

Data Import Tool

CCTS Biostatistics Core

May 2024

Overview

The Data Import Tool allows you to upload data from a CSV file to your REDCap project. This can be useful if data were collected using a spreadsheet or another external process or system.

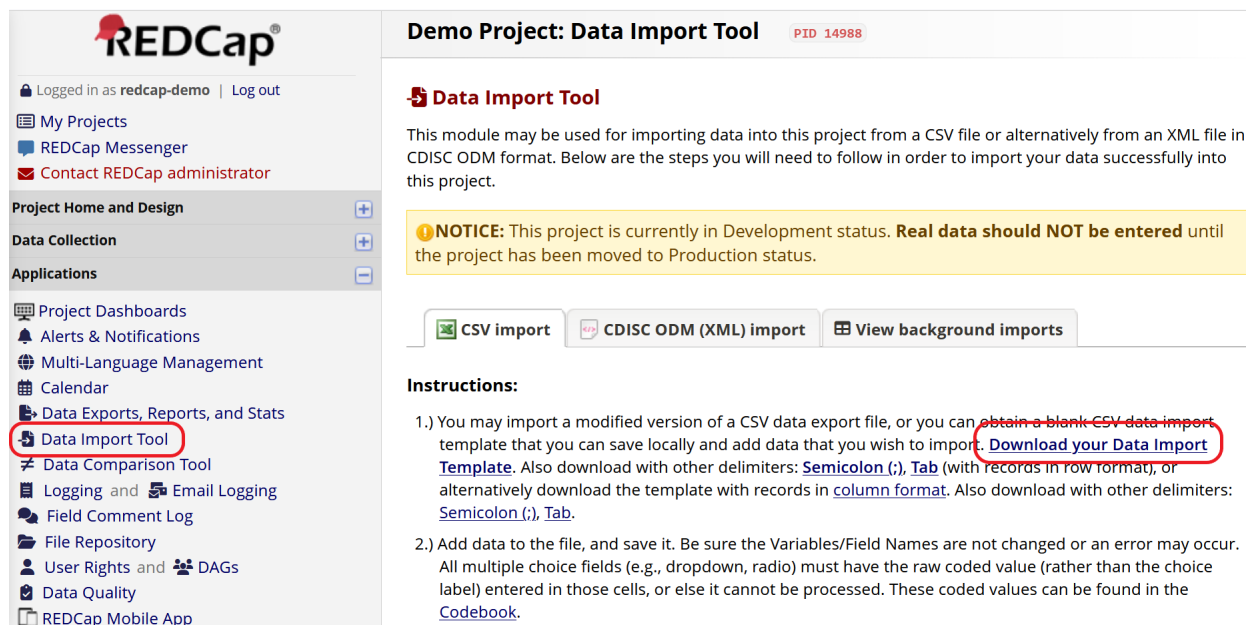
You might also use the Data Import Tool to edit multiple existing records at the same time. To accomplish this, you could download the existing data, add or change specific values, and then import the modified dataset.

Instructions

These instructions assume your REDCap project already contains the instruments and fields you want to populate. If you have not yet built your forms and surveys, you must do that before using the Data Import Tool.

Download Template

Click Data Import Tool in the left menu under Applications. Then download your Data Import Template by clicking the link in the first paragraph under "Instructions."



REDCap
Logged in as redcap-demo | Log out
My Projects
REDCap Messenger
Contact REDCap administrator

Project Home and Design +
Data Collection +
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

Demo Project: Data Import Tool PID 14988

Data Import Tool

This module may be used for importing data into this project from a CSV file or alternatively from an XML file in CDISC ODM format. Below are the steps you will need to follow in order to import your data successfully into this project.

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

CSV import | CDISC ODM (XML) import | View background imports

Instructions:

- 1.) You may import a modified version of a CSV data export file, or you can **obtain a blank CSV data import template** that you can save locally and add data that you wish to import. **Download your Data Import Template**. Also download with other delimiters: Semicolon (;), Tab (with records in row format), or alternatively download the template with records in column format. Also download with other delimiters: Semicolon (;), Tab.
- 2.) Add data to the file, and save it. Be sure the Variables/Field Names are not changed or an error may occur. All multiple choice fields (e.g., dropdown, radio) must have the raw coded value (rather than the choice label) entered in those cells, or else it cannot be processed. These coded values can be found in the Codebook.

