

# ONTARIO INSTITUT BRAIN ONTARIEN INSTITUTE DU CERVEAU

## **Data Quality Framework**

**REDCap Tools & Procedures** 

Version 1.2 August 8, 2024

Written and reviewed by the Brain-CODE Team



### **Table of Contents**

INTRODUCTION	3
STUDY DESIGN	3
MISSING DATA CODING GUIDE	4
STUDY LAUNCH & DATA ENTRY	5
ACCOUNT REQUEST FORM	5
REDCAP MISSING DATA CODE FEATURE	5
MISSING DATA FLAGGING FORM	6
PARTICIPANT STATUS FORM	8
STUDY MONITORING	12
DATA REVIEW FREQUENCY	12
DATA QUALITY RULES	13
DATA RESOLUTION WORKFLOW	15
OBI DATA QUALITY REPORTS	18
HELP & RESOURCES	



### Introduction

Given the importance of high-quality data to research and translation to clinical care, OBI's Data Quality Framework (DQF) intends to set standards across its sponsored research studies and uphold the best data quality practices. REDCap is a tool on the Brain-CODE platform used to build and manage databases for clinical data. REDCap provides many features that help ensure the usability and quality of the data that is collected. This manual describes the REDCap data quality tools and elements of project design and data collection that will help Data Producers meet the requirements of the Data Quality Framework. Usage of these tools will help ensure the study data are trustworthy and shareable, and any changes made are well-documented.

### **Study Design**

As mentioned in the DQF Study Quality Checklist (see Stage 1 document #5), it is important to design REDCap electronic case reports forms (eCRFs) with survey administration, data entry, and future usability of the data in mind. Forms *must* be designed with consideration of participant burden, and should be easy to use and understand. Forms *must* also have consistent <u>variable</u> <u>naming</u> as well as well-designed branching logic, defined data types, and proper field validation. Missing data codes (see <u>below</u>) are *strongly recommended* to be hardcoded within REDCap projects to replace and explain any missing data.

The table below describes various categories of missingness as part of the DQF standards (see Stage 2 document #3) that are meant to be used for partially completed forms within REDCap. With these elements in place, Data Producers are *strongly recommended* to use the Data Quality tools in REDCap to regularly check for any discrepancies within the data (described in the Study Monitoring <u>section</u>). The Missing Data Flagging <u>form</u> (built into the project as an eCRF) is also *strongly recommended* to explain form-level missing data or many unanswered fields within partially completed forms. Additionally, it is *strongly recommended* that Data Producers use a Participant Status <u>form</u> (built into the project as an eCRF) to track participant status and their study progression.



### MISSING DATA CODING GUIDE

Code	Label	Definition
unable_to_perform	Unable to perform	The participant could not participate in the assessment/task due to a physical or cognitive condition.
task_declined	Task declined	The participant chose not to participate in the assessment/task or declined to provide a response.
out_of_time	Ran out of time	The participant was unable to complete the assessment/task within the scheduled timeslot.
admin_error	Administrative error	The administrator of the assessment/ task made a mistake that caused the entry to be invalid or otherwise unobtainable. E.g., miscalibration of equipment by administrators; incorrect administration; scheduling conflicts.
technical_error	Technical error	The equipment or tools (hardware or software) used for data acquisition, collection, or processing failed. E.g., computer crashed; power failure; response button box or keyboard was not properly hooked up or defective.
missing_other	Other	Data is missing for a reason that is not explained by any of the available missing data codes. <i>Please</i> <i>provide a description if possible.</i> *
answer_unknown	Unknown value	The participant could not provide a response to the question because they did not know how to answer it. <i>This code is only to be used if the field</i> <i>does not include a "do not know" option.</i> **
not_applicable	Not applicable	The assessment/task was not performed, or a value could not be derived because it did not apply to the participant. E.g., a questionnaire asks if a participant can organize his/her own medications, but the participant does not take any medications, so the question does not apply. <i>This code is only to be used if the field does not include a "not applicable" option.</i> ***

\*If this code is used, please include a description as a comment within a data query or data verification textbox within the <u>Data Resolution Workflow</u> if possible. For missing forms, please use the text box once this option is selected within the Missing Data Flagging form.

