Package ‘public.ctn0094data’

October 1, 2023

Title De-Identified Data from CTN-0094
Version 1.0.6
Date 2023-09-21
Description These are harmonized datasets produced as part of the Clinical Trials Network (CTN) protocol number 0094. This is a US National Institute of Drug Abuse (NIDA) funded project; to learn more go to <https://ctnlibrary.org/protocol/ctn0094/>. These are datasets which have the data harmonized from CTN-0027 (<https://ctnlibrary.org/protocol/ctn0027/>), CTN-0030 (<https://ctnlibrary.org/protocol/ctn0030/>), and CTN-0051 (<https://ctnlibrary.org/protocol/ctn0051/>).
License MIT + file LICENSE
Encoding UTF-8
Language en-US
LazyData true
Depends R (>= 2.10)
Suggests broom, conflicted, DiagrammeR, dplyr, forcats, ggplot2, ggtthemes, gt, haven, infer, janitor, knitr, kableExtra, magrittr, pkgdown, plyr, psych, purrr, rlang, rmarkdown, roxygen2, rUM, scales, stringr, table1, tibble, tidyverse, testthat (>= 3.0.0), vcdExtra, xfun
RoxygenNote 7.2.3
VignetteBuilder knitr
Config/testthat/edition 3
URL https://ctn-0094.github.io/public.ctn0094data/
BugReports https://github.com/CTN-0094/public.ctn0094data/issues
NeedsCompilation no
Author Raymond Balise [aut, cre] (<https://orcid.org/0000-0002-9856-5901>), Gabriel Odom [aut] (<https://orcid.org/0000-0003-1341-4555>), Anna Calderon [aut] (<https://orcid.org/0000-0002-0139-3841>)
Maintainer Raymond Balise <balise@miami.edu>
R topics documented:

all_drugs ................................. 2
asi ........................................ 4
demographics ............................ 5
detox ........................................ 7
everybody ................................... 8
fagerstrom ................................. 8
first_survey ............................... 9
meta_study_length ........................ 10
meta_substance_groups_uds .......... 12
pain ......................................... 13
psychiatric ............................... 14
qol .......................................... 18
randomization ............................ 18
rbs ......................................... 19
rbs_iv ....................................... 20
screening_date ......................... 22
sex .......................................... 23
site_masked ............................... 30
tlfb ......................................... 30
treatment .................................. 31
uds ......................................... 32
uds_temp .................................. 33
visit ......................................... 34
withdrawal ................................ 38
withdrawal_pre_post ................... 39

Index 41

all_drugs All drugs taken

Description

This is a record of both self-reported drug use and positive drug screening results. See the vignette Harmonization Information for more details.

Usage

data(all_drugs)
Format

A tibble with 307,523 rows and 4 variables:
Type: integer
Description: Patient ID

who
Type: factor (First/Reference level = Acetaminophen)
Description: Name of drug (or alcohol) from self-reported or drug screening
Levels: Acetaminophen, Alcohol, Amphetamine, Antibiotic, Antidepressant, Antiemetic, Antihistamine, Antipsychotic, Barbiturate, Benzodiazepine, Cannabis (THC), Cocaine, Codeine, Deliriant, Depressant, Dextromethorphan, Euphorant, Hallucinogenic, Inhalant, Sedative, Stimulant, Tranquilizer, Tricyclic Antidepressant, Unknown

what
Type: factor (First/Reference level = TFB)
Description: Source of reported drug use. TLFB = Timeline Follow Back; UDS = Urine Drug Screening; UDSAB = Urine Drug Screening & Alcohol Breathalyzer
Levels: TFB, UDS, UDSAB

source
Type: integer
Description: Study day

when

asi

Did patient use intravenous drugs

Description
IV drug use information from the Addiction Severity Index

Usage
data(asi)

Format
A tibble with 3,560 rows and 2 variables:
### used_iv

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Description**

This is baseline demographics. See the vignette *Harmonization Information* for more details.

