

A Gentle Introduction to Data Science Tools for Drug Dependence Research: Use Cases With Examples

Raymond Balise and Gabriel Odom, with Daniel Feaster
Jun 20, 2023

00 - Open Source Computing and R

Open and Reproducible Data Science using R

What could possibly go wrong?

- Your SAS license expired.
- A person needs to be dropped from a study ... please change every number in a paper.
- There is a person who is 1973 years old in the needle exchange database. (Does heroin causes immortality?)
- On revise and resubmit ... reviewer 2 wants you to add their favorite method or describe exactly what was done to process the data.
- What formula does SPSS use to calculate standard errors for mixed effects models?
- Excel is not your friend.

We want tools that ...

- are **free**
- are **open** to explore/expand
- are **trustworthy** (if it is good enough for the FDA and big pharma....)
- facilitate **reproducible** workflows
 - with manuscripts and analysis that are **knit** together
 - in papers with tables and figures which **fix themselves** when the data changes

R or Python

"Fast, Cheap or Good? Pick Two."

– Unknown

We are sure R and Python are cheap.

About R vs. Python ...

- Python is the second best language for everything.
- They both do **any analysis** you can think of.
- They both support **reproducible research**.
- The R community is **friendly**, patient, and community focused.
 - Posit (formerly R Studio) is a B corp.
 - <https://www.datatrail.org/>
- Based on our experience, R is faster to **set up and learn** if you are starting from zero.
- R is the second most referenced language behind SPSS, which was released in 1968.

"R is the most user-hostile programming language money can't buy."

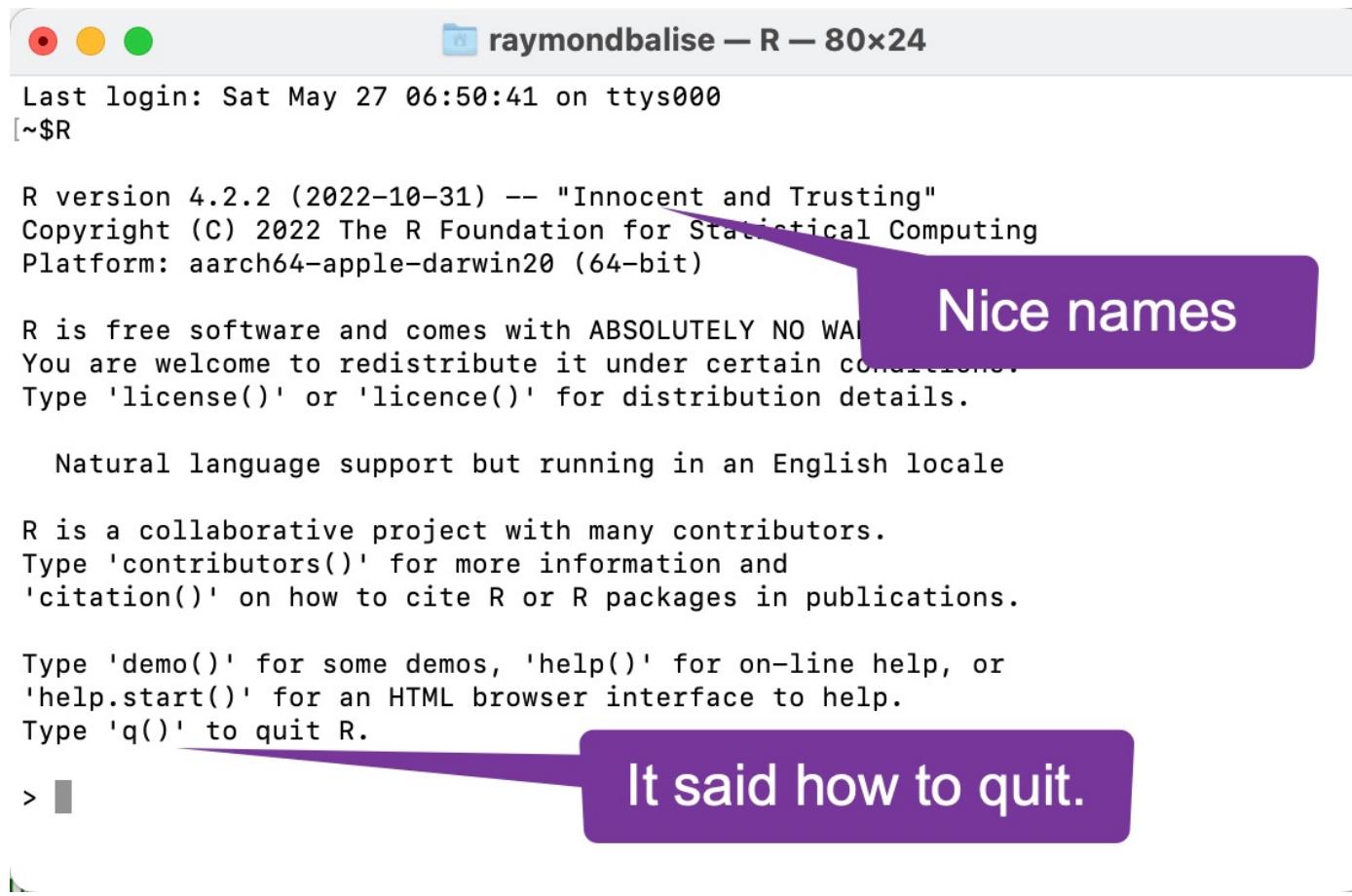
- Balise circa 2010

"I love R."

- Balise circa 2023

What changed?

- In 2010 there were two good things about R.



Last login: Sat May 27 06:50:41 on ttys000
[~\$R

R version 4.2.2 (2022-10-31) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: aarch64-apple-darwin20 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> []

Nice names

It said how to quit.

What changed?

The screenshot shows the RStudio interface with a presentation slide titled "Deidentified Data from CTN-0094" and an R console window.

Top Bar: Menus, Session, Build, Debug, Profile, Tools, Window, Help.

Left Panel (Presentation):

- Introduction:** Analyses were conducted with `r stringr::word(R.Version()$version.string, 1, 3)` with the `tidyverse` (`r packageVersion("tidyverse")`), `rUM` (`r packageVersion("rUM")`), `table1` (`r packageVersion("table1")`) packages used to preprocess and summarize data. [`@R-base`; `@R-tidyverse`; `@tidyverse2019`; `@R-rUM`; `@R-table1`]
- Methods:** Analyses were conducted with `r stringr::word(R.Version()$version.string, 1, 3)` with the `tidyverse` (`r packageVersion("tidyverse")`), `rUM` (`r packageVersion("rUM")`), `table1` (`r packageVersion("table1")`) packages used to preprocess and summarize data. [`@R-base`; `@R-tidyverse`; `@tidyverse2019`; `@R-rUM`; `@R-table1`]
- Results:** R version 4.2.2 (2022-10-31) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin17.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for more information.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Console Output:

```
> victor <- c(1, 2, 3)
> |
```

Environment Tab: We can see what we made.

Help Tab: Help is a click away.

Documentation for package 'public.ctn0094data' version 0.9.0:

- [DESCRIPTION file](#).
- [User guides](#), [package vignettes](#) and other documentation.

Help Pages:

all_drugs	All drugs taken
asi	Did patient use intravenous drugs
demographics	Patient demographics
detox	Start and Stop of Detox
everybody	Everybody with any data
fagerstrom	Fagerstrom Test for Nicotine Dependence
first_survey	First Survey Date
meta_study_length	Metadata About Study Length
meta_substance_groups_uds	Metadata About UDS Groupings
pain	Self-Reported Pain

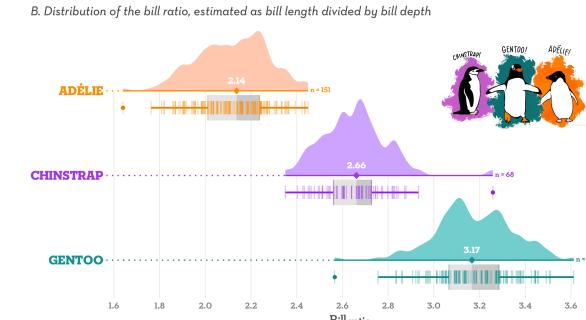
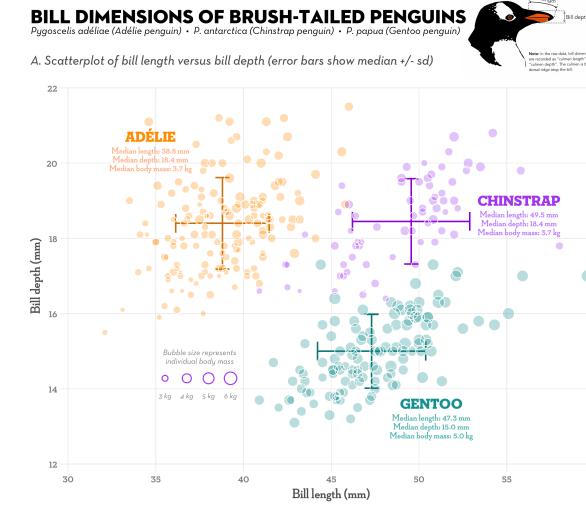
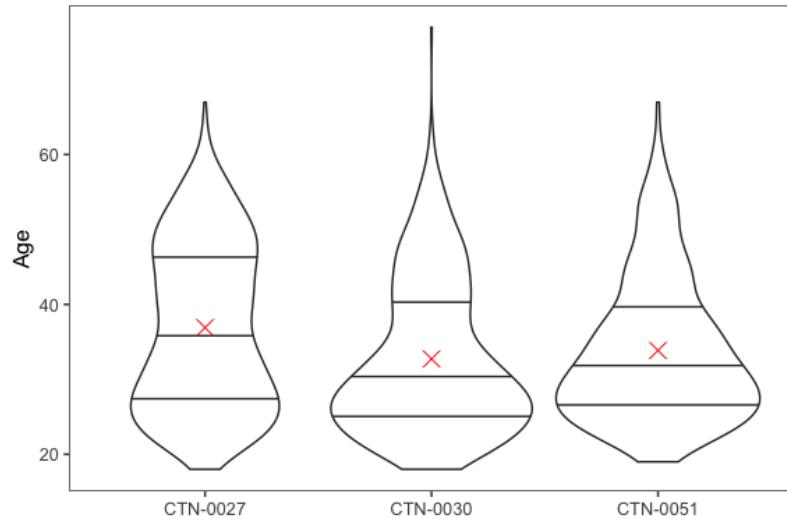
R is Modular



dplyr does data processing with verbs

- Verbs to work on a table of data:
 - `filter()` - pick people
 - `arrange()` - sort/order the people
 - `summarize()` - describe people in groups
 - `select()` - pick a measurement/feature/variable
 - `mutate()` - make a new feature/variable
 - ... lots more
- Verbs to combine data from multiple tables:
 - `inner_join()` - find common people in two datasets
 - ... a bunch more

ggplot2 can make any graphic



Note: In the original data, bill dimensions are recorded as "culmen length" and "culmen depth". The culmen is the dorsal (upper) ridge of a bird's bill.
Visualization: Cedric Scherer • Data: Gorman, Williams & Fraser (2014) DOI: 10.1371/journal.pone.0090081 • Illustrations: Allison Horst

R Plays Well with Others

- R can easily have a **conversation with a REDCap database**.
- R can **read** any kind of data. It has pipes in place to easily read from:
 - Excel
 - SPSS/SAS/Stata/Minitab
 - The US Census
 - CDC Wonder/Epi Info
 - PubMed
 - ClinicalCode.org which has details on ICD-9/10 codes
 - SurveyMonkey and/or Qualtrix
 - ... so much more

01 - Introduction to REDCap

The Plan

- What is REDCap (and why bother)?
- How to build a **basic case report form**
- What is **rational** to do in REDCap
 - Multiple instruments
 - Longitudinal assessments
 - Repeated instruments
- Show the Miami **needle exchange** database (IDEA)
- Using **rUM** to make a **publication-ready paper**
- The **manual export** (and why it is suboptimal)
- **Importing all forms** from a REDCap project into R with **tidyREDCap**
- Using **tidyREDCap summary** functions to add to a paper

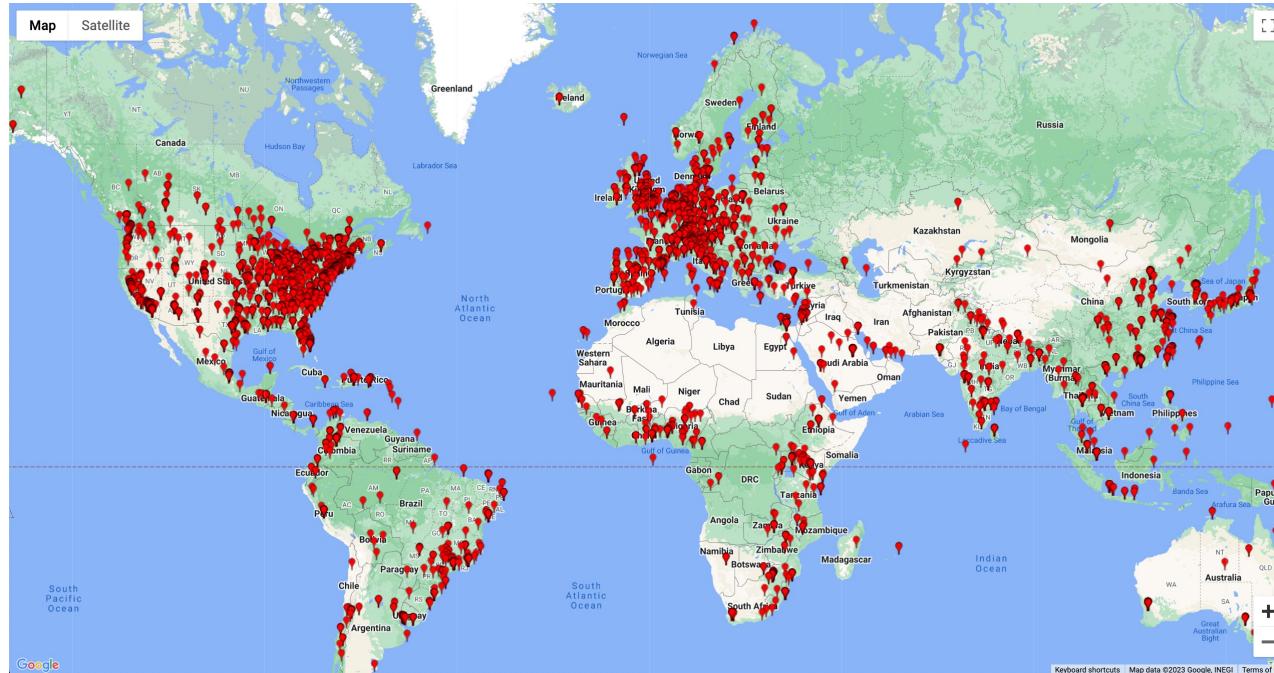
What is REDCap (and why bother)?

What's the big deal?

- Based at Vanderbilt University and supported by the National Institutes of Health (NIH)
- Not open-source (the license specifies noncommercial research use only)
- Continuous development since 2004
- Supports researcher- and participant-facing data entry
- Supports online and offline data collection

By the Numbers

- Free to consortium members (slightly limited by international export controls)
- Wildly popular
 - 6,000+ institutions in more than 150 countries
 - Acknowledged in more than 23,000 articles



Legal

- REDCap is 21 CFR Part 11-ready
- Built-in logging and audit trails
- Designed to support *safe harbor* deidentification

How do I access it?

- Most academic medical centers have it.
- Check the partner list: <https://projectredcap.org/partners/>.
- We will use the locked demonstration project.
 - Normally, you will use point-and-click tools.
 - Today, we will show you things and you will have read-only access to some data.

Workflows

Fundamental Steps

1. Create a project
2. Add/create a data entry form/instrument
3. Collect/enter data
4. View/summarize in REDCap
5. Export for additional magic

But wait... there's more!

- Multi-site support with Data Access Groups
- Built-in, mutliple language support
- Randomization
- Longitudinal projects with study calendars and scheduling
- Repeated instruments for assessing unplanned repeats
 - Describe each "stop by" clinic visit
 - Describe each adverse event
- Public or private dashboards
- Send emails/text
- Extensible via *External Modules*

Create A Project 1

REDCap® Home My Projects + New Project Help & FAQ Training Videos Send-It Messenger Logged in as Profile Log out

+ Create a new REDCap Project

You may begin the creation of a new REDCap project on your own by completing the form below and clicking the Create Project button at the bottom. **Your project will not be created immediately**, but your request will be quickly reviewed by a REDCap administrator, after which you will be notified via email when the project has been created.

Project title: CPDD 2023 Demo

Project's purpose: Practice / Just for fun ▾
How will it be used?

Assign project to a Project Folder?

Project notes (optional):
Description of the project's use or purpose (displayed on the My Projects page)

Project creation option:

Empty project (blank slate)
 Upload a REDCap project XML file (CDISC ODM format) ?
 Use a template (choose one below)

★ Choose a project template

select template	Template title (sorted by title)	Template description
<input type="radio"/>	21 CFR Part 11 eConsent Template	21 CFR Part 11 Validated Template with eConsent Framework enabled. To be used for FDA-regulated studies.
<input type="radio"/>	Basic Demography	Contains a single data collection instrument to capture basic demographic information.
<input type="radio"/>	CTSI - eConsent Template	IRB Approved Consent form Template with eConsent Framework enabled for Non-FDA regulated projects.
<input type="radio"/>	Human Cancer Tissue Biobank	Contains five data entry forms for collecting and tracking information for cancer tissue.
<input type="radio"/>	Longitudinal Database (1 arm)	Contains nine data entry forms (beginning with a demography form) for collecting data longitudinally over eight different events.
<input type="radio"/>	Longitudinal Database (2 arms)	Contains nine data entry forms (beginning with a demography form) for collecting data on

Send Request **Cancel**

REDCap 13.1.29 - © 2023 Vanderbilt University

Create A Project 2

- Use the **Designer** to add your content.

The screenshot shows the REDCap Project Setup interface for the 'CPDD 2023 Demo' project (PID 7421). The left sidebar includes links for My Projects, REDCap Messenger, Contact REDCap administrator, Project Home and Design, Data Collection, Applications, External Modules, Help & Information, and Contact REDCap administrator. The main content area displays project status (Development, 0 completed steps), Main project settings (with options for surveys, longitudinal data collection, and MyCap app), Design your data collection instruments (with links to Online Designer, Data Dictionary, and REDCap Instrument Library), Enable optional modules and customizations (with options for Repeating instruments, Auto-numbering for records, Scheduling module, Randomization module, Designate an email field, and SendGrid Template email services), and Set up project bookmarks (optional). A purple arrow points from the left sidebar to the 'Designer' link in the Project Home and Design section.

Add/Create a Data Entry Instrument/Form

- Add an instrument from the REDCap **Shared Instrument Library** (at Vanderbilt).
- Make **your own** data entry form.
- Do both.

The screenshot shows the REDCap interface with a purple dashed arrow pointing from the top-left towards the 'Online Designer' tab. The 'Online Designer' tab is highlighted in blue. The main content area displays the 'Data Collection Instruments' section. At the top of this section, there are three buttons: '+ Create' (highlighted with a red circle), '+ Import' (highlighted with a red circle), and '+ Upload'. Below these buttons is a table with one row, showing 'My First Instrument' in the 'Instrument name' column, '1' in the 'Fields' column, and a 'View PDF' button. To the right of the table are 'Form options:' and a 'Form Display Logic' button.

REDCap Shared Instrument Library 1

- You can point and click to add one of the thousands of instruments in the REDCap library.

The screenshot shows the REDCap project navigation bar on the left and the main 'Online Designer' interface on the right. The 'Data Collection Instruments' section is highlighted with a purple dashed box and arrow. The 'Create' button is also highlighted with a purple arrow. The interface includes a note about real-time updates for development status and a 'Form options' section with a 'Form Display Logic' link. A table lists the instrument 'My First Instrument' with 1 field.

Project Home and Design

- Project Home · Project Setup
- Designer · Dictionary · Codebook
- Project status: Development

Data Collection

- Record Status Dashboard
- Add / Edit Records
- Show data collection instruments

Applications

- Project Dashboards
- Alerts & Notifications
- Multi-Language Management
- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool
- Data Comparison Tool
- Logging and Email Logging
- Field Comment Log
- File Repository

Project Home · Project Setup · Online Designer · Data Dictionary · Codebook

Create snapshot of instruments · VIDEO: How to use this page · Last snapshot: never [?]

The Online Designer will allow you to make project modifications to fields and data collection instruments very easily using only your web browser. NOTE: While in development status, all field changes will take effect immediately in real time.

Data Collection Instruments

+ Create a new instrument from scratch
Import a new instrument from the official [REDCap Instrument Library](#) ?
Upload instrument ZIP file from another project/user or [external libraries](#) ?

Form options:
[Form Display Logic](#)

Instrument name	Fields	View PDF	Instrument actions
My First Instrument	1	[View]	Choose action ▾

REDCap Shared Instrument Library 2

REDCap Shared Library

The REDCap Shared Library is a repository for REDCap data collection instruments and forms that can be downloaded and used by researchers at REDCap partner institutions. Curated instruments highlighted with a star ★ have been approved for inclusion by the REDCap Library Oversight Committee (REDLOC) after review for research relevance, accuracy in function and coding (see guidelines), and copyright issues. Other instruments and forms are shared by individuals or groups from consortium institutions on "as-is" basis.

