

---

# Clinical Research Center REDCap Database Checklist

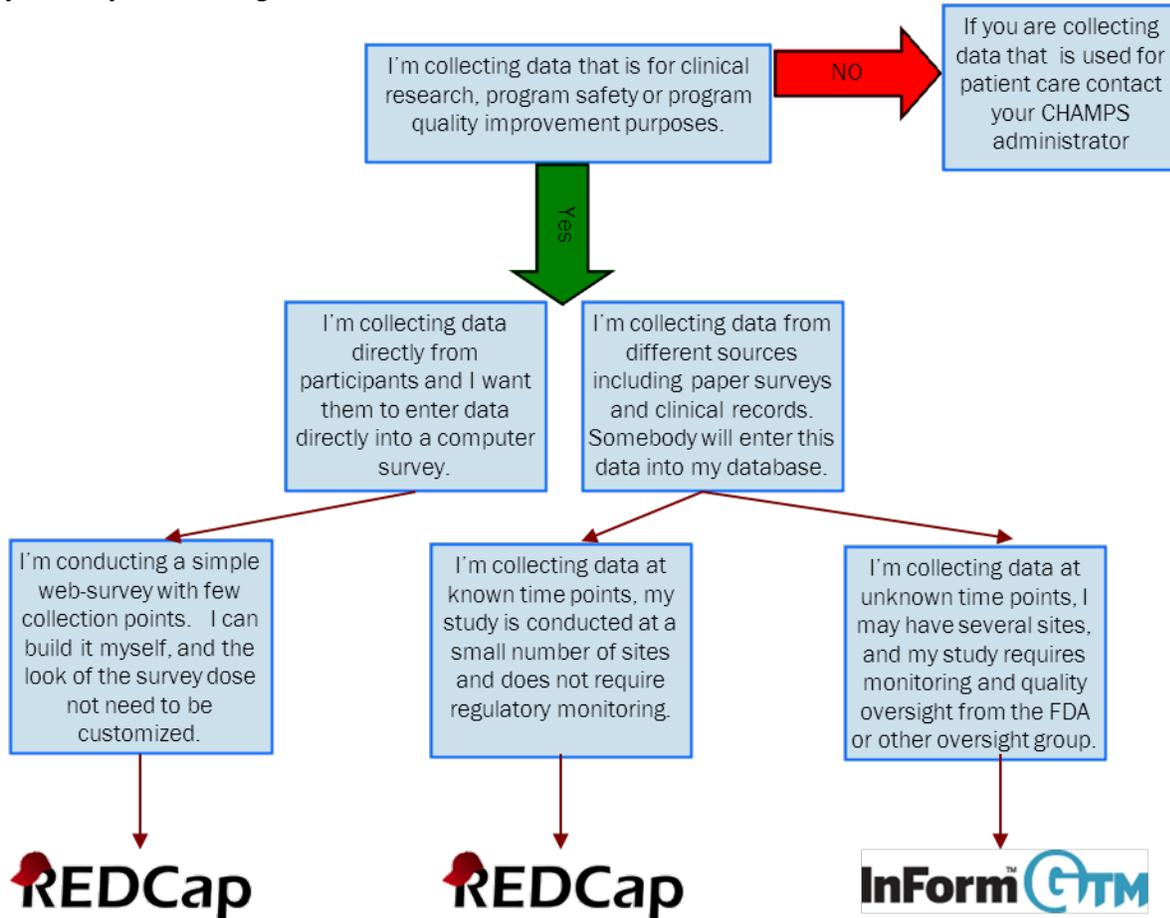
(Helpful tips to guide the database development process)

---

Authors: Chris Botte, Qiaoli Chen, Kimberly Chin, Mohamad Daniar, Jisun Jang, Paul Mitchell

**Use the decision tree below to determine which database is most appropriate for your data.**

If REDCap is most appropriate, then this checklist will help you avoid some common pitfalls on your way to building a database.



For FDA electronic data submissions, variable names are limited to 8 characters.

For more information on FDA e-submissions, please visit:

<http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM199759.pdf>

---

## Clinical Research Center REDCap Database Checklist

(Helpful tips to guide the database development process)

---

### Database Layout

In REDCap, databases that are longitudinal are built differently than cross-sectional databases. Longitudinal databases allow for more than one time-related event (e.g. multiple study visits; a repeated set of labs).

