# MATRIX-007 Study-Specific Procedures (SSP) Manual Section 7 – Data Collection and Management

## TABLE OF CONTENTS

# 7. Purpose of the Data Management Study Specific Procedures

The overall purpose of the Data Management SSP for MATRIX-007 is to explain the process for data collection and management to ensure human subjects protection and reliable study data.

More specifically, the SSP will:

- Explain the roles of each member of the data management team for the study.
- Provide a timeline and general instructions for data collection that happens over the course of the study.
- Explain how to use REDCap for data collection and the quality control and quality assurance functions and processes in REDCap.

The SSP will not provide specific instructions for each Case Report Form (CRF). For specific instructions on how to fill out each CRF, refer to the MATRIX-007 CRF Completion Guide (CCG).

## 7.1. Data Management Team

The main responsibilities of the Data Management Team are to collect data in accordance with the study procedures, specifically the study protocol and Data Management SSP; ensure that data are reviewed for errors and that any identified errors are corrected; and to report on data collection efforts. The Data Management Team has three roles that are to be filled regardless of the specific job title of a given individual: Study Data Manager (study level), Country Data Manager (country level), and Clinical Research Assistant (site-level).

### 7.1.1. Study Data Manager (study-level)

The data management and statistical support team for this study is managed by FHI 360. The primary study data manager of contact will be Sam Nicholes. Send all general questions to snicholes@fhi360.org, remember to CC the MATRIX-007 CRM, Tara McClure (TMcClure@fhi360.org). For urgent questions that need immediate resolution, you may contact Sam Nicholes at +18019608416. Doris Marwanga (DMarwanga@fhi360.org) will be the secondary data manager of contact.

The Data Manager's responsibilities are to be the main administrative manager of the data for MATRIX-007; to write and maintain the Data Management SSP, and the CCG; to program and maintain REDCap functions including the CRFs; generate reports on statistics and data quality across all sites; send queries to the country study coordinators or clinical research assistants; respond to questions and concerns from the Country Data Managers (CDMs) and Clinical Research Assistants (CRA).

### 7.1.2. Country Data Manager (CDM)

The responsibilities of the CDM are to perform audits, systematic reviews, or other "spot checks" to ensure CRF completeness and identify missing, aberrant, or discrepant values; elevate questions and concerns to the data manager that cannot be resolved by the in-country team; receive queries from the data manager; be trained on data management processes; and train clinical research assistants on data management processes. Country teams may choose to have the Study Coordinator perform this role.

### 7.1.3. Clinical Research Assistant (CRA)

CRAs are responsible for data collection, ensuring that CRFs are complete and free of error; receiving queries from the Data Manager, and CDM; resolving missing, aberrant, and discrepant values identified by the Data Manager and CDM; elevating questions and concerns to the CDM; being trained on data management processes.

## 7.2. CARE PrEP Data SharePoint

The study will utilize a SharePoint site, hosted and managed by FHI 360. The SharePoint will operate as a secure, limited access work area for the study team to handle participant data and serve as a study resource repository. The SharePoint is partitioned by Libraries, with access granted based on content and intent of the library workspace. Library content and access is detailed in Table 7-1 below. See SSP section 2 for a description of the user group members. The SharePoint site owners, who are the Study Data Manager and the CRM, manage all permission and content posted on the site. Access it by invitation only; links to files should not be shared with anyone without explicit access granted by the site owners.

**Table 7-1: CARE PrEP Data SharePoint Access**

| Library | Use/Content | User Group Access |
|---|---|---|
| Country/Site Documents | Contains sub folders for each study country used to:<br>• Upload Safety Sub-committee (SSC) queries and SAE/SH reports<br>• Store Participant Tracker<br>• Store electronic source documents (by PID only)<br>• Repository for SOPs | • LOG members (all folders)<br>• SSC members (all folders)<br>• Country team members (PI, SC, CDM, CRA) (access to only respective county folder) |
| Study Resources | Repository for all study materials, training materials issued the study | • LOG members<br>• SSC members<br>• Country team members |
| Study Management | Contains: | • LOG members (all folders) |

| | | |
|---|---|---|
| | • SSC working folder - linked to country team safety folders for queries, SAE, SH reports; working folder for Congenital Anomaly reviews<br>• Study Reports folder - to store generated from REDCap | • SSC members (all folders)<br>• Country team members (only access to Study Reports folder) |
| Data Back-up | Contents raw back-up data files from REDCAP | • ONLY Study Data Manager, CRM, |

### 7.3. REDCap Overview

REDCap is a data collection and storage platform that is open-sourced to organizations, businesses, universities, and non-profits. FHI 360 has available instances of REDCap hosted on secure, US-based servers. FHI 360's instance of REDCap is managed by the department, Information Solutions and Services (ISS) which functions as the organization's principal IT unit. REDCap storage is backed up per ISS policies and regulations.

REDCap was created to primarily serve as a data collection and storage platform for clinical research studies, such as MATRIX-007. It has built-in functions that offer skip patterns; data entry constraints; randomization modules; scheduling modules; API generation; quality control and assurance functions; report generation; records management; data de-identification; and data import and export, etc. The data management team for MATRIX-007 will make use of all necessary functions and features to ensure accurate data entry and data quality conducive to the study's longitudinal design.

The security settings are controlled by the REDCap system administrators at the FHI 360 and therefore are not modifiable by the end users. The user passwords do not expire. After 30 minutes of inaction, the REDCap system will automatically log a user out. After five unsuccessful login attempts, the user's account will be locked and the user will need to contact the administrators to regain access to their account. A user's account will automatically deactivate after 180 days of inactivity.

All research staff who require access to enter or view data within REDCap for their site will need a REDCap user account through the FHI 360 Data Manager. Safety Sub-Committee (SSC) members may also request accounts but are not required to do so. FHI 360 will provide sites with instructions for obtaining accounts.

New employees will also need to request REDCap account access; you cannot allow a new employee to use a former employee's account. If an employee leaves the organization or moves to a different department, FHI 360 Data Manager should be notified so that their access to the MATRIX-007 Study REDCap database can be revoked.

Study teams will be provided with password protected laptops and tablets for access to REDCap. Paper versions of CRFs will also be provided to teams as backups in case of technical or electronic issues and failures. All devices and hardcopies will be kept in a secure unit, such as a locked filing cabinet accessible only to study staff. How each sites manages equipment storage and secure use should be outlined in SOPs.

REDCap allows for mobile and offline data collection via tablets. Per MATRIX-007 protocol, off-site and home visits are allowed. When data collection happens offline via tablets, or on paper copies,

tablet syncing and upload, and electronic data entry from paper copies will occur shortly after the Research Assistant has laptop and internet access.

### 7.4. REDCap User Accounts and Permissions

The Data Manager will provide user accounts for the following persons on country study teams: MATRIX-007 Principal Investors, Country Data Managers (or Study Coordinators who perform this role), and Clinical Research Assistants (and Study Coordinators who perform this role as well).

User access privileges within REDCap are customizable and controlled by the data managers of the study. The study staff responsible for direct data entry are assigned privileges that allow them to create and edit participant records, but they are not able to delete or rename records. User permissions for data collection and access will be managed by the Data Manager. Generally, user permissions will be granted as follows:

**Study Data Manager**: All administrator rights and permissions for MATRIX-007 including access to record level data, ability to read and write data, ability to audit and query data, ability to generate reports, ability to import and export data, and ability to grant and withdraw user permissions.

**Country PI:** Permissions and right to view and edits data, and most importantly to build reports and charts; lock and unlock records.

**Country Data Manager/Study Coordinator**: Administrator permissions similar to the data manager, with the exception of being able to import data into the database, and the ability to grant and withdraw user permissions. These permissions will only pertain to the Country Study Coordinator's respective country.

**Clinical Research Assistant**: Basic user level rights will be granted so that the RA is able to perform the necessary functions as a data collector as well as respond to queries and audits initiated by the Country Study Coordinator or the Data Manager.

