

Section 3: Data Collection

3.	Data Collection.....	3-1
3.1.	Study Population, Sample Size, and Duration of Abstraction Activities	3-1
	Figure 1: Estimated Sample Size	3-2
3.2.	Record Identification and PTID Assignment	3-2
3.3.	Data Collection Procedures.....	3-3
3.3.1	REDCap Database Overview	3-3
3.3.2	REDCap Data Entry	3-3
3.3.3	Data Abstraction Form Completion Instructions	3-5
3.3.4	Data Lock and Signatures by Investigators	3-6
3.4.	Quality Control	3-6
3.5.	Quality Assurance Activities/Management Team Review.....	3-7

3. Data Collection

3.1. Study Population, Sample Size, and Duration of Abstraction Activities

Each site will identify 1-2 healthcare facilities (e.g. a primary care facility and a referral facility, as outlined in site SOPs) to partner with for MTN-042B data collection activities.

All women delivering or receiving immediate postpartum care (within one week of delivery) at the facility(ies) affiliated with each study site for a period of 8 weeks will be included in the data abstraction.

After study activation, each site must define a data collection period in consultation with the MTN-042B management team. This will typically include 8 weeks of data collection as well as a formal pause of approximately 1 week to complete quality assurance (QA) activities. QA activities are described further in SSP Section 3.5 below.

Sites will make every effort to abstract data from all deliveries and women who receive immediate postpartum care during the data collection period. Ideally, chart abstraction should occur soon after delivery (i.e. within one day) and before patient discharge to minimize missing data and avoid sampling bias. Acknowledging this may not always be possible, sites should outline facility-specific approaches and back up plans for abstraction of charts in their MTN-042B SOPs. SOPs should also address plans for deliveries that occur over the weekend. Abstraction activities may extend beyond the defined 8-week data collection period in order to capture data from all records.

Sample sizes are estimated as follows in Figure 1; however, the actual number of records captured may vary:

Figure 1: Estimated Sample Size

Site	Estimated Births in 8-week period (all facilities)
Wits RHI Shandukani	2160
Blantyre	2400
MU-JHU	4000
Zengeza	2400
TOTAL Births	10,960

3.2. Record Identification and PTID Assignment

Sites will outline in SOPs how delivery records and records that include immediate post-partum care within 7 days of delivery from each day at each facility will be identified for inclusion in abstraction activities.

Each site will be provided facility-specific PTID Linkage Logs by the SDMC that have been pre-populated with available study PTIDs. **A new page of the PTID Linkage log should be used for each delivery date.** Staff should enter the delivery date at the top of each page of the linkage log. As needed, a second (or third) page can be used for a given delivery date depending on how many records are identified.

Once a record is identified as eligible for inclusion in abstraction activities, a PTID will be assigned through completion of the next available row on the PTID Linkage Log. Staff should enter the facility or site-assigned record number, and initial and date. Additionally, the PTID Linkage log will pre-select records requiring QC#2, described further in section 3.4 below. Figure 2 is provided as an example. Detailed instructions for completion of the PTID Linkage log are available on the cover page of the log.

Figure 2: Sample PTID Linkage Log

MTN-042B PTID Name Linkage Log Site: Example Site 1 Facility: Example Facility 1		Delivery date: 19-Apr-19			
PTID	Medical Record Number	Clinic Staff Initials	Date (dd-mmm-yy)	QC #2 Clinic Staff Initials	QC #2 date (dd-mmm-yy)
90001	111-11-11111	ABC	19-Apr-19	n/a	n/a
90002	222-22-22222	ABC	19-Apr-19	n/a	n/a
90003	333-33-33333	ABC	19-Apr-19	n/a	n/a
90004	444-44-44444	XYZ	19-Apr-19	n/a	n/a
90005	555-55-55555	XYZ	19-Apr-19	n/a	n/a
90006	666-66-66666	XYZ	19-Apr-19	n/a	n/a
90007	777-77-77777	XYZ	19-Apr-19	n/a	n/a
90008	888-88-88888	XYZ	19-Apr-19	n/a	n/a
90009	999-99-99999	XYZ	19-Apr-19	n/a	n/a
90010	000-00-00000	ABC	19-Apr-19	XYZ	20-Apr-19
90011	111-22-22222	XYZ	20-Apr-19	n/a	n/a

3.3. Data Collection Procedures

3.3.1 REDCap Database Overview

REDCap is the electronic data capture system used by SCHARP to receive and manage study data abstracted from medical records at study sites for MTN-042B. Each site completes study electronic case report forms (eCRFs) by entering data into the REDCap study database. Whenever possible, data will be entered directly into the study database (i.e., eCRF is source) by accessing the MTN-042B REDCap study database over the web in a web browser at:

<https://cdsweb07.fhrc.org/redcap/index.php?>

The web-based REDCap database can be accessed using a computer or a tablet whenever internet is available.