\*\*Please note these scenarios are not considered missing data, as "do not know" is an answer and any items that are "not applicable" are not expected to contain data. It is strongly recommended that these Missing Data Codes



are all hardcoded in REDCap so the missingness data can be captured from the single missing code dropdown menu.

### Study Launch & Data Entry

### ACCOUNT REQUEST FORM

Once the study is launched, ensure that all team members have appropriate user rights according to their role (e.g., opening and closing data queries, locking records, editing completed surveys, etc.). This can be done by the Program Manager or Neuroinformatics Lead submitting requests through the Account Request form on the <u>Brain-CODE website</u> under *Forms*.

*Please view the following articles in the Knowledge Base for more information on <u>user roles</u> and <u>access requests</u>.* 

#### **REDCAP MISSING DATA CODE FEATURE**

After participants have completed their questionnaires, whether it be done directly or with a coordinator or caregiver, there may fields that were left unanswered. Please follow the steps below to fill blank fields with missing data codes. **Note**: within REDCap projects, the missing data codes feature is only meant to be used to replace blank fields within partially completed instruments, NOT incomplete instruments where all fields are blank.

- 1. Within the participant record, select a form.
  - If the form was completed as a survey directly by the participant, the box shown below will appear at the top of the form. Click the **Edit response** button.



- 2. Click the **M** button beside the blank response field. *Should the M button not be visible, please contact the Neuroinformatics Lead as this feature may not have been enabled.*
- 3. Select the applicable missing data code from the dropdown.



Record ID	OBI00_TST_0001	
Postal Code		
Firms Chattan	Mark field as:	
Form Status	[Clear value]	
Complete?	Unable to perform (unable_to_perform)	
	Task declined (task_declined)	
Lock this instrument?	Ran out of time (out_of_time)	
If locked, no user will be able to modify this instrument for this record until someone with Instrument Level Lock/Unlock privileges unlocks it.	Administrative error (admin_error)	
	Technical error (technical_error)	
	Other (provide description) (missing_other)	
	Unknown value (value_unknown)	
	Not applicable (not_applicable)	

For more information on this feature and its set up, please view this <u>article</u> in the Knowledge Base.

#### MISSING DATA FLAGGING FORM

The Brain-CODE team has designed a form to flag missing data within a REDCap project. Firstly, the form can be used to flag missing forms with a missing data code (taken from the <u>coding</u> <u>guide</u>). Secondly, for forms that are not missing but that contain many missing items, a missing data code can be selected which will be applied to all missing items in that form. *Please note that this form does not apply the specified missing data codes to the identified forms; the Missing Data Flagging form data will be reviewed as part of the Data Quality Reporting pipeline external to REDCap*.

This form works in conjunction with the Participant Status <u>form</u> through branching logic to only include forms that are expected to be complete based on the participant's progress through the study protocol. For a form to be shown in the Missing Data Flagging form, two conditions need to be met: 1) the form either has no saved data or a Form Status of 'Incomplete' and 2) the timepoint of the form has been marked as completed in the Participant Status form.

To use this form, please follow the steps below:

1. For each shown form, select the relevant missing data code. If the form contains no data, this missing data code will be used to flag the entire form as missing in the Data Quality Reporting pipeline.



BASELINE_ARM_2			
demographics_adult_v21:	H		
demographics_child_v21:	Not a missing form Unable to perform		
medical_history_v2:	Task declined Out of time		
patient_health_questionnaire_9_phq9:	Technical error		
Form Status	Not applicable		
Complete?	⊢		

- 2. If the form is partially complete (i.e., contains some data but has missing items), the missing data code that is selected from the dropdown menu will be applied to each missing item in the form in the Data Quality Reporting pipeline. This is implemented to reduce the time required to apply a missing code to all missing items in lengthier forms.
- 3. To indicate that a form is not missing but without applying a single missing data code to all missing items in the Data Quality Reporting pipeline, select 'Not a missing form' from the dropdown menu. A warning will appear to remind you that the Form Status should be marked as 'Complete'. Please fill in any missing items with missing codes as described in the previous section before changing the Form Status to 'Complete'.

