE	95	BEATS/MIN	TREATMENT	7	57
VS	09_000748	TREATMENT	PULSE	PULSE	84	BEATS/MIN	TREATMENT	8	63
VS	09_000748	FOLLOW UP	PULSE	PULSE	65	BEATS/MIN	FOLLOW UP	9	69

Alcohol Breathalyzer (ALB)

Web Version: 1.

Segment (PROTSEG):

Visit Number (VISNO):

1. Was an Alcohol Breathalyzer performed? (ABPERFRM)

If "No", specify reason: (ABREASON)

If "Other", specify (AB1OTHSP)

2. Date of assessment: (ABASMTDT)

3. Alcohol Breathalyzer result: (ABRESULT)

4. Is a repeat test required? (ABREPTST)

If "Yes", complete the questions below.

a. Was the repeat Alcohol Breathalyzer performed? (ABREPPRF)

If "No", specify reason: (ABRREASN)

If "Other", specify (AB2OTHSP)

b. Repeat test date: (ABREPD T)

c. Repeat Alcohol Breathalyzer result: (ABREPRES)

Comments: (ALBCOMM)

No Yes

Participant refused to provide sample Study staff error Other

(mm/dd/yyyy)

(.xxx) mg/mL

No Yes

No Yes

Participant refused to provide sample Study staff error Other

(mm/dd/yyyy)

(.xxx) mg/mL

CRF-based SAS File (CTN-0052)

PROT	PATID	VISNO	UDTSTPRF	UDCOLDT	UD1TMP	UD1ADULT	UD1THC	UD1OPI	UD1COC	UD1MDMA	UD1OXY	UD1MTD	UD1BAR
0052	193369	30	1	100	1	0	0	0	0 0	0	0	0	0
0052	193369	29	1	98	1	0	0	0	0 0	0	0	0	0
0052	193369	28	1	93	1	0	0	0	0 0	0	0	0	0
0052	193369	27	1	91	1	0	0	0	0 0	0	0	0	0
0052	193369	26	1	87	1	0	0	0	0 0	0	0	0	0
0052	193369	25	1	85	1	0	0	0	0 0	0	0	0	0
0052	193369	24	1	80	1	0	0	0	0 0	0	0	0	0
0052	193369	23	1	77	1	0	0	0	0 0	0	0	0	0
0052	193369	22	1	72	1	0	0	0	0 0	0	0	0	0
0052	193369	21	1	70	1	0	0	0	0 0	0	0	0	0
0052	193369	20	1	67	1	0	0	0	0 0	0	0	0	0
0052	193369	19	1	65	1	0	0	0	0 0	0	0	0	0
0052	193369	18	1	60	1	0	0	0	0 0	0	0	0	0
0052	193369	17	1	56	1	0	0	0	0 0	0	0	0	0
0052	193369	16	1	51	1	0	0	0	0 0	0	0	0	0
0052	193369	15	1	49	1	0	0	0	0 0	0	0	0	0
0052	193369	14	1	44	1	0	0	0	0 0	0	0	0	0
0052	193369	13	1	42	1	0	0	0	0 0	0	0	0	0
0052	193369	12	1	37	1	0	0	0	0 0	0	0	0	0



Ready to Download Data (cont'd)



- Choose format of data files (if available)
 - ASCII
 - Text only
 - Can be read into Excel, SPSS, SAS, Stata etc.
 - SAS
 - Actually SAS transport files (.xpt)
 - Code for reading into SAS:

```
libname xptlib xport 'Location\dm.xpt';  
libname saslib 'Location';  
proc copy inlib=xptlib outlib=saslib;  
run;
```

!!! User Agreement Highlights !!!

- Not to use the received data, either alone or in conjunction with any other information, in any effort whatsoever to establish identities of any of the subjects from whom the data were obtained
- Acknowledge the CTN database and the specific trials accessed in all presentations/publications: *“The information reported here results from secondary analyses of data from clinical trials conducted as part of the National Drug Abuse Treatment Clinical Trials Network (CTN) sponsored by National Institute on Drug Abuse (NIDA). Specifically, data from CTN-XXXX (actual protocol number and title) were included. CTN databases and information are available at <http://datashare.nida.nih.gov>”*
- Inform the CCTN Data Manager when research paper(s) that include the use of CTN data, is (are) accepted for publication.

Download Data

- After providing your information and agreeing to terms of use click on “Download Data”
- Unzip files

ANALYZE away!!!