**Usage**

demographics

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**who**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**age**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

---

**Type:** integer

**Description:** Patient ID

**who**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**age**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

---

**Type:** factor (First/Reference level = No)

**Description:** Self-reported history of IV drug use

**Levels:** No, Yes

**who**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**age**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

---

**Type:** integer

**Description:** Patient ID

**who**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**age**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

---

**Type:** factor (First/Reference level = No)

**Description:** Self-reported history of IV drug use

**Levels:** No, Yes

**who**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**age**

<table>
<thead>
<tr>
<th>demographics</th>
<th>Patient demographics</th>
</tr>
</thead>
</table>

**Format**

A tibble with 3,560 rows and 9 variables:
demographics

is_hispanic

Type: factor (First/Reference level = Black)

Description: Hispanic heritage

Levels: No; Yes

race

Type: factor (First/Reference level = Full Time)

Description: Self-reported race

Levels: Black, Other Refused/missing, White

job

Type: factor (First/Reference level = No)

Description: Employment status at intake


is_living_stable

Type: factor (First/Reference level = HS/GED)

Description: Education level at intake

Levels: "HS/GED" = High school graduate or GED, "Less than HS" = Less High school and no GED, "Missing", More than HS = "Some education beyond high school"

education

Type: factor (First/Reference level = Married or Partnered)

Description: Marital status at intake
**Levels:**

"" = Not asked, "Married or Partnered", "Never married" "Not answered" = Not asked during at intake, "Separated/Divorced/Widowed"

**marital**

*Type:* factor (First/Reference level = No)

*Description:* Sex (not gender)

*Levels:* No = Not Male); Yes = Is male

**is_male**

---

**Start and Stop of Detox**

*Description*

This is the start and stop date for detox (if known).

*Usage*

`data(detox)`

*Format*

A tibble with 1,316 rows and 3 variables:

*Type:* integer

*Description:* Patient ID

**who**

*Type:* factor (First/Reference level = admission)

*Description:*  Indicator for start or stop of detox

*Levels:* admission, discharge

**what**
when

| everybody | Everybody with any data |

**Description**
This is a list of each person with their original study project.

**Usage**
`data(everybody)`

**Format**
A tibble with 3,560 rows and 2 variables:

- **Type:** integer
  **Description:** Patient ID

who

| project | Fagerstrom Test for Nicotine Dependence |

**Description**
Information on the intensity of physical addiction to nicotine at baseline. See the vignette [Harmonization Information](#) for more details.
Usage

fagerstrom

Format

A tibble with 3,119 rows and 4 variables:

Type: integer

Description: Patient ID

who

Type: factor

Description: No = Is not a smoker, Yes = Is a smoker

Levels: No, Yes

is_smoker

Type: factor

Description: Fagerstrom Test For Nicotine Dependence Score 0-10

Levels: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

ftnd

Type: factor

Description: Cigarettes per day

Levels: , 10 OR LESS, 11-20, 21-30, 31 OR MORE

per_day

<table>
<thead>
<tr>
<th>first_survey</th>
<th>First Survey Date</th>
</tr>
</thead>
</table>
Description

This file contains the dates for the demographics survey in CTN-0027 and CTN-0030 and the medical-psychiatric history for CTN-0051.

Usage

data(first_survey)

Format

A tibble with 3,453 rows and 2 variables:

Type: integer
Description: Patient ID

who

Type: integer
Description: Study day

when

<table>
<thead>
<tr>
<th>meta_study_length</th>
<th>Metadata About Study Length</th>
</tr>
</thead>
</table>

Description

This is information on the duration of CTN-0094 trial periods.

Usage

data(meta_study_length)

Format

A tibble with 16 rows and 6 variables:

Type: factor (First/Reference level = 27)
Description: CTN project number
Levels: 27, 30, 51
project

Type: factor (First/Reference level = Inpatient BUP)

Description: ctn_0027: Methadone, Outpatient BUP
ctn_0030: Outpatient BUP + EMM
ctn_0051: Outpatient BUP + EMM

Levels: Inpatient BUP, Inpatient NR-NTX, Methadone, Outpatient BUP, Outpatient BUP + EMM, Outpatient BUP + SMM

treatment

Type: factor (First/Reference level = 1)

Description: Study phase (needed because of CTN 30)

Levels: 1, 2

phase

Type: factor (First/Reference level = 1)

Description: Treatment stage used with description to capture different treatment phase/stage/period

Levels: 1, 2, 3

stage

Type: integer

Description: Treatment duration in weeks

weeks

Type: factor (First/Reference level = Buprenorphine-naloxone stabilization)

Description: Description of the treatment duration

Levels: Buprenorphine-naloxone stabilization, Buprenorphine-naloxone treatment, Post-medication follow-up, Taper, Treatment
Description

A table of the substances assessed in three clinical trials via urine drug screen (UDS) and their groupings.