You may search below for any available data collection instruments. If you got to this site directly, you will not be able to view the actual shared instruments themselves. This public view listing is for reference only and helps protect the authors' copyright. You will also not see instruments that have been shared locally by REDCap end users if they have not gone through the formal REDLOC curation process. If you arrived here from the REDCap application, you will have the options to view instruments as a webpage, view instruments as a PDF, and import the instruments directly into REDCap. If you wish, you may download a list of all library instruments in Excel/CSV format. If you download and utilize an instrument from the REDCap Shared Library, please cite the RSL manuscript. If you have questions or are experiencing issues, please contact redcap@vumc.org.

[Return to REDCap](#)

Keyword search: depression

Search the library

Logged in as Raymond Balise (University of Miami)

Search options:

- Language: All
- Type: show all
- Minimum downloads: 0
- Recent additions: show all
- Curated by REDLOC? show all

Shared Library

- Search
- Library Metrics**
- My Activity
- Institution Activity
- Consortium Activity
- REDLOC**
- Approval List
- Approval History
- Change Password
- Logout

1 2 3 >>

Found 49 results matching your search

Didn't find what you were looking for? [Suggest a validated instrument for library inclusion](#)

Title	Downloads
► Geriatric Depression Scale GDS Long Form ★	262
► Patient Health Questionnaire Depression Scale (PHQ) ★	728
► Patient Health Questionnaire Depression Scale (PHQ) Scored ★	986
► Geriatric Depression Scale GDS Short Form ★	582
▼ Hamilton Depression Rating Scale (HAM-D) ★	1183

Details:

Institution: REDLOC
Contact: Brenda Minor
Contact email: brenda.l.minor@vumc.org
Submitted by: Brenda Minor
Description: The HAM-D is designed to rate the severity of depression in patients. Although it contains 21 areas, calculate the patient's score on the first 17 answers.
<http://www.ncbi.nlm.nih.gov/pubmed/14399272>
<http://www.assessmentspsychology.com/HAM-D.pdf>

[View as web page](#)
[View as PDF](#)
NOTE: PDFs of non-English instruments may not render correctly here, but will render correctly in REDCap projects.

[Import into my REDCap project](#)

[Admin: View instrument stats](#)

REDCap Shared Instrument Library 3

The screenshot shows the REDCap Online Designer interface. The top navigation bar includes Project Home, Project Setup, Online Designer (selected), Data Dictionary, and Codebook. Below the navigation is a note about creating snapshots and a video link. The main content area is titled "Data Collection Instruments". It features three buttons: Create, Import, and Upload. A "Form options" section includes a "Form Display Logic" link. The table lists two instruments: "My First Instrument" (1 field) and "Hamilton Depression Rating Scale (HAM-D)" (26 fields). A context menu is open over the second row, showing options: Choose action (with Rename, Copy, Delete, and Download instrument ZIP), and a "Choose action" dropdown.

Instrument name	Fields	View PDF	Instrument actions
My First Instrument	1		
Hamilton Depression Rating Scale (HAM-D)	26		

Project Home and Design

- Project Home · Project Setup
- Designer · Dictionary · Codebook
- Project status: Development

Data Collection

- Record Status Dashboard - View data collection status of all records
- Add / Edit Records - Create new records or edit/view existing ones
- Show data collection instruments

Applications

- Project Dashboards
- Alerts & Notifications
- Multi-Language Management
- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool
- Data Comparison Tool
- Logging and Email Logging
- Field Comment Log
- File Repository

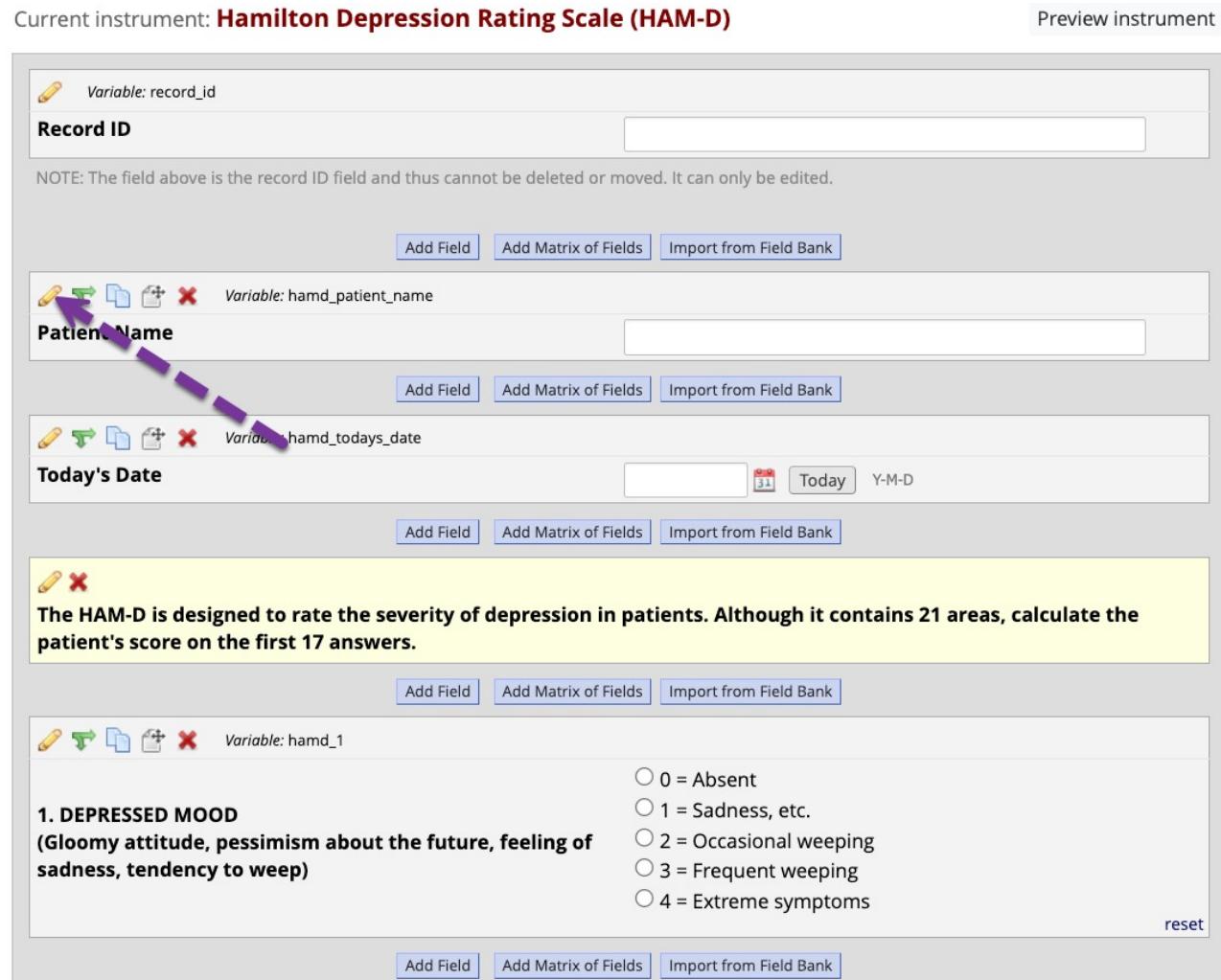
REDCap Shared Instrument Library 4

The screenshot shows the REDCap Online Designer interface. On the left, there is a sidebar with sections for Project Home and Design, Data Collection, and Applications. The Applications section is expanded, showing various tools like Project Dashboards, Alerts & Notifications, Multi-Language Management, Calendar, Data Exports, Reports, and Stats, Data Import Tool, Data Comparison Tool, Logging, Email Logging, and Field Comment Log. The main area has tabs for Project Home, Project Setup, Online Designer (which is active), Data Dictionary, and Codebook. Below the tabs, there is a button for "Create snapshot of instruments" and a link to "VIDEO: How to use this page". A note states: "The Online Designer will allow you to make project modifications to fields and data collection instruments very easily using only your web browser. NOTE: While in development status, all field changes will take effect immediately in real time." The "Data Collection Instruments" section contains buttons for Create, Import, and Upload, and a table listing instruments. The first instrument in the table is "Hamilton Depression Rating Scale (HAM-D)", which is highlighted with a red border and a purple arrow pointing to it from below. The table has columns for Instrument name, Fields, View PDF, and Instrument actions.

Instrument name	Fields	View PDF	Instrument actions
Hamilton Depression Rating Scale (HAM-D)	26		

REDCap Shared Instrument Library 5

Current instrument: **Hamilton Depression Rating Scale (HAM-D)** Preview instrument

 Variable: record_id
Record ID
NOTE: The field above is the record ID field and thus cannot be deleted or moved. It can only be edited.
[Add Field](#) [Add Matrix of Fields](#) [Import from Field Bank](#)
 Variable: hamd_patient_name
Patient Name
[Add Field](#) [Add Matrix of Fields](#) [Import from Field Bank](#)
 Variable: hamd_todays_date
Today's Date  Today Y-M-D
[Add Field](#) [Add Matrix of Fields](#) [Import from Field Bank](#)
 The HAM-D is designed to rate the severity of depression in patients. Although it contains 21 areas, calculate the patient's score on the first 17 answers.
[Add Field](#) [Add Matrix of Fields](#) [Import from Field Bank](#)
 Variable: hamd_1
1. DEPRESSED MOOD
(Gloomy attitude, pessimism about the future, feeling of sadness, tendency to weep)
 0 = Absent
 1 = Sadness, etc.
 2 = Occasional weeping
 3 = Frequent weeping
 4 = Extreme symptoms
reset
[Add Field](#) [Add Matrix of Fields](#) [Import from Field Bank](#)

REDCap Shared Instrument Library 6

The license for the instruments bans meaningful changes but you can change the appearance and set details:

Edit Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

Field Type: Use the Rich Text Editor [\[?\]](#)

Field Label Use the Rich Text Editor [\[?\]](#)

Action Tags / Field Annotation (optional)
Learn about [@ Action Tags](#) or [using Field Annotation](#)

Variable Name (utilized in logic, calcs, and exports) Enable auto naming of variable based upon its Field Label?
ONLY letters, numbers, and underscores

How to use [\[+\] Smart Variables](#) [\[+\] Piping](#) [\[+\] Field Embedding](#)

Validation? (optional)

— OR —

-- select ontology service --

Required?* No Yes
* Prompt if field is blank

Identifier? No Yes
Does the field contain identifying information (e.g., name, SSN, address)?

Custom Alignment Right / Vertical (RV) Align the position of the field on the page

Field Note (optional)
Small reminder text displayed underneath field

Save **Cancel**



REDCap Shared Instrument Library

We Need You!

- We want to add more standardized instruments!
 - They must be free to distribute and have at least one peer-reviewed article with "standardization" information.
- The coding guidelines that the library uses are here:

https://projectredcap.org/wp-content/resources/redcap_library_coding_guidelines.pdf

Create a Form - Question/Field Bank 1

The screenshot shows the REDCap Online Designer interface. At the top, there are navigation tabs: Project Home, Project Setup, Online Designer (which is selected), Data Dictionary, and Codebook. Below the tabs, there are two buttons: 'Create snapshot of instruments' and 'VIDEO: How to use this page'. A note says 'Last snapshot: never ?'. The main content area is titled 'Data Collection Instruments'. It features three buttons: '+ Create' (highlighted with a purple callout '1'), 'Import', and 'Upload'. To the right, there's a 'Form options:' section with a 'Form Display Logic' link. Below these are tables for instruments. The first table has columns: 'Instrument name', 'Fields', 'View PDF', and 'Instrument actions'. It lists 'Hamilton Depression Rating Scale (HAM-D)' with 27 fields. The second table is a modal for creating a new instrument, with columns: 'New instrument name:', 'Fields', 'View PDF', and 'Instrument actions'. It shows 'Case Report Form' in the input field, and 'Create' and 'Cancel' buttons. A purple callout '2' points to the 'Case Report Form' input field.

Create a Form - Question/Field Bank 2

- The NIH Common Data Element Repository allows you to code questions consistently.

This page allows you to build and customize your data collection instruments one field at a time. You may add new fields or edit existing ones. New fields may be added by clicking the **Add Field** buttons. You can begin editing an existing field by clicking on the **Edit** icon. If you decide that you do not want to keep a field, you can delete it. To move fields, simply **drag and drop** a field to a different position within the instrument. Changes will take effect immediately in real time.

Questions from the NIH CDE Repository
(U.S. National Library of Medicine)

Project Home Project Setup Online Designer Data Dictionary Codebook

Create snapshot of instruments VIDEO: How to use this page

Last snapshot: never ?

Return to list of instruments

Current instrument: Case Report Form

Add Field Add Matrix of Fields Import from Field Bank

Preview instrument

Create a Form - Question/Field Bank 3

Inside the NIH CDE Repository, the NCI variables are rich:

Import from Field Bank

Using the Field Bank, search for fields in various catalogs below by selecting a catalog and entering specific keyword. When reviewing the results up of your search, click the "Add Field" button for the field to add that field to the current data collection instrument.

Select a catalog to search: NCI National Cancer Institute

sex

22 fields found for Classification: NIH CDE Repository → NCI - Keyword: sex

1 - 20 of 22

+ Add Field

Person Biological Entity Or Sex Gender Code PCORnet CDM Sex Code

Default field label: Person Biological Entity Or Sex Gender Code PCORnet CDM Sex Code

Classification: NCI

Description: A single human being...Pertaining to biology or to life and living things...An independently existing thing (living or nonliving)...An article used to connect words, phrases, or clauses representing alternatives; used to connect alternative terms for the same thing; used in correlation; used to correct or rephrase what was previously said; otherwise...The assemblage of physical properties or qualities by which male is distinguished from female; the physical difference between male and female; the distinguishing peculiarity of male or female...The assemblage of properties that distinguish people on the basis of their societal roles...A symbol or combination of symbols which is given an arbitrary meaning within a systematized collection of concepts used for data representation...Sex assigned at birth.

Male
Female
Ambiguous
No Information Available
Unknown
Other

+ Add Field

Person Biological Entity Sexual Orientation Code PCORnet CDM Sexual Orientation Code

Default field label: Person Biological Entity Sexual Orientation Code PCORnet CDM Sexual Orientation Code

Classification: NCI

Description: A human being...Pertaining to biology or to life and living things...An independently existing thing (living or nonliving)...The pattern of a person's emotional, romantic, and/or sexual attractions...A symbol or combination of symbols which is assigned to the members of a collection...Sexual orientation.

Something else
Bisexual
Questioning
Queer
Lesbian
Gay
Straight
Multiple Sexual Orientation
No Information Available
Other
Unknown
Response Declined
Asexual

Create a Form - Question/Field Bank 4

Edit Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

Field Type: Multiple Choice - Radio Buttons (Single Answer)

Field Label Use the Rich Text Editor ?

Sex Assigned at Birth Code

Choices (one choice per line) [Copy existing choices](#)

C20197, Male
C16576, Female
C98810, Ambiguous
C53269, No Information
C17998, Unknown
C17649, Other

How do I manually code the choices?

Variable Name (utilized in logic, calcs, and exports)
 Enable auto naming of variable based upon its Field Label?
ONLY letters, numbers, and underscores

How to use [Smart Variables](#) [Piping](#) [Field Embedding](#)

Required?* No Yes
* Prompt if field is blank

Identifier? No Yes
Does the field contain identifying information (e.g., name, SSN, address)?

Custom Alignment Right / Vertical (RV)
Align the position of the field on the page

Field Note (optional)
Small reminder text displayed underneath field

Action Tags / Field Annotation (optional)

Learn about [@ Action Tags](#) or [using Field Annotation](#)

NIH CDE:
<https://cde.nlm.nih.gov/cde/search>

Save **Cancel**

Create a Form - Custom Questions 1

This page allows you to build and customize your data collection instruments one field at a time. You may add new fields or edit existing ones. New fields may be added by clicking the **Add Field** buttons. You can begin editing an existing field by clicking on the **Edit** icon. If you decide that you do not want to keep a field, you can simply delete it by clicking on the **Delete** icon. To reorder the fields, simply **drag and drop** a field to a different position within the form below. NOTE: While in development status, all field changes will take effect immediately in real time.

Current instrument: **Case Report Form**

Return to list of instruments Add Field Add Matrix of Fields Import from Field Bank Preview instrument

Create truly custom questions.

Create a Form - Custom Questions 2

You can make truly custom forms by adding your own field:

Add New Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

---- Select a Type of Field ----

Field Type: Text Box (Short Text, Number, Date/Time, ...)

Field Label

Notes Box (Paragraph Text)
Calculated Field
Multiple Choice - Drop-down List (Single Answer)
Multiple Choice - Radio Buttons (Single Answer)
Checkboxes (Multiple Answers)
Yes - No
True - False
Signature (draw signature with mouse or finger)
File Upload (for users to upload files)
Slider / Visual Analog Scale
Descriptive Text (with optional Image/Video/Audio/File Attachment)
Begin New Section (with optional text)

Use the Rich Text Editor ?

Variable Name (utilized in logic, calcs, and exports)

 Enable auto naming of variable based upon its Field Label?
ONLY letters, numbers, and underscores

How to use [Smart Variables](#) [Piping](#) [Field Embedding](#)

Validation? (optional) ---- None ----
- OR -
 -- select ontology service --

Required?* No Yes
* Prompt if field is blank

Identifier? No Yes
Does the field contain identifying information (e.g., name, SSN, address)?

Custom Alignment Right / Vertical (RV)
Align the position of the field on the page

Field Note (optional)

Small reminder text displayed underneath field

Save **Cancel**

NOTICE: Text can/should have validation.

Create a Form - Custom Questions 3

Add New Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

Field Type: Multiple Choice - Drop-down List (Single Answer) Codebook

Field Label
What is the best symposium (in Plaza Ballroom A) at CPDD 2023? Use the Rich Text Editor ?

Variable Name (utilized in logic, calcs, and exports)
best Enable auto naming of variable based upon its Field Label?

Raw values for choices were added automatically

The choices listed below did not appear to have a raw value listed but only had a label, so a raw value has been provided for them automatically. If you are not satisfied with these auto generated values, you may change them before saving your changes for this field. The choices in the 'Choices' text box have automatically been modified to reflect these changes.

Choices (one choice per line) [Copy existing choices](#)

1, Epidemiology and Public Health Research Methods
2, Oral Communication: Cocaine Science in Humans
3, 10 Years Gone: Remembering Bill Woolverton Through...
4, The Past, Present, and Future of Reinforcer Patholog...
5, A Gentle Introduction to Data Science Tools for Addic...

1 was set as the raw value for **Epidemiology and Public Heal...**
2 was set as the raw value for **Oral Communication: Cocaine...**
3 was set as the raw value for **10 Years Gone: Remembering ...**
4 was set as the raw value for **The Past, Present, and Future...**
5 was set as the raw value for **A Gentle Introduction to Data...**

Enable auto-complete for this drop-down ?

Action Tags / Field Annotation (optional)

Learn about [@ Action Tags](#) or [using Field Annotation](#)

Save **Cancel**

How do I manually code the choices?

Create a Form - Custom Questions 4

Add New Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

Field Type: Multiple Choice - Drop-down List (Single Answer) Codebook

Field Label Use the Rich Text Editor ?

What is the best symposium (in Plaza Ballroom A) at CPDD 2023?

Variable Name (utilized in logic, calcs, and exports)
best Enable auto naming of variable based upon its Field Label?

ONLY letters, numbers, and underscores

How to use [+] Smart Variables [+] Piping [+] Field Embedding

Required?* No Yes
* Prompt if field is blank

Identifier? No Yes
Does the field contain identifying information (e.g., name, SSN, address)?

Custom Alignment Right / Vertical (RV)
Align the position of the field on the page

Field Note (optional) There is a right answer!
Small reminder text displayed underneath field

Enable auto-complete for this drop-down ? How do I manually code the choices?

Action Tags / Field Annotation (optional)

Learn about [@ Action Tags](#) or [using Field Annotation](#)

Save Cancel



Create a Form - Custom Questions 5

Add Field Add Matrix of Fields Import from Field Bank

     Variable: first_name

First name

Add Field Add Matrix of Fields Import from Field Bank

     Variable: last_name

Last name

Add Field Add Matrix of Fields Import from Field Bank

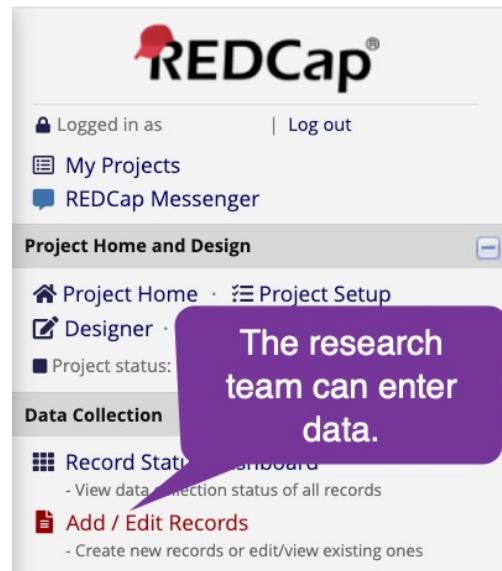
     Variable: best

What is the best symposium (in Plaza Ballroom A) at CPDD 2023? 
There is a right answer.

Add Field Add Matrix of Fields Import from Field Bank

Collect/Enter Data 1

- After you have created your forms, the research team can enter data or you can turn the form into a survey.
- To use surveys, you need to enable them for a project, then tag an individual form as a survey.



The screenshot shows the REDCap interface with three numbered annotations:

- Annotation 1 points to the 'Project Setup' link in the top navigation bar.
- Annotation 2 points to the 'Use surveys in this project?' checkbox in the 'Main project' section.
- Annotation 3, labeled 'Optional', points to the 'I'm done!' button at the bottom of the page.

Collect/Enter Data 2

- The survey needs to be the first questionnaire/instrument.

