



Beginner's Guide



THE UNIVERSITY OF
CHICAGO MEDICINE &
BIOLOGICAL SCIENCES



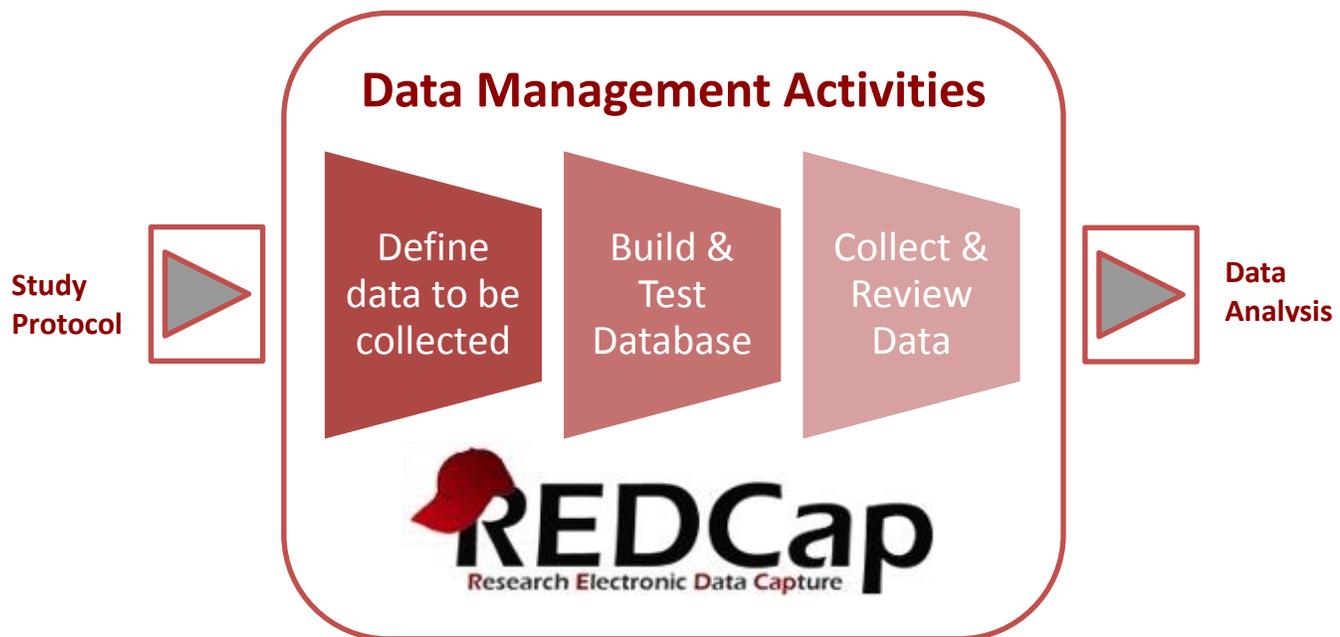
CENTER FOR
RESEARCH
INFORMATICS

Table of Contents

1. INTRODUCTION	3
2. GENERAL INFORMATION	3
2.1 Access to REDCap	3
2.2 Compliance.....	3
2.3 Language for grant and IRB submission.....	4
2.4 Training	4
2.5 Support.....	4
3. HOW TO USE REDCAP	4
3.1 Definition and wording	4
3.2 Project workflow	5
3.3 Build and test a project	5
3.3.1 Prerequisites	5
3.3.2 Build a project.....	6
3.3.3 Test a project	111
3.4 Move a project into production mode.....	111
3.5 Enter, Review and analyze data.....	111
3.5.1 Enter data	111
3.5.2 Review data.....	122
3.5.3 Analyze data.....	133
3.5.4 De-identification methods for data exporting.....	14
3.6 Changes post-production	144
3.7 Close a project.....	144

1. INTRODUCTION

REDCap is a data management tool that should be used primarily for Research and Quality Improvement studies. **REDCap must not be used for collecting clinical data used as primary records for a patient's care.** The purpose of this document is to provide general guidance on the use of REDCap at University of Chicago.



2. GENERAL INFORMATION

2.1 Access to REDCap

If you have a BSDAD or UCHAD account, you may log in to REDCap using either one of those accounts. If you do not have either a BSDAD or UCHAD account (this includes external collaborators from other institutions), a sponsor with a BSD account may request access for you via this [online request form](#). Our CRI Help Desk will confirm the request with the sponsor and then email your login credentials within 24-48 hours.