Figure 1: Data Import Tool

The Data Import Template is a CSV file. It contains column headings that match the field names in your project. The template also includes automatically generated fields that capture information about each form. See below for more about how to use each of these system-generated fields.

- A field will indicate the completion status of each form. For example, if you have an instrument called “Demographics”, the Data Import Template will include a field called `demographics_complete`.
- If your project is longitudinal, with multiple events, the template will include the field `redcap_event_name`.
- If your project includes repeating instruments, the template will include the fields `redcap_repeat_instrument` and `redcap_repeat_instance`.
- If your project uses Data Access Groups, the template will include the field `redcap_data_access_group`.

	A	B	C	D	E	F	G	H	I	J
1	<code>study_id</code>	<code>redcap_event_name</code>	<code>first_name</code>	<code>last_name</code>	<code>dob</code>	<code>sex</code>	<code>demographics_complete</code>	<code>prealbumin</code>	<code>creatinine</code>	<code>lab_data_complete</code>
2										
3										
4										

Figure 2: Data Import Template

Add Data

Fill in the CSV template with the data you want to import, being sure to keep the column names intact. You must retain the record identifier (e.g., `record_id`) as the first column. You can reorder other columns or delete columns that aren't relevant to your import.

Requirements for some different field types are explained below.

Record Identifier The record ID field must be the first column in your file, just as it must be the first field in your project. To edit existing records, match the values in this column to their existing REDCap IDs. To create new records, include a unique value in this column for each unique record.

During the import process, you can choose whether to retain the record IDs in your CSV or let REDCap automatically assign new ID values. Even if you want REDCap to auto-number new records, you must include a unique placeholder as the record ID for each record.

Dates and Date-Times Date values must be formatted as one of the following:

- YYYY-MM-DD (e.g., 2024-01-31)
- MM/DD/YYYY (e.g., 01/31/2024)
- DD/MM/YYYY (e.g., 31/01/2024)

Date-time values must be formatted as one of the following:

- YYYY-MM-DD HH:MM[:SS] (e.g., 2024-01-31 13:59:59)
- MM/DD/YYYY HH:MM[:SS] (e.g., 01/31/2024 13:59:59)
- DD/MM/YYYY HH:MM[:SS] (e.g., 31/01/2024 13:59:59)

Use 24-hour time notation. You must use one of these formats even if your form or survey displays dates in a different format.

Multiple Choice Fields (Select One) For multiple-choice fields (drop-down, radio, yes-no, or true-false), you must import raw coded values, not text labels. For example, consider a multiple-choice field with the following choices:

- 1, Red
- 2, Yellow
- 3, Blue

The corresponding column in your import CSV must contain the values 1, 2, and 3, not the words Red, Yellow, and Blue.

Refer to your project's Codebook for the codes associated with each multiple-choice field. See the [Field Types page](#) for more about multiple-choice fields.

Checkbox Fields (Select All That Apply) REDCap expands the response options for a checkbox field into separate columns, each of which can have a value of 1, meaning checked, or 0, meaning unchecked.

For example, consider a checkbox field called `symptoms` with the following options:

- 1, Cough
- 2, Headache
- 3, Fatigue
- 4, Other

Your Data Import Template will include a column for each choice, called `symptoms___1`, `symptoms___2`, `symptoms___3`, and `symptoms___4`. For each of these columns, include a value of 1 if the option has been selected by a particular record. Otherwise, insert a 0.

Refer to your project's Codebook for the codes associated with each checkbox field. See the [Field Types page](#) for more about checkbox fields.

Calculated Fields Calculated fields will not be included in your Data Import Template, as you cannot manually set the values of these fields. Instead, values will be calculated by REDCap during the import process.

Text fields that use the `@CALCTEXT` or `@CALCDATE` action tags will be included in the Data Import Template, but their values cannot be manually edited during data import. Instead, values will be calculated by REDCap during the import process.