Data Collection Instruments		Form options:		Automated Survey Invitation options:		
		<input checked="" type="checkbox"/> Form Display Logic		<input checked="" type="checkbox"/> Upload or download Auto Invitations ▾		
		<input checked="" type="checkbox"/> Survey Queue		<input checked="" type="checkbox"/> Re-evaluate Auto Invitations		
		<input checked="" type="checkbox"/> Survey Notifications				
Instrument name	Fields	View PDF	Enabled as survey	Instrument actions	Survey-related options	
Hamilton Depression Rating Scale (HAM-D)	27		<input checked="" type="button"/> Enable	Choose action ▾		
Case Report Form	3		<input checked="" type="button"/> Enable	Choose action ▾		

Data Collection Instruments		Form options:		Automated Survey Invitation options:		
		<input checked="" type="checkbox"/> Form Display Logic		<input checked="" type="checkbox"/> Upload or download Auto Invitations ▾		
		<input checked="" type="checkbox"/> Survey Queue		<input checked="" type="checkbox"/> Re-evaluate Auto Invitations		
		<input checked="" type="checkbox"/> Survey Notifications				
Instrument name	Fields	View PDF	Enabled as survey	Instrument actions	Survey-related options	
Case Report Form	4		<input checked="" type="button"/> Enable	Choose action ▾	<input checked="" type="button"/> Survey settings	+ Automated Invitations
Hamilton Depression Rating Scale Hamd	26		<input checked="" type="button"/> Enable	Choose action ▾		

Collect/Enter Data 3

The screenshot shows the 'Project Home and Design' interface. On the left, there's a sidebar with sections like 'Project Home and Design', 'Data Collection', and 'Applications'. The 'Survey Distribution Tools' section under 'Data Collection' is highlighted with a large purple arrow pointing towards it. The 'Survey Distribution Tools' section contains links for creating a public survey link or participant list.

Survey Distribution Tools

[Public Survey Link](#) [Participant List](#) [Survey Invitation Log](#)

Using a public survey link is the simplest and fastest way to collect responses for your survey. You may obtain the survey link below to email it to your participants. Responses will be collected anonymously (unless the survey contains questions asking for identifying data from the participant). **NOTE:** Since this method uses a single survey link for all participants, it allows for the possibility of participants taking the survey multiple times, which may be necessary in some cases.

To obtain the survey link, copy the URL below and paste it into the body of an email message in your own email client. Your email recipient(s) can then click the link to begin taking your survey.

Public Survey URL: <https://redcap.miami.edu/surveys/?s=HCXMHEKEWJXFWMH>

Protect the public survey using the Google reCAPTCHA feature

Link Actions  reCAPTCHA is available. [Customizations](#)

[Open public survey](#) [Get Short Survey Link](#)
[Open public survey + Log out](#) [Create Custom Survey Link](#)
[Send me URL via email](#) [Get Embed Code](#)
[Survey Access Code or QR Code](#)

Collect/Enter Data 4

A screenshot of a web browser window titled "Case Report Form". The URL is <https://redcap.miami.edu/surveys/?s=7XTTW8...>. The page content is as follows:

Case Report Form

Please complete the survey below.

Thank you!

1) First name

2) Last name

3) What is the best symposium (in Plaza Ballroom A) at CPDD 2023?
There is a right answer.

Submit

View/Summarize in REDCap

- You can easily make:
 - public/private dashboards with summary statistics and basic graphics
 - quick **Stats & Charts** reports
 - tabular reports

Dashboards 1

Project Design privileges can create and edit them. A wizard is provided to help you build dashboards. This page provides a step-by-step guide to creating a dashboard. To learn more about the dashboard editor, see the [Dashboard Editor Help](#) page to help you easily construct the syntax to add Smart Functions, Smart Tables, or Smart Charts to your dashboards. A wizard is provided to help you build dashboards. This page provides a step-by-step guide to creating a dashboard. To learn more about the dashboard editor, see the [Dashboard Editor Help](#) page to help you easily construct the syntax to add Smart Functions, Smart Tables, or Smart Charts to your dashboards.

Name it

Public or private

Dashboard content:

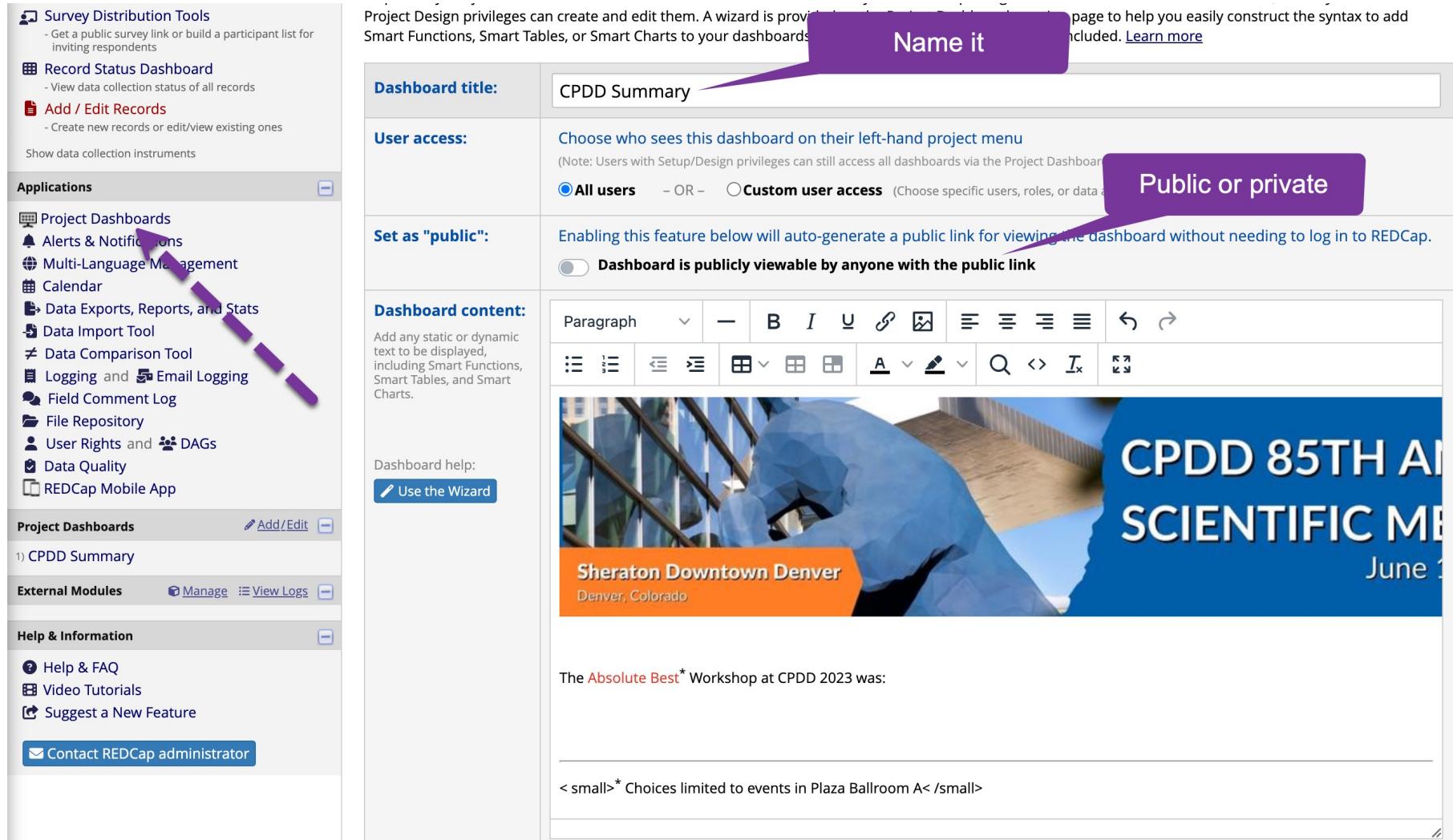
Add any static or dynamic text to be displayed, including Smart Functions, Smart Tables, and Smart Charts.

Dashboard help:

[Use the Wizard](#)

The **Absolute Best*** Workshop at CPDD 2023 was:

<small>* Choices limited to events in Plaza Ballroom A</small>



Dashboards 2

Survey Distribution Tools
- Get a public survey link or build a participant list for inviting respondents

Record Status Dashboard
- View data collection status of all records

Add / Edit Records
- Create new records or edit/view existing ones

Show data collection instruments

Applications

- Project Dashboards
- Alerts & Notifications
- Multi-Language Management
- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool
- Data Comparison Tool
- Logging and Email Logging
- Field Comment Log
- File Repository
- User Rights and DAGs
- Data Quality
- REDCap Mobile App

Project Dashboards [Add/Edit](#)

- 1) CPDD Summary

External Modules [Manage](#) [View Logs](#)

Help & Information

- Help & FAQ
- Video Tutorials
- Suggest a New Feature

[Contact REDCap administrator](#)

Project Design privileges can create and edit them. A wizard is provided on the Project Dashboard creation page to help you easily construct the syntax to add Smart Functions, Smart Tables, or Smart Charts to your dashboards, and a list of helpful examples is also included. [Learn more](#)

Dashboard title: CPDD Summary

User access: Choose who sees this dashboard on their left-hand project menu
(Note: Users with Setup/Design privileges can still access all dashboards via the Project Dashboards page.)
 All users - OR - Custom user access (Choose specific users, roles, or data access groups who will have access)

Set as "public": Enabling this feature below will auto-generate a public link for viewing
 Dashboard is publicly viewable by anyone with the public link

Dashboard content: Add any static or dynamic text to be displayed, including Smart Functions, Smart Tables, and Smart Charts.

Dashboard help: [Use the Wizard](#)

Point-and-click formatting REDCap.

The Absolute Best* Workshop at CPDD 2023 was:

Sheraton Downtown Denver
Denver, Colorado

Paste in images

<small>^{*} Choices limited to events in Plaza Ballroom A</small>

Dashboards 3

Survey Distribution Tools
- Get a public survey link or build a participant list for inviting respondents

Record Status Dashboard
- View data collection status of all records

Add / Edit Records
- Create new records or edit/view existing ones

Show data collection instruments

Applications

- Project Dashboards
- Alerts & Notifications
- Multi-Language Management
- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool
- Data Comparison Tool
- Logging and Email Logging
- Field Comment Log
- File Repository
- User Rights and DAGs
- Data Quality
- REDCap Mobile App

Project Dashboards Add/Edit

- 1) CPDD Summary

External Modules Manage View Logs

Help & Information

- Help & FAQ
- Video Tutorials
- Suggest a New Feature

Contact REDCap administrator

Project Design privileges can create and edit them. A wizard is provided on the Project Dashboard creation page to help you easily construct the syntax to add Smart Functions, Smart Tables, or Smart Charts to your dashboards, and a list of helpful examples is also included. [Learn more](#)

Dashboard title: CPDD Summary

User access: Choose who sees this dashboard on their left-hand project menu
(Note: Users with Setup/Design privileges can still access all dashboards via the Project Dashboards page.)
 All users - OR - Custom user access (Choose specific users, roles, or data access groups who will have access)

Set as "public": Enabling this feature below will auto-generate a public link for viewing the dashboard without needing to log in to REDCap.
 Dashboard is publicly viewable by anyone with the public link

Dashboard content: Add any static or dynamic text to be displayed, including Smart Functions, Smart Tables, and Smart Charts.

Dashboard help:
[Use the Wizard](#)

The Absolute Best* Workshop at CPDD 2023 was:

* Choices limited to events in Plaza Ballroom A

HTML code if you want

Dashboards 4

The screenshot shows a 'Smart Functions' creation wizard window overlaid on a dashboard titled 'CPDD Summary'. The dashboard has a sidebar with 'Dashboard content' and 'Dashboard help' sections, and a button 'Use the Wizard'.

Click to add statistics and graphics

Creating Smart Functions, Smart Tables, and Smart Charts

This window guides you through creating a new Smart Function, Smart Table, or Smart Chart. It includes steps for choosing a Smart Variable, selecting fields, applying optional filters, and copying generated code.

Step 1) Choose a Smart Variable to create:
bar-chart

Step 2) Choose a field to utilize:
best "What is the best symposium (in Plaza)"

Set details

Step 3) Optional data filtering and other settings:
By default, Smart Functions, Smart Tables, and Smart Charts will utilize all the data from *all records* in the project. However, you can utilize a subset of the data in the project by limiting them to a specific report's data (using a unique report name), to records belonging to one or more DAGs, and/or to data in specific events (if the project is longitudinal).

Filter the data using a report: -- no filtering by report --

Copy the code

Step 4) Copy the generated Smart Variable syntax and paste it in your project dashboard:

[bar-chart:best]

Copy to clipboard

Close

Dashboards 5

The screenshot shows a Microsoft Word document with a ribbon menu at the top. The ribbon has tabs for Home, Insert, Page Layout, References, Mailings, and Review. Below the ribbon is a toolbar with various icons for text and image editing. The main content area features a large blue banner for the "CPDD 85TH ANNUAL SCIENTIFIC MEETING" scheduled for "June 1-4, 2023". To the left of the banner is a photograph of a modern building with a blue, angular facade. An orange callout box contains the text "Sheraton Downtown Denver" and "Denver, Colorado". A purple speech bubble points from the text "[bar-chart:best]" to the statement "I pasted the Wizard's code.". At the bottom of the page, there is a small note: "* Choices limited to events in Plaza Ballroom A".

The Absolute Best* Workshop at CPDD 2023 was:

[bar-chart:best]

I pasted the Wizard's code.

* Choices limited to events in Plaza Ballroom A

Dashboards 6

Contact REDCap administrator

Project Home and Design

- Project Home · Project Setup
- Designer · Dictionary · Codebook
- Project status: Development

Data Collection

- Survey Distribution Tools
 - Get a public survey link or build a participant list for inviting respondents
- Record Status Dashboard
 - View data collection status of all records
- Add / Edit Records
 - Create new records or edit/view existing ones

Show data collection instruments

Applications

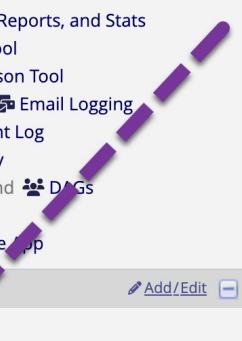
- Project Dashboards
- Alerts & Notifications
- Multi-Language Management
- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool
- Data Comparison Tool
- Logging and Email Logging
- Field Comment Log
- File Repository
- User Rights and D&Gs
- Data Quality
- REDCap Mobile App

Project Dashboards

1) CPDD Summary

External Modules

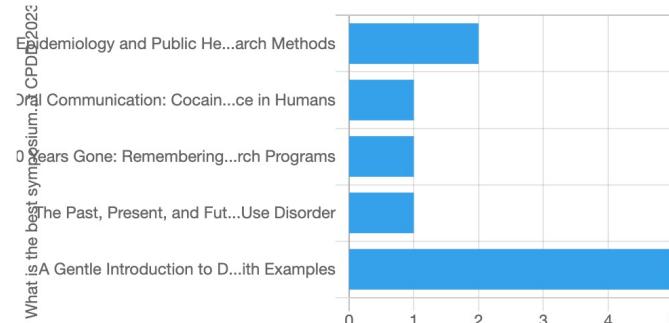
Manage View Logs



CPDD Summary

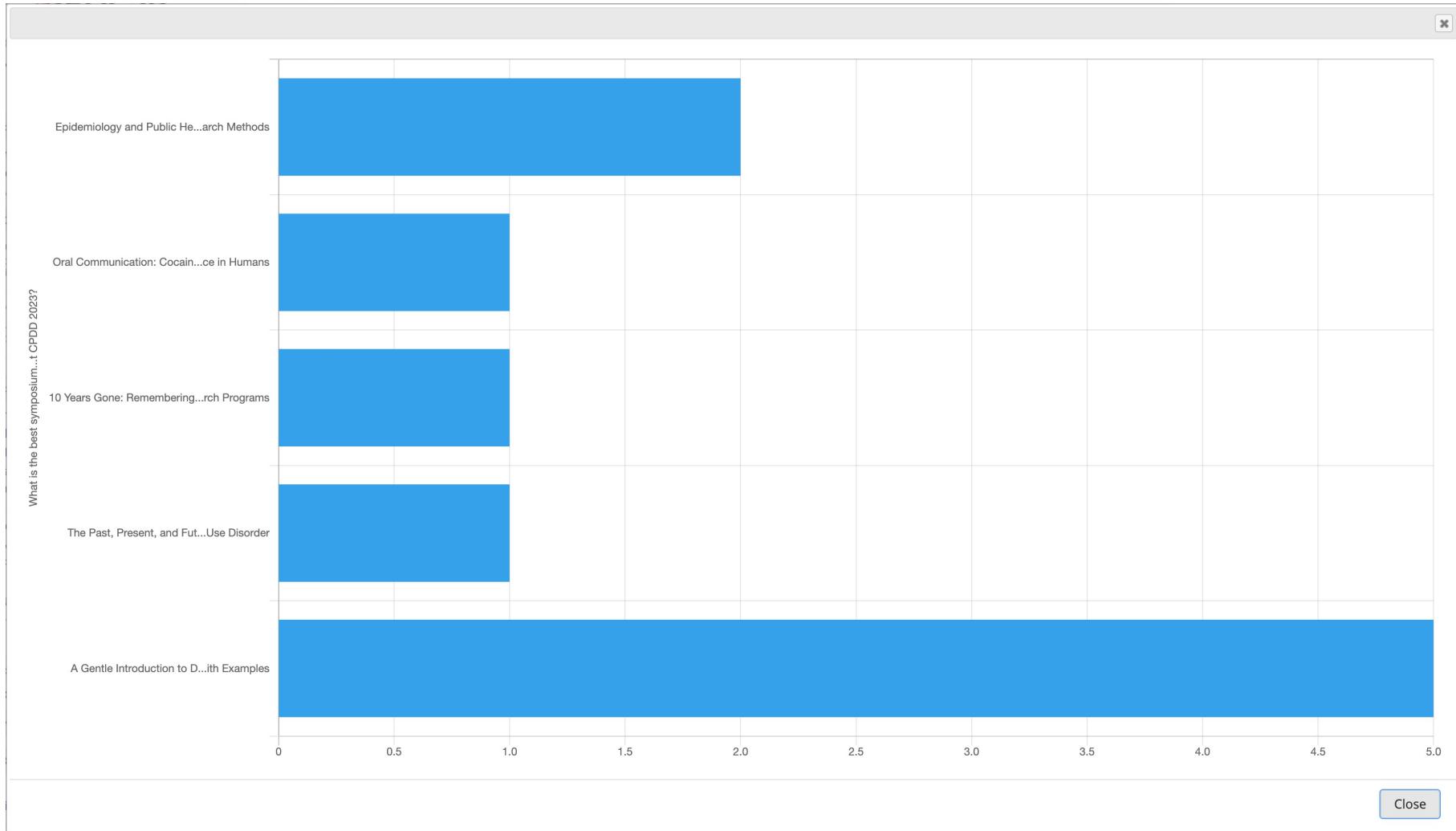


The **Absolute Best*** Workshop at CPDD 2023 was:



* Choices limited to events in Plaza Ballroom A

Dashboards 7



Stats & Charts 1

Data Exports, Reports, and Stats

VIDEO: How to use Data Exports, Reports, and Stats

+ Create New Report My Reports & Exports Other Export Options

This module allows you to easily view reports of your data, inspect plots and descriptive statistics of your data, as well as export your data to Microsoft Excel, SAS, Stata, R, or SPSS for analysis (if you have such privileges). If you wish to export your *entire* data set or view it as a report, then Report A is the best and quickest way. However, if you want to view or export data from only specific instruments (or events) on the fly, then Report B is the best choice. You may also create your own custom reports below (if you have such privileges) in which you can filter the report to specific fields, records, or events using a vast array of filtering tools to make sure you get the exact data you want. Once you have created a report, you may view it as a webpage, export it out of REDCap in a specified format (Excel, SAS, Stata, SPSS, R), or view the plots and descriptive statistics for that report.

My Reports & Exports

	Report name	View/Export Options	Management Options	Report ID ? (auto-generated)	Unique report name (auto-generated)
A	All data (all records and fields)	View Report Export Data Stats & Charts			
B	Selected instruments (all records)	Make custom selections			
	+ Create New Report				

First name (*first_name*)

Total Count (N)	Missing*
10	0 (0.0%)

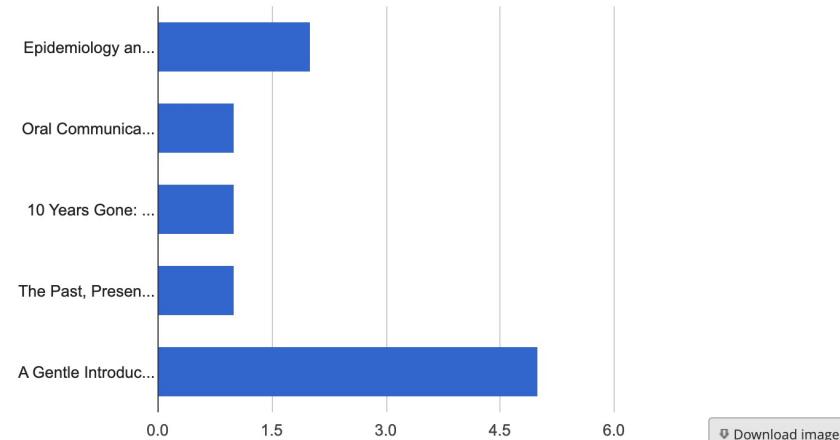
Last name (*last_name*)

Total Count (N)	Missing*
10	0 (0.0%)

What is the best symposium (in Plaza Ballroom A) at CPDD 2023? (best) [Refresh Plot](#) | [View as Bar Chart](#) ▾

Total Count (N)	Missing*	Unique
10	0 (0.0%)	5

Counts/frequency: Epidemiology and Public Health Research Methods (2, 20.0%), Oral Communication: Cocaine Science in Humans (1, 10.0%), 10 Years Gone: Remembering Bill Woolverton Through His Impact on our Current Research Programs (1, 10.0%), The Past, Present, and Future of Reinforcer Pathology Theory in Substance Use Disorder (1, 10.0%), A Gentle Introduction to Data Science Tools for Addiction Research: Use Cases With Examples (5, 50.0%)



Tabular Reports 1

R/Medicine 2023 Demo PID 6754

Data Exports, Reports, and Stats

VIDEO: How to use Data Exports, Reports, and Stats

+ Create New Report My Reports & Exports Other Export Options

This module allows you to easily view reports of your data, inspect plots and descriptive statistics of your data, as well as export your data to Microsoft Excel, SAS, Stata, R, or SPSS for analysis (if you have such privileges). If you wish to export your *entire* data set or view it as a report, then Report A is the best and quickest way. However, if you want to view or export data from only specific instruments (or events) on the fly, then Report B is the best choice. You may also create your own custom reports below (if you have such privileges) in which you can filter the report to specific fields, records, or events using a vast array of filtering tools to make sure you get the exact data you want. Once you have created a report, you may view it as a webpage, export it out of REDCap in a specified format (Excel, SAS, Stata, SPSS, R), or view the plots and descriptive statistics for that report.