REDCap is a web-based application and can be used from anywhere, outside the University network, and from any device (laptop, smartphone, tablets, iPad). The application URL is: <https://redcap.uchicago.edu/>

2.2 Compliance

REDCap is designed with built-in features to address confidentiality and compliance requirements.

- a. Access control - Access is based on the University's BSD and UCHAD Active Directories. External collaborators who were granted access are added to the BSD Active Directory.
At the project level, access control is fully configurable: a project owner can authorize study staff access, as required per IRB protocol. Identifiers can be flagged and removed from data exports, depending on user profiles. See section 3.3.4 for more details on user rights.
- b. Audit trail - Every operation performed in REDCap is tracked and can be easily retrieved at the project level. (See logging feature in REDCap).

- c. Centralization of data on local servers - The REDCap application and data are both hosted on University secure servers. A backup is performed nightly.

2.3 Language for grant and IRB submission

Boilerplates for citing REDCap, or for grant & IRB documentation regarding REDCap are available at:

<http://cri.uchicago.edu/faq/#acknowledgeredcap>

2.4 Training

There is no formal training required to use REDCap but video tutorials and help text are available within each module in REDCap. Also, the 'Help & FAQ' tab in REDCap contains a lot of useful information. Additional help is available on the University's help site at: <http://cri.uchicago.edu/redcap-training/>

2.5 Support

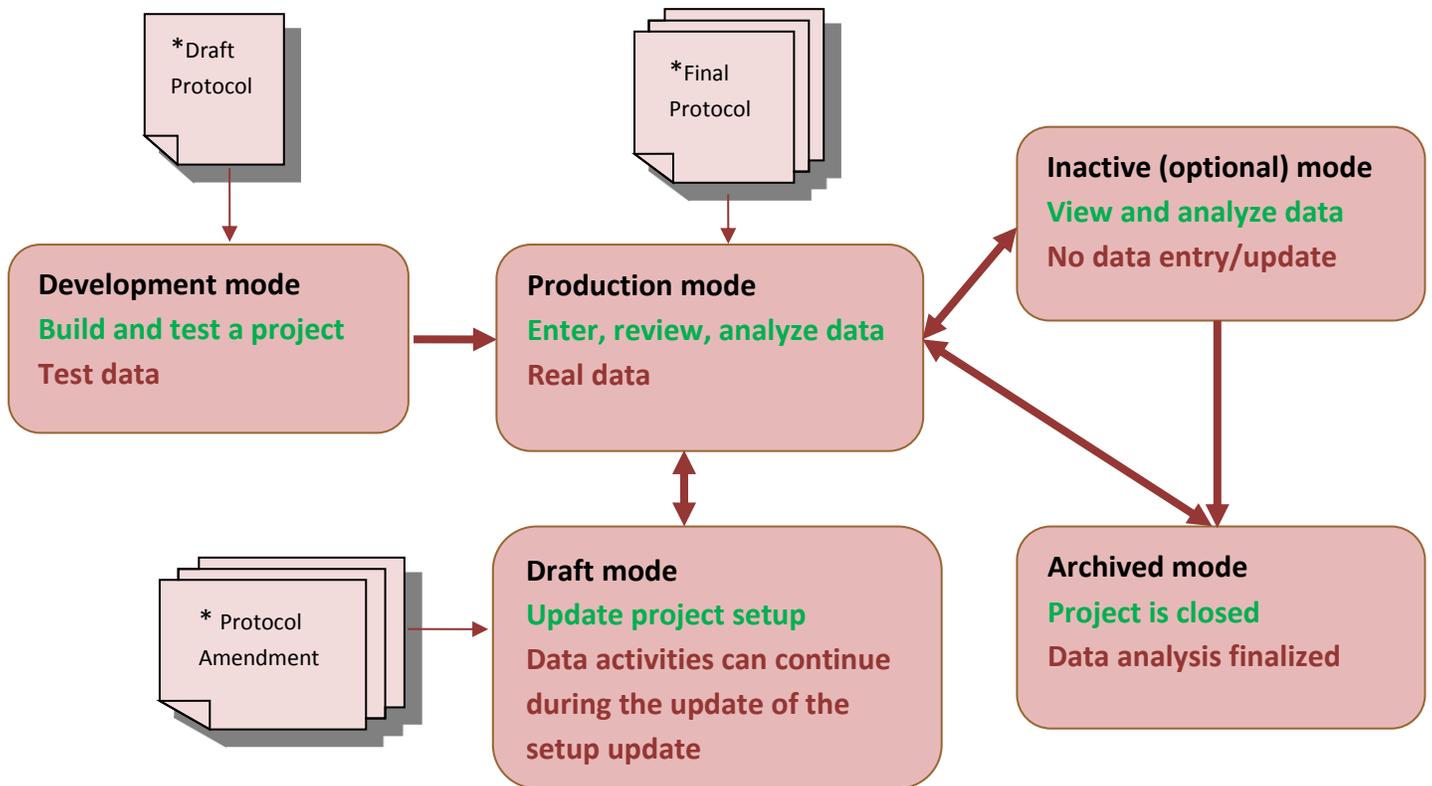
The REDCap application is developed and deployed by Vanderbilt University. The University's Center for Research Informatics (CRI) is responsible for the maintenance and the administration of the application. CRI provides support to all users: one-on-one consultation, answers to questions/issues, centralization of project change requests, guidance on data management best practices, and downtime and upgrade communications.