**SSC Members:** Read-only access so that they are unable to change any data elements.

**Table 7-2: REDCap User Access Rights**

| User Rights and Access | Study Data Manager/Data Management HQ | Country Data Manager | Clinical Research Assistant | Principle Investigator |
|---|---|---|---|---|
| Project Design and Setup | ☑ | | | |
| User Rights | ☑ | | | |
| Data Access Groups | ☑ | | | |
| Data Viewing Rights | View and Edit | View and Edit Country DAGs | View and Edit Site DAG | View and Edit |
| Data Export Rights | Full data set | No Access | No Access | No Access |
| Alerts & Notifications | ☑ | ☑ | ☑ | ☑ |
| Reports & Report Builder | ☑ | | | ☑ |
| Stats & Charts | ☑ | | | ☑ |
| Calendar & Scheduling | ☑ | ☑ | ☑ | |

| | | | | |
|---|---|---|---|---|
| Data Import Tool | ☑ | | | |
| Data Comparison Tool | ☑ | | | |
| Logging | ☑ | | | |
| File Repository | ☑ | ☑ | | ☑ |
| Record Locking Customization | ☑ | | | ☑ |
| Lock/Unlock Records | ☑ | | | ☑ |
| Data Quality (create/edit rules) | ☑ | ☑ | | |
| Data Quality (execute rules) | ☑ | ☑ | | |
| Data Resolution Workflow | Open, close, and respond to queries | Open, close, and respond to queries | Respond to queries | |
| API | ☑ | | | |
| REDCap Mobile App | ☑ | ☑ | | |
| Create Records | ☑ | | | |
| Rename Records | ☑ | | | |
| Delete Records | ☑ | | | |

## 7.5. Data Collection

CRFs have been programmed into the MATRIX-007 REDCap for direct data entry. PDFs of the CRFs are available to view or download within the MATRIX-007 Project in REDCap. In some cases, the downloadable PDF will not provide the most accurate view of the CRF as it would appear for direct entry. The eCRFs in REDCap have skip patterns programmed, as well as visit-based logic, allowing some questions to only appear when entering data for the relevant visits. Paper CRFs that are a better representation of the eCRF view have been created. PDFs of these CRFs are available to download on the matrix4prevention.org website and can be used in place of eCRFs when direct entry into the REDCap is not feasible. Staff should consult the CRF Completion Guidelines (CCG) for direct data entry as well as when utilizing a paper CRF.

Data collection is planned to occur in three countries: Kenya, Lesotho, and Zimbabwe, with each country further divided into sites. Each site will have a partition (Data Access Group or DAG) of the REDCap database to store records of enrolled participants. Only study staff that are assigned to the site can access the records for their designated DAG on REDCap. Some study staff might be granted access to other in-country DAGs on a case-by-case basis. Country Study Coordinators will have access to all DAGs for their respective country.

The Data Manager and Country Data Manager will initiate and oversee routine quality control and assurance processes including report generation, audits, and spot checks to ensure data quality.

### 7.5.1. CRF completion

To the extent possible, site staff should utilize direct data entry into the REDCap study database so that the eCRF serves as the source document. Direct data entry is especially encouraged for data collection based on participant self-report. Paper CRF completion requires an intermediate step of data management and quality assurance (QA/QC) review into site data management workflows. In addition, paper CRF completion introduces the risk of data transcription errors and could contribute

to database data entry errors. When paper CRFs are utilized, they are to be retained in the participant's paper chart/file. When direct data entry is not possible due to the nature of the source data - for example, when the source data is a local lab results report - site staff should enter study data into the REDCap study database based on the site-specific source document.

When paper forms are used as source, data should be entered into the appropriate REDCap eCRF as soon as possible. Ideally, completion of all required eCRFs for a given visit will occur within 1–2 business days of the visit, though up to 7 days is acceptable, especially if data abstraction from medical records is needed. The REDCap dashboard shows a list of forms that are available at each study visit.

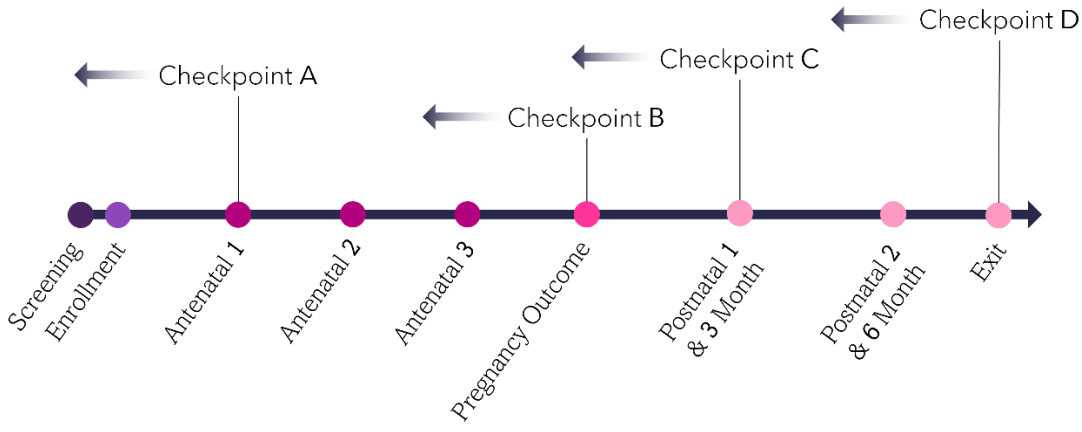
## 7.6. Data Quality Control and Quality Assurance

### 7.6.1. Data Collection Time and Checkpoints

As a participant progresses through the study, it is good practice to ensure that CRFs are completed and checked for errors. Figure 7-1 proposes a timeline by visit and highlights five checkpoints indicating that any CRFs that were started leading up to that checkpoint should be complete and free from errors. For example, when a participant has her first antenatal visit, the CRA should ideally go through all CRFs that were started before the first antenatal visit and ensure that the CRFs are completed in full and free from error, or if there is CRFs pending completion the CRF should be completed as soon as possible per availability of the needed data. In this example it would be preferred that the CRA perform the quality check before the first antenatal visit. This timeline represents an ideal workflow. Adaptation to the workflow is welcome and encouraged with the end goal being quality data. When drafting the Site-Specific SOP for data management, please take the timeline and checkpoints into consideration to adequately address quality assurance needs.

**Figure 7-1: Timeline of study visits.**

*Checkpoints highlight key points along the study timeline when CRFs up to that checkpoint should be complete and error free.*




### 7.6.2. Data Resolution Workflow

REDCap allows users to audit data quality using the Data Resolution Workflow. Various training videos are available online (see links at end of section) and on REDCap to learn about the Data

Resolution Workflow. The Data Manager will provide detailed training on how to implement the Data Resolution Workflow for the study, which will be recorded and available on the CARE PrEP Data SharePoint

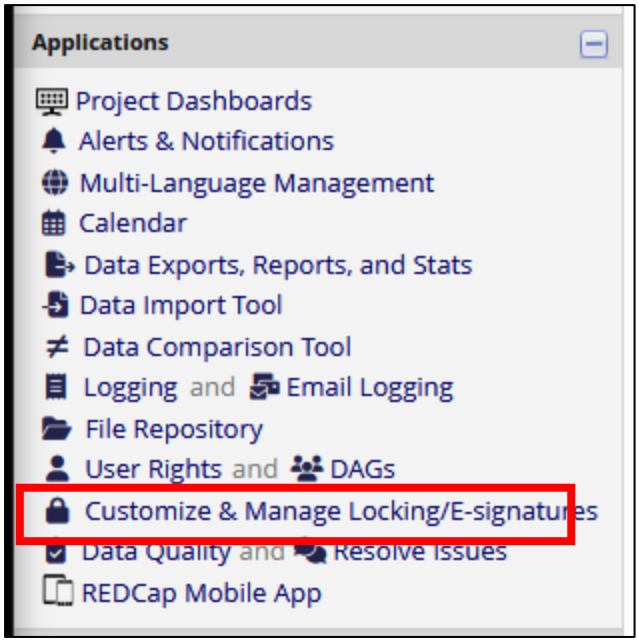
### 7.6.3. Data Quality Rules Module

REDCap also has premade data quality rules that can be executed to look for missing, aberrant, and other discrepant values. Various training videos are available online (see links at end of section) and on REDCap to learn about the Data Quality Rules Module. The Data Manager will provide detailed training on how to implement the Data Quality Rules Module for the study.

The Data Resolution Workflow and Data Quality Rules are advanced REDCap features. The Data Manager will provide further training on the use and implementation of these features for the study. Please reach out with questions as needs arise.

