

REDCap Guide and Best Practices

Table of Contents

- Why Use REDCap 2
- REDCap Terminology 3
- REDCap Best Practices 6
 - Before You Get Started..... 6
 - Data Collection Instruments 6
 - Variable Names 7
 - User Rights 7
 - Testing 8
 - Managing the Project or Database 8

Why Use REDCap?

REDCap provides a secure web application for building and managing online surveys and databases. It enables rapid project development and provides users with advanced functionality, complete autonomy and control of projects, and direct exports to common statistical packages.

Accessibility

REDCap can be accessed from anywhere with internet access. Users can build their projects and enter data from anywhere in the world.

Ease of Use

REDCap has an intuitive user interface with video tutorials imbedded throughout the application to provide additional insight into the functionality. Additionally, all projects have a link to the REDCap Help & FAQ page, which consists of information covering a range of topics from project design and data entry to applications and customizations.

Secure

REDCap offers 3-tiered access with 1) system login; 2) project-level access; and 3) defined user rights within each project. All data is stored behind the institution's firewall.

Highly Customizable

Each project creator has full control over the design and setup of his/her project. Through functionality like survey creation, longitudinal project setup, repeating instruments and events, and advanced field specifications, REDCap can be designed to collect data for any clinical research or quality improvement project and many operational projects.

Autonomy

Project teams have complete autonomy and control over the REDCap project development process and designation of project-level user rights.

Wide User Base

REDCap is used by over 2,000 institutions around the world, who are all part of the REDCap Consortium, an online community where REDCap programmers and administrators can share ideas and ask questions. The Consortium allows for a breadth of support, training and knowledge-sharing.

REDCap Terminology

Archiving: Archiving a project takes the project offline and removes it from your My Projects list. Archived projects can only be accessed by clicking *Show Archived Projects* at the bottom of the My Projects page. A user can un-archive a project at any time and bring it back online.

Arms: Arms are groups of events (see Event definition). You may want to employ multiple arms when using different treatment groups. Each arm can have as many events as you wish.

API: The REDCap API is an interface that allows external applications to connect to REDCap remotely, and is used for programmatically retrieving or modifying data or settings within REDCap, such as performing automated data imports/exports from a specified REDCap project.

Branching Logic: Branching Logic allows you to hide fields under certain conditions. If branching logic is defined, a field will only be visible if the conditions provided are true. For example, you may want to hide the question “Are you pregnant?” unless “Female” is selected for a field that asks for gender.

Codebook: The Codebook is a read-only version of the project’s Data Dictionary and serves as a quick reference for viewing field attributes without having to download and interpret the Data Dictionary.

Data Access Groups: Data Access Groups (DAGs) allow you to restrict user access to a subset of data within the project. For example, if your project is for a multi-site study, you can create a Data Access Group for each site and assign the appropriate users to each DAG so that they can only see study participant data from their site.

Data Collection Instrument: A Data Collection Instrument is a data entry form. Building a REDCap project is all about customizing the contents of the data collection instruments. You will define how many instruments there are and what they are named, what data is captured, and how your data is stored and displayed.

Data Comparison Tool: The Data Comparison Tool allows a user to compare two records currently in the project and displays a comparison table, showing the differences between the two records.

Data Dictionary: The Data Dictionary is a specifically formatted CSV (comma delimited) file that you can use in place of the Online Designer to create and edit your data collection instruments and fields.

Data Import Tool: The Data Import Tool allows a user to import data from a CSV or XML file into a REDCap project. The data must be formatted according to the project’s Data Import Template in order to import successfully.

Data Quality: The Data Quality application allows a user to execute data quality rules upon the project data to check for discrepancies. The application provides some pre-defined data rules that users can run, and users may also customize their own rules.

Development Mode: Development Mode is the default status of all newly created REDCap projects. Projects are in Development when users are still working through the design of data collection instruments and/or testing their instruments by entering mock data. Project changes (via the Online Designer or the Data Dictionary) in Development Mode take effect immediately and do not require review from an Administrator.