Usage

data(meta_substance_groups_uds)

Format

A tibble with 13 rows and 4 variables:

Type: factor (First/Reference level = Alcohol)

Description: substances screened by UDS per study; names drawn from trial case report forms

Levels: Alcohol, Amphetamine, Barbiturate, Benzodiazepine, Buprenorphine, Cannabinoids, Cocaine, Methadone, Methamphetamine, Opiate 2000, Opiate 30, Oxycodone, Propoxyphene

Substance

Type: factor (First/Reference level = Alcohol)

Description: Derived label used to group substances of interest, or "NO" if the substance was not screened for in ctn_0027

Levels: Alcohol, Amphetamine, Benzodiazepine, Cocaine, Methadone, NO, Opioid, THC

CTN-0027

Type: factor (First/Reference level = Amphetamine)

Description: Derived label used to group substances of interest, or "NO" if the substance was not screened for in ctn_0030

Levels: Amphetamine, Benzodiazepine, Buprenorphine, Cocaine, Methadone, NO, Opioid, THC

CTN-0030

Type: factor (First/Reference level = Amphetamine)
CTN-0051

Details

This table indicates which substances were screened in each trial. The first column (substance) is drawn from labels which appear in case report forms for the three clinical trials. The remaining three columns hold "NO" if a substance was not screened in that trial, or a grouping label indicating what type of drug was screened. The ungrouped data can be found in all_drugs.

For example, "Opiate 300 ng" and "Oxycodone" are assigned to the grouping label "Opioid", and they were assessed in each clinical trial (so none of the rows show "NO"). In contrast, while "Opiate 2000 ng" is also assigned to the grouping label "Opioid", it was neither assessed in ctn_0027 nor ctn_0030; thus, the grouping label is "NO" for these trials. For more details, see the Harmonization Information vignette.

Description

This is self-reported pain from the SF-36 (ctn_0027 and ctn_0030) and EuroQoL (ctn_0051). See the Harmonization Information vignette for more details.

Usage

data(pain)

Format

A tibble with 3,082 rows and 3 variables:

Type: integer

Description: Patient ID

who

Type: factor (First/Reference level = Missing)

Description: Pain severity
Levels: Missing, No Pain, Severe Pain, Very mild to Moderate Pain

pain

Type: integer
Description: Study day

when

Description
Information on psychiatric symptoms and diagnoses. The same constructs were measured using different instruments. For example, the Addition Severity Index (ASI) asks "Have you had a significant period of time (that was not a direct result of drug/alcohol use) in which you have experienced hallucinations - saw things or heard voices that other people did not hear or see?" and the medical and psychiatric history evaluation asks about schizophrenia. The definitions of substance abuse have changed in the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders. Also see the Harmonization Information vignette for more details.

Usage
data(psychiatric)

Format
A tibble with 3,560 rows and 16 variables:

Type: integer
Description: Patient ID

who

Type: factor (First/Reference level = No)
Description: Medical and psychiatric history interview indicates schizophrenia
Levels: No, Yes
has_schizophrenia
Type: factor (First/Reference level = No)
Description: Medical and psychiatric history interview indicates major depression
Levels: No, Yes

has_major_dep
Type: factor (First/Reference level = No)
Description: Medical and psychiatric history interview indicates bipolar disorder
Levels: No, Yes

has_bipolar
Type: factor (First/Reference level = No)
Description: Medical and psychiatric history interview indicates anxiety panic disorder
Levels: No, Yes

has_anx_pan
Type: factor (First/Reference level = No)
Description: Medical and psychiatric history interview indicates brain damage
Levels: No, Yes

has_brain_damage
Type: factor (First/Reference level = No)
Description: Medical and psychiatric history interview indicates epilepsy
Levels: No, Yes

has_epilepsy
psychiatric

**depression**

*Type:* factor (First/Reference level = Yes)