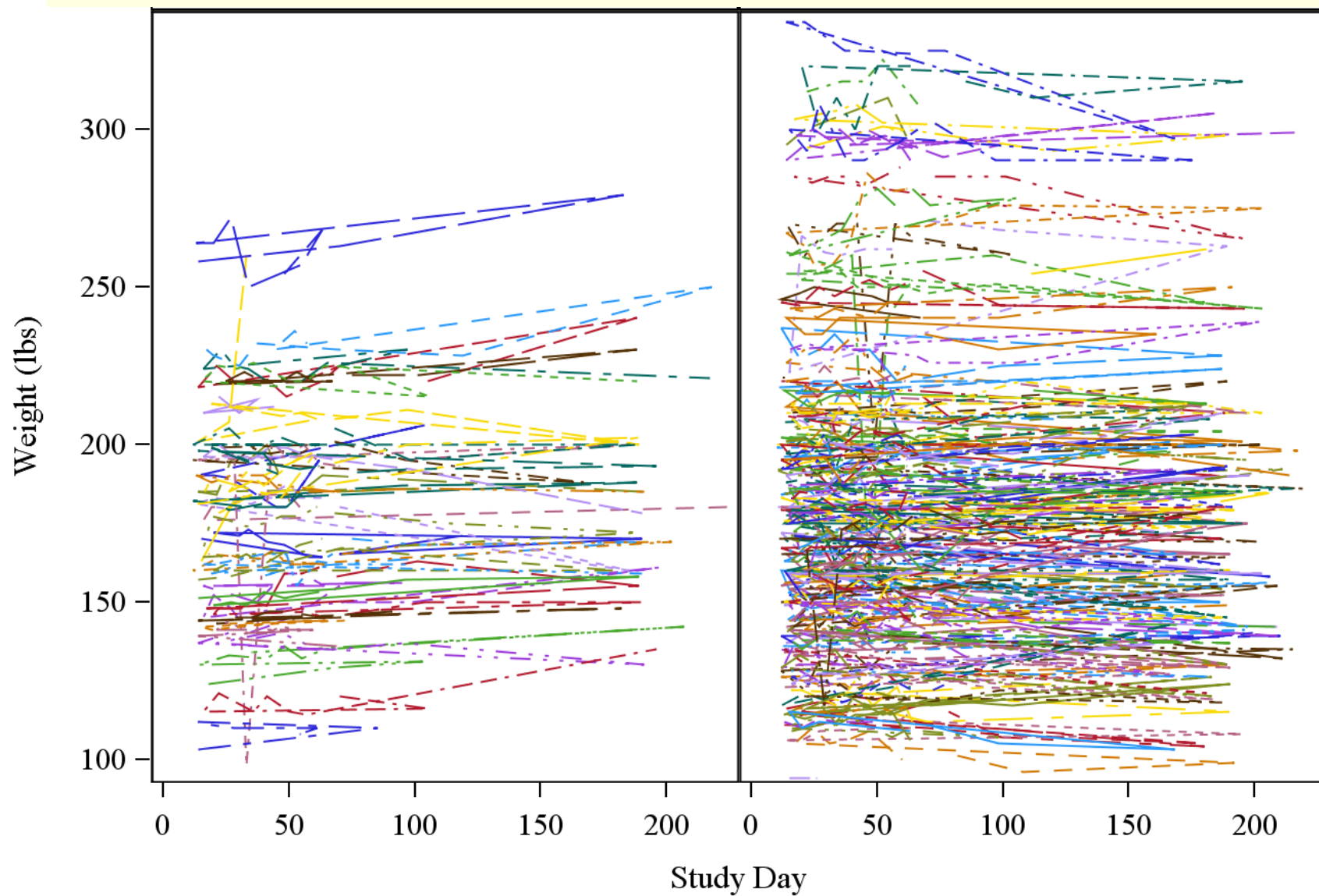
RECOMMENDATIONS

Recommendations

- Pay special attention to CDISC data
 - “Epoch” and “Visit Name” usually must be combined or use study day
 - “Domain” is defined on CRFs
- Raw data
 - Outliers
 - Note how missing data is coded
- Be very careful when merging data sets
- Be mindful of statistical issues with secondary analyses and their interpretation

Upcoming Presentation of Datashare Analyses

- Eating Disorders
- Monday 7-9pm Ballroom C
- Presents analysis of body weight longitudinally from four studies downloaded from Datashare



**THANK YOU
&
GOOD LUCK!!!**