BASELINE_ARM_2				
demographics_adult_v21:	B Not a missing form			
Return to this form and change the form status to "Complete"				
demographics_child_v21:				
medical_history_v2:				
patient_health_questionnaire_9_phq9:				
Form Status				
Complete?	) Incomplete V			



### PARTICIPANT STATUS FORM

The 'Participant Status' form was designed by the Brain-CODE team to track participant status within a study. The form allows Data Producer team members to mark when a participant has completed all required assessments/tasks according to the study protocol and, if the study has multiple visits, which visits have been completed. The form will be customized according to the number of visits for the given study. This form is intended to be updated should a participant's status change (e.g., withdrawal), and if there are multiple timepoints, after a participant has completed a subsequent visit. Please follow the steps according to the participant lifecycle below:

 Enrollment: once a participant has enrolled in the study and their consent has been collected, they are part of active data collection. The 'Active' option is to be selected under participant status and a banner appears labelled as ACTIVE. At this stage, leave all fields blank and the Form Status as 'Incomplete' to indicate that this form will be returned to (enter additional comments as necessary).

	Participant Status: <u>ACTI</u>	<u>VE</u>
The part	ticipant is still actively enrolle	d in the study.
What is the status of the participant?		
* must provide value		$\sim 6$
Active		
O Complete		
Withdrew consent     Withdrew by DI		
O Screen fail		
Active (undergoing study protocol). Complete (completed all	timenoints) Withdrawn (no longer na	rese reset of study) Lost to follow-up (upreachable)
Active (undergoing study protocol), complete (completed and		(unreachable)
Which timepoints has the participant comple	eted?	
baseline_arm_1		
□ followup_arm_1		
Please check box if no more data is expected to be collected f	for the given visit	
Additional notes:		
	20	
		Expan
Form Status		ar i por i
Complete?		Incomplete 🗸

2. Active data collection in longitudinal studies: once a participant has completed the required assessments/tasks for a particular timepoint, return to the form and mark the respective checkbox. Keep the **Form Status** marked as **'Incomplete'**.



Participant State	is: <u>ACTIVE</u>
The participant is still actively	enrolled in the study.
What is the status of the participant?	θ
* must provide value	$ ightarrow \mathbb{M}$
Active	
○ Complete	
○ Withdrew consent	
O Withdrawn by Pl	
O Lost to follow-up	
O Screen fail	rosot
Active (undergoing study protocol), Complete (completed all timepoints), Withdrawn (no	longer part of study), Lost to follow-up (unreachable)
Which timepoints has the participant completed?	H Q
✓ baseline_arm_1	
□ followup_arm_1	
Please check box if no more data is expected to be collected for the given visit	
Additional notes:	H
Additional notes.	$\sim M$
	Expand
Form Status	
Complete?	🛞 Incomplete 🗸

- 3. *Study protocol complete*: once a participant has gone through the entirety of the study protocol, mark all necessary checkboxes for completed visits and change the participant status to 'Complete'. The banner will then change to be labelled as **COMPLETE**. Mark the **Form Status** as **'Complete'**. In addition:
  - a. If a participant is missing particular assessments/tasks within a visit, refer to the Missing Data Flagging form.



	Participant Status: COMPLETE	
The partici	pant has completed the entire study protocol.	
What is the status of the participant?		
* must provide value		ightarrow M
○ Active		
Complete		
O Withdrew consent		
O Withdrawn by PI		
○ Lost to follow-up		
○ Screen fail		
Active (undergoing study protocol), Complete (completed all t	timepoints). Withdrawn (no longer part of study). Lost to follow-up (unreach	reset
Date of study completion:		H P M
Which timepoints has the participant comple	eted?	H P M
✓ baseline arm 1		
✓ followup_arm_1		
Please check box if no more data is expected to be collected f	for the given visit	
Additional notes:	- M	
		Expand
Form Status		
Complete?	🖰 Complete 🗸	

- 4. *Withdrawal*: should a participant withdraw from the study (whether of their own accord or deemed necessary by the principal investigator) or be lost to follow-up, select appropriate option. The banner will be labelled as **WITHDRAWN**. Provide the date of withdrawal and select a reason for withdrawal from the dropdown. If any visits had been completed prior to withdrawal or being lost to follow-up, check off the necessary timepoints. Mark the **Form Status** as **'Complete'**.
  - a. If the participant has withdrawn and has requested for their data to be removed from Brain-CODE, please contact the <u>Brain-CODE Help Desk</u>.
  - b. If a participant does not provide their data and does not respond to follow-ups (according to site/study procedures) from the study team, the participant is considered as **'lost to follow-up'**. Should a participant become responsive again, please update the form and status accordingly.