My Reports & Exports

	Report name	View/Export Options	Management Options	Report ID (auto-generated)	Unique report name (auto-generated)
A	All data (all records and fields)	View Report Export Data Stats & Charts			
B	Selected instruments (all records)	Make custom selections			
	+ Create New Report				

Tabular Reports 2

Data Exports, Reports, and Stats

[VIDEO: How to use Data Exports, Reports, and Stats](#)

+ Create New Report | My Reports & Exports | Other Export Options | View Report: Selected instruments (all records)

Number of results returned: 10
Total number of records queried: 10
Report execution time: 0 seconds

Stats & Charts | Export Data | Print Page

Live filters: Record ID ▾

Selected instruments (all records)

Search

Record ID record_id	Survey Identifier redcap_survey_identifier	Survey Timestamp case_report_form_timestamp	First name first_name	Last name last_name	What is the best symposium (in Plaza Ballroom A) at CPDD 2023? best	Complete? case_report_form_complete
1		06-15-2023 08:39	Raymond	Balise	A Gentle Introduction to Data Science Tools for Addiction Research: Use Cases With Examples (5)	Complete (2)
2		06-15-2023 08:39	R	Balise	A Gentle Introduction to Data Science Tools for Addiction Research: Use Cases With Examples (5)	Complete (2)
3		06-15-2023 08:39	Ray	Balise	A Gentle Introduction to Data Science Tools for Addiction Research: Use Cases With Examples (5)	Complete (2)
4		06-15-2023 08:40	Me	Balise	A Gentle Introduction to Data Science Tools for Addiction Research: Use Cases With Examples (5)	Complete (2)
5		06-15-2023 08:40	Momma	Balise	A Gentle Introduction to Data Science Tools for Addiction Research: Use Cases With Examples (5)	Complete (2)
6		06-15-2023 08:41	Not	Balise	Epidemiology and Public Health Research Methods (1)	Complete (2)
7		06-18-2023 15:25	Not	Ray	Epidemiology and Public Health Research Methods (1)	Complete (2)
8		06-18-2023 15:25	Also not	Ray	Oral Communication: Cocaine Science in Humans (2)	Complete (2)
9		06-18-2023 15:25	Still not	Ray	10 Years Gone: Remembering Bill Woolverton Through His Impact on our Current Research Programs (3)	Complete (2)
10		06-18-2023 15:25	Maybe	Ray	The Past, Present, and Future of Reinforcer Pathology Theory in Substance Use Disorder (4)	Complete (2)

Export for Additional Magic 1

- You can manually export any tabular report.

The screenshot shows the REDCap Project Home and Design interface. On the left, there's a sidebar with sections for Project Home and Design, Data Collection, and Applications. The Applications section is expanded, showing options like Project Dashboard, Alerts & Notifications, Multi-Language Management, Calendar, Data Exports, Reports, and Stats (which is highlighted with a purple arrow), Data Import Tool, Data Comparison Tool, Logging, and Email Logging. The main content area is titled "Data Exports, Reports, and Stats". It has tabs for "+ Create New Report", "My Reports & Exports" (which is selected and highlighted with a purple arrow), and "Other Export Options". Below the tabs is a descriptive text about the module's functionality. The "My Reports & Exports" section contains a table with columns for Report name, View/Export Options, Management Options, Report ID (auto-generated), and Unique report name (auto-generated). There are two rows in the table:

	Report name	View/Export Options	Management Options	Report ID (auto-generated)	Unique report name (auto-generated)
A	All data (all records and fields)	View Report Export Data Stats & Charts			
B	Selected instruments (all records)	Make custom selections			
	+ Create New Report				

Export for Additional Magic 2

- Export a single form.

The screenshot shows the REDCap interface for managing data exports. On the left, a sidebar menu includes 'Project Home and Design' (with 'Project Home' selected), 'Data Collection' (with 'Survey Distribution Tools' and 'Record Status Dashboard'), and 'Applications' (with various tools like Project Dashboards, Alerts & Notifications, Multi-Language Management, etc.). The main content area is titled 'Data Exports, Reports, and Stats' and features a 'VIDEO: How to use Data Exports, Reports, and Stats'. It has tabs for '+ Create New Report', 'My Reports & Exports' (selected), and 'Other Export Options'. A descriptive text explains the module's purpose: exporting data to Microsoft Excel, SAS, Stata, R, or SPSS. Below this is a table titled 'My Reports & Exports' with columns for Report name, View/Export Options, Management Options, Report ID, and Unique report name. Two rows are shown: Row A ('All data (all records and fields)') and Row B ('Selected instruments (all records)'). Both rows include 'View Report', 'Export Data', and 'Stats & Charts' buttons. Row B also includes a dropdown for 'Instruments' (set to 'Case Report Form') and a 'Create report' button. A 'Create New Report' button is at the bottom of the table.

Export for Additional Magic 3

The screenshot shows the REDCap interface with a modal dialog titled "Exporting 'All data (all records and fields)'".

Choose export format:

- CSV / Microsoft Excel (raw data)
- CSV / Microsoft Excel (labels)
- SPSS Statistical Software
- SAS Statistical Software
- R Statistical Software
- Stata Statistical Software
- CDISC ODM (XML)

De-identification options (optional):

The options below allow you to limit the amount of sensitive information that you are exporting out of the project. Check all that apply.

- Remove all tagged identifier fields (tagged in Data Dictionary)
- Hash the Record ID field (converts record name to an unrecognizable value)

Additional export options:

- Export survey identifier field and survey timestamp field(s)?

Advanced data formatting options - Export blank values for gray Form Status:

All Form Status fields with a gray status icon can be exported either as a blank value or as "0" (Incomplete). Hint: Blank values are recommended if the data will be imported back into REDCap, in which this preserves the gray status icons for all the imported records.

Export gray Form Status fields with value of "0" ▾

Set CSV delimiter character:

Set the delimiter used to separate values in the CSV data file (only valid for CSV Raw Data and CSV Labels export formats):
(comma) - default ▾

Force all numbers into a specified decimal format?

You may choose to force all data values containing a decimal to have a specified decimal character (comma or period/full stop). This will be applied to all calculations and number-validated text values in the export file.

Use fields' native decimal format (default) ▾

Note: Your data formatting selections above will be remembered in the future and will be pre-selected upon your next export.

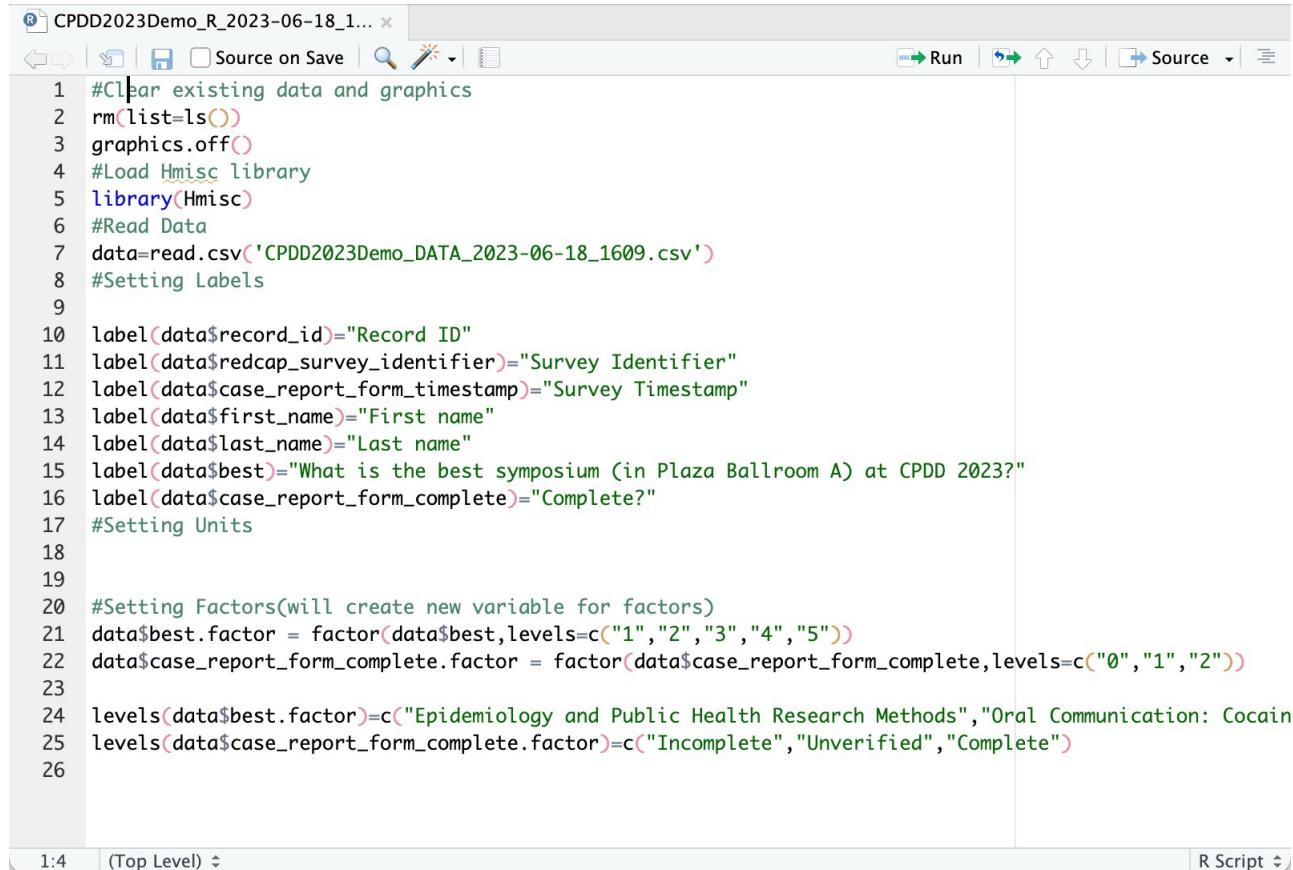
Buttons:

- Export Data
- Cancel

Export for Additional Magic 4

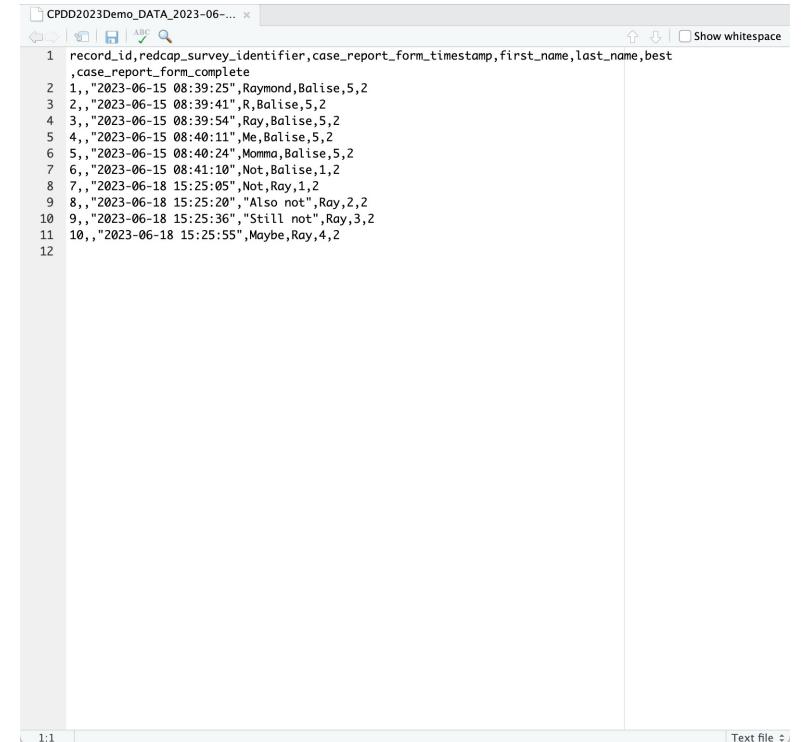
The screenshot shows the REDCap interface for the "R/Medicine 2023 Demo" project (PID 6754). The left sidebar contains navigation links for Project Home and Design, Data Collection, Applications, and Help & Information. The main content area displays the "Data Exports, Reports, and Stats" section. A modal window titled "Data export was successful!" is open, providing instructions on how to download the data. It mentions that the data export was successful and the data is now ready to be downloaded. It instructs users to click the download icon(s) below to download their data file. It also notes that if exporting to a specific statistical analysis package, they will additionally need to download the syntax file provided for that stats package. Below this, a "Citation Notice" section asks users to cite the REDCap project when publishing manuscripts, with citation information and template methods available here. At the bottom of the modal, there is an "R Statistical Software" section with instructions to use the command `read.csv('filename')` to read in the data file. There are download icons for R and DATA CSV, and a "Send file" link. The background of the main interface shows a "My Reports & Exports" section with a table for Report ID and Unique report name.

Export for Additional Magic 5



The screenshot shows an RStudio interface with an R script titled "CPDD2023Demo_R_2023-06-18_1...". The script contains code for clearing existing data, loading the Hmisc library, reading a CSV file, setting labels and units, and creating factor variables. The code is color-coded for syntax.

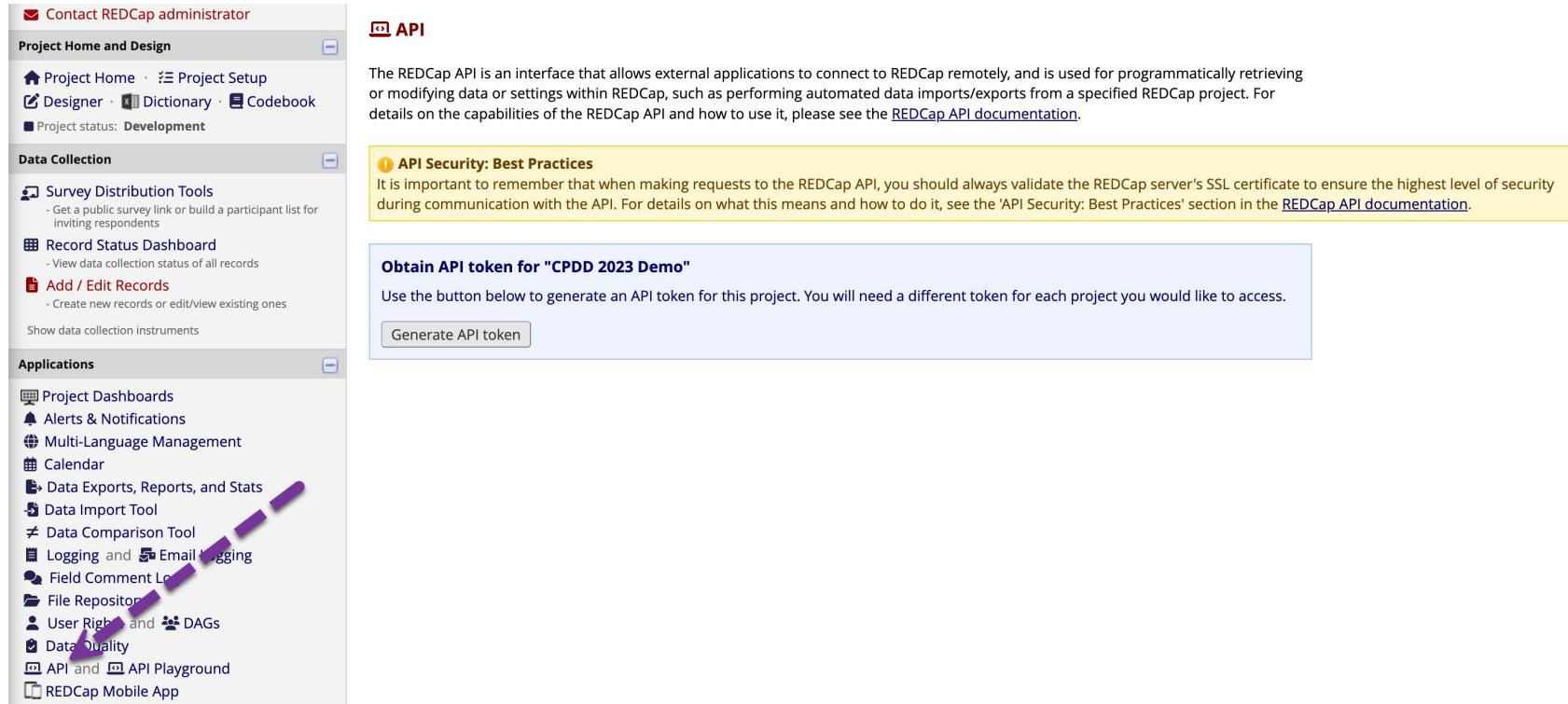
```
1 #Clear existing data and graphics
2 rm(list=ls())
3 graphics.off()
4 #Load Hmisc library
5 library(Hmisc)
6 #Read Data
7 data=read.csv('CPDD2023Demo_DATA_2023-06-18_1609.csv')
8 #Setting Labels
9
10 label(data$record_id)="Record ID"
11 label(data$redcap_survey_identifier)="Survey Identifier"
12 label(data$case_report_form_timestamp)="Survey Timestamp"
13 label(data$first_name)="First name"
14 label(data$last_name)="Last name"
15 label(data$best)="What is the best symposium (in Plaza Ballroom A) at CPDD 2023?"
16 label(data$case_report_form_complete)="Complete?"
17 #Setting Units
18
19
20 #Setting Factors(will create new variable for factors)
21 data$best.factor = factor(data$best,levels=c("1","2","3","4","5"))
22 data$case_report_form_complete.factor = factor(data$case_report_form_complete,levels=c("0","1","2"))
23
24 levels(data$best.factor)=c("Epidemiology and Public Health Research Methods","Oral Communication: Cocain")
25 levels(data$case_report_form_complete.factor)=c("Incomplete","Unverified","Complete")
26
```



The screenshot shows an RStudio interface with a CSV file titled "CPDD2023Demo_DATA_2023-06-18_1609.csv". The file contains 12 rows of data with columns: record_id, redcap_survey_identifier, case_report_form_timestamp, first_name, last_name, best, and case_report_form_complete.

record_id	redcap_survey_identifier	case_report_form_timestamp	first_name	last_name	best	case_report_form_complete
1	"2023-06-15 08:39:25"		Raymond	Balise	5	2
2	"2023-06-15 08:39:41"		R	Balise	5	2
3	"2023-06-15 08:39:54"		Ray	Balise	5	2
4	"2023-06-15 08:40:11"		Me	Balise	5	2
5	"2023-06-15 08:40:24"		Momma	Balise	5	2
6	"2023-06-15 08:41:10"		Not	Balise	1	2
7	"2023-06-18 15:25:05"		Not	Ray	1	2
8	"2023-06-18 15:25:20"		Also not	Ray	2	2
9	"2023-06-18 15:25:36"		Still not	Ray	3	2
10	"2023-06-18 15:25:55"		Maybe	Ray	4	2
11						
12						

API Export for Better Magic 1



The screenshot shows the REDCap API documentation page. The left sidebar contains navigation links for 'Project Home and Design', 'Data Collection', and 'Applications'. A purple arrow points from the 'API' link in the sidebar to the 'API' section in the main content area. The main content area starts with a heading 'API' and a brief description of what the API is used for. Below this is a yellow box containing 'API Security: Best Practices' information. At the bottom of the main content area is a light blue box titled 'Obtain API token for "CPDD 2023 Demo"'. It contains a button labeled 'Generate API token'.