See the [Field Types page](#) for more about calculated fields. See the [Action Tags page](#) for more about the `@CALCTEXT` and `@CALCDATE` action tags.

Form Status Your Data Import Template will include automatically generated fields to capture the completion status of each instrument. Column names follow the pattern `form_name_complete`. Form status is coded as a multiple-choice field with the following options:

- 0, Incomplete
- 1, Unverified
- 2, Complete

Longitudinal Projects with Named Events If your project is longitudinal, with multiple named events, your template will include a column called `redcap_event_name`. You must include data for different events on different rows of your CSV. On each row, be sure to include the record ID in the first column and the unique event name in the `redcap_event_name` column. Find event names on the project's Codebook or Define My Events pages.

The example below illustrates an import file for a project with four events. The unique event names are `enrollment_arm_1`, `visit_1_arm_1`, `visit_2_arm_1`, and `visit_3_arm_1`. An instrument called Demographics is completed during the enrollment event, while a form called Lab Data is completed during all four events. The image shows data for three records.

- Each row includes the record ID (`study_id`) in Column A.
- Rows 2, 6, and 10 contain data for the enrollment event. The unique event name, `enrollment_arm_1`, is specified in Column B. Data from the Demographics form (columns C through K) and the Lab Data form (Columns L, M, and N) are filled in.

- The other rows contain data for visits 1, 2, and 3, with the event name for each row specified in Column B. The Demographics form was not completed during these events, so columns C through K are blank. The Lab Data form was completed during these events, so Columns L, M, and N contain data.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	study_id	redcap_event_name	first_name	last_name	dob	ethnicity	race_1	race_2	race_3	sex	demographics_complete	prealbumin	creatinine	lab_data_complete
2	1	enrollment_arm_1	Lorena	Randal	1952-10-12	1	0	1	0	0	2	23.14	0.84	2
3	1	visit_1_arm_1										22.71	0.84	2
4	1	visit_2_arm_1										23.47	0.90	2
5	1	visit_3_arm_1										21.60	0.82	2
6	2	enrollment_arm_1	Charley	Pitts	1962-02-23	0	1	0	1	1	2	20.22	0.98	2
7	2	visit_1_arm_1										18.78	1.00	2
8	2	visit_2_arm_1										20.05	0.99	2
9	2	visit_3_arm_1										20.34	1.03	2
10	3	enrollment_arm_1	Eun-Woo	Munson	1954-09-25	0	0	0	1	1	2	16.55	1.07	2
11	3	visit_1_arm_1										14.98	1.11	2
12	3	visit_2_arm_1										16.27	1.03	2
13	3	visit_3_arm_1										16.24	1.03	2

Figure 3: Longitudinal Data Format

Repeating Instruments If your project has repeating events or instruments, your template will include fields called `redcap_repeat_instrument` and `redcap_repeat_instance`. Data from different instances must be included on different rows of your CSV.

- For rows that correspond to non-repeating instruments, leave the `redcap_repeat` fields blank.
- For rows that correspond to repeating instruments, fill in the name of the instrument in the `redcap_repeat_instrument` field. Also fill in the number of the instance in the `redcap_repeat_instance` field. If you're adding data about new instances and don't want to manually number them, you can instead include the word "new" in the `redcap_repeat_instance` column.

The example below illustrates an import file for a project with a non-repeating instrument called Demographics and a repeating instrument called Visits. The image shows data for three records.

- Each row includes the record ID (`record_id`) in Column A.
- Rows 2, 7, and 11 contain data for fields on the non-repeating Demographics form (columns D through G). The `redcap_repeat` columns (B and C) are blank.
- The other rows contain data for fields on the repeating Visits form (columns H through J). The `redcap_repeat` fields contain the instrument name (Column B) and the instance number (Column C).

