

Naming and Coding Best Practices in REDCap

One benefit of REDCap is the ability to name and code variables when building the project. Intentional variable naming and coding can have many benefits, including:

- Decreasing the time needed to clean data
- Ensuring consistency across similar question types
- Improving clarity across team members

We recommend following these best practices for variable naming and coding.

This document is primarily based on a presentation at the 2019 REDCapCon in Vancouver, British Columbia, Canada. The presentation, “Variable Naming and Choice Coding Consistency,” was given by Ceci Chamorro (Duke University), Janet Keener (East Tennessee State University), and Alexander Maclay (Sheppard Pratt Health System).

Instrument Naming

Having clear instrument names can help with long-term data management and clarity across team members.

- Instrument names should include:
 - Full instrument name
 - Abbreviated name
 - Version number
 - For example: Activities-Specific Balance Confidence Scale – ABC-6
- If the name is too long, be as descriptive as possible while staying within the size limitations.
 - Instead of: Patient-Reported Outcomes Measurement Information System (PROMIS): Upper Extremity Function 8 Version 2
 - Use: PROMIS Upper Extremity 8v2
 - To change an instrument’s name, select the “Choose Action” dropdown menu next to the instrument in question and select “Rename.”

The screenshot displays the 'Data Collection Instruments' management interface. At the top, there are sections for 'Add new instrument:' (with 'Create', 'Import', and 'Upload' buttons), 'Survey options:' (with 'Survey Queue', 'Survey Login', and 'Survey Notifications' buttons), and 'Automated Survey Invitation options:' (with 'Upload or download Auto Invitations' and 'Re-evaluate Auto Invitations' buttons). Below these is a table of instruments:

Instrument name	Fields	View PDF	Enabled as survey	Instrument actions	Survey-related options
Demographic Survey	7			Choose action ▾	Survey settings + Automated Invitations
Patient Health Questionnaire 9 - PHQ-9	14			Rename Copy Delete Download instrument ZIP	Settings + Automated Invitations

Variable Naming

Required Characteristics of Variable Names

- Can contain lower-case letters, numbers, and underscores
- Cannot begin with a number
- Must be unique within a project
- Variable names **cannot** be changed after a project has moved to production
 - Be careful **not** to change variable names if using the data dictionary to make modifications to an in-production project

Recommended Variable Naming Strategies

- Keep variable names under 26 characters.
 - Some statistics packages cut off variable names at 26 characters.
- For standardized questionnaires:
 - Use instrument abbreviation and the question number, e.g.
 - The first question on the Activities-Specific Balance Confidence Scale – 6: abc1
 - The seventh question on the Patient Health Questionnaire – 9: phq7
- For non-standardized questionnaires:
 - Begin variable name with a brief description of the instrument (2-6 characters)
 - Use common abbreviations or intuitive naming
 - Separate the instrument name from the descriptor with an underscore
 - demo_dob (demographic form, date of birth)
 - fd_time_bfast (food diary, time ate breakfast)
 - Yes/no questions can use a leading verb that indicates what “yes” means
 - health_pregnant vs. health_ispregnant

Variable Design and Coding

There are many field types in REDCap. This section will focus on text fields and multiple choice questions. For more information about and best practices for other field types, see the [“Survey Design Best Practices”](#) document.

Validating Text Fields

- When using text fields, utilize field validation when appropriate
 - Common field validations are email, phone number, integer, number, and an array of date-time options
- Apply minimums and maximums to fields validated as integers, numbers, or date-times
 - Dynamic ranges can be set by piping values from previous fields and/or using “today” or “now” on date-time fields
 - Use the action tag @FORCE-MINMAX to enforce the minimum and maximum values
- Convert notes boxes and unvalidated text boxes to multiple choice questions, if possible

Coding Multiple Choice Questions

- Use externally established and validated codes, when applicable. Established codes are often related to the final scoring of a scale.
 - Patient Health Questionnaire – 9 coding:
 - 0, Not at all
 - 1, Several days
 - 2, More than half the days
 - 3, Nearly every day
- Response sets with more than two options often begin with 1.
 - Likert scale for level of agreement:
 - 1, Strongly disagree
 - 2, Disagree
 - 3, Neither agree nor disagree
 - 4, Agree
 - 5, Strongly agree
- For “I don’t know”/“Not applicable” responses, use codes well outside of a normal range.
 - 999, Not applicable
- Dichotomous options should be coded where the affirmative = 1 and the negative = 0
 - REDCap comes with two dichotomous variables (yes/no and true/false) where 1 = “yes” or “true” and 0 = “no” or “false”; the codes cannot be modified for these field types.
 - If you need to use a radio button to add additional options, such as “maybe” or “I don’t know,” ensure that coding is consistent with existing dichotomous variables.
 - 0, No
 - 1, Yes
 - 998, I don’t know
 - 999, Not applicable