*Description:* Addiction Severity Index-Lite Follow-up depression: P4 = "Experienced serious depression-sadness, hopelessness, loss of interest, difficulty with daily function?" See the Details section below

*Levels:* Yes, No, Not answered, Missing,

**anxiety**

*Type:* factor (First/Reference level = Yes)

*Description:* Addiction Severity Index-Lite Follow-up anxiety: P5 = "Experienced serious anxiety/tension, uptight, unreasonably worried, inability to feel relaxed?" See the Details section below

*Levels:* Yes, No, Not answered, Missing,

**schizophrenia**

*Type:* factor (First/Reference level = No)

*Description:* DSM-4 opioid abuse or dependence diagnosis or DSM-5 "opioid use disorder" (OUD) diagnosis

*Levels:* No, Yes

**has_opiates_dx**

*Type:* factor (First/Reference level = No)

*Description:* DSM-4 alcohol abuse or dependence diagnosis or DSM-5 "alcohol use disorder" (AUD) diagnosis

*Levels:* No, Yes

**has_alcol_dx**

*Type:* factor (First/Reference level = No)
Description: DSM-4 amphetamine and similar sympathomimetics abuse or dependence diagnosis or DSM-5 amphetamine-type substance use disorder

Levels: No, Yes

has_amphetamines_dx

Type: factor (First/Reference level = No)

Description: DSM-4 cannabis abuse or dependence diagnosis or DSM-5 cannabis use disorder

Levels: No, Yes

has_cannabis_dx

Type: factor (First/Reference level = No)

Description: DSM-4 cocaine abuse or dependence diagnosis or DSM-5 cocaine use disorder

Levels: No, Yes

has_cocaine_dx

Type: factor (First/Reference level = No)

Description: DSM-4 sedatives abuse or dependence diagnosis or DSM-5 sedative hypnotic/anxiolytic use disorder

Levels: No, Yes

has_sedatives_dx

Details

Note that the data from the ASI (depression, anxiety, and schizophrenia) contains four levels plus NA. People who did not take the ASI are coded with NA. People who took the ASI but are completely missing an answer to a question are coded as "Missing". Others who are known to not answer (i.e., refused to answer) are coded with "Not Answered".

Source


**Description**

This is quality of life data from the PhenX Quality of Life survey (see [https://www.phenxtoolkit.org/protocols/view/221302](https://www.phenxtoolkit.org/protocols/view/221302)). This was used by the Clinical Trials Network protocol CTN-0051.

**Usage**

data(qol)

**Format**

A tibble with 657 rows and 2 variables:

- **Type:** integer
  - **Description:** Patient ID

- **who**
  - **Type:** factor (First/Reference level = No)
  - **Description:** Are you currently homeless or living in a shelter?
  - **Levels:** No, Yes

- **is_homeless**

**randomization**

**Description**

This is the information on the treatment group. Note that CTN30 had two randomization events.

**Usage**

data(randomization)
Format

A tibble with 4,691 rows and 4 variables:

- **Type**: integer
  
  **Description**: Patient ID

- **who**
  
  **Type**: factor (First/Reference level = Inpatient BUP)
  
  **Description**: What treatment is prescribed?
  
  **Levels**: Inpatient BUP, Inpatient NR-NTX, Methadone, Outpatient BUP, Outpatient BUP + EMM, Outpatient BUP + SMM

- **treatment**
  
  **Type**: factor (First/Reference level = 1)
  
  **Description**: Indicator of which randomization. Needed because CTN 30 has two randomization dates.
  
  **Levels**: 1, 2

- **which**
  
  **Type**: integer
  
  **Description**: Study day

- **when**

---

**rbs**  

**Risk Behavior Survey**

---

**Description**

This is the drug use information on from the RBS. Questions ask, "How many days did you use ___ in the last 30 days?" Days were categorized in ctn_0051. See the Harmonization Information vignette for more details.
Usage
data(rbs)