	A	B	C	D	E	F	G	H	I	J
1	record_id	redcap_repeat_instrument	redcap_repeat_instance	first_name	last_name	sex	demographics_complete	visit_date	weight	visits_complete
2	1			Pelles	Kumar	0	2			
3	1	visits	1					2020-05-19	164.94	2
4	1	visits	2					2020-09-19	163.3	2
5	1	visits	3					2021-02-14	166.68	2
6	1	visits	4					2021-07-12	166.23	2
7	2			Geraldine	Tyson	0	2			
8	2	visits	1					2020-06-12	174.28	2
9	2	visits	2					2020-12-18	173.38	2
10	2	visits	3					2021-07-29	183.74	2
11	3			Fannie	Montaña	0	2			
12	3	visits	1					2020-06-19	128.48	2
13	3	visits	2					2020-11-27	130.92	2
14	3	visits	3					2021-06-03	132.61	2
15	3	visits	4					2021-10-16	132.4	2

Figure 4: Repeating Data Format

Data Access Groups If your project uses Data Access Groups, your Data Import Template will include a column called `redcap_data_access_group`. Use this column to assign or reassign a record to a DAG. Include the unique DAG name, available on the project's Data Access Groups page.

Upload File

When you're ready to import data, return to the Data Import Tool page. Review the import settings and change them as needed. Then click **Upload File**. Each import setting is explained below.

The screenshot shows the following settings:

- Choose an import option:** Import in real time
- Select your CSV data file:** Choose File (No file chosen)
- Display the data comparison table?:** Yes, display uploaded data prior to importing
- Rename records?:** No, use the record name provided
- Overwrite data with blank values?:** No, ignore blank values in the file
- File format settings:**
 - CSV delimiter of data file: Comma (,)
 - Format for date/datetime values: MM/DD/YYYY or YYYY-MM-DD
 - Records in file are formatted as: Rows

Upload File

Figure 5: Import Settings

Choose an import option Small datasets can usually be uploaded in real time. For larger datasets, you might want to choose “Import as background process” to avoid page timeouts while the dataset is transferred. With this option, REDCap will email you when the data import process is complete.

Select your CSV data file Click **Choose File**, find the CSV file that contains your data, and click **Open**.

Display the data comparison table? We recommend that you choose “Yes” in most cases. The data comparison table will allow you to review changes before they become final.

Rename records? If your CSV file contains new or altered data for records that already exist in REDCap, you must include the correct record IDs in the first column of your file. In this case, select “No, use the record name provided” to ensure the new values are attached to the correct record IDs.

If your upload should create new records and your project has auto-numbering enabled, you may want to select “Yes, rename all record names”. This will tell REDCap to assign a new automatically generated ID to each new record. Note that your CSV file must still include the record ID column, with unique values differentiating each unique record, even if those values will be overwritten by REDCap’s auto-assigned record ID.

Overwrite data with blank values? Select “No” if blank values in your CSV should be ignored during import (common). Select “Yes” if blank values should overwrite existing data in REDCap.

File format settings Confirm that the correct date/datetime format is selected. In most cases, the default options for delimiter and record format will be appropriate.

Review the Data Display Table

If you have opted to view the Data Display Table, you will see a summary of your data import before it's finalized. New values will appear in black text. Existing values that will be overwritten will appear in red text. Existing data that will not change will appear in gray text.

If the summary is acceptable, click **Import Data**. Otherwise, modify your CSV file and repeat the import process.

DATA DISPLAY TABLE													
study_id	redcap_event_name	first_name	last_name	dob	ethnicity	race__1	race__2	race__3	sex	demographics_complete	prealbumin	creatinine	lab_data_complete
1 (existing record)	enrollment_arm_1	Lorena	Randal	1952-10-12	1	0	1	0	0	2 (0)	23.14	0.84	2
1 (existing record)	visit_1_arm_1					0	0	0			23.33 (23.14)	0.84	2
1 (existing record)	visit_2_arm_1					0	0	0			23.47	0.90	2
2 (new record)	enrollment_arm_1	Charley	Pitts	1962-02-23	0	1	0	1	1	2	20.22	0.98	2
2 (new record)	visit_1_arm_1					0	0	0			18.78	1.00	2
2 (new record)	visit_2_arm_1					0	0	0			20.05	0.99	2
2 (new record)	visit_3_arm_1					0	0	0			20.34	1.03	2

Do you wish to import the new data (displayed above) into the project?
(Click the button below to import the data.)

Import Data [Cancel](#)

Figure 6: Data Display Table