Participant Status: <u>WITHDRAWN</u>				
The participant has witho	drawn from the study.			
What is the status of the participant?	H			
* must provide value	$\sim M$			
O Active O Complete				
Withdrew consent				
O Withdrawn by PI				
○ Lost to follow-up				
O Screen fail	reset			
Active (undergoing study protocol), Complete (completed all timepoints), Withdrawn (	(no longer part of study), Lost to follow-up (unreachable)			
Date of withdrawal:				
Today Y-M-D				
Reason for withdrawal of consent (select primary reason):	B → Medical reasons unrelated to reported condi			
Which timepoints has the participant completed?	H P			
✓ baseline arm 1				
✓ followup_arm_1				
Please check box if no more data is expected to be collected for the given visit				
Additional notes:				
	70			
	Expand			
Form Status				
Complete?	😑 Complete 🗸			

5. *Additional features*: using the Participant Status form will also allow for customizations to form display logic within the REDCap project that will disable the relevant forms for participants that were marked as withdrawn or lost to follow-up. Therefore, coordinators will be able to easily identify participants that are not expected to provide more data.



Participant Status: WITHDRAWN						
The participant has withdrawn from the study.						
What is the status of the participant?				θ		
* must provide value				$\supset M$		
Withdrew consent						
O Withdrawn by PI						
O Lost to follow-up						
Active (undergoing study protocol), Complete (completed all timepoints), Wit	thdrawn (no longer pa	art of study), Lost to follo	w-up (unreachable)	reset		
Date of withdrawal:				H M		
2024-07-22 🛅 Today Y-M-D						
Reason for withdrawal of consent (select primary reason	n):	Prefer not to answ	ver	~		
Which timepoints has the participant completed?				H P M		
✓ baseline_arm_2						
□ followup1_arm_2						
□ followup2_arm_2						
Please check box if no more data is expected to be collected for the given vis	sit					
Additional notes:						
	$\supset M$					
				Expand		
Data Collection Instrument	baseline	followup1	followup2	closeout		
Demographics Adult V21						
Demographics Child V21						
Kindlr						
Medical History V2 (survey)						
Patient Health Questionnaire 9 Phq9 (survey)	$\bigcirc$					
Participant Status Form				۲		
Missing Data Flagging Form						

### **Study Monitoring**

#### DATA REVIEW FREQUENCY

While data is being collected and entered, it is important that Neuroinformatics Leads (or Data Monitors) are reviewing the study data regularly. The Neuroinformatics Lead *must* conduct a final quality check before notifying the OBI Data Quality team that the data is ready for review.



### DATA QUALITY RULES

To accurately identify any missing data or errors, it is *strongly recommended* to use the Data Quality Rules tool. *Please note that these tools do not include any data entered in the Missing Data Flagging form*. Please follow the steps below to check for the predefined discrepancies and any custom rules that were created:

1. The Data Quality page will display a Data Quality Rules dashboard with a predefined set of discrepancies that are searched for and the ability to add any new rules that are necessary to the project at the bottom.

				Upload or do	wnload Data Quality Rul	les 🗢
Dat	a Qualit	y Rules	Execut	e rules: All	All except A&B Clear	
		, ,		A	pply to: All Records 🗙	
	Rule #	Rule Name	Rule Logic (Show discrepancy only if)	Real-time execution ?	Total Discrepancies	Delete rule?
	А	Missing values*	-		Execute	
	В	Missing values* (required fields only)	•		Execute	
	С	Field validation errors (incorrect data type)	-		Execute	
	D	Field validation errors (out of range)	-		Execute	
	E	Outliers for numerical fields (numbers, integers, sliders, calc fields)**			Execute	
	F	Hidden fields that contain values***	-		Execute	
	G	Multiple choice fields with invalid values	•		Execute	
	Н	Incorrect values for calculated fields	-		Execute	
	1	Fields containing "missing data codes"	-		Execute	
	Add					
		Enter descriptive name for new rule (e.g., Participants below age 18)	Enter logic for new rule (e.g., [age] < 18) How do Luse special functions?	Execute in real time on data entry forms ?		