Contact REDCap administrator

Project Home and Design

- Project Home · Project Setup
- Designer · Dictionary · Codebook
- Project status: Development

Data Collection

- Survey Distribution Tools
 - Get a public survey link or build a participant list for inviting respondents
- Record Status Dashboard
 - View data collection status of all records
- Add / Edit Records
 - Create new records or edit/view existing ones

Show data collection instruments

Applications

- Project Dashboards
- Alerts & Notifications
- Multi-Language Management
- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool
- Data Comparison Tool
- Logging and Email Logging
- Field Comment Log
- File Repository
- User Rights and DAGs
- Data Quality
- API and API Playground
- REDCap Mobile App

02 - Using REDCap, R, and Markdown to Produce Publications

Infectious Disease Elimination Act

Intersecting Epidemics

- HIV
- hepatitis C
- overdose deaths

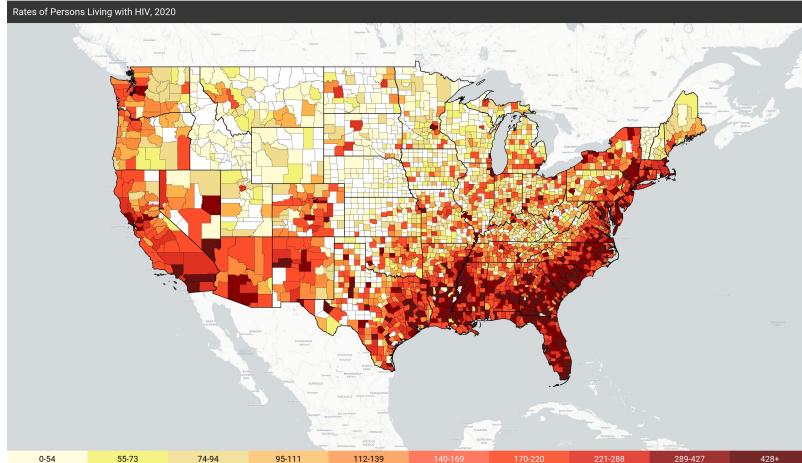
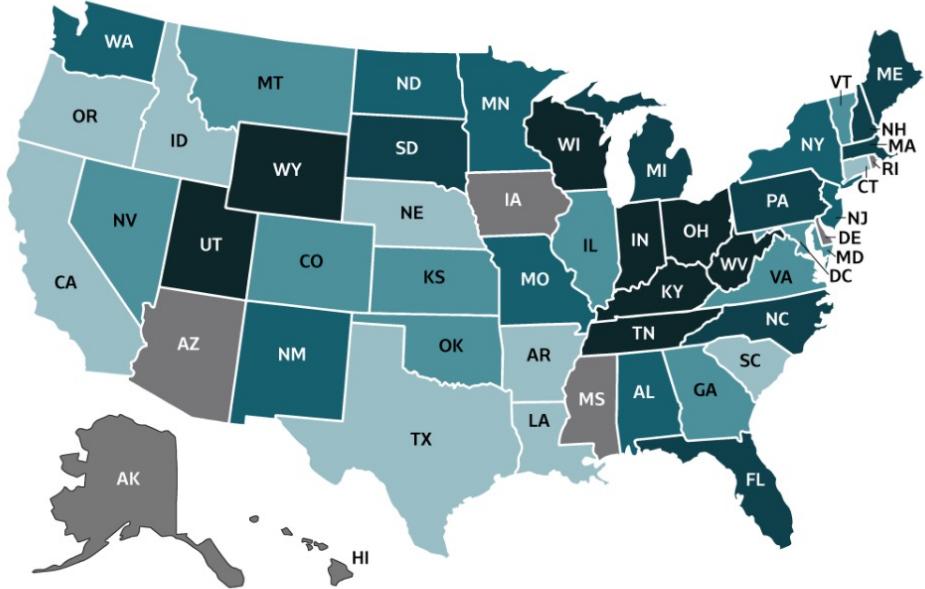


Figure 3.3. Rates of reported acute hepatitis C, by state or jurisdiction — United States, 2018



Color Key	Cases/100,000 Population	States
Light Blue	0-0.3	AR, CA, CT, ID, LA, NE, OR, SC, TX
Moderate Blue	>0.3-0.9	CO, GA, IL, KS, MD, MT, NV, OK, VT, VA
Dark Blue	>0.9-1.3	AL, MN, MO, NJ, NM, NY, ND, WA
Very Dark Blue	>1.3-2.2	FL, ME, MA, MI, NH, NC, PA
Black	>2.2-4.0	IN, KY, OH, SD, TN, UT, WV, WI, WY
Grey	Data Not Available	AK, AZ, DE, DC, HI, IA, MS, RI

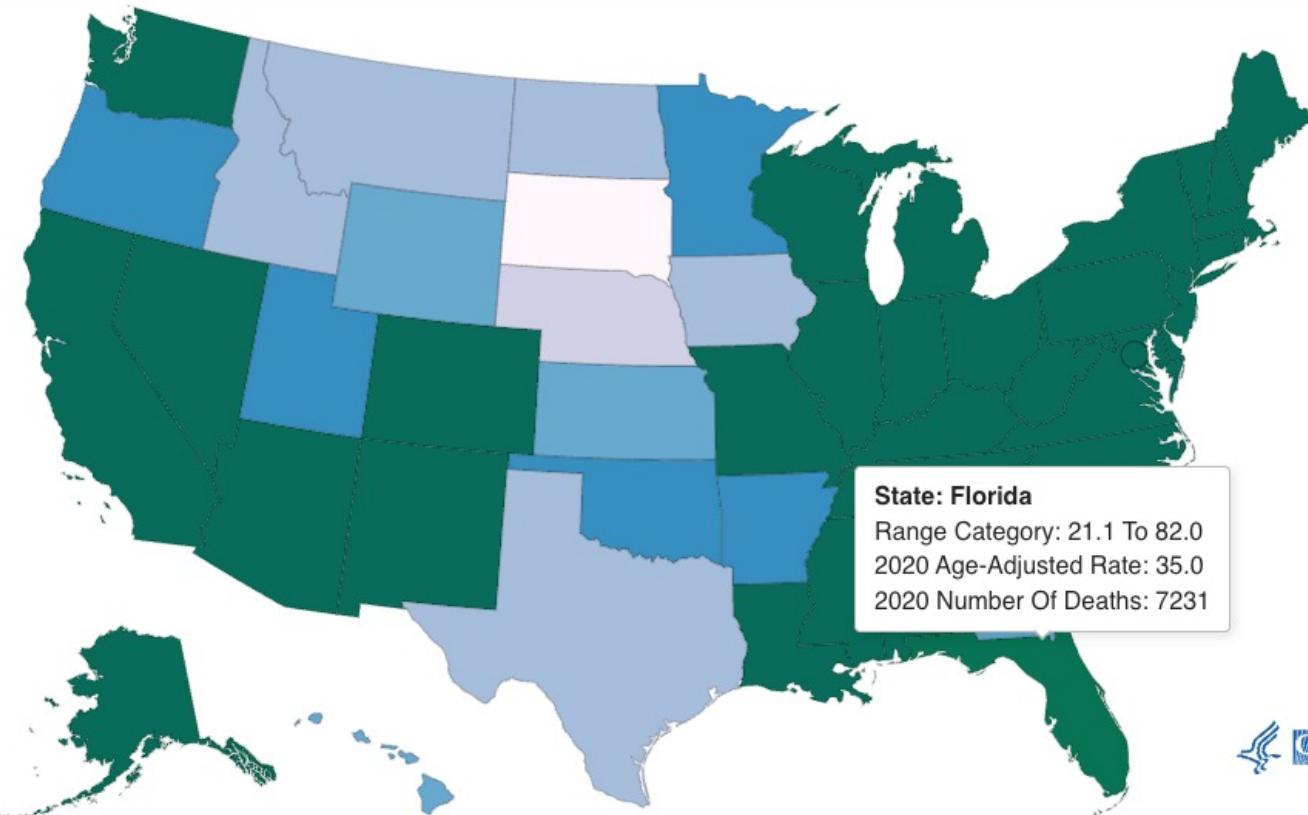
Source: CDC, National Notifiable Diseases Surveillance System.

* Data not shown to protect privacy because of a small number of cases and/or a small population.

** State health department, per its HIV data re-release agreement with CDC, requested not to release data to AIDSVu. See Data Methods for more information.

NOTE: There are no county-level maps for Alaska, District of Columbia, and Puerto Rico because there are no counties in these states.

Number and Age-adjusted Rates of Drug Overdose Deaths by State, US 2020



Range Category

6.9 to 11.0

11.1 to 13.5

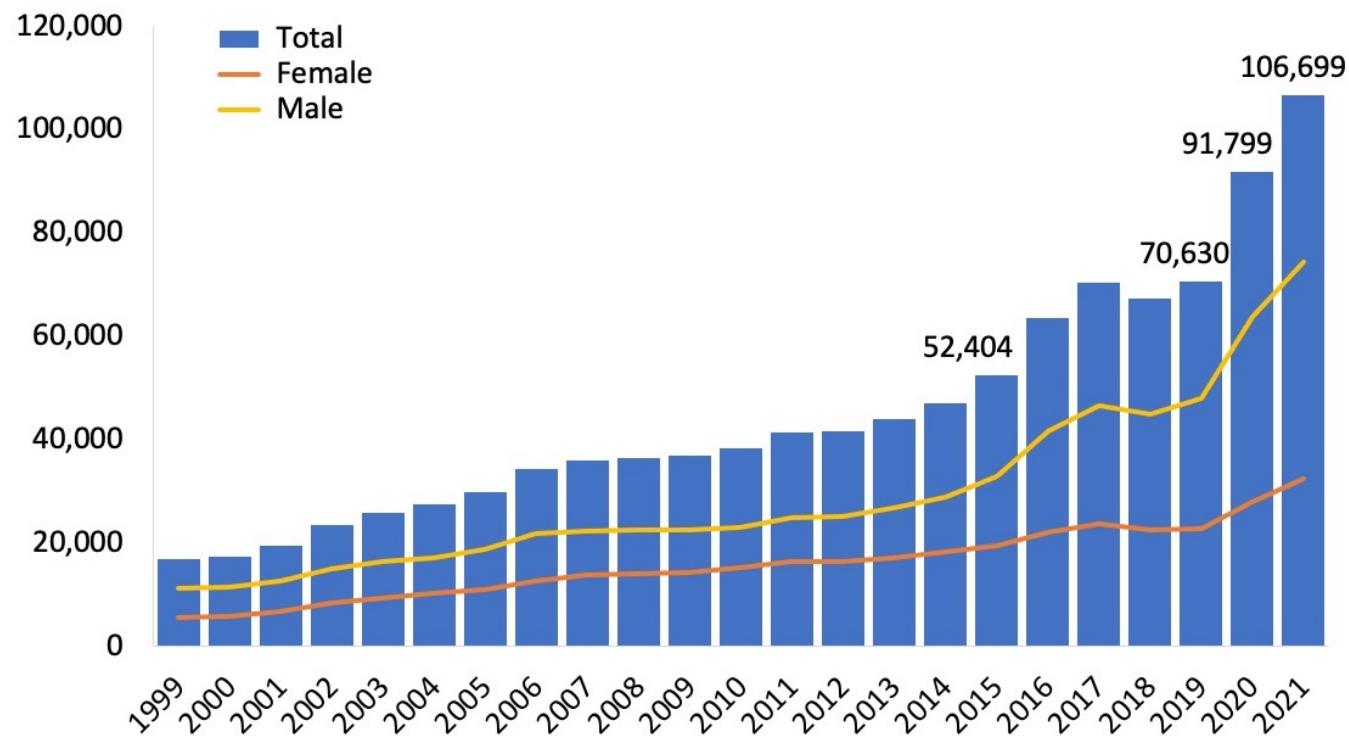
13.6 to 16.0

16.1 to 18.5

18.6 to 21.0

21.1 to 82.0

Figure 1. National Drug-Involved Overdose Deaths*, Number Among All Ages, by Gender, 1999-2021



*Includes deaths with underlying causes of unintentional drug poisoning (X40–X44), suicide drug poisoning (X60–X64), homicide drug poisoning (X85), or drug poisoning of undetermined intent (Y10–Y14), as coded in the International Classification of Diseases, 10th Revision. Source: Centers for Disease Control and Prevention, National Center for Health Statistics. Multiple Cause of Death 1999-2021 on CDC WONDER Online Database, released 1/2023.

Harm Reduction

<https://www.youtube.com/watch?v=ldTxMIblePk>



National Institute on Drug Abuse (NIDA) Videos

Addressing Barriers to Care

https://www.youtube.com/watch?v=2d_3DqMEdEg

What is Harm Reduction?

<https://www.youtube.com/watch?v=ikmKxgCTXFA>

Caring for People Who Use Drugs

<https://www.youtube.com/watch?v=7RI72zvO2dU>

IDEA Syringe Services Program

- The Infectious Disease Elimination Act (IDEA) of 2016 was signed into law on July 1, 2016.
- This policy authorized the University of Miami to open the **first** legal syringe services program (or **needle exchange**) in the state of **Florida**
 - decriminalized possession and distribution of syringes for those engaged in the program
 - 5-year, unfunded pilot project, set to end in July, 2021
- Prior to implementation, REDCap was used to build a data collection system to evaluate the pilot program with the sole purpose of **providing data to the legislature** to enhance advocacy efforts for statewide expansion of the program.

Infectious Disease Elimination Act - REDCap

Current Database

- More than 2,100 people have been helped.
- Lessons Learned
 - Services variables were added.
 - Questions were added and some dropped.
 - Not all variables were validated.
- Using new features in REDCap

New Version for the Other Sites

- a single, fixed/mobile project for each county
- multiple instruments in the project
- follows `tidyverse` naming convention
- removes dropped content
- validation on all text
- numeric calculated fields
 - for needle supply by size
- calculated text and hidden fields using action tags
- automatic alerts for testing (HIV/Hep C)
 - referrals (smart variables for latest info)
- added supply tracking (not shown here)
- greatly improved aesthetics & organization
 - Custom HTML/CSS
 - Field embedding
- Data Access Groups for locations in a county

Show and Tell - Designer

The screenshot shows the REDCap Project Home and Design interface. On the left, there is a sidebar with sections for Project Home and Design, Data Collection, and Applications. A purple arrow points from the 'Designer' link in the Project Home and Design section towards the 'Online Designer' tab at the top. The top navigation bar includes tabs for Project Home, Project Setup, Online Designer (which is highlighted in blue), Data Dictionary, and Codebook. Below the tabs, there is a button to 'Create snapshot of instruments' and a link to a 'VIDEO: How to use this page'. A note states: 'The Online Designer will allow you to make project modifications to fields and data collection instruments very easily using only your web browser. NOTE: While in development status, all field changes will take effect immediately in real time.' The main content area is titled 'Data Collection Instruments' and contains three buttons: '+ Create' (a new instrument from scratch), '+ Import' (a new instrument from the official REDCap Instrument Library), and '+ Upload' (instrument ZIP file from another project/user or external libraries). A table lists existing instruments: 'Enrollment' (98 fields) and 'Daily' (129 fields). Each row has a 'View PDF' button and a 'Choose action' dropdown.

Instrument name	Fields	View PDF	Instrument actions
Enrollment	98		Choose action
Daily	129		Choose action

Show and Tell - Repeated

The screenshot shows the REDCap Project Home and Design interface. The main navigation bar includes Project Home, Project Setup, Other Functionality, and Project Revision History. The Project status is listed as Development, and 6 of 7 steps have been completed.

Main project settings

- Enable Use surveys in this project? (checkbox)
- Enable Use longitudinal data collection with defined events? (checkbox)
- Modify project title, purpose, etc. (button)

VIDEO: How to create and manage a survey

Design your data collection instruments

- Complete! (green checkmark icon)
- Add or edit fields on your data collection instruments. This (online method) or by uploading a Data Dictionary (offline method). [Instruments](#) OR [Download the current Data Dictionary](#)
- Go to [Online Designer](#) or [Data Dictionary](#)
- Have you checked the [Check For Identifiers](#) page to ensure all identifiers are unique?
- Learn how to use [Smart Variables](#), [Piping](#), and [Action Tags](#).

Enable optional modules and customizations

- Complete! (green checkmark icon)
- Modify Repeatable instruments (checkbox selected)
- Enable Auto-numbering for records (checkbox)
- Enable Scheduling module (longitudinal only) (checkbox)
- Enable Randomization module (checkbox)
- Enable Designate an email field for communications (checkbox)
- Additional customizations (button)

Repeatable instruments

An excellent way to collect repeating data in REDCap is to use repeatable instruments and/or repeatable events. This is sometimes called one-to-many data collection. Some examples may include but are not limited to the following: data from multiple visits or observations, concomitant medications, adverse events, or repetitive surveys (daily, weekly, etc.).

Below you can specify a data collection instrument to be infinitely repeatable, which means that an instrument can be repeated over and over again (a different number of times for each record) even without enabling REDCap's longitudinal module. Once an instrument is set to repeat, you will see options on the Record Home Page to add another instance of the instrument for the currently selected record. All instances of a repeating instrument will then be displayed as a table near the bottom of the Record Home Page, thus allowing viewing of the instances and easy navigation within them.

Repeat this instrument?	Instrument name	Custom label for repeating instruments (optional) (i)
<input type="checkbox"/>	Enrollment	<input type="text"/>
<input checked="" type="checkbox"/>	Daily	<input type="text"/>

Save Cancel

Show and Tell - Add Typical

The screenshot shows the 'Project Home and Design' interface. The left sidebar contains the following sections and items:

- Project Home and Design**
 - Project Home · Project Setup
 - Designer · Dictionary · Codebook
 - Project status: Development
- Data Collection**
 - Record Status Dashboard
 - View data collection status of all records
 - Add / Edit Records
 - Create new records or edit/view existing ones
 - Show data collection instruments
- Applications**
 - Project Dashboards
 - Alerts & Notifications
 - Multi-Language Management
 - Calendar
 - Data Exports, Reports, and Stats
 - Data Import Tool
 - Data Comparison Tool
 - Email Logging
 - Field Comment Log
 - File Repository
 - User Rights and DAGs

Add / Edit Records

You may view an existing record/response by selecting it from the drop-down lists below. To create a new record/response, click the button below.

NOTICE: This project is currently in Development status. **Real data should NOT be entered** until the project has been moved to Production status.

Total records: 22

Choose an existing ID Code

-- select record --

+ Add new record

Data Search

Choose a field to search

(excludes multiple choice fields)

All fields

Search query

Begin typing to search the project data, then click an item in the list to navigate to that record.

Normally projects sequentially assign *record_id* values.

Show and Tell - Add

The screenshot shows the 'Project Home and Design' interface. On the left, there's a sidebar with sections for 'Project Home and Design', 'Data Collection', and 'Applications'. Under 'Data Collection', the 'Add / Edit Records' option is selected. Under 'Applications', several tools like Project Dashboards, Alerts & Notifications, and Data Exports are listed. The main content area is titled 'Add / Edit Records' and contains a notice about the project status. It includes fields for choosing or entering an ID code and a data search section.

Add / Edit Records

You may view an existing record/response by selecting it from the drop-down lists below. To create a new record/response, type a new value in the text box below and hit Tab or Enter. To quickly find a record without using the drop-downs, the text box will auto-populate with existing record names as you begin to type in it, allowing you to select it.

NOTICE: This project is currently in Development status. **Real data should NOT be entered** until the project has been moved to Production status.

Total records: 23

Choose an existing ID Code

-- select record --

Enter a new or existing ID Code

Data Search

Choose a field to search

(excludes multiple choice fields)

All fields

Search query

Begin typing to search the project data, then click an item in the list to navigate to that record.

IDEA uses previously
printed random ID
Codes (cards).

Show and Tell - Enrollment

Project Home and Design

- Project Home · Project Setup
- Designer · Dictionary · Codebook
- Project status: Development

Data Collection

- Record Status Dashboard
 - View data collection status of all records
- Add / Edit Records
 - Create new records or edit/view existing ones

ID Code 8977

Select other record

Applications

- Project Dashboards
- Alerts & Notifications
- Multi-language Management

Record Home Page

Record "8977" is a new ID Code. To create the record and begin entering data for it, click any gray status icon below.

The grid below displays the form-by-form progress of data entered for the currently selected record. You may click on the colored status icons to access that form/event.

Legend for status icons:

- | | | | | |
|--|------------|--|----------------------------|--|
| | Incomplete | | Incomplete (no data saved) | |
| | Unverified | | Many statuses (all same) | |
| | Complete | | | |

NEW ID Code 8977

Data Collection Instrument	Status
Enrollment	
Daily	

Show and Tell - Enrollment

REDCap®

Logged in as | Log out

My Projects REDCap Messenger

Project Home and Design

- Project Home · Project Setup
- Designer · Dictionary · Codebook
- Project status: Development

Data Collection

- Record Status Dashboard
- View data collection status of all records
- Add / Edit Records
- Create new records or edit/view existing ones
- ID Code 8977** Select other record
- Data Collection Instruments:
 - Enrollment
 - Daily +
- Lock entire record

Applications

- Project Dashboards
- Alerts & Notifications
- Multi-Language Management
- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool

University of Miami
Office of the Vice Provost for Research

Actions: [Modify instrument](#) [Download PDF of instrument\(s\)](#) [Video: Basic data entry](#)

Enrollment

Editing existing ID Code **8977**.

