

Comprehensive Tips and Resources for Effective Database Design

Last time, we discussed the naming convention for your files and documents. During the first quantitative lecture about the data cycle, we also briefly discussed the structure of the database. Here, we will describe the basic principles of database design for a small project with some useful tips, using the RedCap platform as an example.

Note:

You can listen to our L&L lectures online here [WHRI Lunch & Learn Series – Women's Health Research Institute](#)

Visit our “Stats corner” at the end of our June e-blast for previously published tips in data management and analysis [eBlast Archive – Women's Health Research Institute \(whri.org\)](#)

Tip one: Key steps for Database Design

Before starting with RedCap or any other platform to design your database:

1. Determine the purpose of your database: Most databases will be tied to your research project, grant proposal, or ethics.
2. Find and organize the required information: Decide on the data you want to collect and organize it accordingly.
3. Divide the information into tables: Use the principles of a rational database that we discussed during the lecture.
4. Transform the information into columns (variables): Decide on the data types (e.g., number, text, boolean, date, etc.) for each column.
5. For larger projects with multiple tables, decide on the primary (unique) keys: Primary keys can be a single variable or a combination of multiple variables (matrix). The primary key will uniquely identify each row in the table.
6. Use a flowchart to establish relationships between the tables using primary keys for linkage.
7. Validate and test your design to ensure that no information is missing: Use your research proposal and ethics guidelines to ensure that only relevant information is collected.

Tip two: Sequential Group Allocation and Data Collection for Comparative Projects

For projects involving the comparison of multiple groups (e.g., treatment/control, different study arms, etc.), it is recommended to allocate participants to the groups based on appropriate criteria before collecting any additional information. Subsequently, collect additional information for all participants or specific groups of participants at a later stage.

Real life example:

We want to design RedCap data base for the “*Patient Characteristics of Subpopulations Experiencing Recurrent Pregnancy Loss: A Comparison of Spontaneous Conceptions, Euploid Embryo Transfers, and Biochemical Losses*” project. Based on the project proposal we first need to identify women with Recurrent Pregnancy loss – defined as two+ pregnancy losses and later allocate them to the sub-groups based on each sub-group definition

Tip three: Prepare Your Database Requirements Document

This document should include detailed instructions on how database should be prepared. I usually create my own table. Alternatively, you can use RedCap data dictionary format to design your database [How to Use a REDCap Data Dictionary \(illinois.edu\)](#)

Variable name: Use self-explanatory short name for each variable. It is a good idea to distinguish between types of variables within the name. For example, `_n` for numeric, `_b` for boolean, `_c` for categorical and so on.

Variable Name	Variable Type	RedCap field type	Rules	Notes to programmer/field validation
Loss_n	Continuous	Text	Loss_n>=2	Restrict to numbers only

*do NOT use more than 26 characters for your variable name

Tip Four: Testing your Data Base

After the database is created in RedCap or any other platform, it is crucial to test it thoroughly to ensure that the correct information is being recorded, in accordance with your requirement document, survey instruments, and proposal.

Fortunately, there are numerous free and useful tools available online for designing RedCap databases. Here are several examples and other useful links to explore:

What is RedCap?

Research Electronic Data Capture is a secure and easy-to-use web application for building and managing online surveys and databases. It is specifically designed to collect research data. No background knowledge or technical experience is needed to use REDCap, giving researchers the ability to directly

control and manage their own projects. It can be used for the research registries, randomized clinical trials, quality Improvement (QI)/ Quality Assurance (QA) projects, and etc.

[Conducting Surveys \(phsa.ca\)](#)

[BCCHR REDCap-PRO](#)

rc.bcchr.ca/redcap/hooks/portal/access.php

[UBC ARC REDCAP | arc.ubc.ca](#)

[Resources – REDCap \(projectredcap.org\)](#)

[REDCap Training – CENTER FOR RESEARCH INFORMATICS \(uchicago.edu\)](#)