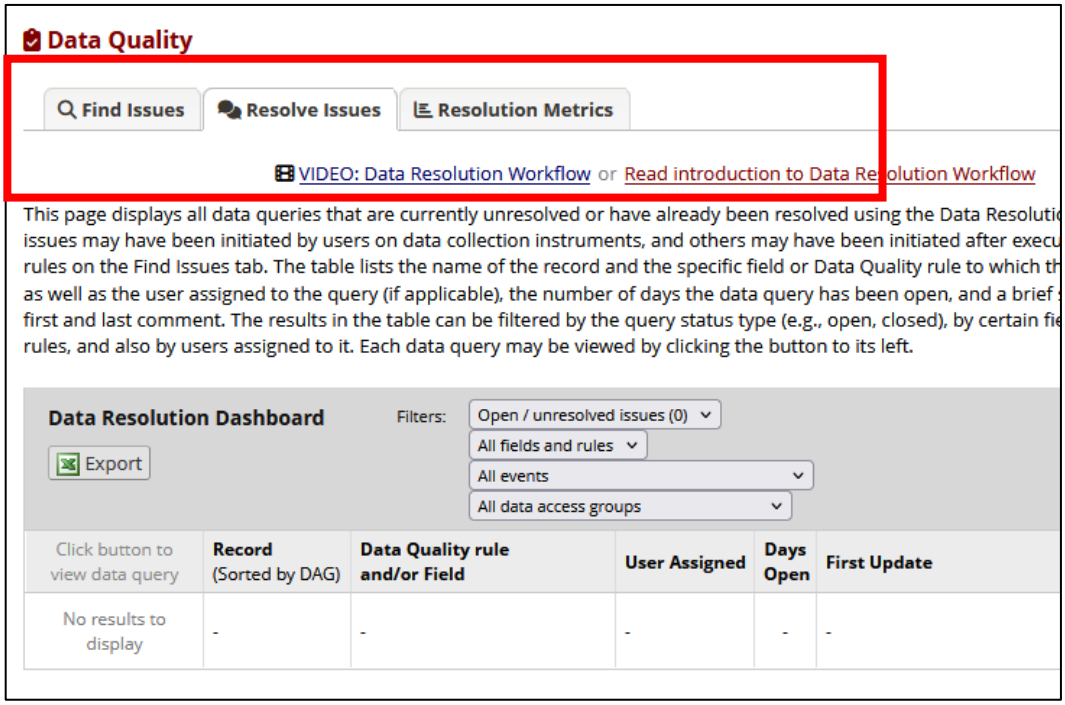
To ensure data cleanliness, every effort should be made by the site to address all system queries at the time of data entry (i.e., as soon as they are generated), and resolve all manual queries no later than 7 days after the query is placed.

### 7.6.4. Training Videos

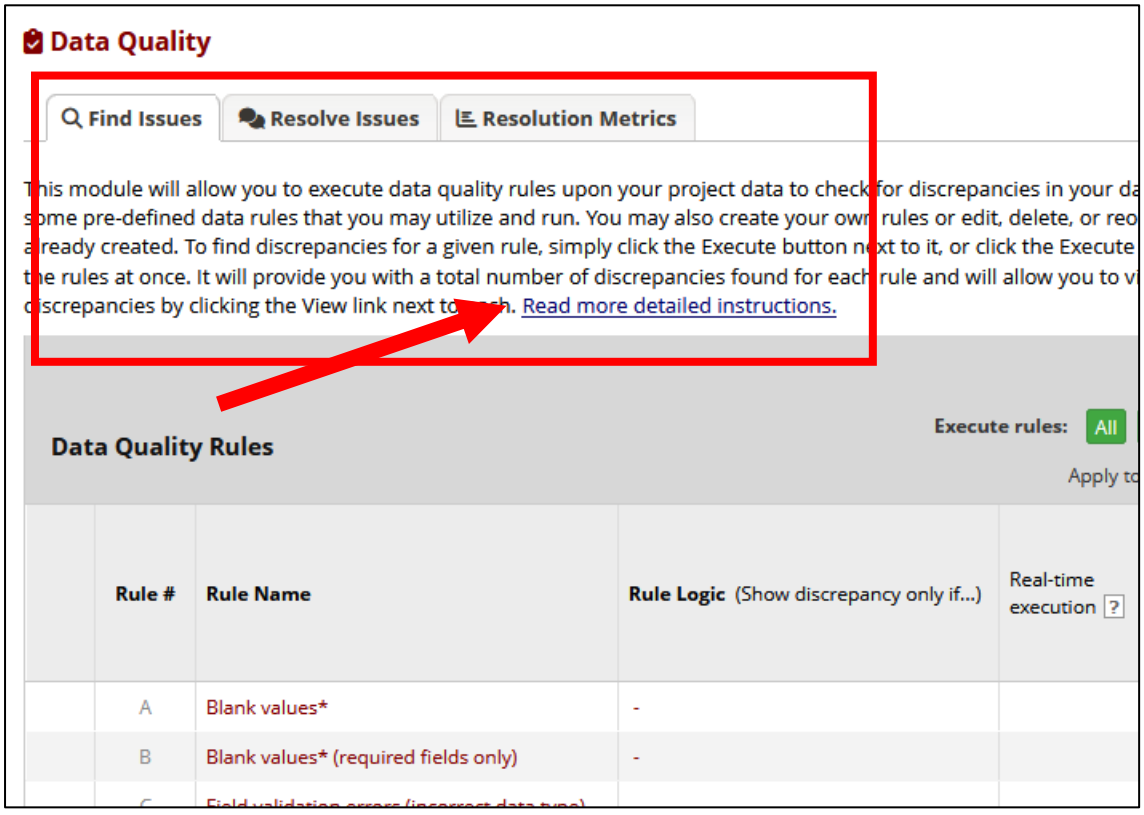



REDCap has links to training videos and instructions for the Data Resolution Workflow and the Data Quality Rules Module. To access them, login to REDCap and using the left-hand toolbar select Resolve Issues for the Data Resolution Workflow; select Data Quality for the Data Quality Rules Module.

For the Data Resolution Workflow, the video and instructions are found under the **Resolve Issues** tab.




For the Data Quality Rules module, more instructions are found under the **Find Issues** tab.




Additionally, YouTube offers tutorials on Data Quality Control and Assurance Processes:

**Data Resolution Workflow**

https://www.youtube.com/watch?v=uoFzol_-ldM

https://www.youtube.com/watch?v=KDDOtDBhmH0

https://www.youtube.com/watch?v=6iLHRaTnx2E

**Data Quality Rules**

https://www.youtube.com/watch?v=LgmFqfBEbHQ&t=5s

https://www.youtube.com/watch?v=1T_2xMTQCv4

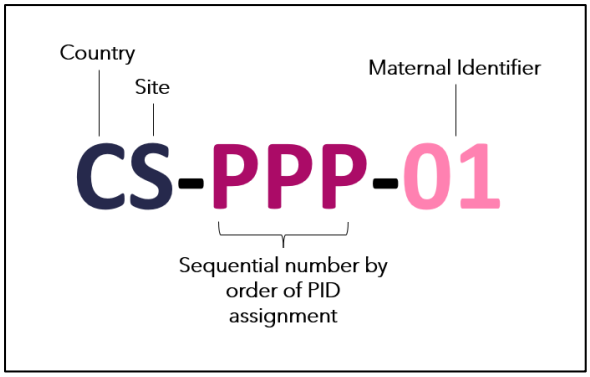
## 7.7. Participant Identifiers (PIDs) and Records in REDCap

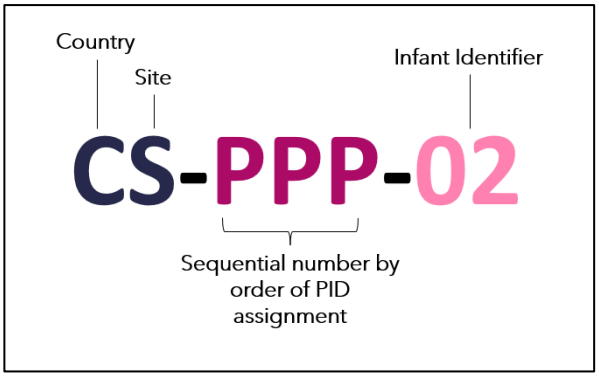
### 7.7.1. PIDS

A list of Participant Identifiers (PIDs) will be generated and uploaded to REDCap before data collection starts. The CRA will track and assign a PID to each maternal participant that is screened for enrollment and each infant participant that enrolls using the **Assign PID CRF** on REDCap. PIDs will be made up of a character followed by 6 digits, for example: Z1-012-01. The leading character identifies to the country: K-Kenya, L-Lesotho, Z-Zimbabwe, this is followed by the first digit that identifies the site. The next set of three digits refer to the participant and will follow a sequential order. The last two digits are used to identify if a participant is a mother or an infant. A mother's PID will end with 01. Infant PIDs will be the same PID as their associated maternal participant but end sequentially with 02, 03, 04, etc, a depending on the number of infants born in a given pregnancy. Maternal participant will be assigned an PID at the enrollment visit immediate after providing consent for the study. *Note: some participants who consent for the study may not enroll (screen fail), but their PIDs are not reused.* Only infants born alive with properly completed consent on file will be assigned a PID; this will occur at the next study visit following the outcome of the birth.