ID Code	8977
To rename the record, see the record action drop-down at top of the Record Home Page .	
 MILLER SCHOOL of MEDICINE	
Miami IDEA Exchange - Enrollment	
Date	<input type="text" value="2023-04-04"/> <input type="button" value="Today"/> Y-M-D
Enrollment Location:	<input type="text" value="Mobile"/>
If this is a mobile visit, where did it occur?	<input type="text" value="Miami Beach"/>
Required Demographics	
Name Initials	<input type="text" value="..."/>

Save & Exit Form
Save & ...
- Cancel -

Show and Tell - Daily

REDCap®

Logged in as | Log out

My Projects REDCap Messenger

Project Home and Design

- Project Home · Project Setup
- Designer · Dictionary · Codebook
- Project status: Development

Data Collection

- Record Status Dashboard
- Add / Edit Records
- ID Code 8977** Select other record
- Data Collection Instruments:
 - Enrollment
 - Daily +
- Lock entire record

Applications

- Project Dashboards
- Alerts & Notifications
- Multi-Language Management
- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool

University of Miami
Office of the Vice Provost for Research

IDEA - R Medicine 2023 PID 6753

Actions: [Modify instrument](#) [Download PDF of instrument\(s\)](#) [Video: Basic data entry](#)

Daily

Current instance: 1

Editing existing ID Code 8977. (Instance #1)

ID Code	8977
 MIAMI	MILLER SCHOOL of MEDICINE
Miami IDEA Exchange - Daily Log	
Date (Daily)	* must provide value <input type="text" value="2023-04-04"/> <input type="button" value="Today"/> Y-M-D
Daily Location	* must provide value <input type="text" value="Mobile"/>
If this is a mobile daily visit, where did it occur?	* must provide value <input type="text" value="Miami Beach"/>
What type of encounter was this?	<input checked="" type="radio"/> In-person <input type="radio"/> Telehealth <input type="radio"/> Phone <input type="radio"/> Text

Save & Exit Form
Save & ...
- Cancel -

Show and Tell - Naloxone

Naloxone Information

Was person consented for naloxone? * must provide value Yes No

Was naloxone distributed? * must provide value No Yes

Required Referral Information

Referrals (Recommendation) * must provide value

- HIV Test
- HCV Test
- Referral to HIV treatment
- Referral to HCV treatment
- Referral to PrEP
- Referral to Primary Care or FQHC
- Referral to CMHC
- Referral to Detox
- Referral to Residential treatment
- Referral to Shelter
- COVID-19 Test
- IDEA Wound Care
- Other Outpatient treatment
- Referral to other research studies
- Not Referred

Agrees to HIV test today? * must provide value Yes

HIV Test Result * must provide value Positive Negative

Save & Exit Form
Save & ...
- Cancel -

Before and After Field Embedding

Naloxone Information		
Was person consented for naloxone?	<input checked="" type="radio"/> Yes <input type="radio"/> No	reset
Was naloxone distributed?	<input checked="" type="radio"/> Yes <input type="radio"/> No	reset
Type of Narcan Given - Nasal	<input checked="" type="checkbox"/> Nasal	
Amount of Narcan - Nasal	1 <input type="button" value="▼"/>	
Nasal: Distributed to...	<input checked="" type="radio"/> Participant <input type="radio"/> Family Member	reset
Is this a narcan (nasal) refill?	<input checked="" type="radio"/> Yes <input type="radio"/> No	reset
Reason for Narcan Nasal Refill	<input checked="" type="radio"/> Due to loss <input type="radio"/> Due to use	reset
Date of loss - Nasal	2023-05-20 <input type="button" value="Today"/> Y-M-D	
Type of Narcan Given - Auto-Injector	<input checked="" type="checkbox"/> Auto-Injector	
Is this a narcan (injectable) refill?	<input checked="" type="radio"/> Yes <input type="radio"/> No	reset
Type of Narcan Given - Injectable	<input checked="" type="checkbox"/> Injectable	
Amount of Narcan given - Auto-Injector	1 <input type="button" value="▼"/>	
Amount of Narcan given - Injectable	1 <input type="button" value="▼"/>	
Auto-Injector: Distributed to...	<input checked="" type="radio"/> Participant	

Naloxone Information							
Was person consented for naloxone?	<input checked="" type="radio"/> Yes <input type="radio"/> No	reset	Was naloxone distributed?	<input checked="" type="radio"/> Yes <input type="radio"/> No	reset		
Instructions: Check the Type of Naloxone Distributed and Fill in the Details to include the amount distributed, to whom distributed, whether it was a refill, the reason for the refill, the date of loss (if applicable), and the number of reversals (if applicable).							
*** IMPORTANT: Record units <u>NOT</u> Boxes (1 BOX = 2 UNITS) ***							
* must provide value							
Type	Amount (Units)	To Who	Refill	Refill Reason	Date of Loss	Reversals	
<input checked="" type="checkbox"/> Nasal	1 <input type="button" value="▼"/>	<input checked="" type="radio"/> Participant <input type="radio"/> Family Member	<input checked="" type="radio"/> Yes <input type="radio"/> No	<input checked="" type="radio"/> Due to loss <input type="radio"/> Due to use	2023-05-20 <input type="button" value="Today"/> Y-M-D		
<input type="checkbox"/> Auto-Injector	1 <input type="button" value="▼"/>	<input checked="" type="radio"/> Participant <input type="radio"/> Family Member	<input checked="" type="radio"/> Yes <input type="radio"/> No	<input checked="" type="radio"/> Due to loss <input type="radio"/> Due to use	2023-05-20 <input type="button" value="Today"/> Y-M-D		
<input checked="" type="checkbox"/> Injectable	1 <input type="button" value="▼"/>	<input checked="" type="radio"/> Participant <input type="radio"/> Family Member	<input checked="" type="radio"/> Yes <input type="radio"/> No	<input checked="" type="radio"/> Due to loss <input type="radio"/> Due to use	2023-05-20 <input type="button" value="Today"/> Y-M-D		
Total	3 <input type="button" value="▼"/> View equation						

Integrating with R

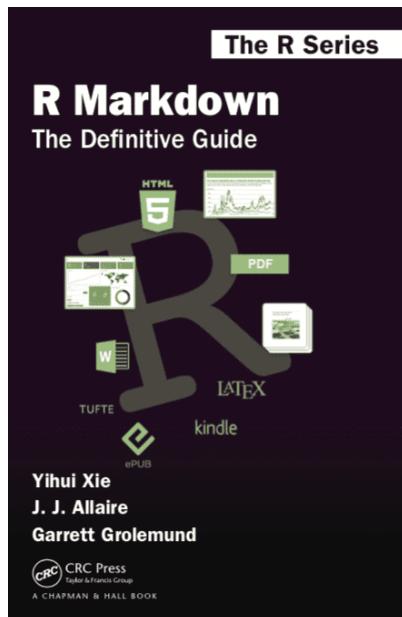
rUM

- I teach in a medical school (i.e., smart people who are not programmers).
- I teach `r` at UM, so to support my students, I give them `rUM`.
- It has menu-driven tools to add R Markdown code or to add complete projects.
- The projects contain:
 - Quarto or R Markdown paper shells for manuscripts
 - full-featured YAML header
 - code to add used packages to the paper bibliography
 - an aggressive `.gitignore` file
 - a data folder

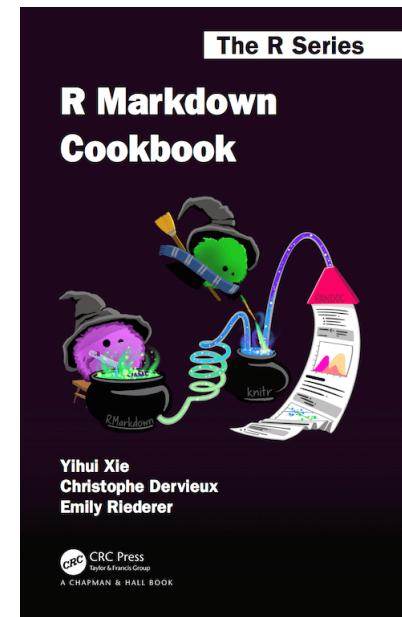
rUM Uses Posit's Brilliance

- rUM uses the Quarto and R Markdown magic that Yihui Xie and his colleagues at Posit/RStudio and around the world have created.

<https://quarto.org/>



[R Markdown: The Definitive Guide](#)



[R Markdown Cookbook](#)

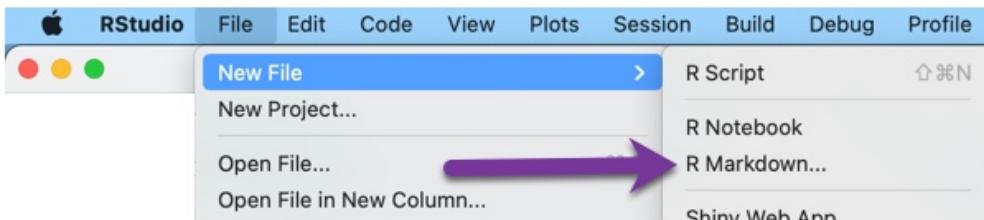
Using rUM

- Install with `remotes::install_cran("rUM")` or `install.packages("rUM")`.
- Completely stop and restart all your instances of RStudio. There are hooks into the graphical user interface, so you need a FULL restart. Expect problems if you just try to restart R.

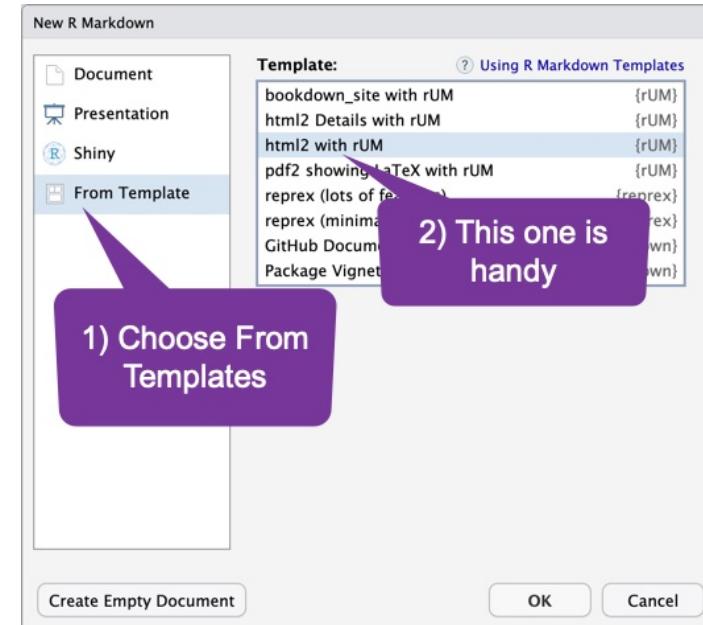
Add a Splash of rUM

- rUM has a bunch of headers that include YAML and setup instructions. A header is setup code that "always" appears at the top of a program.

1) From the RStudio file menu, choose R Markdown...



2) Choose From Template and then choose "html2 with rUM":



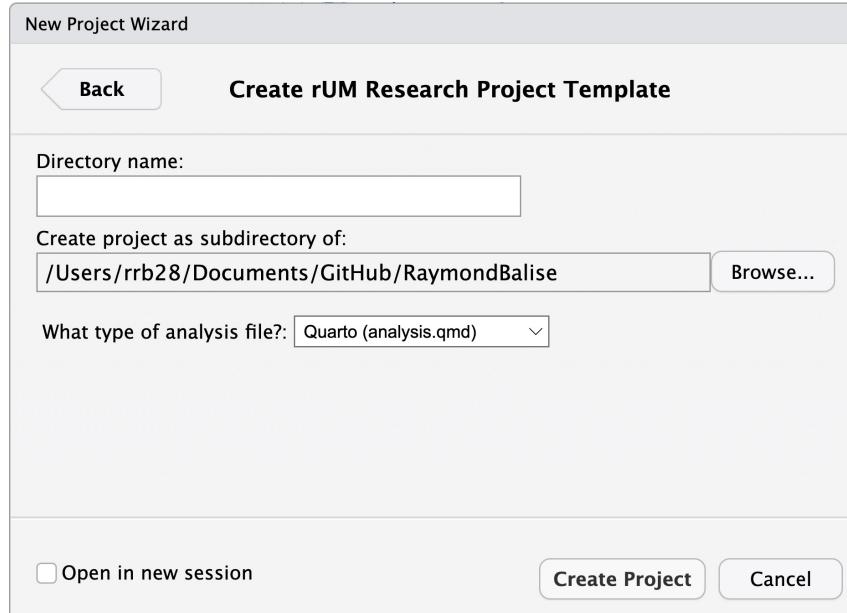
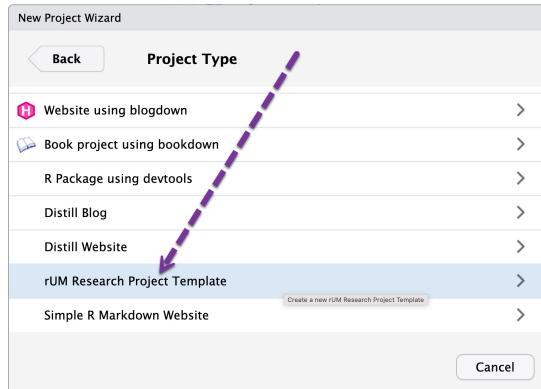
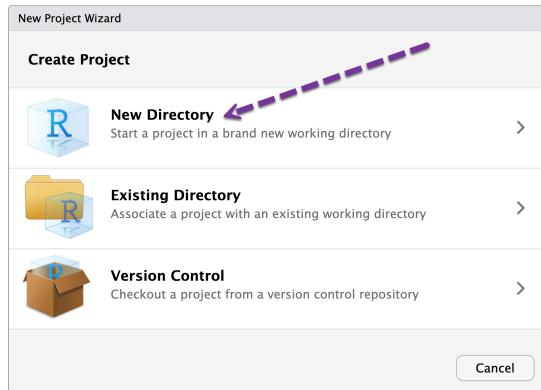
```
---
```

```
title: "html2"
author: "name"
date: "`r Sys.Date()`"
output:
  bookdown::html_document2
---
```

```
```{r setup, include=FALSE}
knitr::opts_chunk$set(
 # These options can be set to FALSE (capitalization matters).
 echo = TRUE, # show code
 message = TRUE, # show messages
 warning = TRUE, # show warnings
 error = TRUE, # show errors
 comment = "", # don't show ## with printed output
)
```



# Make a rUM-Infused Project



```
rUM::make_project(
 path,
 type = c("Quarto (analysis.qmd)", "R Markdown (analysis.Rmd)")
)
```

The sections (with code) are waiting.

An aggressive .gitignore file.

Format for your favorite journal.

Untitled1 x analysis.qmd x

Source Visual Render on Save ABC Render Run Outline

```

```

```
title: "your_title_goes_here"
author: "your_name_goes_here"
date: `r Sys.Date()`
format:
 html:
 self-contained: true
knitr:
 opts_chunk: ##### set global options #####
 collapse: true # keep code from blocks together (if shown)
 echo: false # don't show code
 message: true # show messages
 warning: true # show warnings
 error: true # show error messages
 comment: "" # don't show ## with printed output
 R.options:
 digits: 3 # round to three digits
editor: visual
```

(Top Level) Quarto

Console Terminal Background Jobs

R 4.2.2 · ~/Desktop/blah/

```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

Environment History Connections Tutorial

Files Plots Packages Help Viewer Presentation

New Folder New Blank File Delete Rename More

Home > Desktop > blah

Name	Size	Modified
..	1.4 KB	Apr 6, 2023, 7:51 AM
.gitignore	780.7 KB	Apr 6, 2023, 7:58 AM
analysis.html	1.5 KB	Apr 6, 2023, 7:58 AM
analysis.qmd	276 B	Apr 6, 2023, 7:51 AM
blah.Rproj	4.6 KB	Apr 6, 2023, 7:58 AM
data	1 B	Apr 6, 2023, 7:51 AM
packages.bib	5.5 KB	Apr 6, 2023, 7:51 AM
R		
references.bib		
the-new-england-journal-of-medicine.csl		

# R Markdown

```

```

```
title: "your_title_goes_here"
author: "your_name_goes_here"
date: "`r Sys.Date()`"
output:
 bookdown::html_document2:
 number_sections: false
bibliography: [references.bib, packages.bib]
csl: the-new-england-journal-of-medicine.csl

```

```
{r setup, echo=FALSE}
knitr::opts_chunk$set(
 ##### set global options #####
 echo = FALSE, # don't show code
 collapse = TRUE, # keep code from blocks together (if shown)
 message = TRUE, # show messages
 warning = TRUE, # show warnings
 error = TRUE, # show error messages
 comment = "" # don't show ## with printed output
)

R's default rounding is to show 7 digits. This rounds results to 3 digits.
options(digits = 3)
```

```
{r tidyverse, echo=FALSE}
library(conflicted)
conflict_prefer("filter", "dplyr", quiet = TRUE)
conflict_prefer("lag", "dplyr", quiet = TRUE)
suppressPackageStartupMessages(library(tidyverse))

suppress ``summarise()`` has grouped output by `` messages
options(dplyr.summarise.inform = FALSE)
```

# Quarto

```

```

```
title: "your_title_goes_here"
author: "your_name_goes_here"
date: "`r Sys.Date()`"
format:
 html:
 self-contained: true
knitr:
 opts_chunk: ##### set global options #####
 collapse: true # keep code from blocks together (if shown)
 echo: false # don't show code
 message: true # show messages
 warning: true # show warnings
 error: true # show error messages
 comment: "" # don't show ## with printed output
 R.options:
 digits: 3 # round to three digits
 editor: visual
bibliography: [references.bib, packages.bib]
csl: the-new-england-journal-of-medicine.csl

```

```
{r}
#| label: tidyverse
#| echo: false

library(conflicted)
conflict_prefer("filter", "dplyr", quiet = TRUE)
conflict_prefer("lag", "dplyr", quiet = TRUE)

suppressPackageStartupMessages(library(tidyverse))

suppress ``summarise()`` has grouped output by `` messages
options(dplyr.summarise.inform = FALSE)
```

# The Default Ouput

your\_title\_goes\_here

AUTHOR

your\_name\_goes\_here

PUBLISHED

May 19, 2023

## Introduction

## Methods

Analyses were conducted with R version 4.3.0 with the `tidyverse` (2.0.0), `rUM` (1.0.2), `table1` (1.4.3) packages used to preprocess and summarize data.<sup>1–5</sup>

## Results

## Conclusion

NEJM reference style

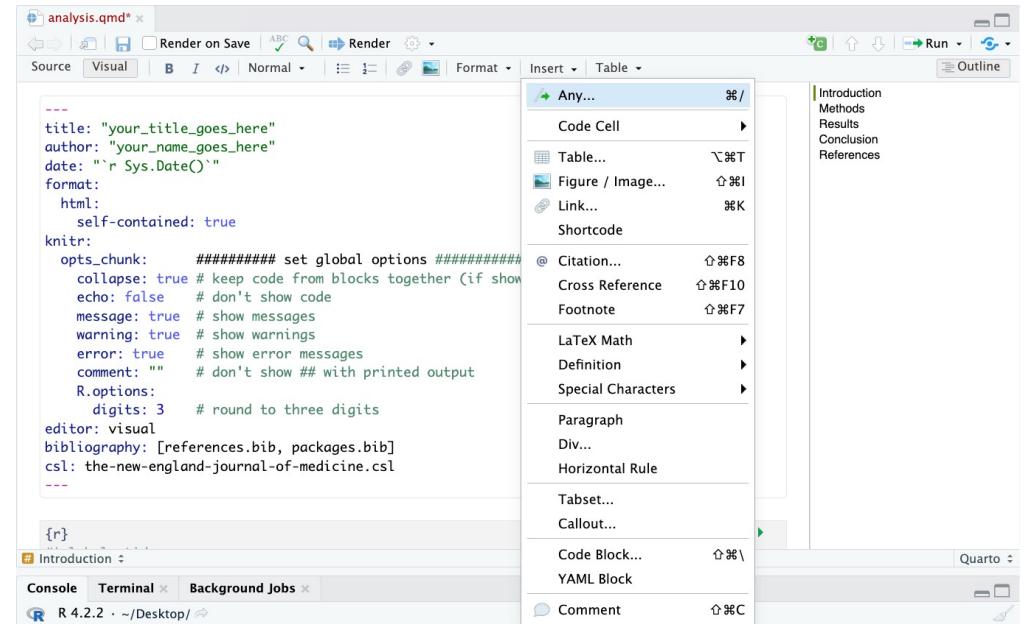
## References

1. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2023. Available from: <https://www.R-project.org/>
2. Wickham H. Tidyverse: Easily install and load the tidyverse. 2023.
3. Wickham H, Averick M, Bryan J, et al. [Welcome to the tidyverse](#). Journal of Open Source Software 2019;4(43):1686.
4. Balise R, Odom G, Grealis K, Cardozo F. rUM: R templates from the university of miami. 2023.
5. Rich B. table1: Tables of descriptive statistics in HTML [Internet]. 2023. Available from: <https://github.com/benjaminrich/table1>

# Use the Visual Editor to Do Stuff

- point-and-click formatting
- basic tables
- images
- add references/citations

In the Visual Editor, typing  / on Mac or cmd / on Linux/Windows is very useful.