In the event that the web-based REDCap study database cannot be accessed due to internet outages the REDCap Mobile app can be used to collect data offline.

In the rare even that both the web-based REDCap study database and the REDCap Mobile app cannot be accessed due to internet outages and technical difficulties at the study site, data will be entered on a paper CRF first, then entered into the study database based on the paper CRF source document. SCHARP will post a CRF (pdf) file, representing output from the REDCap study database eCRF, on the following MTN-042B Atlas web page:

<https://atlas.scharp.org/cpas/project/MTN/042b/begin.view?>

Each site's DoD log designates the site staff members responsible for entering data into the study database. Staff must have MTN-042B training documented prior to being delegated responsibilities for the study. SCHARP grants designated site staff access with specific user permissions to the study database.

Detailed guidance on data collection, entry, navigation and general use of REDCap Mobile is provided in section 4 of the SSP.

Site staff should contact the study Clinical Data Manager with any questions related to study data collection and management, or any time a site has technical questions or problems related to access or use of the REDCap software.

3.3.2 REDCap Data Entry

To the extent possible, site staff should utilize direct data entry into the REDCap study eCRF using web-based REDCap. Direct entry into web-based REDCap sends data to and receives data from the server in real-time.

Site staff are encouraged to use the REDCap Mobile App only when internet connectivity is not available as the data downloaded and upload procedures introduce the risk of data loss. See SSP section 4 for further information on using the REDCap Mobile App.

Site staff are encouraged to avoid paper CRF completion, as it represents an additional, intermediate step of data management and QA/QC review into site data management workflows. In addition, paper CRF completion introduces the risk of data transcription errors, and could contribute to data entry errors. In the rare event that a site utilizes paper CRFs as source documents, any changes to the paper CRFs must be entered into the REDCap eCRF.

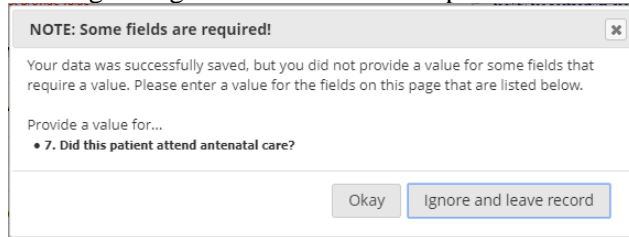
The Form Status field on the eCRF should be used to indicate the completion status of the form. If the data is incomplete and the data abstractor needs to review the chart again, the data abstractor should set the Form Status field to "Incomplete" (indicated by a red icon). If

the data is complete, and the PTID was not designated as requiring double data abstraction, the Form Status field should be set to “Complete” (indicated by a green icon). If the data is complete and the PTID requires QC review step #2, the Form Status field should be set to “Unverified” (indicated by a yellow icon).

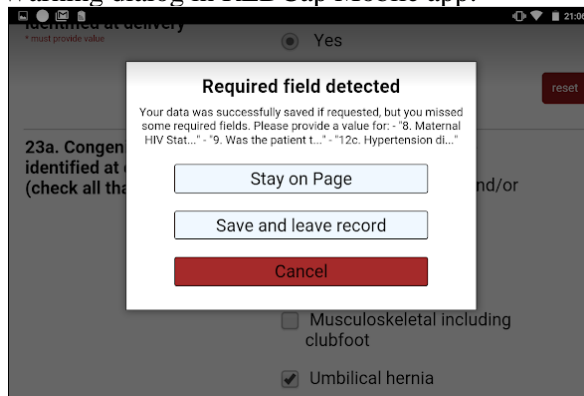
3.3.2.1. Required fields

Many fields on the eCRF are required. The text “* must provide value” will be displayed in red font below a required field. If a required field is left blank, the user will be prompted with a warning dialog when the form is saved:

Warning dialog in web-based REDCap:



Warning dialog in REDCap Mobile app:



The user should review the required fields that are missing data.

When using web-based REDCap:

- If the data is available, the user should select “**okay**” to return to the data entry page and complete the fields.
- If the data is not available, but it is possible that the data will become available, the staff member will select “**okay**” to return to the data entry page and ensure the Form Status is set to “**Incomplete**”.
- If the data is not available, and it will not become available (or it has been confirmed that the Form Status is set to “Incomplete”, the staff member can select “**Ignore and leave record**” to save the form without the required data.