3. HOW TO USE REDCAP

3.1 Definition and wording

- **Arms:** group events into 'arms'; there may be one or more arms/groups per project. Each arm can have as many events as needed.
- **Branching logic:** branching Logic may be employed when fields/questions need to be hidden for data entry under certain conditions.
- **Calendar:** project calendar to help organize the patients' scheduling and keep track of any upcoming events.
- **CRF:** Case Report Form.
- **Data Dictionary:** Excel file, containing the list of fields of a given project and their associated attributes.
- **Events:** allow for the utilization of data collection forms multiple times for any given project record (often used when collecting longitudinal data). An 'event' may be a temporal event in the course of a project, such as a participant visit or a task to be performed.
- **Field (or variable or question)** may be used interchangeably and represent unique items of data to be collected and analyzed.
- **File repository:** repository that stores and retrieves files and documents used for a project. Whenever a data export is performed, the resulting data and syntax files are also stored in the file repository of the project.
- **Instrument:** survey page or data entry form.
- **Logging:** module that lists all changes made to this project, including data exports, data changes, and the creation or deletion of users (audit trail).
- **Project status:** development, production, draft, inactive and archived.
- **Project type:** single survey or data entry form(s) or a mix of a single survey / data entry form(s).
- **Record label:** information/variables added to the unique ID of the study (e.g. pat_id) to help select the right record during data entry. For example, date of birth or last name can be added as record labels when selecting a subject for data entry. Record labels are displayed only and have no impact when exporting data.

3.2 Project workflow



* if IRB approval required

3.3 Build and test a project

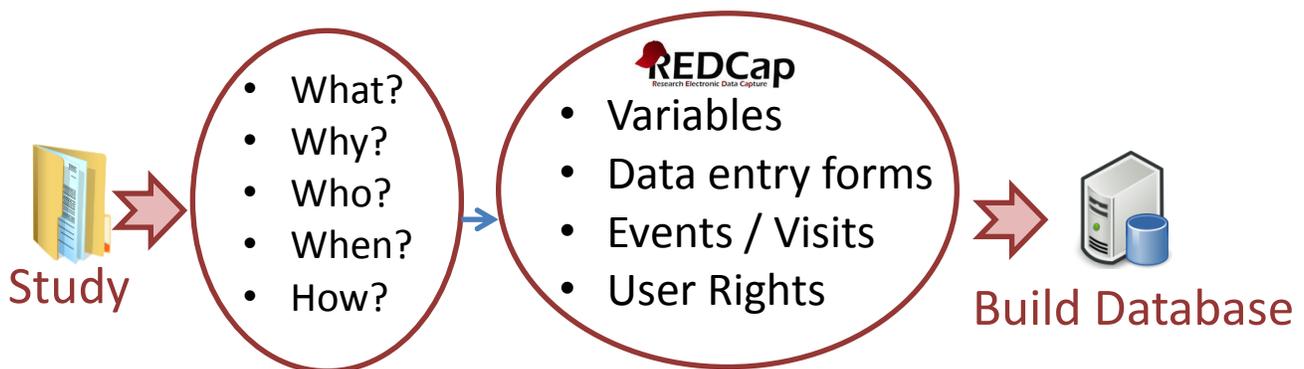
3.3.1 Prerequisites

If collecting data for the purpose of human subject's research, review and approval of the project by the Institutional Review Board is required. Note, the PI name and IRB# must be recorded in REDCap when creating a research project.