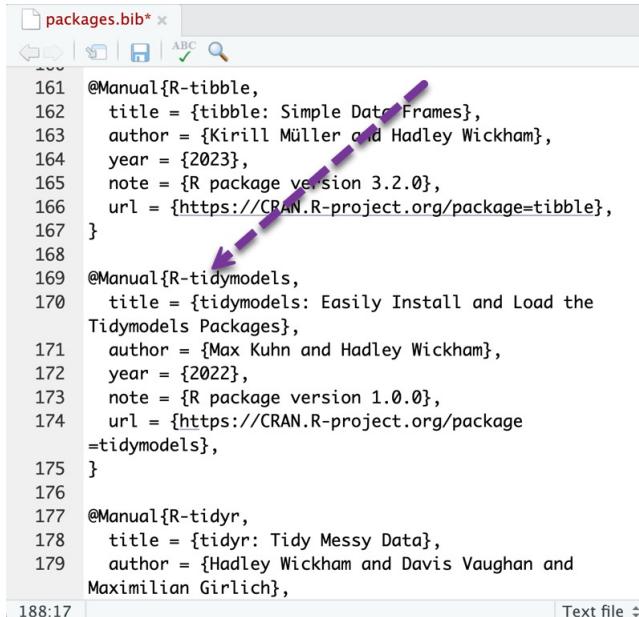
# The Bibliography

- The bibliography section of the template has code to find the packages used in the R Markdown or Quarto file and add them to the "packages.bib" file.

```
automatically create a bib database for loaded R packages & rUM
knitr::write_bib(
 c(
 .packages(),
 "rUM",
 "table1"
),
 "packages.bib"
)
```

# Use Package

- Say you add `library(tidymodels)` to your analysis file.
- After you "knit" or "render" once, the package details will be added to the "packages.bib" file.



```
packages.bib* x
161 @Manual{R-tibble,
162 title = {tibble: Simple Data Frames},
163 author = {Kirill Müller and Hadley Wickham},
164 year = {2023},
165 note = {R package version 3.2.0},
166 url = {https://CRAN.R-project.org/package=tibble},
167 }
168
169 @Manual{R-tidymodels,
170 title = {tidymodels: Easily Install and Load the
171 Tidymodels Packages},
172 author = {Max Kuhn and Hadley Wickham},
173 year = {2022},
174 note = {R package version 1.0.0},
175 url = {https://CRAN.R-project.org/package
176 =tidymodels},
177
178 @Manual{R-tidyr,
179 title = {tidyr: Tidy Messy Data},
180 author = {Hadley Wickham and Davis Vaughan and
181 Maximilian Girlich},
```

# Reference It

- You can copy and paste the reference examples that are in the *Methods* (or point and click) to add it to the text in your methods section.

```
@Manual{R-tidymodels,
 title = {tidymodels: Easily Install and Load the
Tidymodels Packages},
 author = {Max Kuhn and Hadley Wickham},
 year = {2022},
 note = {R package version 1.0.0},
 url = {https://CRAN.R-project.org/package
=tidymodels},
}
```

Analyses were conducted with `r stringr::word(R.Version()\$version.string, 1, 3)` with the `tidyverse` (`r packageVersion("tidyverse")`), `rUM` (`r packageVersion("rUM")`), `table1` (`r packageVersion("table1")`) packages used to preprocess and summarize data. [R-base; R-tidyverse; tidyverse2019; R-rUM; R-table1]. Modeling was done with the `tidymodels` ecosystem. [R-tidymodels]

## Methods

Analyses were conducted with R version 4.2.2 with the `tidyverse` (2.0.0), `rUM` (1.0.2), `table1` (1.4.3) packages used to preprocess and summarize data.<sup>1-5</sup>. Modeling was done with the `tidymodels` ecosystem.<sup>6</sup>

## Results

## Conclusion

### References

1. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2022. Available from: <https://www.R-project.org/>
2. Wickham H. Tidyverse: Easily install and load the tidyverse [Internet]. 2023. Available from: <https://CRAN.R-project.org/package=tidyverse>
3. Wickham H, Averick M, Bryan J, et al. [Welcome to the tidyverse](#). Journal of Open Source Software 2019;4(43):1686.
4. Balise R, Odom G, Greolis K, Cardozo F. rUM: R templates from the university of miami. 2023.
5. Rich B. table1: Tables of descriptive statistics in HTML [Internet]. 2023. Available from: <https://github.com/benjaminrich/table1>
6. Kuhn M, Wickham H. Tidymodels: Easily install and load the tidymodels packages [Internet]. 2022. Available from: <https://CRAN.R-project.org/package=tidymodels>

# Reference Style

By default, the style is set to be New England Journal of Medicine. That is controlled by one line in the YAML header

```
csl: the-new-england-journal-of-medicine.csl
```

and a details file that you can get from here: <https://www.zotero.org/styles>

## Zotero Style Repository

Here you can find [Citation Style Language](#) 1.0.2 citation styles for use with [Zotero](#) and other CSL 1.0.2–compatible software. For more information on using CSL styles with Zotero, see the [Zotero wiki](#).

Style Search      Format: numeric  
JAMA      Fields: medicine  
 Show only unique styles

2 styles found:

- [JAMA \(The Journal of the American Medical Association\)](#) (2013-06-03 16:58:03)
- [JAMA Dermatology](#) (2013-04-29 23:55:49)

Save the CSL file into the project directory and tweak the YAML.

# .gitignore

- I teach workflows that use git and GitHub to students in the medical school. So [rUM](#)-infused projects include my very aggressive **.gitignore** file. It attempts to block common dataset formats (Excel, Python, R, SAS, SPSS, STATA, text, and zip files). Take a look and tell me if I am missing anything.

```
Version 2022-08-12

History files
.Rhistory
.Rapp.history

Session Data files
.RData
*.RData

User-specific files
.Ruserdata

Example code in package build process
*-Ex.R

Output files from R CMD build
/*.tar.gz
```

# REDCap Processing with tidyREDCap

# What tidyREDCap Does

## 1. Easy imports into R from REDCap

- `import_instruments()` imports all data with labels (i.e., the text that appears on the survey page).

## 2. Summarizes categorical variables

- `make_choose_one_table()` makes a summary of a choose-one question.
- `make_choose_all_table()` shows counts of all the options chosen from choose-all questions.
- `make_binary_word()` shows response patterns from choose-all questions.

<https://raymondbalise.github.io/tidyREDCap/>

# Import Everything with `tidyREDCap::import_instruments()`

- `tidyREDCap` allows you to pull in all the instruments by adding a `tidyREDCap::import_instruments()` call. All you need is the **URL** for your copy of REDCap and your **API** token.
- **Do not save your API token in your code.** It is way too easy to accidentally give away your password.
- There are good options to securely store your API keys. If you work on an secure/encrypted machine, a fair alternative is to store your API keys in your *R environment* file. I use: `usethis::edit_r_environ()`

# Store Your API Token

- Copy the token (password text) from REDCap.

The screenshot shows the REDCap Applications page. On the left, there's a sidebar titled "Applications" with various project management tools listed. On the right, under the heading "Your API token", it says: "The API token below is **ONLY for you** and will work **ONLY with this project**. This token allows special access to REDCap data and should **NOT be shared with others**. If you think your token has been compromised, then please contact your REDCap administrator immediately AND either delete or regenerate your token by using the buttons below." Below this text is a large, highlighted API token string: "CZ0H2PQD0L7R2PXXBXKXKXJ1D0". At the bottom, there are two buttons: "Delete token" and "Regenerate token". A large purple arrow points from the top-left towards the API token string.

- Open your R environment file with `usethis::edit_r_environ()` and give your token a meaningful name. Use an = and quote the string.

The screenshot shows an RStudio code editor with a file named ".Renviron\*" open. The code contains a single line of R code: `fake_idea_data_cpdd = "CZ0H2PQD0L7R2PXXBXKXKXJ1D0"`. The RStudio interface includes a toolbar with "Run" and "Run Script" buttons, and a status bar at the bottom.

# Add Code to Load

- Typically you would just do:

```
tidyREDCap::import_instruments(
 "https://redcap.miami.edu/api/",
 Sys.getenv("nacho_anxiety_key")
)
```

- Because this project uses a custom subject ID name and numbering, do:

```
run me once
tidyREDCap::import_instruments(
 url = "https://redcapdemo.vanderbilt.edu/api/",
 token = Sys.getenv("fake_idea_data_cpdd"),
 record_id= "id_code", first_record_id = "07"
)
```

# Loaded Data

The screenshot shows the RStudio interface with the 'Environment', 'History', 'Connections', 'Git', and 'Tutorial' tabs at the top. Below the tabs, there are buttons for 'Import Dataset' and 'Memory'. The 'R' dropdown is set to 'Global Environment'. The 'Data' section lists two datasets:

Dataset	Description
daily	26 obs. of 190 variables
enrollment	23 obs. of 127 variables

- The data is imported and the variables have labels:

	id_code	date Date	is_mobile Enrollment Location:	is_mobile_where If this is a mobile visit, where did it occur?	initials Name Initials
1	07	2016-07-05	Fixed	NA	AB
2	15	2016-09-09	Fixed	NA	EF
3	23	2016-11-16	Fixed	NA	UJ

	id_code	redcap_repeat_instance	da_date Date (Daily)	da_is_mobile Daily Location	da_is_mobile_where If this is a mobile daily visit, where did it occur?
13	45		1	2016-02-23	Fixed
14	45		2	2016-07-22	Fixed
15	45		3	2017-03-02	Mobile
16	52		1	2016-12-25	Mobile
17	52		2	2017-08-21	Mobile
18	65		1	2016-10-10	Fixed
19	65		2	2017-01-14	Mobile
20	65		3	2017-09-04	Mobile

# The Origin of tidyREDCap

- Novices could not process the export of choose-all-that-apply questions.
    - For example, the export has a column for Carfentanil, Cocaine, Crack, etc.
  - How do we get the count of all the chosen things?
  - How can we look at co-occurring patterns?

injection_drugs__1	injection_drugs__2	injection_drugs__3
In the past 30 days, which injection drugs have you used?: ...	In the past 30 days, which injection drugs have you used?: ...	In the past 30 days, which injec
Unchecked In the past 30 days, which injection drugs have you used?: Carfentanil	Unchecked	Unchecked
Unchecked	Checked	Unchecked
Unchecked	Unchecked	Unchecked
Unchecked	Checked	Unchecked
Unchecked	Unchecked	Unchecked
Unchecked	Checked	Unchecked
Unchecked	Unchecked	Unchecked
Unchecked	Checked	Unchecked

# **make\_choose\_one\_table()**

```
library(tidyREDCap)
enrollment |>
 make_choose_one_table(injection_drugs____1) |>
 knitr::kable()
```

## In the past 30 days, which injection drugs have you used?: Carfentanil

<b>Response</b>	<b>n</b>	<b>percent</b>
Checked	1	3%
Unchecked	33	97%

# make\_choose\_all\_table()

- Merge all the choices into a single column and do a count:

```
library(tidyREDCap)
enrollment |>
 make_choose_all_table("injection_drugs__") |>
 knitr::kable()
```

What	Count
Carfentanil	1
Cocaine	5
Crack	2
Fentanyl	9
Heroin	12
Hormones	0
Methamphetamine	6
Pain Killers	0
Prescription Opioids	1
	0

# make\_binary\_word() 1

- Use \_ or a letter to show co-occurring patterns:

Var	Abbreviation	Drug
1	R	Carfentanil
2	C	Cocaine
3	K	Crack
4	F	Fentanyl
5	H	Heroin
7	M	Methamphetamine
8	P	Pain Killers
9	O	Prescription Opioids
10	S	Speedball

pattern	n	percent
R_____	1	2.94%
_C_F____S	1	2.94%
_C_F_____	1	2.94%
_C__HM__S	1	2.94%
_C__H__OS	1	2.94%
_C___M___	1	2.94%
_K_____	2	5.88%
__F_M__S	1	2.94%
__F_____	6	17.65%
__H__S	4	11.76%
__H_____	6	17.65%
___M___	3	8.82%
_____S	1	2.94%
_____	5	14.71%

# make\_binary\_word() 2

```
suppressPackageStartupMessages(library(tidyverse))

labels <-
 c("R", "C", "K", "F", "H", "M", "P", "O", "S")

pattern <-
 enrollment |>
 select(
 injection_drugs__1:injection_drugs__5,
 injection_drugs__7:injection_drugs__10
) |>
 make_binary_word(the_labels = labels)

janitor::tabyl(pattern) |>
 janitor::adorn_pct_formatting(digits =2) |>
 knitr::kable()
```

pattern	n	percent
R_____	1	2.94%
_C_F____S	1	2.94%
_C_F_____	1	2.94%
_C_HM__S	1	2.94%
_C_H_OS	1	2.94%
_C_M____	1	2.94%
_K_____	2	5.88%
_F_M__S	1	2.94%
_F_____	6	17.65%
_H__S	4	11.76%
_H_____	6	17.65%
_M____	3	8.82%
_____S	1	2.94%
_____	5	14.71%

# But wait there is more!

- `tidyREDCap` also has tools for converting checked/unchecked/unknown into yes/no responses.

```
library(tidyREDCap)
make_yes_no(enrollment$injection_drugs___1) |>
 tidyREDCap::make_choose_one_table()
```

```
Response n percent
No or Unknown 33 97%
Yes 1 3%
```

```
make_yes_no_unknown(enrollment$injection_drugs___1) |>
 tidyREDCap::make_choose_one_table()
```

```
Response n percent
No 33 97%
Yes 1 3%
Unknown 0 0%
```

# Thank you!

- IDEA
  - Hansel Tookes, MD MPH, University of Miami
  - Tyler Bartholomew, PhD, University of Miami
  - Belen del Sol Hervera, MPH, University of Miami
- [rUM](#)
  - Gabriel Odom, PhD, Florida International University
  - Kyle Grealis, MS, University of Miami
  - Francisco Cardoza, (almost PhD) University of Miami
  - Frank Gutierrez, MS, University of Miami
- [tidyREDCap](#)
  - Gabriel Odom, PhD, Florida International University
  - Anna Calderon, MS, University of Miami
  - Layla Bouzoubaa, MSPH, Drexel University
  - Wayne DeFreitas, MS, University of Miami
  - Lauren Nahodyl, MS, University of Miami
  - Lionel Henry, Posit
  - Davis Vaughan, Posit
- All of this work
  - Daniel Feaster, PhD, University of Miami
  - The R and R/Medicine community
  - Lori Balise

# Funding That Made This Work Possible

- Healing Communities Study: Developing and Testing an Integrated Approach to Address the Opioid Crisis-New York State. National Institute on Drug Abuse, 1 UM1 DA049415
- CTN-0094 Individual Level Predictive Modeling of Opioid Use Disorder Treatment Outcome. Florida Node Alliance of The Drug Abuse Clinical Trials Network (NIDA) UG1 DA013720
- University of Miami Center for HIV and Researching Mental Health (CHARM) NIH 1P30MH116867-01A1
- University of Miami, Sylvester Comprehensive Cancer Center
- Florida International University, Stempel College of Public Health

# 03 - Helper Modules for R

# Packages

- Because of R's modular design it is easy to build packages for **domain specific** problems.
- R packages
  - typically contain **new code**/functions for data processing/summary/inference
  - have **standardized documentation**
  - can contain **datasets**
  - can contain **long form documentation** or even papers

# Years Cleaning and Harmonizing Data then a Package

- CTN-0094 processed data from the **three largest clinical trials** on Medication Assisted Treatment (**MAT**) for Opioid Use Disorder (**OUD**): CTN-0027, CTN-0030, CTN-0051.
- The documentation was suboptimal (**PDFs**) and the data required many hundreds of fixes.
  - **Murphy's Law** for datasets
  - True problems - impossible and incorrect **dates**
  - Problematic values - **Free text** drug names
- Data scientists with decades of experience at **Columbia and University of Miami** worked independently and meeting to resolve differences.

public.ctn0094data

# The Package

- 25 datasets with documentation
- 3 Long-form documentation vignettes

The screenshot shows the RStudio interface with the following details:

- Toolbar:** Files, Plots, Packages, Help, Viewer, Presentation.
- Search Bar:** R: Deidentified Data from CTN-0094 - Find in Topic
- Title:** Deidentified Data from CTN-0094
- R Logo:** A large blue 'R' inside a grey circle.
- Navigation:** Two small circular icons with arrows.
- Text:** Documentation for package 'public.ctn0094data' version 0.9.0.
- Links:**
  - DESCRIPTION file.
  - User guides, package vignettes and other documentation.
- Section:** Help Pages
- Table of Contents:** A list of dataset names and their descriptions.

<a href="#">all_drugs</a>	All drugs taken
<a href="#">asi</a>	Did patient use intravenous drugs
<a href="#">demographics</a>	Patient demographics
<a href="#">detox</a>	Start and Stop of Detox
<a href="#">everybody</a>	Everybody with any data
<a href="#">fagerstrom</a>	Fagerstrom Test for Nicotine Dependence
<a href="#">first_survey</a>	First Survey Date
<a href="#">meta_study_length</a>	Metadata About Study Length
<a href="#">meta_substance_groups_uds</a>	Metadata About UDS Groupings
<a href="#">pain</a>	Self-Reported Pain
<a href="#">psychiatric</a>	Psychiatric History
<a href="#">qol</a>	Quality of Life
<a href="#">randomization</a>	Randomization Data
<a href="#">rbs</a>	Risk Behavior Survey
<a href="#">rbs_iv</a>	Risk Behavior Survey IV drug use information
<a href="#">screening_date</a>	Screening date information
<a href="#">sex</a>	Sexual Activity in Risk Behavior Survey
<a href="#">site_masked</a>	Site Regrouped
<a href="#">tlfb</a>	Timeline Followback (TLFB) drug use information
<a href="#">treatment</a>	Amount of Study Drug Per Day
<a href="#">uds</a>	Urine Drug Screening (UDS) Results
<a href="#">uds_temp</a>	Urine Drug Screening (UDS) Temperature
<a href="#">visit</a>	Patient Visit Data
<a href="#">withdrawal</a>	Patient Withdrawal Symptoms Per Day
<a href="#">withdrawal_pre_post</a>	Patient Withdrawal Symptoms Pre and Post Induction

# Manual Pages

all\_drugs {public.ctn0094data}

R Documentation

## 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

all\_drugs

### Format

A tibble with 307,523 rows and 4 variables:

#### who

Type: integer

Description: Patient ID

#### what

Type: factor

Description: Name of drug (or alcohol) from self-reported or drug screening

Acetaminophen, Alcohol, Amphetamine, Antibiotic, Antidepressant, Antiemetic,

# Vignettes

## Vignettes and other documentation



Long explanations  
and papers.

Notice the  
code!

### Vignettes from package 'public.ctn0094data'

[public.ctn0094data::final](#)

Final Report CTN-0094

[HTML](#) [source](#) [R code](#)

[public.ctn0094data::harmonization](#)

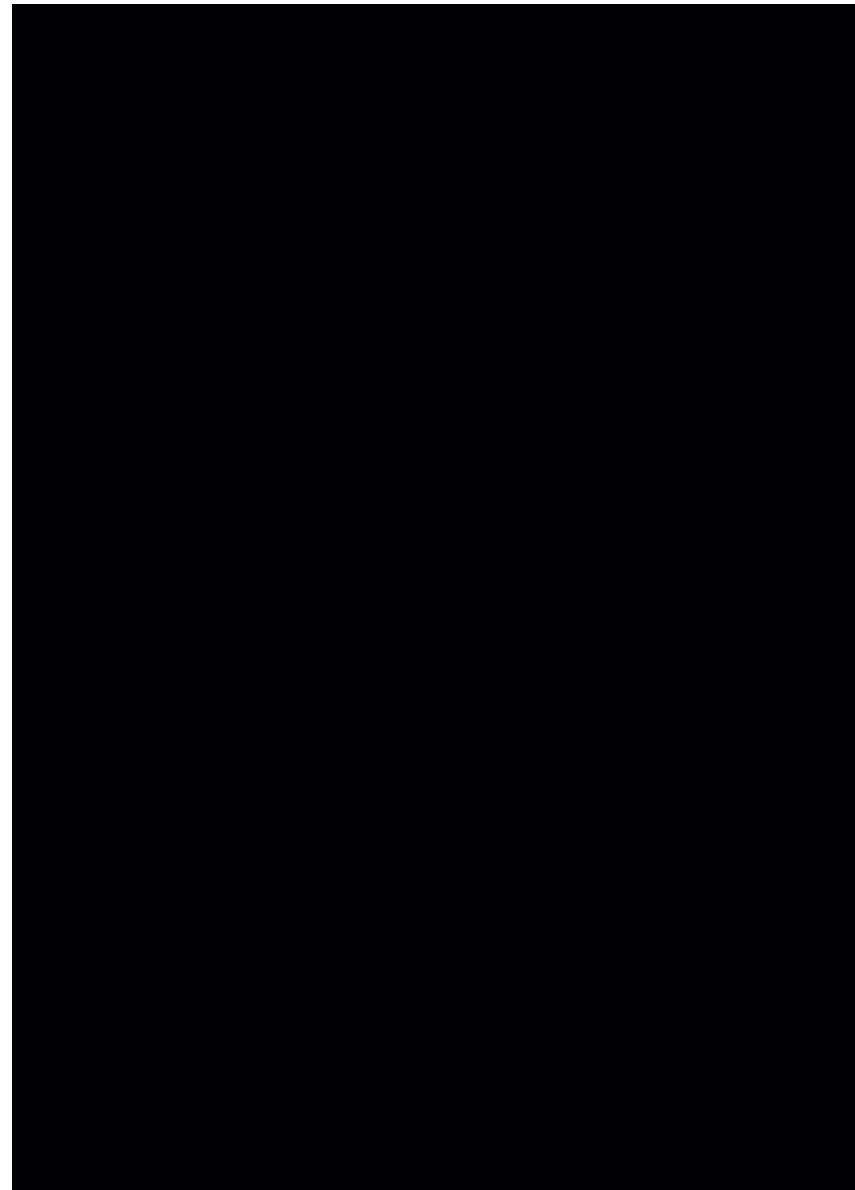
Harmonization Information

[HTML](#) [source](#) [R code](#)

[public.ctn0094data::highStakesManuscript](#)

The Harmonization of “High Stakes”  
Clinical Trial Data: Medication Assisted  
Treatment for Opioid Use Disorder

[HTML](#) [source](#) [R code](#)



# Vignettes - Study Report - The Table Code

# Table 1

Harmonized demographics data from the three National Drug Treatment Clinical Trials Network (CTN) opioid use disorder treatment trials (CTN-0027, CTN-0030, CTN-0051, , N= `r nrow(analysis\_demographics)`)

```
```{r}
analysis_demographics %>%
  select(-who) %>%
  mutate(Sex = fct_drop(Sex)) %>%
  mutate(Sex = fct_drop(Ethnicity)) %>%
  table1(
    ~ Sex + Age + Ethnicity + Race + Education + `Relationship Status` +
      Employment + `Usual Living Arrangements` | project,
    data = .
  )
````
```

# Table 2

Harmonized Urine Drug Screen (UDS) data from the three National Drug Treatment Clinical Trials Network (CTN) opioid use disorder treatment trials

# Vignettes - Harmizonation Details

---

## Harmonization Information

One of the objectives for CTN-0094 was to harmonize data from three clinical trials, ctn\_0027, ctn\_0030, and ctn\_0051. This vignette describes harmonization details and the identification/fixing of problematic values in the trial data. Every dataset in this package has its own documentation. To help protect the anonymity of the study participants two steps were taken. First, the study site information was modified (see the documentation for [site\\_masked](#) for details). Second, all dates have been replaced by the number of days relative to study consent. Therefore, some information, like the day of drug use in the month before enrollment is stored as negative numbers. Below, you will find additional details on the harmonization process. Section headings correspond to data sets.