Format
A tibble with 15,410 rows and 4 variables:

who
Type: factor (First/Reference level = cocaine)
Description: What drug was used: "cocaine" = Cocaine by itself "heroin" = Heroin by itself "opioid" = Have you ever used Other Opiates "speed" = Have you ever used Amphetamines (Speed, Methamphetamine, Crank) "speedball" = Heroin and cocaine mixed
Levels: cocaine, heroin, opioid, speed, speedball

what
Type: factor (First/Reference level = No)
Description: Is there a self-reported history of use?
Levels: No, Yes

did_use
Type: integer
Description: Number of days out of 30 that the drug was used

days

rbs_iv
Risk Behavior Survey IV drug use information

Description
This is aggregated data in IV drug use. See the Harmonization Information vignette for more details.
Usage
data(rbs_iv)

Format
A tibble with 3,560 rows and 10 variables:

- **who**
  - Type: integer
  - Description: Patient ID

- **days**
  - Type: integer
  - Description: Maximum number of days of IV drug use across all injected drug

- **max**
  - Type: integer
  - Description: Indicator of total IV drug exposure for the most used IV drug

- **amount**
  - Type: factor (First/Reference level = No)
  - Description: Did you share needles in the last 30 days?
  - Levels: No, Yes

- **shared**
  - Type: integer
**screening_date**

*Description:* Number of days out of last 30 when cocaine was injected

**cocaine_inject_days**

*Type:* integer

*Description:* Number of days out of last 30 when heroin was injected

**heroin_inject_days**

*Type:* integer

*Description:* Number of days out of last 30 when speedball was injected

**speedball_inject_days**

*Type:* integer

*Description:* Number of days out of last 30 when opioid was injected

**opioid_inject_days**

*Type:* integer

*Description:* Number of days out of last 30 when speed was injected

**speed_inject_days**

---

<table>
<thead>
<tr>
<th>screening_date</th>
<th>Screening Date Information</th>
</tr>
</thead>
</table>

**Description**

The information on the screening data and baseline drug screening data is complex and inconsistent across studies. This file has information on dates around the screening visit and baseline timeline follow back assessments.

**Usage**

`data(screening_date)`
sex

**Format**

A tibble with 3,430 rows and 3 variables:

- **Type:** integer
- **Description:** Patient ID

**who**

- **Type:** integer
- **Description:** The day of screening

**screening_day**

- **Type:** integer
- **Description:** Best guess at the end of TLFB

**day_zero**

---

**sex**  
*Sexual Activity in Risk Behavior Survey*

**Description**

Information on sexual activities from the (RBS)

**Usage**

data(sex)

**Format**

A tibble with 3,299 rows and 40 variables:

- **Type:** integer
- **Description:** Patient ID
who

Type: factor (First/Reference level = 0)

Description: Total sex partners: 0 = 0 1 = 1 2 = more than one

Levels: 0, 1, 2

txx_prt

Type: factor (First/Reference level = 0)

Description: Total male sex partners: 0 = 0 1 = 1 2 = more than one

Levels: 0, 1, 2

txx_mprr

Type: factor (First/Reference level = 0)

Description: Total female sex partners: 0 = 0 1 = 1 2 = more than one

Levels: 0, 1, 2

txx_fprr

Type: factor (First/Reference level = 0)

Description: Men who have sex with women (MSW), sex partners: 0 = 0 1 = 1 2 = more than one

Levels: 0, 1, 2

msw_npt

Type: integer

Description: MSW count of sex

msw_freq

Type: integer
Description: MSW count of protected (with condom use) sex

**msw_pxx**

*Type:* integer

Description: MSW count of unprotected sex

**msw_uxx**

*Type:* factor (First/Reference level = 0)

Description: Men have sex with men (MSM), sex partners: 0 = 0 1 = 1 2 = more than one

*Levels:* 0, 1, 2

**msm_npt**

*Type:* integer

Description: MSM count of sex

**msm_frq**

*Type:* integer

Description: MSM count of protected (with condom use) sex

**msm_pxx**

*Type:* integer

Description: MSM count of unprotected sex

**msm_uxx**

*Type:* factor (First/Reference level = 0)

Description: Women have sex with men (WSM), sex partners: 0 = 0 1 = 1 2 = more than one