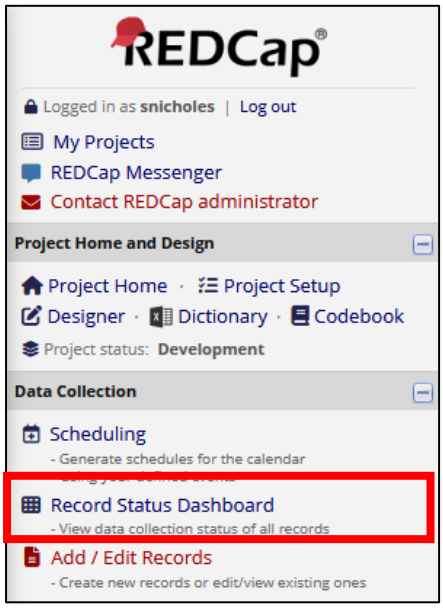
Examples:

- A mother is the first participant to be assigned a PID at site 2 in Kenya. Her PID will be: K2-001-01. She then gives birth to twins; their PIDs will be K2-001-02 and K2-001-03.
- A mother is the 132$^{nd}$ participant to be assigned a PID at Site 1 in Lesotho. Her PID will be L1-132-01. She gives birth to one infant; its PID will be: L1-132-02.
- A mother is the 216$^{th}$ participant to be assigned a PID at site 2 in Zimbabwe. Her PID will be Z2-216-01. She then gives birth to one infant; its PID will be: Z1-216-02. She becomes pregnant a second time within the study period and decides to enroll again at Zimbabwe site 2 as the 374$^{th}$ participant. She will be assigned a new PID: Z1-374-01. She again gives birth to one infant whose PID will be: Z1-374-02. When she enrolls for the second time, her first PID, 32-216-01 is recorded on the Eligibility CRF so that her two instances of enrollment can be accounted for in analysis.



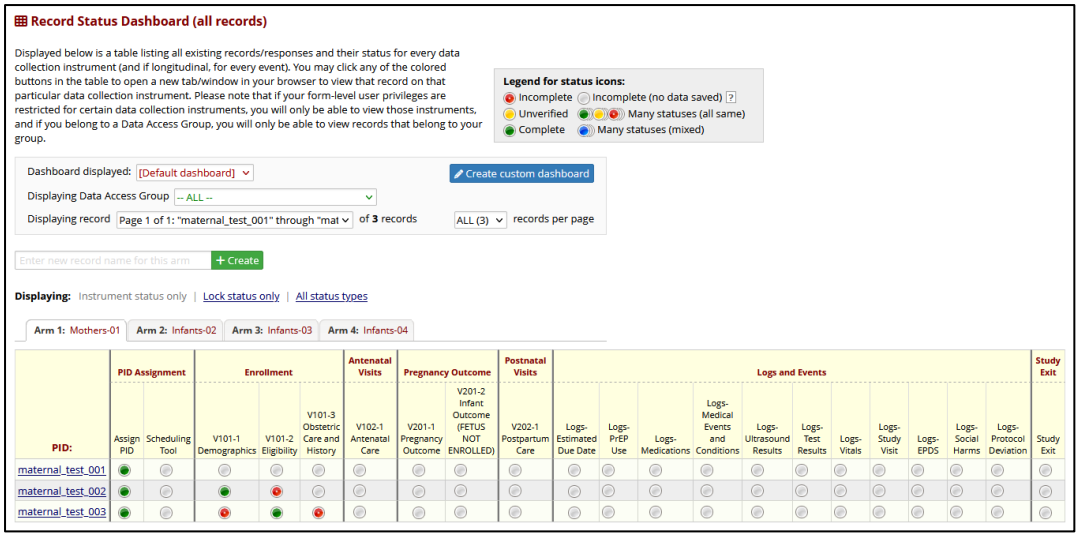

### 7.7.2. Maternal and Infant Record Arms

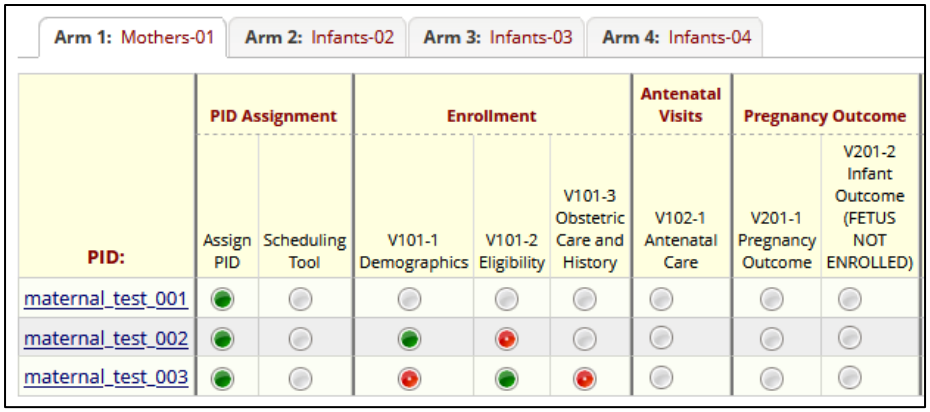


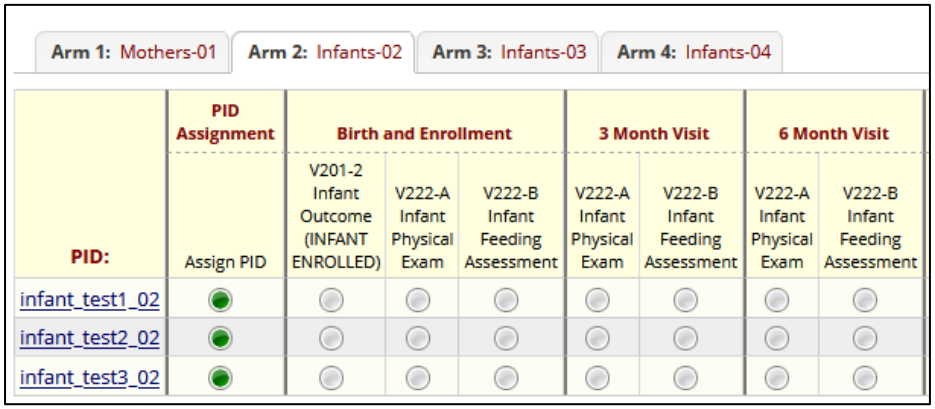

When documenting and looking up participant records on REDCap, the main location to look is the **Record Status Dashboard**, which can be accessed using the left-hand toolbar under the section **Data Collection**.

The Record Status Dashboard provides a listing of all records that you have access to. The Data Manager has access to all records across the study. The CDM has access to all records within their country. The CRA have access to the records for the site to which they are assigned and operate.

The records are further divided up by maternal records, and infant records. At the top of the records listing, you will see tabs that say: **Arm 1: Mothers-01; Arm 2: Infants-02**; **Arm 3: Infants-03**; and **Arm 4: Infants-04**. Maternal records are found under the Arm 1 tab. Infant records are found under the tabs for Arms 2 – Arms 4. If the infant PID ends in 02, the record is found under Arm 2. If the infant PID ends in 03, the record is found under Arm 3, and similarly for records ending in 04 which are found under Arm 4.

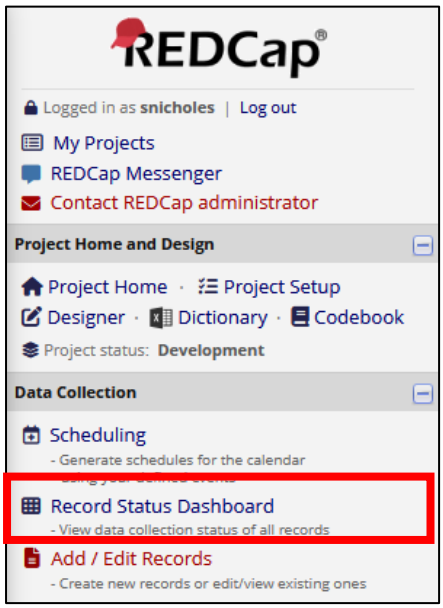


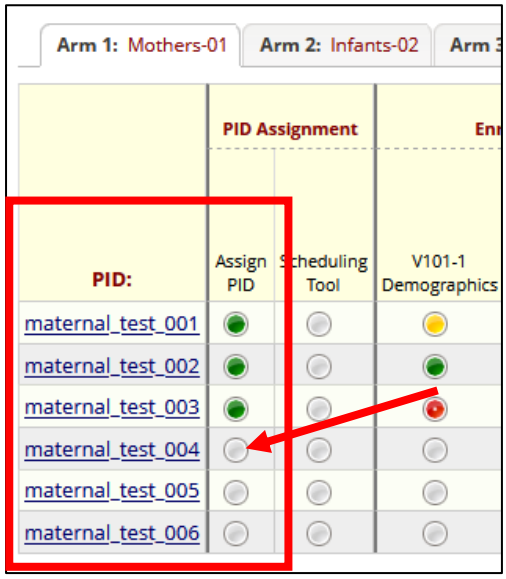




### 7.7.3. Assigning Maternal PIDs

Mothers are assigned a PID at the beginning of the Enrollment Visit once study consent is provided (see SSP section 5 for order of Enrollment Visit procedures). Once a PID is assigned, it is not re-assigned even if they screen fail. The PID stays with them through study participation until they exit.