### The `all_drugs` dataset

The `all_drugs` dataset is an agglomeration of all self-reported drugs, drugs found in urine drug screening and alcohol screening in ctn\_0027, ctn\_0030, and ctn\_0051. This data is the result of extensive preprocessing of free text to harmonize drug names, but drugs were not collapsed into groups. For example, the many descriptions, abbreviations, and spellings of variants of suboxone (e.g., "street suboxone", "bup/nx", "buxnx", "bupnx", "pbupnx", "bupxx") were harmonized into a single "suboxone" group but suboxone was not collapsed with other buprenorphine formulations.

While there were **many** spellings and text variants (including mg and location where drug was administered), the list in [Table 1](#) summarizes the additional text and the changes were made to the free text entries. Many free text entries included the combination of two or more drugs. In these cases, a record was created for each drug. For example, the free text entry of 'Amitriptyline & Trazadone' was converted two records: 'Tryclic-antidepressant' and 'Antidepressant'.

Table 1: Recoded Free Text Descriptions of Drugs.

| Original Text               | Final Text                                  |
|-----------------------------|---------------------------------------------|
| 'Acid'                      | 'Hallucinogen'                              |
| 'Adderall'                  | 'Amphetamine'                               |
| 'Ambien'                    | 'Sedative-Hypnotic'                         |
| 'Amitriptyline & Trazadone' | 'Tryclic-antidepressant' & 'Antidepressant' |
| 'Angel Dust'                | 'PCP'                                       |
| 'Ativan'                    | 'Benzodiazepine'                            |

# Vignettes - Harmizonation Manuscript

## The Harmonization of “High Stakes” Clinical Trial Data: Medication Assisted Treatment for Opioid Use Disorder

CTN-0094 Team

2023-05-13 at 10:57 am

- NOTE: The Abstract is at the bottom of this document.
- Introduction
- Methods
  - Source Data Studies
  - Sample Size and Study Population
  - Survey Domains and Harmonization Techniques
  - Database Structure
  - Social and Demographic Characteristics
    - Age
    - Gender and Sex
    - Race and Ethnicity
    - Additional Sociodemographic Information
  - Timeline Follow-Back (TLFB)
  - Urine Drug Screening (UDS)
  - Nicotine Use

DOPE

# Drug Ontology Parsing Engine is DOPE

- In the CTN-0094 Timeline Follow-back data, there were free text fields for drugs.
- We saw brand, generic and street names: "Acid", "Angel Dust", "Ativan", etc.
- We combined data from the DEA, IQVIA and the "no slang" online street name dictionary.
- The DOPE function `lookup()` allows us to query various drug name databases for matching names, classes, and categories.
- The DOPE package is here: <https://ctn-0094.github.io/DOPE/>

# Parsing Free-Text Fields with `lookup()`

```
suppressPackageStartupMessages(library(DOPE)) # for the lookup() function
c("Acid", "Angel Dust", "Ativan", "Xylazine") %>%
 lookup() %>%
 kable()
```

| original_word | class        | category       | synonym    |
|---------------|--------------|----------------|------------|
| acid          | hallucinogen | lsd            | acid       |
| angel dust    | hallucinogen | pcp            | angel dust |
| ativan        | depressant   | benzodiazepine | ativan     |
| xylazine      | NA           | NA             | NA         |

- For CTN-0094 data analysis (including work on polysubstance use), we grouped reported substances by the `class` field
- Note that Xylazine doesn't exist in our database yet, so let us know if you're interested and want to [get involved](#).



# Drug Ontology Parsing Engine (DOPE)



The goal of DOPE is to provide a structured vocabulary and tools to look up details on drugs tracked by the DEA.  
The data structure is:



## Links

Download from CRAN at  
[https://cloud.r-project.org/  
package=DOPE](https://cloud.r-project.org/package=DOPE)

Browse source code at  
<https://github.com/CTN-0094/DOPE/>

Report a bug at  
[https://github.com/CTN-0094/DOPE/  
issues](https://github.com/CTN-0094/DOPE/issues)

## License

Full license

MIT + file LICENSE

## Developers

Raymond Balise  
Author, maintainer

Layla Bouzoubaa  
Author 

Gabriel Odom  
Author 

Nathaniel Castor  
Author

## Dev status

CRAN 2.1.0

lifecycle maturing

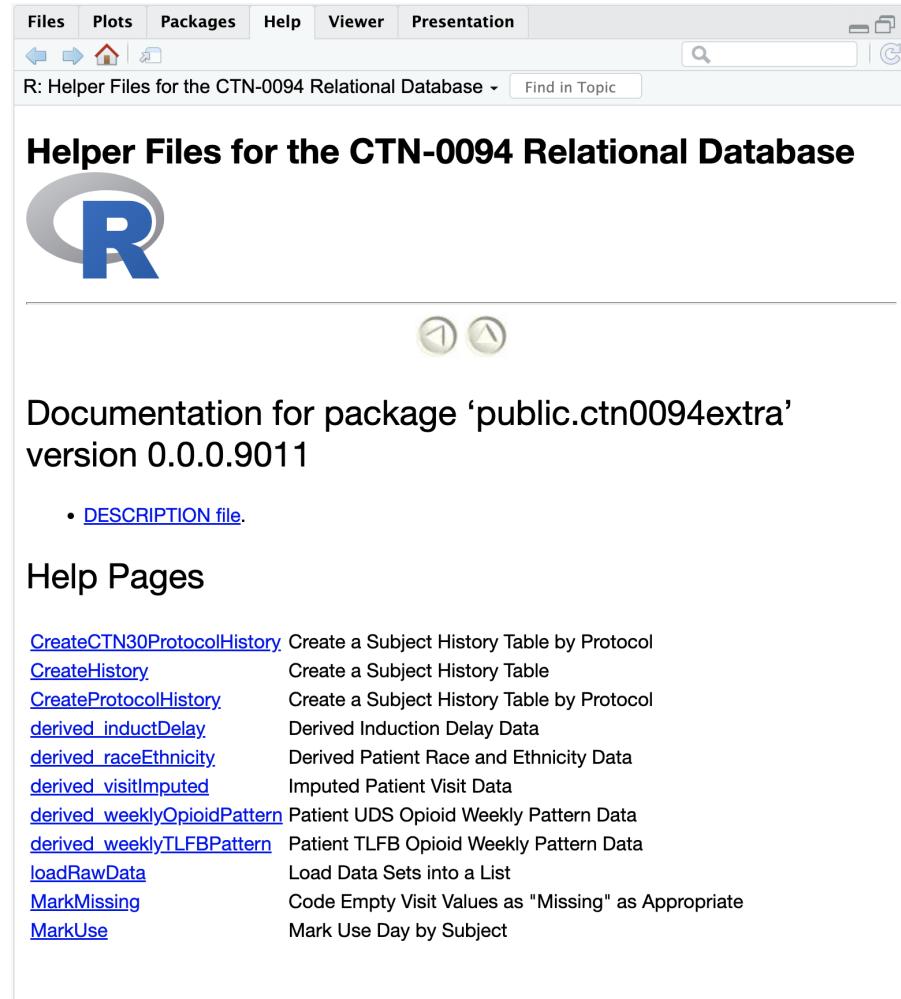
# We need help with **DOPE!**

- This is an **unfunded** passion project. We would love to find funding for this work.
- We want **more dictionaries** of slang terms.
- We need a **contact at the DEA** to help us sort through their publications.
- We need subject matter experts!

public.ctn0094extra

# The CTN-0094 "Extra" Package

- The `public.ctn0094data` package is a relational database.
- This database does not have "engineered" variables, so we found ourselves constantly re-calculating certain measurements for our analyses.
- The `public.ctn0094data` package contains "helper" datasets and functions.



The screenshot shows the R Help Browser interface. The title bar reads "R: Helper Files for the CTN-0094 Relational Database". The main content area is titled "Helper Files for the CTN-0094 Relational Database" and features the R logo. Below the title, there are two small circular icons with arrows. The text "Documentation for package 'public.ctn0094extra'" is followed by "version 0.0.0.9011". A bullet point lists "• [DESCRIPTION file](#)". The section "Help Pages" is expanded, showing a list of functions with their descriptions:

|                                             |                                                     |
|---------------------------------------------|-----------------------------------------------------|
| <a href="#">CreateCTN30ProtocolHistory</a>  | Create a Subject History Table by Protocol          |
| <a href="#">CreateHistory</a>               | Create a Subject History Table                      |
| <a href="#">CreateProtocolHistory</a>       | Create a Subject History Table by Protocol          |
| <a href="#">derived_inductDelay</a>         | Derived Induction Delay Data                        |
| <a href="#">derived_raceEthnicity</a>       | Derived Patient Race and Ethnicity Data             |
| <a href="#">derived_visitImputed</a>        | Imputed Patient Visit Data                          |
| <a href="#">derived_weeklyOpioidPattern</a> | Patient UDS Opioid Weekly Pattern Data              |
| <a href="#">derived_weeklyTLFBPattern</a>   | Patient TLFB Opioid Weekly Pattern Data             |
| <a href="#">loadRawData</a>                 | Load Data Sets into a List                          |
| <a href="#">MarkMissing</a>                 | Code Empty Visit Values as "Missing" as Appropriate |
| <a href="#">MarkUse</a>                     | Mark Use Day by Subject                             |

A screenshot of a data visualization interface titled "filter(public.ctn0094data::all\_drugs... x)". The table has four columns: "who", "what", "source", and "when". The data shows various substance uses over time for a single participant. The "when" column includes some numerical values and a few entries with a small gray box containing the number "3".

|    | who | what           | source | when |
|----|-----|----------------|--------|------|
| 1  | 2   | Benzodiazepine | UDS    | 0    |
| 2  | 2   | Opioid         | UDS    | 0    |
| 3  | 2   | Oxycodone      | TFB    | 0    |
| 4  | 2   | Oxycodone      | UDS    | 0    |
| 5  | 2   | Thc            | UDS    | 0    |
| 6  | 2   | Fentanyl       | TFB    | 1    |
| 7  | 2   | Oxycodone      | TFB    | 1    |
| 8  | 2   | Thc            | TFB    | 1    |
| 9  | 2   | Oxycodone      | TFB    | 2    |
| 10 | 2   | Benzodiazepine | TFB    | 3    |
| 11 | 2   | Oxycodone      | TFB    | 3    |
| 12 | 2   | Alcohol        | TFB    | 4    |
| 13 | 2   | Benzodiazepine | TFB    | 4    |
| 14 | 2   | Oxycodone      | TFB    | 5    |
| 15 | 2   | Oxycodone      | TFB    | 6    |
| 16 | 2   | Benzodiazepine | UDS    | 7    |
| 17 | 2   | Opioid         | UDS    | 7    |
| 18 | 2   | Oxycodone      | UDS    | 7    |
| 19 | 2   | Thc            | UDS    | 7    |
| 20 | 2   | Alcohol        | TFB    | 9    |
| 21 | 2   | Alcohol        | TFB    | 10   |
| 22 | 2   | Alcohol        | TFB    | 12   |
| 23 | 2   | Thc            | UDS    | 13   |
| 24 | 2   | Alcohol        | TFB    | 16   |
| 25 | 2   | Alcohol        | TFB    | 18   |
| 26 | 2   | Benzodiazepine | UDS    | 20   |
| 27 | 2   | Thc            | UDS    | 20   |
| 28 | 2   | Alcohol        | TFB    | 23   |
| 29 | 2   | Thc            | TFB    | 23   |
| 30 | 2   | Thc            | TFB    | 25   |

# Summarizing Drug Use Patterns

- The data shown are only a few records for a single trial participant. **Is the treatment working?**
- Drawing a quick clinical conclusion from this data is nearly impossible.
- The `public.ctn0094extra` package contains functions to turn the long-form data on the left into a compact summary of patient use.
- Substance use pattern "words" are a compact summary of their opioid use over the course of the clinical trial.

# Substance Use Pattern "Words"

- We summarize a participant's use of a class of substances for a single week in the trial protocol via the following "letters":
  - **+**: positive for the substance(s) by urine screen (or participant self-report/TLFB, if such data are of interest) at a specific visit or in a specified window of time (a day, week, etc.)
  - **-**: negative for the substance(s)
  - **0**: patient failed to provide a urine sample
  - **\***: inconclusive results or mixed results (e.g., patient provided more than one urine sample in the time interval and the results did not agree)
  - **\_**: no specimens required per study design (weekends, holidays, pre-randomization period, alternating visit days/weeks, clinic excused visits)
- For subject 2, their **opioid use pattern** during treatment is: +----○○-○-○-○+○

# Properties of "Words"

Because these patterns are being used to replace aspects of patient medical charts in analysis, they must have the following properties:

1. (*Machine-Readable*) It can be directly parsed by a **computer**.
2. (*Human-Readable*) It can be quickly and easily interpreted by a **human**.
3. (*Sufficient*) It represents **all of the same information** about the [univariate] sequence of interest that would be present in the full data.

# Summarized Opioid Use

We can compactly represent the substance use pattern (in this case opioid use pattern) for all the participants across various trial protocol phases in an easy-to-read form. Here are 10 real patients we chose as an example:

```
A tibble: 10 × 2
Subject Treatment UDS
<dbl> <chr>
1 4 -----
2 13 -----o-oooooooooooo
3 17 ---+*++++++-++++++-+-
4 33 +---o-o-o-o+---o---o---o-o-o-o-o-
5 163 -o---o---o---o+-----
6 210 -++++++-+-----+
7 233 *++++++-+o++++++-+o
8 242 -----
9 1103 +---oo---o-+-o-----o-o-oo++o
10 2089 +----+-----o-
```

- Which of these patients succeeded in treatment?
- Which patients failed?

CTNote

# The CTNote Package

- The `CTNote` package contains outcomes, treatment success measures, and endpoints used in Clinical Trials Network (CTN) studies.
- The three vignettes form a library of 50+ coded trial endpoints used in various OUD clinical trials since the early 1970s.
- The standard wisdom is to group these algorithms into metrics of **Abstinence, Relapse, and Use Reduction**
- <https://ctn-0094.github.io/CTNote/>

The screenshot shows the R help viewer interface. The title bar reads "R: CTN Outcomes, Treatments, and Endpoints". The main content area is titled "CTN Outcomes, Treatments, and Endpoints" and features the R logo. Below the title, there are two small circular icons with arrows. The text "Documentation for package 'CTNote' version 0.1.3" is displayed. A bullet point lists a link to the "DESCRIPTION file". The "Help Pages" section lists several functions with their descriptions:

|                                           |                                                           |
|-------------------------------------------|-----------------------------------------------------------|
| <a href="#">collapse_lattice</a>          | Combine Multiple Simple Study Design "Lattices"           |
| <a href="#">count_matches</a>             | Count Periods of Substance Use / Abstinence               |
| <a href="#">detect_in_window</a>          | Detect Visits before Matching a Fuzzy Sub-Pattern         |
| <a href="#">detect_subpattern</a>         | Detect a Consecutive Sub-Pattern                          |
| <a href="#">egOpioidsCTN0094</a>          | Opioid Use by Study Day for Example CTN-0094 Participants |
| <a href="#">impute_missing_visits</a>     | Naively Impute Missing Visits                             |
| <a href="#">measure_abstinence_period</a> | Measure the Length of the Longest Abstinent Period        |
| <a href="#">measure_retention</a>         | Measure Length of Use Pattern before Dropout              |
| <a href="#">outcomesCTN0094</a>           | All Treatment Outcomes for CTN-0094 Participants          |
| <a href="#">recode_missing_visits</a>     | Recode Missing or Ambiguous UDS in a Subject Use Pattern  |
| <a href="#">view_by_lattice</a>           | View a Pattern through a Study Design "Lattice"           |
| <a href="#">weight_positive_visits</a>    | Weight Visits in a Subject Use Pattern                    |

# Why So Many Endpoints?

```
 opioidPattern_df
```

```
A tibble: 10 × 2
Subject Treatment UDS
<dbl> <chr>
1 4 -----
2 13 -----
3 17 ---+*++++++-++++++-+
4 33 +---o-o-o-o+---o---o---o-o-o-o-
5 163 -o---o---o---o+-----
6 210 -++++++-+-----+
7 233 *++++++-+o++++++-+o
8 242 -----
9 1103 +---oo---o---+---o-----o-o-oo++o
10 2089 +---+-----o-
```

- When you decided who "succeeded" and who "failed", did you look at:
  - End-of-study strict abstinence?
  - Not relapsing while in treatment?
  - Longest time of abstinence?
  - Treatment retention?
  - Greatest proportion of negative UDS?
- Does a missed visit imply substance use?
- Did we all agree on how to measure success?

# Defining Treatment Endpoints in Code

- Prof. Laura Brandt spent two years cataloging every opioid use disorder clinical trial outcome used between 1970 and 2020.
- We partnered with Brandt to create a library of computer code in the R language to explicitly define and calculate each of these outcomes given a patient-specific use pattern "word".
- The code libraries are published here:
  - [Abstinence outcomes](#)
  - [Relapse outcomes](#)
  - [Use Reduction outcomes](#)

# Example Outcome Code and Output

Schottenfeld et al. (2005)

**Definition:** Proportion of negative UOS; exclude missing UDS

```
outcomesRed_df <-
 outcomesRed_df %>%
 rowwise() %>%
 # Exclude missing
 mutate(
 usePatternPresent = recode_missing_visits(
 usePatternUDS,
 missing_becomes = "")
)
) %>%
Count negative
 mutate(
 Rd_schottenfeld_2005 = count_matches(
 use_pattern = usePatternPresent,
 match_is = "-"
 # Mixed results weeks count as half of a negative week
 mixed_results_are = "*", mixed_weight = 0.5,
 proportion = TRUE
)
) %>%
 select(who, Rd_schottenfeld_2005) %>%
 left_join(outcomesRed_df, ., by = "who")
```

This is a measure of opioid use reduction  
(where smaller numbers are better).

```
A tibble: 10 × 3
who usePatternUDS
<dbl> <chr>
1 1 0ooooooooooooooo
2 4 -----0-0-0
3 13 -----0-000000000
4 17 -++*++++++-----
5 163 -0---0---0+----
6 210 -++++++-----
7 233 *++++++0++++++0
8 242 -----
9 1103 +--+0--0-++-0-----0-0-00++0
10 2089 +----0-
```

Rd\_schottenfeld\_2005  
<dbl>

| who | usePatternUDS               | Rd_schottenfeld_2005 |
|-----|-----------------------------|----------------------|
| 1   | 0ooooooooooooooo            | 0                    |
| 2   | -----0-0-0                  | 1                    |
| 3   | -----0-000000000            | 1                    |
| 4   | -++*++++++-----             | 0.239                |
| 5   | -0---0---0+----             | 0.95                 |
| 6   | -++++++-----                | 0.542                |
| 7   | *++++++0++++++0             | 0.0227               |
| 8   | -----                       | 1                    |
| 9   | +--+0--0-++-0-----0-0-00++0 | 0.769                |
| 10  | +----0-                     | 0.783                |

## Discuss:

- Is patient 2089 really doing worse than patient 0013?
- If "end of treatment abstinence" were used, would patient 0004 still be a success?



# Introduction to CTNote: The CTN Outcomes, Treatments, and Endpoints Library

The package `CTNote` exists as a comprehensive implementation of *outcomes*, *treatments*, and *endpoints* used in clinical trials to evaluate the efficacy of medication-assisted treatment of substance use disorders (SUDs). The functions in this package are programmatic building blocks with which to construct algorithms which calculate single-value summaries of clinical trial participants' substance use patterns. We assume that these substance use patterns are symbolized as a "word" (see below). The vignettes for this package include a standard and code-based library of algorithms for treatment outcome definitions useful to evaluate medication-based treatments for SUDs.

## Installation

You can install the released version of `CTNote` from [CRAN](#) with:

```
install.packages("CTNote")
```

## Use Pattern "Words"

If you wish to apply the functions in this package to create algorithms to calculate endpoints for your data, the participants' substance use patterns must be stored in the "substance use pattern word" format shown below.

We define the following five-value legend:

### License

[Full license](#)

[GPL-3](#)

### Table of contents

[Installation](#)

[Use Pattern "Words"](#)

[Example](#)

### Developers

[Gabriel Odom](#)

Author, maintainer

[Laura Brandt](#)

Author

[Raymond Balise](#)

Author

[More about authors...](#)

### Funding

NIMHD FIU-RCMI Pilot

AWD000000009108; NIDA

UG1DA013035, U10DA13720

### Citation

Wrapping Up

# Tools We Learned About

- **REDCap** can help you collect structured data; we recommend using REDCap well to prevent "garbage in: garbage out" in your analyses.
- The R package [rUM](#) (available on CRAN) can set up a research project directory complete with a manuscript shell on your computer and is a valuable tool to help teach new researchers how to organize files for a reproducible manuscript.
- [Quarto](#) is the next generation of literate programming and reproducible scientific manuscript preparation. It comes packaged with the newest release of RStudio (or available as a stand-alone).
- The R packages [public.ctn0094data](#), [public.ctn0094extra](#), [DOPE](#), and [CTNote](#) can be helpful to a) show examples of clinical trials data, b) clean interview notes and other free-text fields where drugs are discussed, and c) summarize long-form UDS/Timeline Followback data and calculate endpoints from these summaries.

# Thank you!

Questions?