*Levels:* 0, 1, 2
**wsm_npt**

*Type:* integer

*Description:* WSM count of sex

**wsm_frq**

*Type:* integer

*Description:* WSM count of protected (with condom use) sex

**wsm_pxx**

*Type:* integer

*Description:* WSM count of unprotected sex

**wsm_uxx**

*Type:* integer

*Description:* TOTAL count of sex

**txx_frq**

*Type:* integer

*Description:* TOTAL count of protected (with condom use) sex

**txx_pxx**

*Type:* integer

*Description:* TOTAL count of unprotected sex

**txx_uxx**

*Type:* integer
<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mvw_frq</td>
<td>integer</td>
<td>Men vaginal sex with women count of sex</td>
</tr>
<tr>
<td>maw_frq</td>
<td>integer</td>
<td>Men anal sex with women count of sex</td>
</tr>
<tr>
<td>mim_frq</td>
<td>integer</td>
<td>Men insertive sex with men count of sex</td>
</tr>
<tr>
<td>mrm_frq</td>
<td>integer</td>
<td>Men receptive sex with men count of sex</td>
</tr>
<tr>
<td>wvm_frq</td>
<td>integer</td>
<td>Women vaginal sex with men count of sex</td>
</tr>
<tr>
<td>wam_frq</td>
<td>integer</td>
<td>Women anal sex with men count of sex</td>
</tr>
<tr>
<td>wam_frq</td>
<td>integer</td>
<td>Men vaginal sex with women count of protected sex</td>
</tr>
</tbody>
</table>
**mvw_pxx**

*Type:* integer

*Description:* Men anal sex with women count of protected sex

**maw_pxx**

*Type:* integer

*Description:* Men insertive sex with men count of protected sex

**mim_pxx**

*Type:* integer

*Description:* Men receptive sex with men count of protected sex

**mrm_pxx**

*Type:* integer

*Description:* Women vaginal sex with men count of protected sex

**wvm_pxx**

*Type:* integer

*Description:* Women anal sex with men count of protected sex

**wam_pxx**

*Type:* integer

*Description:* Men vaginal sex with women count of unprotected sex

**mvw_uxx**

*Type:* integer

*Description:* Men anal sex with women count of unprotected sex
**maw_uxx**

*Type:* integer

*Description:* Men insertive sex with men count of unprotected sex

**mim_uxx**

*Type:* integer

*Description:* Men receptive sex with men count of unprotected sex

**mrm_uxx**

*Type:* integer

*Description:* Women vaginal sex with men count of unprotected sex

**wvm_uxx**

*Type:* integer

*Description:* Women anal sex with men count of unprotected sex

**wam_uxx**

*Type:* integer

*Description:* Total count of sex partners

**t_p**

*Type:* integer

*Description:* Total count of female sex partners

**t_fp**

*Type:* integer
**Description:** Total count of male sex partners

**t_mp**

**site_masked**

### Site Regrouped

**Description**

This is study site (clinic/research site number) information. Large study sites were split into smaller groups and small study sites were grouped into larger sites. The regrouped study sites are each approximately 100 subjects. This was done to protect the anonymity of the study sites.

**Usage**

data(site_masked)

**Format**

A tibble with 3,560 rows and 2 variables:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>integer</td>
<td>Patient ID</td>
</tr>
</tbody>
</table>

**who**

**Type:** factor (First/Reference level = 2)

**Description:** Study Site Regrouped

**Levels:** 2, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36

**site_masked**

---

**tlfb**

**Timeline Followback (TLFB) Drug Use Information**

---
Description

This is self-reported drug use. The values are the result of extensive processing of free text as well as structured values. These substances are group grouped as opioids: Codeine, Fentanyl, Hydromorphone, Merperidine, Oxycodone, Oxymorphone, Propoxyphene. **NOTE: Records where people self-reported the study drug after it was prescribed have been removed from this file.** The all_drugs dataset contains these nebulous records. See the vignette Harmonization Information for more information.