2. To find any discrepancies for each individual rule, click the **Execute** button for that rule or click **All** in the top-right corner to execute all rules. REDCap will calculate each rule and provide you with a total number of discrepancies found.



				Upload or do	wnload Data Q	uality Rules	s 🗢
Dat	a Qualit	y Rules	✓ Processing Complete! Execut	e rules: All	All except A&B	Clear	
				A	pply to: All Red	cords 🗸	
	Rule #	Rule Name	Rule Logic (Show discrepancy only if)	Real-time execution ?	Total Discre	pancies	Delete rule?
	А	Missing values*	-		2 <u>ex</u>	port   <u>view</u>	
	В	Missing values* (required fields only)	-		0		
	С	Field validation errors (incorrect data type)	-		0		
	D	Field validation errors (out of range)			0		
	E	Outliers for numerical fields (numbers, integers, sliders, calc fields)**	-		0		
	F	Hidden fields that contain values***	-		0		
	G	Multiple choice fields with invalid values	-		0		
	н	Incorrect values for calculated fields	-		1 <u>ex</u>	port   <u>view</u>	
	I.	Fields containing "missing data codes"			0		
	Add						
		Enter descriptive name for new rule (e.g., Participants below age 18)	Enter logic for new rule (e.g., [age] < 18) How do I use special functions?	Execute in real time on data entry forms ?			

3. To view the list of results, click the **View** button in the **Total Discrepancies** column for a given rule.

Rule: Missing values*         Discrepancies found: 1         Image: CSV (2 exclusions not displayed - view)				
Record	Discrepant fields with their values	Status	Resolve issue 😡	
test_001 Baseline	"Nome:" test_name = [ <u>no data]</u>	Missing value	💭 0 comments	

- 4. The **Resolve Issue** column brings up the Data Resolution Workflow pop up (described in the next section), which allows project users to open, respond to, or close a data query which is used for documenting details of the data issue, including the origin of the issue, who resolved the issue, and how it was resolved.
- 5. Please resolve these issues with the necessary actions, such as following up with the participant, using missing data codes, etc.

For more information on this tool, please view this <u>article</u> in the Knowledge Base.



#### DATA RESOLUTION WORKFLOW

The Data Resolution Workflow (DRW) module works with the Data Quality Rules tool and enables users to open a workflow for tracking and documenting the process of resolving issues with data in the project. It is *strongly recommended* to use this tool to identify and resolve any errors or missing values within your team. Please follow the steps below:

1. As mentioned earlier, during your quality checks, you may have found an error or blank field that hadn't been accounted for. Once identified, click the grey balloon next to the field.

Record ID	OBI00_TST_0001
Postal Code	
Form Status	
Complete?	B Complete V

2. This prompts the DRW pop-up to appear that allows you to open a query. You can assign another user to the query to notify them of the issue to resolve, and leave a comment.

a Data Resolution Workflow			
EB <u>VIDEO: Data Resolution Workflow</u> This pop-up displays the Data Resolution Workflow for the specified record for a given field and/or Data Quality rule. Users with appropriate user privileges may open data queries to begin a documented process of resolving an issue with the data. Opened data queries may thus be responded to by users with appropriate privileges, and then they may be closed once the issue has been resolved. All data queries can also be viewed on the Resolve Issues page in this project.			
Record ID: <u>OBI00_TST_0001</u> Event: Baseline (Arm 1: Arm 1) Field: test5_pstl_cde ("Postal Code") Status:  Mot Opened			
Date/Time	User	Comments and Details	
07/19/2024 9:35am	jzych2	<ul> <li>○ Verified data value</li> <li>─ OR</li> <li>○ Open query</li> <li>Assign query to a user (optional): select user </li> <li>✓ Notify this user of their assignment using: Email</li> <li>□ REDCap Messenger</li> </ul> Comment: Please use a missing data code.	
		Open query Cancel	

3. Once a query is opened, the grey balloon beside the field changes to yellow with a red exclamation mark.



Record ID	OBI00_TST_0001
Postal Code	)
Form Status	
Complete?	B Complete V

4. To respond or close the query, click on the yellow balloon. This will prompt the DRW pop-up to appear again. You can respond to the query using the **Choose Response** dropdown, or if the issue was resolved, leave a comment regarding the action taken and close the query.