Define data to be collected before starting building a project.

The information that is going to be defined during this step will determine how to build the database.



General recommendations:

- Protocol determines what data should be collected on the CRF.
- Data not needed should not be collected: allow for efficient and complete data processing, analysis and reporting.
- Avoid duplication: do not record the same information into multiple variables for the same event: e.g. DOB and age.
- Provide instructions to reduce misinterpretations.
- Be clear and concise with data questions.
- Provide units to ensure comparable values.
- Request minimal free text responses: provide pre-defined “choices”.
- Use “None” and “Not done” to help tracking missing data.
- Determine if data are collected once or multiple times during the project.
- Determine if data are used across studies and organizations.
- Determine how data will be collected: manual key entry, imported from Excel, imported from another study.
- Determine who will enter data: patients, nurses, study staff, etc.

3.3.2 Build a project

3.3.2.1 Define a project type

There are three project types available in REDCap:

- | | | |
|--|---|--|
| a. Data collection performed only by study team |  | Data entry forms |
| b. Data collection performed only by subjects / study participants |  | Survey |
| c. Screening phase before enrolling subjects |  | Survey + Data entry forms (survey will be used to screen subjects and data entry forms to collect study data). |

There are two collection formats available for data entry forms:

- Classic: one record per patient
- Longitudinal: one record per patient per event, with the possibility of defining multiple arms

3.3.2.2 Create instruments

When creating a new project in REDCap, a demographics form is automatically added to the project. This form may be modified or deleted by the project owner.

A given field cannot be used in more than one instrument. The instruments are meant to organize data collection only; they have no impact on the structure of the database or on data extraction.

Create instruments and regroup the fields by instrument from a data collection logic standpoint:

- By modules : i.e. demographics, vitals, medication, etc
- By type of persons who collect data
- By data acquisition type: manual data entry or data import
- Regroup identifiers (PHI) on the same instrument to better protect subject’s data
- In a longitudinal study, fields that are entered at the same event or within the same arm are grouped together. Instruments may be used at multiple events/arms

Longitudinal study example:

Data Collection Instrument	Events for Arm 1: Study Cases		
	Event 1 (1)	Event 2 (2)	Event 3 (3)
Demographics	●		
Baseline Data Initial Visit	●		
Follow-Up Visit	●	●	●
Summary	●		

Data collection instruments may be also downloaded from the public Shared Library of instruments available within REDCap. If a public instrument is used in a given project, the content **may not be modified**.

3.3.2.3 Create fields

A field is defined by the following attributes: *type, label, name, validation, required, identifier, note*. In addition, a branching logic can be applied to specify whether or not a question will be displayed, depending on values entered in previous question(s).

The screenshot shows the 'Add Field Here' dialog box in REDCap. A green arrow labeled 'Branching Logic' points to a green branching logic icon (a right-pointing arrow) in the top left corner of the dialog. The variable name is 'discharge_date'. The field label is 'Date of discharge'. The field type is set to 'Date (M-D-Y)'. There are 'Add Field Here' buttons at the top and bottom of the dialog.

Field attributes

The screenshot shows the 'Edit Field' dialog box in REDCap. The field type is 'Text Box (Short Text)'. The field label is 'Date of birth'. The variable name is 'birth_date'. The validation is set to 'Date (M-D-Y)'. The minimum date is '06-12-1993' and the maximum date is '06-12-2011'. The 'Required?' checkbox is checked. The 'Identifier?' checkbox is checked. The custom alignment is 'Right / Vertical (RV)'. The field note is 'Small reminder text displayed underneath field'. There are 'Save' and 'Cancel' buttons at the bottom.