Build your REDCap database to reflect your study design:

- Cross-sectional: a single study visit with data collected at a defined time
- Longitudinal: multiple study visits or a single study visit with a panel of data collected more than once per subject

For longitudinal databases only:

- Utilize [Define My Events](#) to define visits/events.
- Use [Designate Instruments for My Events](#) to select forms/instruments/CRFs appearing at each visit/event.

For REDCap data export:

- A cross-sectional design will generate a simple table output with exactly one record per subject. Below is sample output for **3 subjects**.

Cross Sectional Example Output					
study_id	first_name	last_name	visit_date	height	weight
P001	John	Smith	10/1/2012	200	100
P002	Susan	Jones	10/3/2012	212	95
P003	William	Bell	10/1/2012	185	76

- Longitudinal designs will create a single table output with a subject record for each event in the study. The output below represents **repeated height and weight measurements for the same 3 subjects noted above**.

Longitudinal Example Output						
study_id	redcap_event_name	first_name	last_name	visit_date	height	weight
P001	Baseline	John	Smith			
P001	Visit 1			10/1/2012	200	100
P001	Visit 2			10/3/2012	210	105
P002	Baseline	Susan	Jones			
P002	Visit 1			10/1/2012	210	90
P002	Visit 2			10/3/2012	212	95
P003	Baseline	William	Bell			
P003	Visit 1			10/1/2012	185	76
P003	Visit 2			10/14/2012	188	80



---

**Clinical Research Center**  
**REDCap Database Checklist**  
(Helpful tips to guide the database development process)

---

**Variable Coding**

- Code responses as numbers: *[Number code], [Label]*.

For example:

<i>Have you experienced any symptoms in the past 24 hours?</i>	<i>What is the highest level of education you have obtained (choose one)?</i>
1, Yes 0, No	1, High school diploma 2, Associate's degree 3, Bachelor's degree 4, Master's degree 5, Doctorate

The comma in REDCap is used to separate the number code from the label.

- Use consistent coding throughout the database to denote special cases.

The following are commonly used:

- 9 = missing
- 8 = don't know/refused
- 3 = N/A

The coding used for these special cases should be well outside the range of values possible for each variable.

- Include the units of measurement and number of decimal places wherever appropriate.
- Use branching (skip) logic to prevent data that is not applicable from being entered. For example, use branching logic to prevent questions related to pregnancy from appearing for male subjects.
- Use consistent formatting for dates throughout the database. For example, if you use the REDCap format M-D-Y (month/day/year) for one date, be sure to use this same date format for all dates.
- Use validation to specify the number of decimal points, range of expected data, and type of variable (character vs. numeric). The validation function provides feedback to the user when data entered is inconsistent with what is expected.
- Limit the use of open-ended questions for quantitative research.

---

**Clinical Research Center**  
**REDCap Database Checklist**  
(Helpful tips to guide the database development process)

---

**Variable Coding cont.**

- Be aware that checkboxes (“check all that apply”) have several limitations:
1. When you assign the question a variable name (e.g. X), REDCap will then automatically generate variables for each checkbox by appending a numeric suffix to this variable name. So if there are three checkboxes, REDCap will generate variables X\_\_1, X\_\_2, and X\_\_3 corresponding to the first, second, and third checkbox.
  2. It won't be possible to differentiate between “no” and “unknown/missing” responses when a box is not checked.
  3. You should not use negative response codes for checkboxes, since the automatic variable that REDCap will generate will contain a negative sign (e.g. X\_\_-8), which is an illegal variable name in several software systems, including SAS, SPSS...
- Code ethnicity (Hispanic/Latino) independently from race, and be sure it appears prior to the race question.

For example:

<p><b>What is your ethnicity?</b></p> <p>Hispanic Yes.....1 No.....0</p> <p><b>What is your race?</b></p> <p>White Yes.....1 No.....0</p> <p>Black Yes.....1 No.....0</p> <p>Asian Yes.....1 No.....0</p> <p>Other Yes.....1 No.....0</p>
---