Draft Mode: Draft Mode is a status in the Online Designer or Data Dictionary that a user must enter in order to make project changes in Production. Once a user enters Draft Mode, the user can make any project changes in the same way that one does in Development, but those changes must be submitted to a REDCap Administrator for review before they take effect in the project.

Event: An event is a scheduled (or unscheduled) occurrence in a longitudinal project during which data is captured using your data collection instruments. An event may be a temporal event in the course of your project, such as a participant visit or a task to be performed. Events allow for the utilization of data collection instruments multiple times for any given project. For example, if a study is collecting the same vital signs at each follow-up visit, then each visit could be its own event in the REDCap project.

Field: A field captures a singular data entry, such as age or height. In the Online Designer or Data Dictionary, a user defines what kind of data is collected via the creation of fields.

Form Status: The form status is the status of a record's data form completeness. Form statuses are color-coded such that red is *incomplete*, yellow is *unverified*, and green is *complete*.

Inactive Status: A project can be moved to inactive status if data collection is complete. Making a project inactive disables most project functionality while all collected data remains accessible. Once inactive, the project can be moved back to Production at any time.

Logging: The Logging application provides an audit trail of modifications in the project, including data exports, data changes, and the creation or deletion of users.

Longitudinal Data Collection: Longitudinal data collection allows instruments to be utilized repeatedly over a specific number of times (for example, using an instrument named "Vitals" over five visits for the same subject). The longitudinal format also allows one to utilize the scheduling module.

Online Designer: The Online Designer is where a user creates data collection instruments and makes modifications to existing fields using only a web browser. The instruments displayed in the Online Designer directly define the data entry webpages. Building and managing instruments through the Online Designer defines how you interact with the webpages during data entry.

Production Mode: Production Mode is the status that a user should move the project to once all forms have been built, testing is complete, and the user is ready to begin collecting real data. Project changes (via the Online Designer or Data Dictionary) in Production Mode do not take effect immediately and require review from an Administrator.

Randomization: Randomization is a process that assigns participants by chance (rather than by choice) into specific groups. The REDCap randomization module allows a user to implement a defined randomization model within the project.

Raw (Coded) Value: A raw value is a numerical or alphabetical value assigned to a multiple choice text label. Each multiple choice label must have an associated raw value. A user can choose his/her own raw value for each multiple choice option or let REDCap auto-assign this value. A raw value must be unique within each question (i.e., you cannot use the raw value of '1' for two different multiple choice options in the same field). The raw value does not affect the order that choices are displayed in the field.

Record: A record is a set of information for a unique participant or subject. Each record consists of a number of fields (data elements), which can be spread across multiple forms.

Record ID: The Record ID is the unique identifier for each record in the project. This field must be the first field in the first instrument of your project. It cannot be moved or deleted, but can be renamed (for example, to Study ID or Participant ID).

Record Status Dashboard: The Record Status Dashboard provides a tabular view of all existing records and their associated form status (complete, incomplete, unverified). Users can use the Record Status Dashboard to navigate to a record's home page or to a record's data collection instrument. Users can also create and save custom versions of the Record Status Dashboard, similar to reports.

REDCap Shared Library: The REDCap Shared Library is a repository of REDCap data collection instruments that can be downloaded and used by researchers at REDCap partner institutions.

Repeatable Events: A repeatable event is an event that can be repeated as many times as needed during data entry. This allows a user to repeat an entire event of instruments together in unison. Each record in the project can have a different number of repeated events. This can be especially useful for a study that needs to collect data at participant follow-up visits, but the number of follow-up visits differs per participant.

Repeatable Instruments: A repeatable instrument is a data collection instrument that can be repeated as many times as needed during data entry. For example, if each record in a project represents a person, then one might have a repeating instrument for medications or adverse events. Each record in the project can have a different number of repeated instances of any given instrument.

Reports: Reports are custom, dynamic tables in the REDCap project that you can build to view a subset of your data.

Scheduling: The Scheduling module can generate schedules for your project calendar that are auto-generated from project-defined events (e.g., visits, time-points). Scheduling is only available for projects using longitudinal data collection.

User Rights: The User Rights application is where the project creator defines the customized privileges that each project member has in terms of project design, data entry, and data access.