Label: will be displayed on the data entry forms/survey, reports, PDFs

Variable Name: will be column name in the export file for data analysis

Validation: to define data type and improve data quality: format will be checked during data entry & data import

Required?: a value will be required during data entry/data import

Identifier?: must be set to Yes for all PHI to be able to de-identify data exports and limit access to PHI within REDCap

Field Note: to guide data entry

Field names

- With some meaning to identify the content easily
 - Avoid using Q1, Q2, Q3
 - Use underscore for spaces
 - Do not use '\$' or '.'
 - Examples:
 - Use the word 'date' for all date fields
 - Use the word 'time' for all time fields
- MUST begin with a letter, otherwise data exports will not work when using SPSS
- Cannot end with a period
- Reserved keywords cannot be used (ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, and WITH)
- Should be kept short whenever possible (less than 26 characters)
 - To avoid spelling errors during programming (analysis)
 - To display as column headers in reports and data exports
- For repeating measurements, use the same name and add number (example: weight1, weight2, weight3...)

Available field types:

Field Type:	--- Select a Type of Field ---
Question Number <small>Displayed only on</small>	Text Box (Short Text)
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/File Attachment)
	Begin New Section (with optional text)

The collection of free text must be limited:

- Use categorical responses (drop-down, radio button or checkbox). REDCap will associate a number (code) with each answer (label). Consider using checkboxes to collect fields such as medications, symptoms and complications. REDCap will create a yes/no column in the dataset for each value that can be checked.
- Use pre-defined data type (text field with validation), i.e. yes/no, true/false, number, integer, date, datetime, time, zip code, email, phone.

Define range minimum / maximum to allow REDCap to perform basic data validation / quality control.

All dates in a given study must be collected using the same format.

Limit the use of calculated fields in REDCap:

- Calculations are useful when the resulting value is used to take some action regarding the subject
- Other calculations should be done outside of REDCap

Sample display of field types:

Editing existing Subject ID 1

Event Name: **Screening Visit (Arm 1: Screening)**

Subject ID 1

Date 30-09-2011 Today *date*

New Section

Radio
 First Answer
 Second Answer *radio*
 Third Answer reset value

Drop-down list
select one answer only

First Answer
Second Answer *drop down*
Third Answer
Fourth Answer
Fifth Answer

Check all that apply *checkbox*
 First Answer
 Second Answer
 Third Answer

New section

Yes No
 Yes
 No *yes/no* reset value

Time 08:35 Now *time*

True False
 True *true/false*
 False reset value

Scale
low medium high
 slider / visual analog reset value

Form Status

Complete?

Lock this record for this form?
If locked, no user will be able to edit this record on this form until someone with Lock/Unlock privileges unlocks it.
 Lock

Save Record

3.3.2.4 Define user rights

The rights of each user who is granted access to a project must be defined individually, depending on the user's profile. The table below lists the rights that are recommended per profile. A given person may have more than one profile. It is recommended to check all access rights on a regular basis during the study.

Profile Function	Project Setup	Data entry	Data review	Data Analysis
Calendar	x	x	x	
Data Export	x		x 	X 
Data Import	x	X 		
Data Comparison (for double data entry)		x 	x 	
Logging (audit trail) <i>(to be limited)</i>	x			
File Repository	x	x	x	x
User Rights <i>(to be limited)</i>	x			
Data Access Groups (to be used when multiple data entry sites)		x		
Graphical Data View & Stats	x	x	x	x
Data Quality	x	x	x	
Reports & Report Builder	x	X	X	
Project Design & Setup	x			
Record locking customization (if used)	x			
Lock/Unlock Records (if used)		x	x	
*Records (C_create, D_delete, R_rename) <i>(D to be limited)</i>	C,D,R (for testing)	C	R,D	
Data Entry Rights (none, read, edit)	Edit (for testing)	Edit 	Read or Edit 	



Users who may not view PHI will have access only to:

1. Calendar
 2. De-identified data export
 3. File repository
 4. Data Entry Rights section: *The 'No Access' box must be checked for all forms that contain PHI. The 'Data entry rights' section determines the access to data when using the Calendar, building reports and viewing data with the Graphical Data View tool.* Also, if access is given to some data entry forms, the record label (optional customization in REDCap) must not contain PHI.
 5. Data Quality
 6. Data exports, Reports, and Stats
- All other rights should be unchecked.

3.3.3 Test a project



The following tasks are highly recommended before moving a project to production status!