When a mother is ready to be screened assign her a PID, login to REDCap, and using the left-hand toolbar navigate to the **Record Status Dashboard**.




Under **Arm 1: Mothers-01**, there is a listing of PIDs in table format. PIDS that have been assigned will have the Assign PID CRF completed as indicated by the green bubble. Look for the next available PID as indicated by a grey bubble where the Assign PID has not been completed. In the example shown to the left, maternal_test_004 is the next available PID to assign the participant.

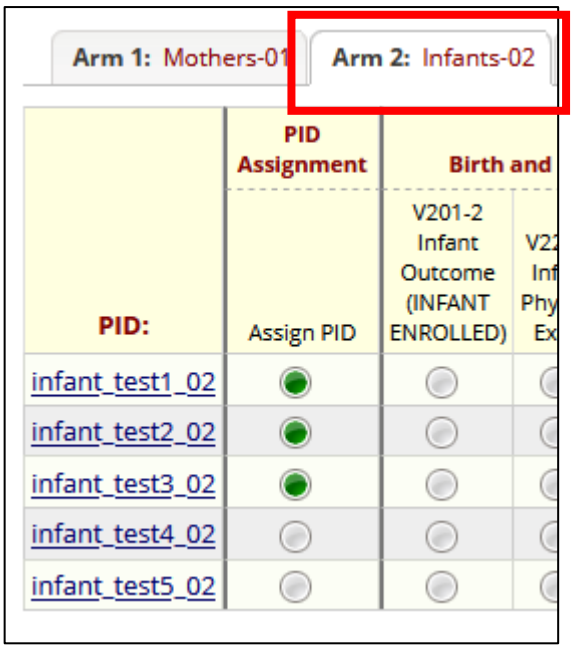
Refer to the MATRIX-007 CCGs for the Assign PID CRF for more details on the CRF.

### 7.7.4. Assigning Infant PIDs

Infants are assigned PIDs only if they are born alive and enrolled in the study. Infants that experience neonatal death are also assigned a PID. Fetuses/infants that are not born alive or consent has been withdrawn before birth will are not assigned a PID. See SSP section 5 for details on infant enrollment procedures.

Similar to how maternal PIDs are accessed, you log into REDCap, and access the **Records Status Dashboard**. Infant PIDs also use the **Assign PID CRF,** found under the infant record arms on the Record Status Dashboard.

Infant Record Arms: Arm 2: Infants-02; Arm 3: Infants-03; and Arm 4: Infants-04.




Find the PID that corresponds with the mother's PID and assign it to the infant.

Example: The mother's PID is **L1-032-01**. You would find the infant PID **L1-032-02** under **Arm 2: Infants-02** and assign the PID to the infant. If the mother gave birth to twins, you would subsequently navigate to **Arm 3: Infants-03**, find the PID **L1-032-03** and assign it to the twin.

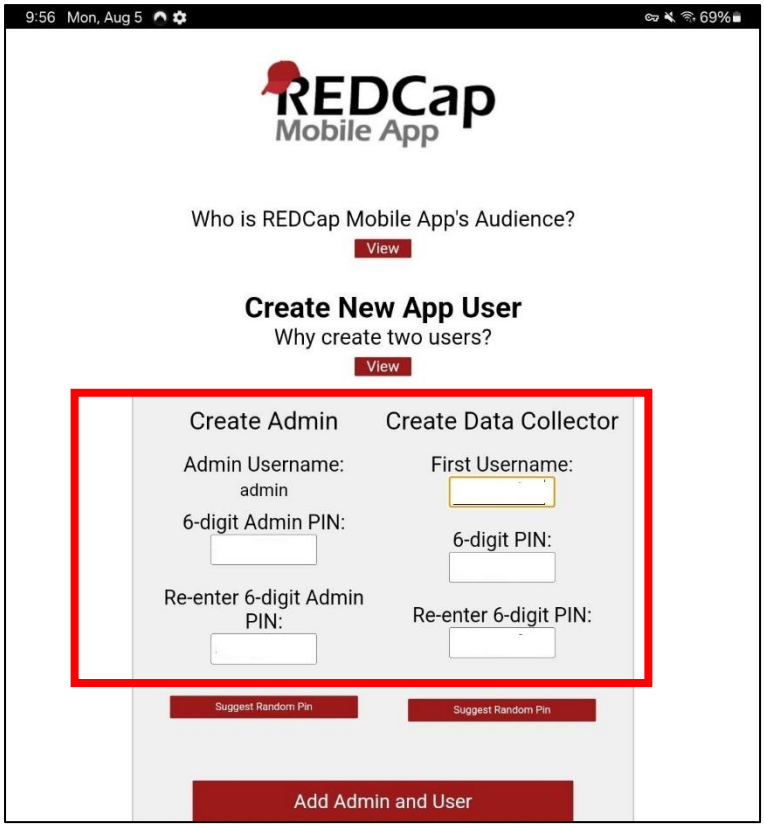
## 7.8. Use of Tablets or Mobile Devices for Data Collection

### 7.8.1. REDCap Mobile App – Internet Connection Required

REDCap allows data collection to happen offline, away from the study site, for example for a home or off-site visit. This is done with a tablet or other mobile device that is 889 compliant with the REDCap Mobile App. Initial installation and app setup needs to be done while connected to the internet. Data collection can happen without internet connection.

The REDCap Mobile App can be downloaded from the device's app store.

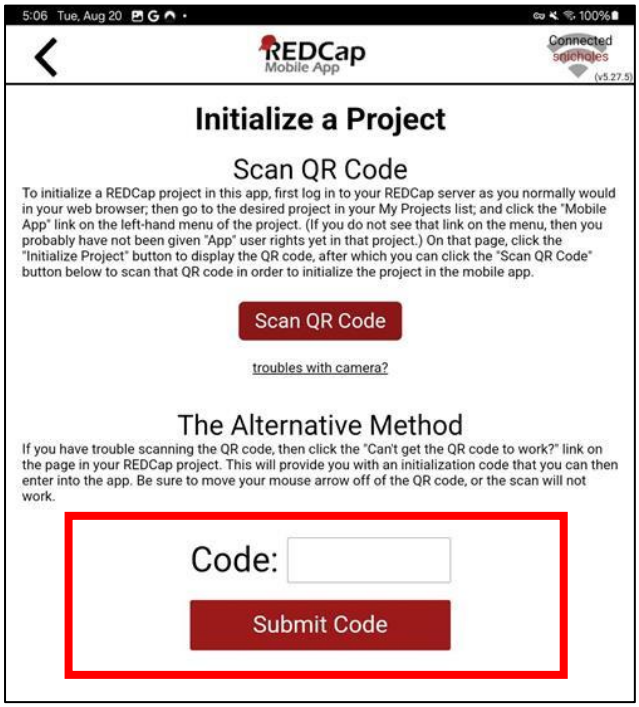
Once downloaded, the app will prompt the creation of credentials for an **Admin** and **Data Collector**. The **Admin** credentials require a PIN only. The **Data Collector** credentials require a username and PIN.





Once the credentials are created you will be presented with the home screen.

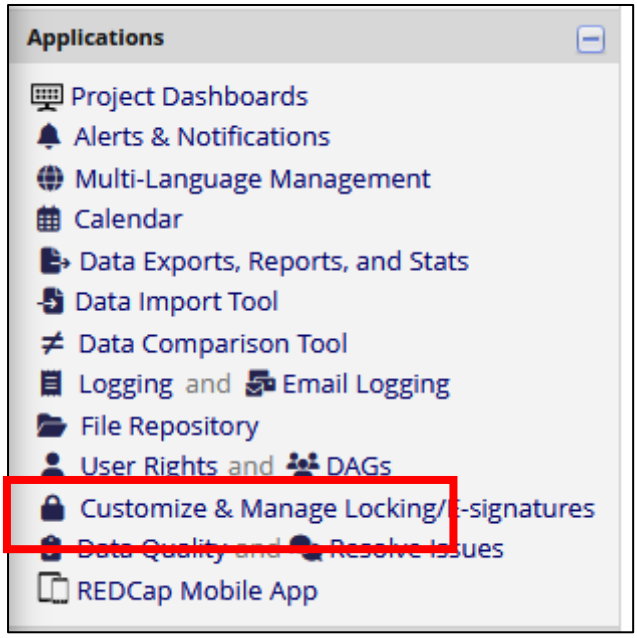
To finish setting up the app, you must click "Set Up Mobile Project.