Variable Name: The variable name is the backend name of the field that is stored in the REDCap database. It is not visible during data entry. Variable names must be unique within a project.

REDCap Best Practices

Before You Get Started

Before you begin your REDCap project build, ensure you already know the following:

1. The end goals of your project
2. The specific data you are going to capture
3. The type of study you are performing (longitudinal, multi-site, etc.)
4. Who will comprise your research or project team

The specific data you are going to capture

Carefully consider how the data will be used. Data that are not needed should not be collected. For research projects, the protocol determines what data can be collected.

It may be important to consult with a statistician before building your project. Statisticians can provide guidance on how best to setup your project (classic vs. longitudinal) and how to organize your instruments. If you are already working with a statistician, it will be important to consult with him/her on the naming convention for your variable names and your multiple choice raw values.

The type of study you are performing (longitudinal, multi-site, etc.)

Determine if data are collected once or at multiple time points during the project. Are time points associated with a particular set of visits/events or on an as needed basis? REDCap makes it easy to create a project with only one event or a longitudinal project with multiple events. In longitudinal projects, you only need to create the form once and then later specify that the form will be used to collect data for multiple data collection periods (events).

REDCap also allows for user rights restrictions based on Data Access Groups (DAGs). These can be useful if you are conducting a multi-site study and only want users to have access to records from their site.

Who will comprise your research or project team

Determine if data are used across studies and organizations, and determine who will enter the data (patients, nurses, study staff, etc.).

Data Collection Instruments

Design your data collection instruments with the end user in mind, and think about ways to help the user efficiently enter data. Group variables together that will follow the data entry workflow, and use field types that minimize changing from keyboard to mouse. For example, you can enter a dropdown field option by typing the first character of the label, allowing you to tab through the data entry fields, while radio buttons require using the mouse to select an option.

It is also important to describe the input data fields so that data entry staff are sure about what they are entering. Use clear field labels and field notes to describe exactly what should be entered. Below are some general best practices to keep in mind.

1. Avoid duplication. Do not record the same information into multiple variables for the same event (e.g., date of birth and age, height on two separate forms).
2. Be clear and concise with data questions.

3. Indicate units to ensure comparable values are collected and use field validation whenever possible.
4. Define range minimum/maximum when possible to allow REDCap to perform basic data validation and quality control.
5. Ensure all dates in a project are collected using the same validation format.
6. Request minimal free text responses. Instead, provide pre-defined multiple choice fields. Open text fields are very difficult and time-consuming to analyze.
7. For inexact dates, collect day, month and year separately. For example, if patients will rarely know the exact date they started a medication, consider creating three fields so that you can capture *month* and *year* and leave *day* blank if they do not remember.
8. Indicate specific, standardized choices for reasons why data is missing (e.g., "Patient Refused to Answer," "Not Applicable," "Don't Know").
9. Be consistent with assigning raw values. For example, if "Yes" is coded as a '1' in one field, it should be coded as '1' in all other project fields.
10. Use case and punctuation consistently across all field labels.
11. Don't make the question text (field label) too long if you can help it. This text will be used as column headings in data exports.
12. Use the descriptive text field to provide additional information or instructions throughout a form.
13. When a free text field with no validation is necessary (e.g., patient name), use a field note to provide guidance to the data entry staff or survey respondent to specify how you want the data formatted (e.g., "Last, First").
14. Flag your REDCap identifiers appropriately (this is critical when building a clinical research project). There are 18 data elements that must be marked as identifiers in a REDCap clinical research project, as per HIPAA: name, fax number, phone number, email address, account numbers, social security number, medical record number, health plan number, certificate/license number, URL, IP address, vehicle identifiers, device ID, biometric ID, full face/identifying photo, other unique identifying number/characteristic/code, postal address, date precision beyond year.

Variable Names

Variable names are critical in the data analysis process. If you export your data to a statistical software program, the variable names are what you or your statistician will use to conduct the analysis. Below are some best practices to keep in mind when considering variable names.