🗣 Data Resolution Workflow				
EB VIDEO: Data Resolution Workflow This pop-up displays the Data Resolution Workflow for the specified record for a given field and/or Data Quality rule. Users with appropriate user privileges may open data queries to begin a documented process of resolving an issue with the data. Opened data queries may thus be responded to by users with appropriate privileges, and then they may be closed once the issue has been resolved. All data queries can also be viewed on the Resolve Issues page in this project. Record ID: <u>OBIO_TST_0001</u>				
Event: Baseline (Arm 1: Arm 1) Field: test5_pstl_cde ("Postal Code") Status: 🙀 Open / Unresolved (unresponded)				
Date/Time	User	Comments and Details		
07/19/2024 9:36am	jzych2	Comment: "Please use a missing data code."		
07/19/2024 9:36am	jzych2	<ul> <li>Reply with response: choose response Upload file (optional): Upload document</li> <li>OR OR</li> <li>Close the query</li> <li>Comment:</li> <li>Followed up with participant. Entered 'task_declined' missing data code.</li> </ul>		
		Close the query Cancel		

5. Once a query has been resolved and closed, the yellow balloon beside the field now has a green check mark.

Record ID	OBI00_TST_0001		
Postal Code	task_declined		
20 (a)	Task declined (task_declined)		
Form Status			
Complete?	Complete 🗸		



6. To view the query history or to reopen the query if need be, click the yellow balloon.

#### 🗣 Data Resolution Workflow

x

#### E VIDEO: Data Resolution Workflow

This pop-up displays the Data Resolution Workflow for the specified record for a given field and/or Data Quality rule. Users with appropriate user privileges may open data queries to begin a documented process of resolving an issue with the data. Opened data queries may thus be responded to by users with appropriate privileges, and then they may be closed once the issue has been resolved. All data queries can also be viewed on the Resolve Issues page in this project.

#### Record ID: <u>OBIO0\_TST\_0001</u> Event: Baseline (Arm 1: Arm 1) Field: test5\_pstl\_cde ("Postal Code") Status: 🥪 Closed / Resolved

Date/Time	User	Comments and Details
07/19/2024 9:36am	jzych2	Action: Opened query Comment: "Please use a missing data code." <b>₹</b> Assign query to a user
07/19/2024 9:39am	jzych2	Action: Closed query Comment: "Followed up with participant. Entered 'task_declined' missing data code."
07/19/2024 9:41am	jzych2	Data Changes Made: test5_pstl_cde = 'task_declined'
07/19/2024 9:41am	jzych2	Reopen the closed query     Comment:
		Reopen query Cancel

 Alternatively, data queries can be accessed from the Resolve Issues page on the lefthand menu. Here, the Data Resolution Dashboard shows all queries (whether open or closed). You can view a data query by selecting the comments button with the yellow balloon in the first column.

Data Resolution	n Dashboard	Filters:	Open / unresolved issues (2) V All fields and rules V All events V User assigned (all users) or not assigned V				
Click button to view data query	Record	Data Quality rule and/or Field		User Assigned	Days Open	First Update	Last Update
Ga 3 comments	OBIO0 TST 000 Baseline	Field: pstl_cd (Postal Code)	e	-	0	jzych2 (02/09/2024 4:31pm): "Please use a missing data code"	jzych2 (02/09/2024 4:45pm): "test"

For more information on this tool, please view this <u>article</u> in the Knowledge Base.



### **OBI Data Quality Reports**

Upon receival of data from Data Producers, the OBI Data Quality team will conduct a secondary REDCap quality check. As part of this process, they will review the results of the Data Quality Rules tool, the Data Resolution Dashboard, Participant Status form, and the Missing Data Flagging form for all records, before generating the Data Quality Report that is to be shared with the Data Producer. If the secondary check identifies any major data quality issues, the OBI Data Quality team will notify the Data Producer, allowing for the Data Producer team members to resolve the issues.

### **Help & Resources**

For account reactivation, please contact your Neuroinformatics Lead or Program Manager.

For password reset and technical assistance, please contact the Brain-CODE help desk at <u>help@braincode.ca</u>.

### **General Assistance and Questions:**

Please refer to Brain-CODE for Knowledge Base articles, training videos, and help desk ticket submission:

https://www.braincode.ca/content/help-resource