Usage

data(tlfb)

Format

A tibble with 237,778 rows and 3 variables:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>integer</td>
<td>Patient ID</td>
</tr>
</tbody>
</table>

who

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor</td>
<td>Text description of drugs; for more information, see Harmonization Information</td>
</tr>
</tbody>
</table>

| Levels     | Alcohol, Amphetamine, Analgesic, Antibiotic, Antidepressant, Antiemetic, Antihistamine, Antipsychotic, Ben

what

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>integer</td>
<td>Study Day</td>
</tr>
</tbody>
</table>

when

<table>
<thead>
<tr>
<th>treatment</th>
<th>Amount of Study Drug Per Day</th>
</tr>
</thead>
</table>

Description

The doses (usually in milligrams) of the study drug administered to each subject by day. When the study drug is listed as an injection, then the amount is recorded as 1.
Usage
data(treatment)

Format
A tibble with 216,242 rows and 3 variables:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>integer</td>
<td>Patient ID</td>
</tr>
</tbody>
</table>

who

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>integer</td>
<td>The amount of drugs received on a day. Value is 1 for injections and mg otherwise</td>
</tr>
</tbody>
</table>

amount

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>integer</td>
<td>Study day</td>
</tr>
</tbody>
</table>

when

uds  Urine Drug Screening (UDS) Results

Description
UDS findings. Note that oxycodone and propoxyphene are grouped into the opioid category. ctn_0051 did not specifically screen for propoxyphene. For details on substances screened, see the vignette Harmonization Information.

Usage
data(uds)

Format
A tibble with 42,906 rows and 3 variables:
uds_temp

Type: integer
Description: Patient ID

who

Type: factor (First/Reference level = Alcohol)
Description: Name of drug identified
Levels: Alcohol, Amphetamine, Benzodiazepine, Buprenorphine, Cocaine, Mdma/Hallucinogen, Methadone, Opioid, Sedatives, Thc

what

Type: integer
Description: Study day

when

uds_temp

Urine Drug Screening (UDS) Temperature

Description
This is information on whether the urine temperature was in the acceptable range.

Usage
data(uds_temp)

Format
A tibble with 36,680 rows and 3 variables:

Type: integer
Description: Patient ID

who

Type: factor (First/Reference level = 0)
**Visit**

Description: was the temperature $92°F \leq X \leq 96°F \text{ OR } 33.3°C \leq X \leq 35.5°C$? $0 = "No", 1 = "Yes", 99 = "Unknown"

Levels: 0, 1, 99

was_temp_ok

Type: integer

Description: Study day

when

---

**visit**

*Patient Visit Data*

**Description**

This contains planned visits. Not all appointments were kept. Indicator variables show reasons for a missed appointment (if known). This data is not simple. There are more than 1300+ dates duplicated with different visit types. There are week 24 visits that happen at the wrong time (e.g., the date of week 24 is the same as week 8). There are also many cases where two adjacent visits happen on the same day (e.g., both week 7 and week 8 have their visits on the same day). There are both visit and no visit reports on the same day for some people. There are nearly 850 “Cross Active Study” visits and all but 43 happen on the same date as another visit. Many variables have a “1” indicating "Yes". Other are NA because we don’t know if those values are real "No" or actually "Unknown". **Proceed with great caution.**

<table>
<thead>
<tr>
<th>Visit Type</th>
<th>ctn_0027</th>
<th>ctn_0030</th>
<th>ctn_0051</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;BASELINE&quot;</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>&quot;WK__&quot;</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>&quot;Cross Active Study&quot;</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>&quot;Cross Active Study&quot;</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>&quot;P1____&quot;</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>&quot;P2____&quot;</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>&quot;EOT&quot;</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>&quot;M1F&quot;</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>&quot;M3F&quot;</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

**Usage**

data(visit)
Format

A tibble with 53,899 rows and 19 variables:

Type: integer
Description: Patient ID

who
Type: factor (First/Reference level = BASELINE)
Description: Indicator of the visit type. For example, "WK12" is week 12 of ctn_0027 or "P2Wk13" is week 13 of phase 2 of ctn_0030.
Levels: BASELINE, Cross Active Study, EOT, M1F, M3F, P1Finl, P1Unsc, P1Wk10, P1Wk12, P1Wk1A, P1Wk1B, P1Wk2, P1Wk3, P1Wk4, P1Wk6, ... WK14, WK15, WK16, WK17, WK18, WK19, WK2, WK20, WK21, WK22, WK23, WK24, WK28, WK3, WK32, WK4, WK5, WK6, WK7, WK8, WK9

visit
Type: factor (First/Reference level = )
Description: Disposition of appointment - visit, no visit or MISSING. Note there are records with no disposition.
Levels: , final, MISSING, no visit, visit