The app will prompt you to enter a code.

The code needs to be generated using the online version of REDCap. See the next section to generate the code.

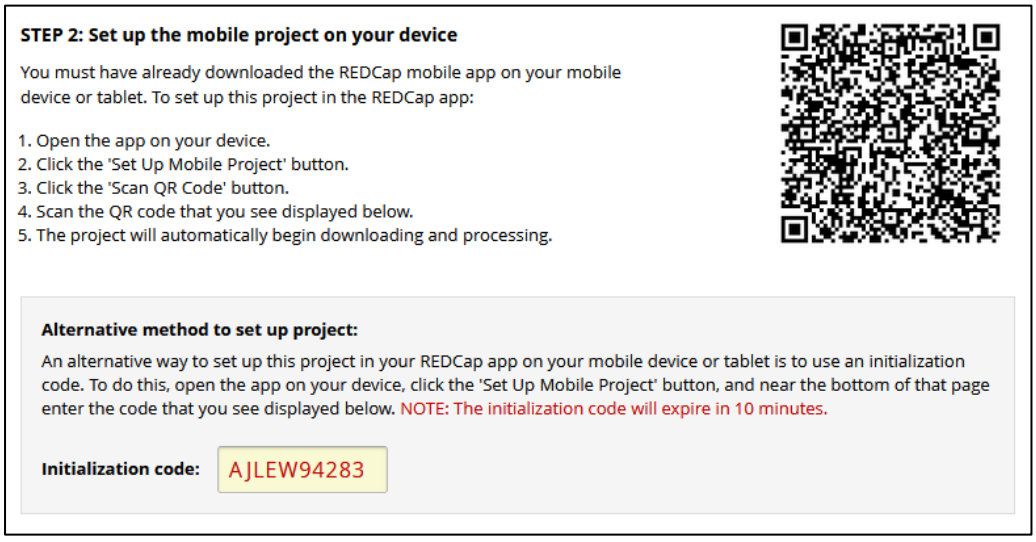
### 7.8.2. Generate Code for REDCap Mobile App

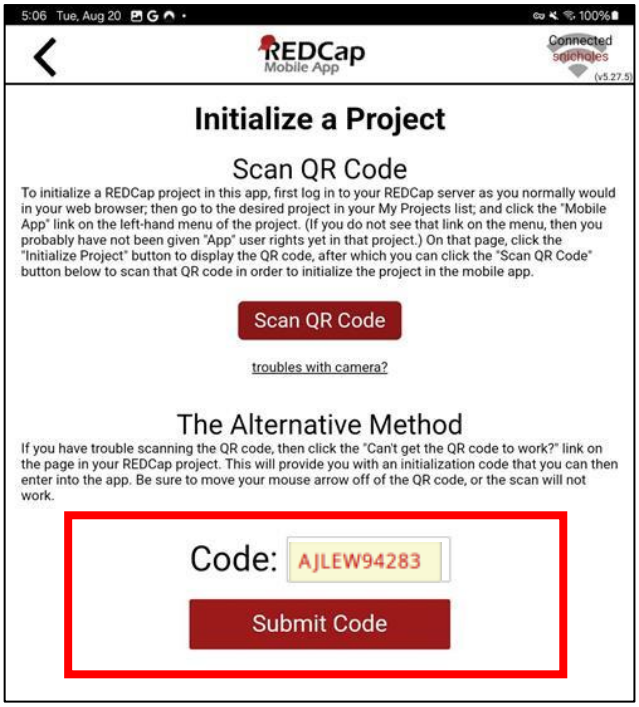



To generate the code, login into REDCap online and using the left-hand toolbar, under the applications section, click **REDCap Mobile App**.

You will find the steps describing the mobile app setup. Below the QR code is a link, **Can't get the QR code to work?**


An **initialization code** will be generated below the QR code. This is the code that needs to be entered into the tablet.





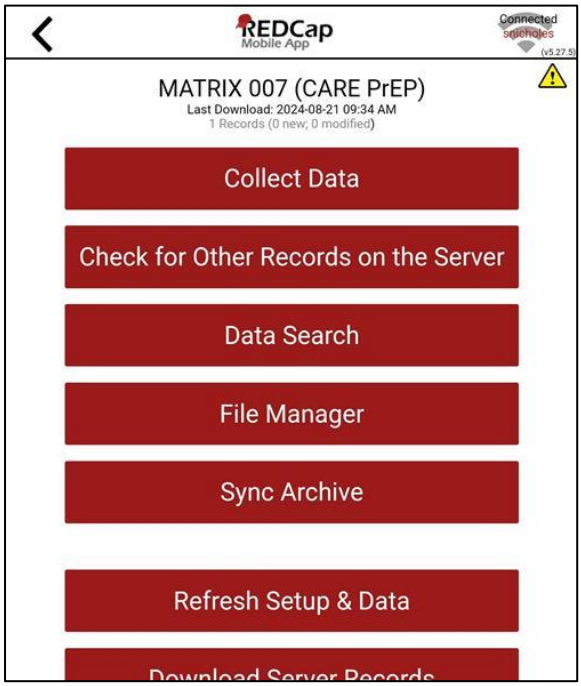

On the tablet: Click **Submit Code** once the code is entered.

The tablet will show the syncing status with the database and confirm once syncing is complete.

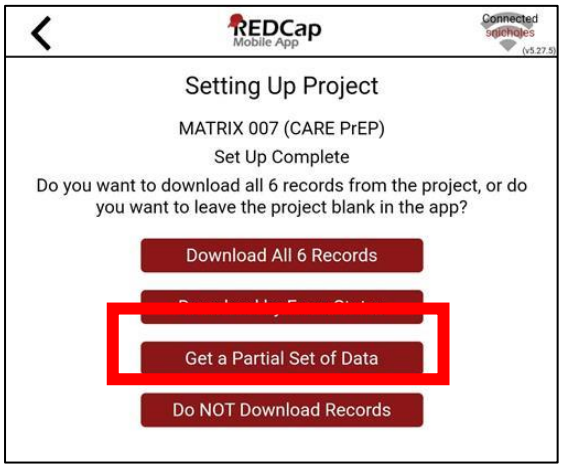
### 7.8.3. Download records to the REDCap Mobile App for offline data collection – Internet Connection Required

This step is done when the project is initially set up on the tablet and is the first thing you will do before leaving to do off-site visits.

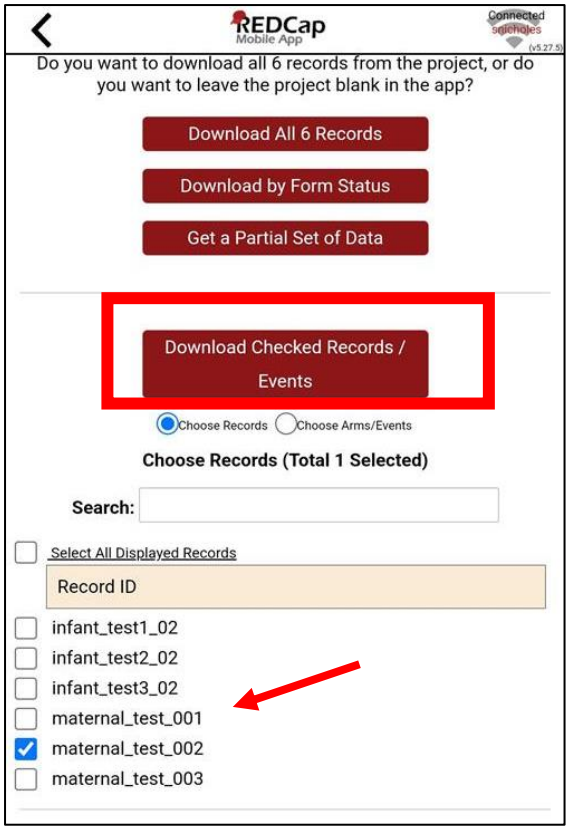

On the home screen, select **My Projects**




On the project menu, select **Refresh Setup & Data**. The mobile app will then sync with the REDCap server.




Next, select **Get a Partial Set of Data**.