- **Enter test data:** using all fields in all instruments and all events to validate instruments and event definition, branching logic and calculated fields. Test data must include different cases that will allow testing all scenarios of branching logic and calculated fields.
- **Review test data:** open data entry forms, create reports, export data and send the blank CRF to all co-investigators for review.
- **Have a statistician review your database:**
 - Send a PDF version of the blank CRF.
 - Send the data dictionary of your database to the statistician. This will clearly communicate any defined branching logic --- not communicated in the raw data file or meta-data formatting accessible through the "Data Export" application. This will also clearly communicate the formulas from calculated fields.
 - Have the Statistician perform a data export to ensure that they do not extract identifiers. They can give you feedback regarding the overall design of your database as well as the definition of each field. It is important to think through the planned statistical analysis before collecting any data. A statistician can make sure that you are collecting the fields you need, in the format you need them in, in order to perform the statistical analyses you need.

3.4 Move a project into production mode

Once all steps listed in the above sections are performed successfully, the project may be moved into production mode. Test data can be deleted automatically when moving the project into production mode (check off the red box).

3.5 Enter, Review and analyze data

3.5.1 Enter data

Every record is identified by a unique key: Study ID (aka Record ID or Subject ID). This key identifier must be the first question on the first form (or survey). It is used by REDCap as the key variable to link all forms/events for a particular subject.

If multiple sites are involved in a study (and the Data Access Group feature is not being used), it is recommended to use conventions in order to group subjects by site. For example: 1 digit for the site and 3 digits for the subject number:

Site 1, subject 1: Study ID = 1001

Site 2, subject 1: Study IDd = 2001

Data can be key-entered into REDCap, using regular data entry via Data Entry Forms. Double data entry is an option available in REDCap. When this option is enabled, two data entry people will enter data, but instead of saving their data as normal, the Study ID (or unique record identifier) will have either a "--1" or "--2" appended to it, thus marking it as the record entered by data entry person 1 or 2. This is how the data entry people can enter data for the same subject/patient without having a duplicate record. Because all the records entered by the two users will end in the "--#" format, those users can only view the data that they themselves have entered. The "--#" format on the records are hidden from those two users, so to them everything appears as it normally would. However, Reviewers will be able to see the "--#" appended to the records when viewing them. When Reviewers choose to merge the two records entered by the data entry people, the new "third" record is saved without the "--#" appended to the end of the record name. The double data entry feature can be turn on by the REDCap Administrator only.

Data can be imported into REDCap from a CSV file, using the data import tool. REDCap generates a template with all columns. Data must be copied into this template; columns may be removed but the first column must always be kept (e.g. study_id). During data import, data can be added or modified but cannot be deleted.

3.5.2 Review data

The data review should start at the beginning of data collection to detect any potential data entry errors and to update data entry instructions, if needed.

There are different tools available in REDCap to review data:

- a. **Automatic checks:** during data entry or data import: REDCap will check automatically during data entry or data import for correct formatting of dates, times, numbers, ZIP codes, email addresses...and missing values for required fields.
- b. **Data review tracking:** each data entry form has a status that can be used to track the progress of data entry and/or data review: Incomplete, Unverified and Complete. In addition, the Lock/Unlock feature can be used to prevent accidental data modification or deletion. The e-signature is also an option that can serve as the equivalent of a handwritten signature. Example of use:

Data Collection Instrument	Events for Arm 1: Study Cases		
	Event 1 (1)	Event 2 (2)	Event 3 (3)
Demographics			
Baseline Data Initial Visit			
Follow-Up Visit			
Summary			

- Not entered
- Partially entered
- Complete
- Verified
- Validated

Note: the use of these tracking tools can be customized for each data entry form and project, depending on the needs of the project.

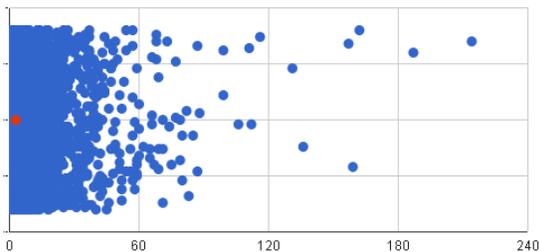
- c. **Data Reports and Exports:** will allow users to create, save, export and print customized data reports. Data reports can be exported to Excel. Once a report is saved, it can be seen by all users who are granted access to the project and the Data Report module of the project. For example, a report can be created to check the chronological orders of dates across events.
- d. **Data Stats:** This module can display any project data (numeric and coded data) in aggregate graphical format and as descriptive statistics. If clicking any of the points on the plot, the data entry form of that exact record will open. Example:

Length of therapy (days): [Refresh Plot](#)

Total (N)	Missing	Unique	Min	Max	Mean	StDev	Percentile						
							.05	.10	.25	.50 Median	.75	.90	.95
10360	70 (0.7%)	90	0.00	214.00	5.20	8.91	1.00	1.00	2.00	3.00	5.00	10.00	17.00

Lowest values: 0, 0, 0, 0, 0

Highest values: 157, 159, 162, 187, 214



- e. **PDFs:** data for one or all records may be downloaded into a single PDF file.
 - Contains the actual page format as seen on the data entry page or survey
 - Includes all data or data from a single instrument
 - May be used if data need to be shared within a short time period
- f. **Data Quality module:** The aim of this module is to help quickly find discrepancies and errors in project data. This module contains rules that can be customized for the project; these rules can be executed at any time during the study. Users can exclude any results that they do not wish to see in future results.

3.5.3 Analyze data

When data review is finalized, the project may be moved into ‘inactive’ mode where you can query/view the data but cannot add/modify/delete the data. The database is ready for statistical analysis.

The data export function will allow exporting data in various statistical formats, with the option to select all data or partial data. Formats available are: SPSS, SAS, R, STATA. The ‘Send-it’ function in REDCap can be used to send files that contain sensitive data. **Send-It** is a secure data transfer application that allows users to upload a file (up to 260 MB in size) and then allows multiple recipients to download the file in a secure manner. Each recipient will receive an email containing a unique download URL, along with a second follow-up email with the password for downloading the file (for greater security). The file will be stored securely and then later removed from the server after the specified expiration date. Send-It is the perfect solution for anyone wanting to send files that are too large for email attachments or that contain sensitive data.

3.5.4 De-identification methods for data exporting

REDCap provides advanced de-identification options that can be optionally used when exporting data, such as removing known identifier fields, removing invalidated text fields, notes fields, or date fields, date shifting and hashing of the record names. These options provide greater security and data protection when exporting sensitive data out of REDCap.

Date shifting: As part of the de-identification options in REDCap, all dates in the database may be shifted so that they do not reflect the actual dates. If the option to have all dates shifted is chosen, they will be shifted by a consistent length of time for each record, thus preserving the interval between dates. Date shifting leaves the database record intact and will not affect the actual saved dates in the database. The shifted dates will always be the same for each record when performing a data export. The dates are shifted back in time up to 364 days, in which the shifted value is based on the Study ID using a complex algorithm. Because the shifted value is based on the record name, the amount that dates are shifted is unique for each record, but the shifted amount remains the same within each record to maintain internal consistency. Date shifting prevents any dates from being used as identifiers for each database record while preserving the interval between dates.

3.6 Changes post-production

Changes to the project design after it is moved to Production are possible. Changes that may potentially affect/change collected data will be submitted for review. The REDCap Administrator will review changes and they will become effective only after they are approved by the REDCap Administrator. Any changes to the structure of the database, such as removal of forms, or modifying events in a longitudinal project can only be done by the REDCap Administrator.

Altering a database that is in Production can cause data loss. To protect data:

- (a) Do not change existing variable names, or data stored for those variables will be lost. To restore data that have been lost in this way, revert to previous variable name(s).
- (b) Do not change existing form names via a data dictionary upload, or form completeness data will be lost. Form names may be changed within the Online Designer without data loss.
- (c) Do not modify the codes or answers for existing dropdown, radio, or checkbox variables or existing data will be lost or confused. It is only acceptable to add choices to a dropdown, radio, or checkbox field.

3.7 Close a project

Once the data analysis is complete, the project in REDCap should be moved into 'archive' mode. This will take the project offline, and you and other users will no longer be able to access it normally. The project can be brought back online at any time after being archived by returning to My Projects page and clicking on 'Show Archived Projects' at the bottom. It is good practice to remove all user rights from the project except for one user because if the project needs to be put back in production status, the access will be limited to one person.

Acknowledgement:

The contact of this document was provided by The Children's Hospital of Philadelphia Center for Biomedical Informatics.