what
Type: factor (First/Reference level = 1)
Description: UNKNOWN
Levels: 1

is_no_note
Type: factor (First/Reference level = 1)
Description: Patient died
Levels: 1

is_dead
visit

Type: factor (First/Reference level = 1)
Description: Patient did not show with no explanation
Levels: 1

is_no_show

Type: factor (First/Reference level = 1)
Description: Patient could not afford to get to appointment
Levels: 1

is_no_funding

Type: factor (First/Reference level = 1)
Description: Patient withdrew from the study
Levels: 1

is_left_study

Type: factor (First/Reference level = 1)
Description: Patient is incarcerated
Levels: 1

is_in_jail

Type: factor (First/Reference level = 1)
Description: Patient forgot appointment
Levels: 1

is_forgot

Type: factor (First/Reference level = 1)
**Description:** Patient hospitalized during appointment

**Levels:** 1

**is_in_hospital**

**Type:** factor (First/Reference level = 1)

**Description:** Patient reported being too sick to attend

**Levels:** 1

**is_illness**

**Type:** factor (First/Reference level = 1)

**Description:** Patient moved from the study area

**Levels:** 1

**is_moved**

**Type:** factor (First/Reference level = 1)

**Description:** Patient dropped for non-compliance

**Levels:** 1

**is_missing_14_consecutive**

**Type:** factor (First/Reference level = 1)

**Description:** UNKNOWN

**Levels:** 1

**is_window**

**Type:** factor (First/Reference level = 1)

**Description:** Patient reports being unable to attend
Levels: 1

is_unable

Type: factor (First/Reference level = 1)

Description: Patient reports being on vacation

Levels: 1

is_on_vacation

Type: factor (First/Reference level = 1)

Description: Other reason given

Levels: 1

is_other

Type: integer

Description: Study day

when

| withdrawal | Patient Withdrawal Symptoms Per Day |

Description

CTN 27 and 30 use the Clinical Opiate Withdrawal Scale (COWS). CTN 51 uses SOWS. See harmonization vignette for more details.

Usage

data(withdrawal)

Format

A tibble with 14,983 rows and 3 variables:
withdrawal_pre_post

Type: integer

Description: Patient ID

who

Type: factor (First/Reference level = 0)

Description: 0 = "None" 1 = "mild" 2 = "moderate" 3 = "severe"

Levels: 0, 1, 2, 3

withdrawal

Type: integer

Description: Study day

when

Source


withdrawal_pre_post Patient Withdrawal Symptoms Pre and Post Induction

Description

This is a information on the severity of withdrawal symptoms.

Usage

data(withdrawal_pre_post)

Format

A tibble with 4,805 rows and 4 variables:
**who**

*Type:* factor (First/Reference level = post)

*Description:* Indicator of induction day type: "pre" or "post"

*Levels:* post, pre

**what**

*Type:* factor (First/Reference level = 0)

*Description:* 0 = "None" 1 = "mild" 2 = "moderate" 3 = "severe"

*Levels:* 0, 1, 2, 3

**withdrawal**

*Type:* integer

*Description:* Day of assessment

**when**
Index

* datasets
  all_drugs, 2
  asi, 4
  demographics, 5
  detox, 7
  everybody, 8
  fagerstrom, 8
  first_survey, 9
  meta_study_length, 10
  meta_substance_groups_uds, 12
  pain, 13
  psychiatric, 14
  qol, 18
  randomization, 18
  rbs, 19
  rbs_iv, 20
  screening_date, 22
  sex, 23
  site_masked, 30
  tlfb, 30
  treatment, 31
  uds, 32
  uds_temp, 33
  visit, 34
  withdrawal, 38
  withdrawal_pre_post, 39

all_drugs, 2, 13
asi, 4

demographics, 5
detox, 7
everybody, 8
fagerstrom, 8
first_survey, 9

meta_study_length, 10
meta_substance_groups_uds, 12

pain, 13
psychiatric, 14
qol, 18
randomization, 18
rbs, 19
rbs_iv, 20
screening_date, 22
sex, 23
site_masked, 30
tlfb, 30
treatment, 31
uds, 32
uds_temp, 33
visit, 34
withdrawal, 38
withdrawal_pre_post, 39