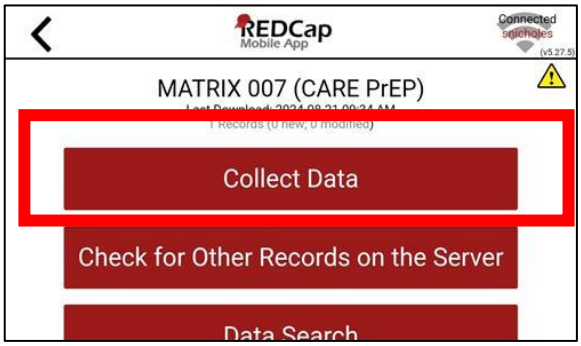

A list of all available records will be listed on the screen:

1. Select only those records for the participant(s) that you will be seeing that day by checking the box next to the PID.

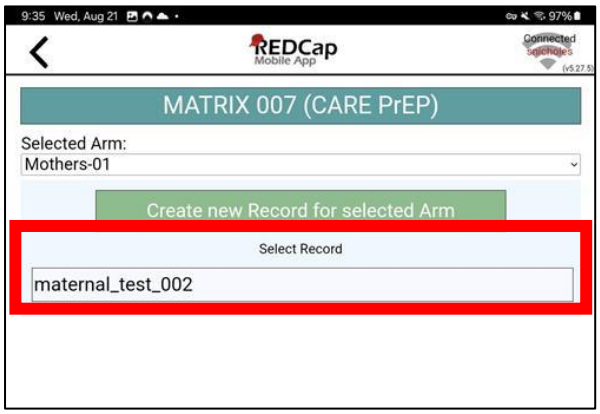
2. Select **Download Checked Records/Events.**

After selecting the records and downloading them, you will return to the project menu and are ready for data collection.

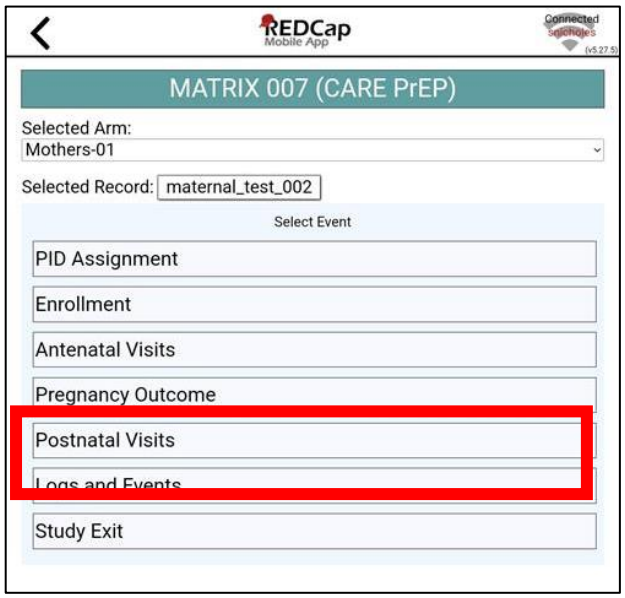
### 7.8.4. Offline Data Collection – Internet Connection NOT required




When you are ready to collect data during the study visit, navigate to the project menu and select **Collect Data**.

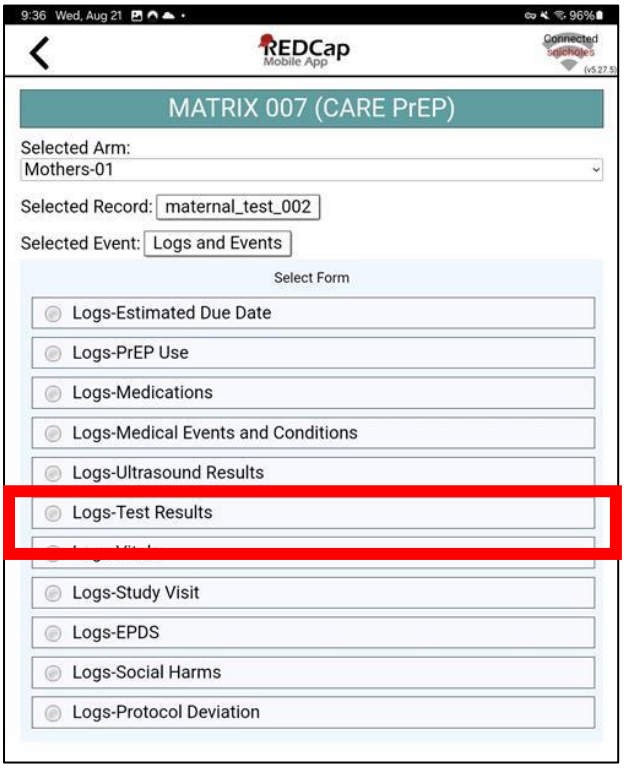



Select the PID for the participant that you are visiting.



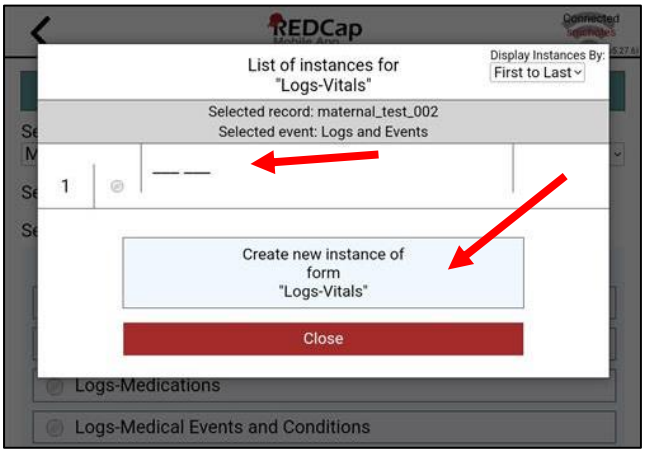

Similar to REDCap online, the CRFs are organized by event. Select the event where the desired CRF is organized.

For this example, we want to fill out the **Vitals CRF** which is found under **Logs and Events**.




Next, select the CRF.

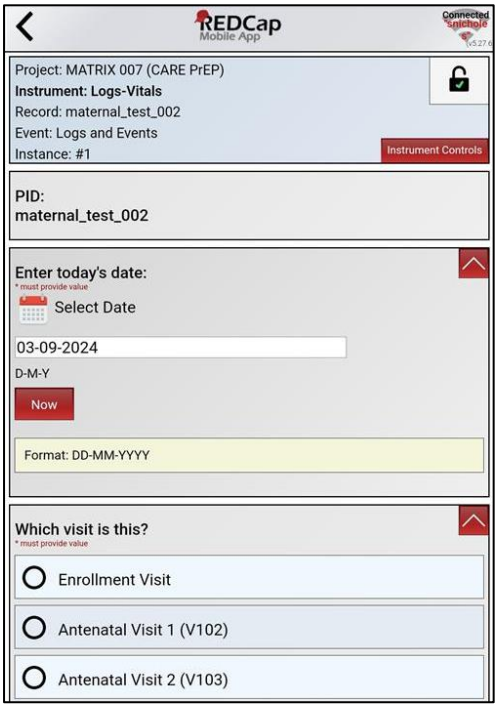
(**Logs-Vitals**, for this example).




Because this CRF is repeatable instrument, or a log, you can select to fill out an existing instance, or create a new instance.

To add data to an existing instance, select the numbered instance from the listing.

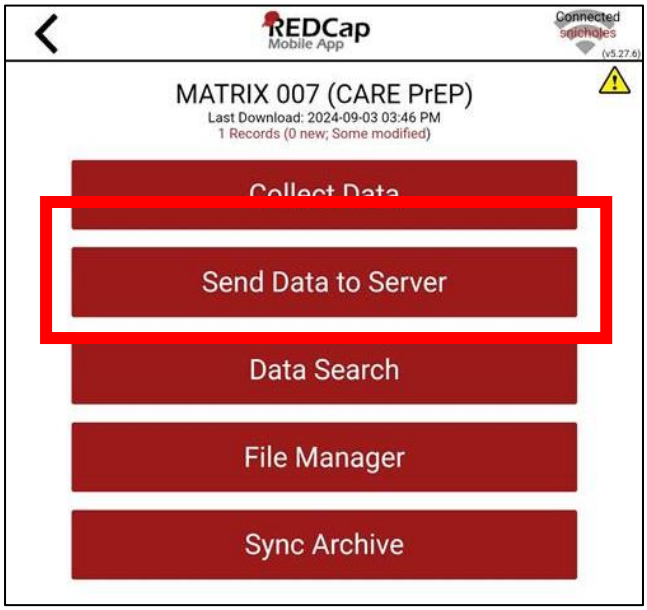
To create a new instance, select Create ne instance of form _____




Once the CRF opens, you may proceed with data collection similar to how it is done online.