1. Keep variable names as short as possible, while still maintaining clarity in what the field is. The recommended variable name length is 3 to 10 characters.
2. Avoid using names like Q1, Q2, Q3. This does not tell us what the field/question is.
3. Only use underscores when absolutely necessary. This can quickly add to the length of the name.
4. For large, complex projects, consider developing a naming convention for your variable names (e.g., use a prefix to indicate which form the variable is in, use the abbreviation 'dt' for any variables that ask for a date).
5. Be aware of the limitations of your statistical software. Naming rules vary depending on software and version.

User Rights

User rights define the customized privileges that each project member has in terms of project design, data entry, and data access. It is important that project members are only given access to what they need and nothing more. Below are some best practices to keep in mind regarding user rights.

1. Use Data Access Groups (DAGs) if you are collecting data at more than one study site.
2. Don't give a team member more access rights than s/he needs.
3. If you have a large number of users, it may be helpful to create User Roles to simplify management of user rights.
4. Delete users from your project when they leave the project team.
5. Do not give someone access to the User Rights module unless you want them to have full autonomy in the project. Giving them access to User Rights enables them to modify their own rights as well as the rights of others on the project.

Testing

Test all forms thoroughly before moving your project to Production Mode. Test data should include different scenarios that will allow testing all possibilities with branching logic, calculated fields, survey invitations, etc. It is sometimes best to test your instruments as you build them, rather than waiting until the end to test everything, as this can get overwhelming. Some particular things to look for when testing:

1. **Branching logic** – Ensure your branching logic is functioning properly. This requires entering test data for all possible responses that affect the logic.
2. **Calculated fields** – Ensure your calculated fields are calculating correctly.
3. **Data Exports** – Export your test data to ensure you are comfortable with how it is formatted before you begin collecting real data.
4. **Third Party Input** – Get a third party to enter data as well. Someone who was not involved in building the database will find errors that may otherwise be missed.

Managing the Project or Database

Equally as important as the project design and build phase is the management of a production REDCap project. The REDCap project manager is responsible for the quality and security of the project data. REDCap has many tools available for successful data management.

Data Exports, Reports, and Stats

REDCap allows for the creation of an infinite number of custom reports, which are saved and viewed from within the application. Project Managers should consider using this feature to easily view a subset of the project data and run counts on particular variables (e.g., number of males vs. females enrolled). Each report also has access to a view of bar or pie charts and descriptive statistics of the data. Custom reports can also be used to export a subset of data to Microsoft Excel, SAS, Stata, R, or SPSS for further analysis.

Data Quality

Project Managers should become familiar with the REDCap Data Quality module, which allows for real-time execution of pre-defined and custom-built rules to check for discrepancies in the data. Pre-defined rules include missing values, data out of range, outliers for numerical fields, multiple choice fields with invalid values and incorrect values for calculated fields. It may be useful to run data quality checks on a

regular cadence during the data collection process, particularly if several team members are entering data.

Audit Trails

REDCap collects an audit trail of all project changes via the Logging application. Project Managers can use this application to pinpoint what happened to missing data (i.e., was it deleted by a user) or who made changes to project fields. The Logging application can also be used to monitor project activity on a regular basis for compliance purposes. The audit trail can be exported to Microsoft Excel.

Record Locking

Record locking freezes data, ensuring that users don't accidentally modify it without authorization. This controls data entry and is commonly used to prepare for statistical analysis or to archive final datasets. Locking permissions are controlled within the User Rights application. Unauthorized users cannot modify a locked instrument.

Project Archival or Deletion

Once a project has ended and all project data has been exported to the appropriate application for analysis, the project should either be archived or deleted. This can be done from within Project Setup, by clicking on the 'Other Functionality' tab.

Deleting a project will permanently remove it from the REDCap system, and all of its data will also be deleted. Projects should be deleted if you are confident that you will not ever need to access the project again.

Archiving a project will remove the project from your My Projects list, but will not remove it from the REDCap database. Archived projects can still be accessed, and moved back to Production, by clicking *Show Archived Projects* at the bottom of the My Projects page. Projects should be archived if you think you may need to access the project at a later date.

Projects can also be moved to Inactive Status. This is typically done after data collection but before the team officially closes out the project, often during the data analysis phase. Moving a project to Inactive Status will disable most project functionality, but will keep all collected data intact. Inactive projects can be moved back to Production Mode at any time.