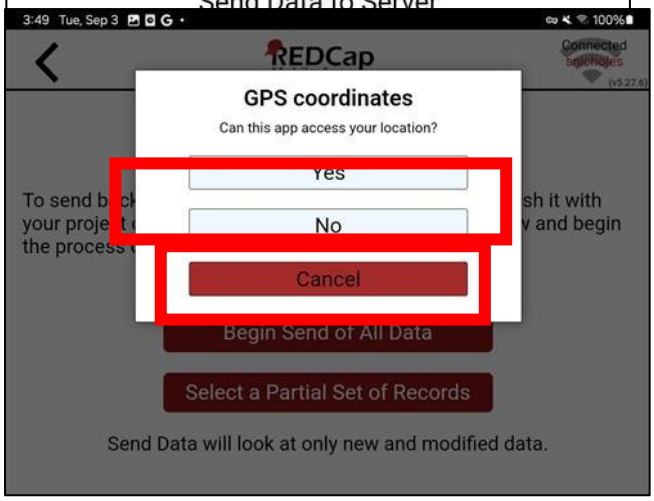
### 7.8.5. Sync tablet and upload records after offline data collection – Internet Connection Required

Once you are done doing data collection offline and again have internet connection, you need to sync the table with the database and upload the records. **This step must be done when you are finished with data collection on the tablet for the day, otherwise there is a risk of data loss.**




Navigate to the project menu and select **Send Data to Server**.


Next, click Begin Send of All Data.

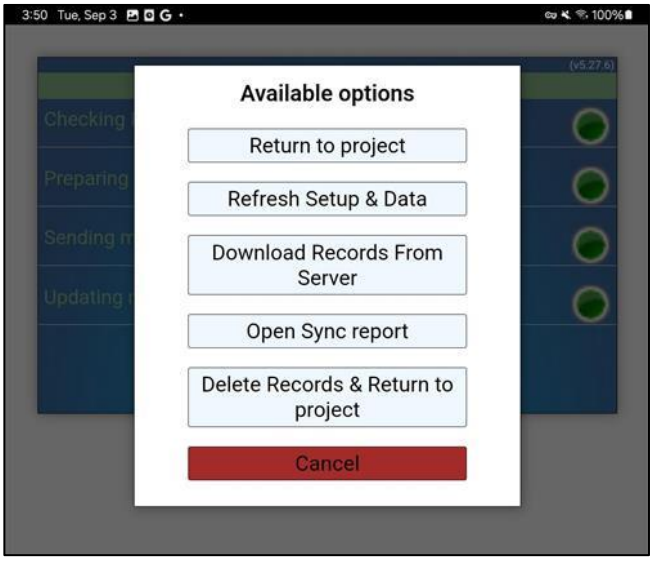



REDCap will ask you if you want to upload GPS coordinates of the tablet with the data. Select **No**.


REDCap may ask you to confirm the proceed with syncing the records. Select **Proceed with sync**.

REDCap will confirm that the sync is complete. If any errors occur, check internet connection. If internet connection is not the problem then contact the Study Data Manager (Sam Nicholes). Do not use the tablet until the error is resolved.


Otherwise, select **Options**.



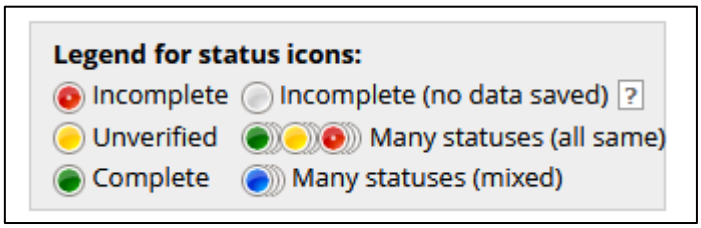

On the pop-up menu, select **Return to project**. Data syncing is complete.

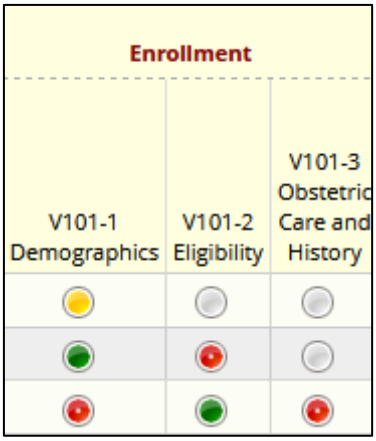
## 7.1. CRF Completion Status

At the end up each CRF is a drop-down menu to indicate the completion status of the CRF. There are 3 options, Incomplete, Unverified, and Complete. If the CRF is marked as Complete, this indicates that the CRF has been filled out to completion and should be free from missing values and other errors. The current process for marking the completion status is as follows: When a CRF is filled out in entirety, or as much as possible, the CRA should mark is as Unverified. A secondary CRA should then follow-up and perform a quality assurance check on the CRF to check for errors. Once the quality assurance check is complete and any errors are corrected, the CRF should be marked as Complete.

If a CRF is not done, or needs to be revisited or updated at a later date or study visit, the CRA should mark it as Incomplete until it is filled out to completion.


When viewing the records on the dashboard, a bubble will reflect the completion status of the CRF. A red bubble with a white dot in the center shows an Incomplete status. A yellow bubble indicates the CRF is unverified. A green bubble indicates a complete CRF. For CRFs with multiple instances (a repeatable form), a blue bubble indicates mixed statuses among the instances of that CRF.




## 7.2. Site-Specific Standard Operating Procedures for Data Management

As a condition for study activation each country data management team will need to have a Country/Site-Specific Standard Operating procedures (SOPs) for Data Management. This SOP needs to be approved by the Data Managers at FHI 360: Sam Nicholes and Doris Marwanga.

It is assumed that updates and modifications will need to be made to the SOP over the course of the study. When changes are made the expectation is that the Data Manager will have been notified and have approved the changes, and that all country level team members have been notified and trained on the changes.

Elements that need to be included and addressed in the SOP:

- Timing of data entry into the study database, including abstraction of medical records, and data collection done at home visits.
- Quality Assurance processes and schedule.

- Data Reports processes and schedule.
- Data requests from CATALYST
- Participant study file organization, storage, and security
- Participant confidentiality.
- Overall data storage, security, and management in relation to laptops, tablets, mobile devices, and REDCap login credentials.
- Contingency plans in case of interrupted access to REDCap, SharePoint, and other electronic based systems.

## 7.3. Data Management Plan

### 7.3.1. Metadata and Documentation

The goal of MATRIX-007 is to provide robust, quality data on the clinical observations and outcomes of pregnant women exposed to PrEP and their infants. Proper metadata and documentation will be among the deliverables to meet this goal. This will consist of the generation of codebooks, and README files to help orient secondary users to the data set; how variables were defined and operationalized; important dates of creation, cleaning, and revision; data set authors; and other necessary information to provide context to the data collected over the study period.

One of REDCap's features allows users to generate a codebook of variables, questions, response options, and other points of data control and input. This will be used in conjunction with packages common to statistical software to generate codebooks, dictionaries, and other documentation to help provide context and definitions for the dataset.

### 7.3.2. Ethics and Legal Compliance

As FHI 360 is the sub-awarded organization by MATRIX to manage study operations, including data collection FHI 360's policy, which is based on the European Union's General Data Protection Regulation, will serve as the basis of the legal requirements for MATRIX-007 data management.

## 7.4. Personal Identifiable Information

Personal Identifiable Information (PII) will be collected during the study period but will not be stored on REDCap. Paper-based forms will be used to record PII and will be stored in a secure location at each study site, such as a locked filing cabinet; only authorized study staff will have access. The site-based MATRIX-007 Name Linkage Log is the sole point where PII can be connected to the participant identifiers (PIDs) that are used on REDCap for record identification. A paper-based country-specific Participant Locator Information Form will be used to collect and maintain PII necessary for contacting and tracking participants. See SSP Section 1 for further information on secure storage or participant information and data.

Sensitive data such as pregnancy information, HIV infection, PrEP use, medical history, etc. will be kept on REDCap separate from PII. As mentioned above, REDCap will be stored on a secure, US based server under the management of FHI 360's ISS department.

## 7.5. Data Ownership and Sharing

MATRIX Prime is the owner of all primary data collected for MATRIX-007, with FHI 360 as the sub-awarded organization tasked with study and data management. MATRIX-007 will request a

specified set of data from the MOSAIC CATALYST Study on participants who are enrolled in both CATALYST and MATRIX-007. MATRIX-007 and CATALYST have a data sharing agreement allowing MATRIX-007 to gain access to CATALYST data on PrEP exposure, demographics, and pertinent clinical data for CATALYST participants that are enrolled in MATRIX-007.

MATRIX-007 does not anticipate preparing or entering into other data sharing agreements, beyond the goal of providing PrEP safety data for pregnant women and infants to organizations such as USAID and the WHO.

### 7.5.1. Process for Requesting Data from CATALYST

DEVELOPMENT PENDING