When using REDCap Mobile app:

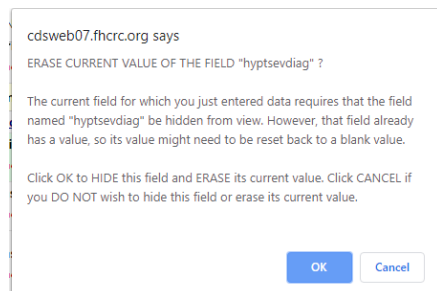
- If the data is available, the user should select “**cancel**” to return to the data entry page and complete the fields.
- If the data is not available, but it is possible that the data will become available, the staff member will select “**Stay on Page**” to return to the data entry page and ensure the Form Status is set to “**Incomplete**”.

- If the data is not available, and it will not become available, the staff member can select **“Save and leave record”** to save the form without the required data.

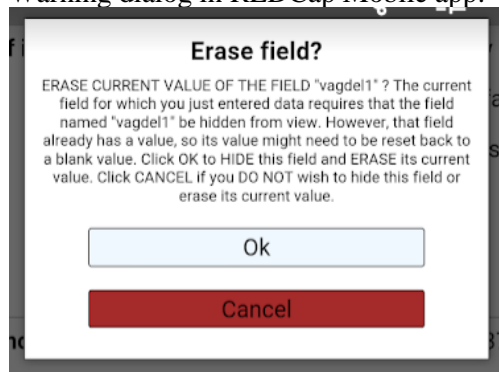
3.3.2.2. Skip patterns

Some fields will be hidden or visible based on the response selected for another field (referred to as a trigger). If the response to a trigger is changed, it is possible that a field with data will become hidden. If this happens, a warning dialog will appear and alert the user that data is present in a field that will be hidden:

Warning dialog in web-based REDCap:



Warning dialog in REDCap Mobile app:



The user should select **“OK”** if the field can be hidden and the data in it can be cleared. If the user selects **“Cancel”**, the field will remain visible and the data will not be cleared. The user will select **“Cancel”** to return to the form and correct the data.

3.3.2.3. Other tips on entering data:

- Click in a textbox to reveal the onscreen keyboard.
- Click in the dot or box of a multiple choice field to select that option.
- Click in a dropdown list, to reveal the choices. Scroll with your finger. Some dropdowns may be slow if they have a large number of choices.

3.3.3 Data Abstraction Form Completion Instructions

Detailed form completion instructions for the data abstraction form is provided in the CRF Completion Guidelines (CCG) document. The instructions document required fields, skip patterns, and include guidance on completion of the study CRF.

The current version of the MTN-042B CCGs is located on the MTN-042B ATLAS webpage:

<https://atlas.scharp.org/cpas/project/MTN/042b/begin.view?>

Efforts should be made to abstract complete and accurate data from the patient chart. In the event staff identify discrepant information in the patient record, clinical discretion should be used to determine the most appropriate CRF entry. If information is not available in the patient record, 'not documented' response option should be used, where appropriate.

3.3.4 Data Lock and Signatures by Investigators

Each record will be locked by staff responsible for data collection/entry throughout the data abstraction period. Records can only be locked using the web-based REDCap site. When entering data using the web-based REDCap site, staff should lock the record at the same time it is marked as "Complete". Records cannot be locked using the REDCap Mobile app; any records entered using the REDCap Mobile app should be locked at a later time by logging into the web-based REDCap site. Sites will regularly review the "Record Status Dashboard" and lock any unlocked records that have been marked as "Complete", as described in the Site-Specific SOPs.

Once data abstraction activities are complete, each site investigator or designee is expected to complete a single paper sign-off form to attest that all study data at his/her site has been collected in a manner consistent with the study guidance documents (e.g. protocol, SSP and CCGs) and site specific SOPs. SCHARP will inform site investigators when investigator sign off of study data should be completed.

3.4. Quality Control

As described in the site's SOP, each site must perform Quality Control (QC) review steps. Each record abstracted will require QC Review Step #1. A subset of records will require QC Review Step #2. Records requiring QC Review Step #2 will be selected to represent approximately 10% of all charts reviewed, with no more than 5 charts requiring QC#2 each day. Charts requiring QC#2 review will be pre-selected on the PTID Linkage Log as every 10th PTID. If/when 5 PTIDs have been flagged for QC#2 for the day, sites may indicate "N/A" on the PTID Linkage Log to indicate the daily required QC targets have been met. Sites are permitted to conduct QC#2 on more than 10% of charts at their discretion. Should sites choose to do this, procedures for selection and documenting additional QC#2 should be outlined in site SOPs.

Below are specific review guidelines that should be followed for these QC review steps (details should be outlined in site specific SOPs).

MTN-042b QC Review Step #1

- Review the eCRF for completeness and accuracy based on information available in the patient chart.
- Completes the "Form Status" item

QC #1 if paper CRFs used

- Checks the paper CRF s/he completes for completeness and accuracy based on patient chart
- Using the paper CRF, s/he transcribes data from the paper CRF to the eCRF
- Checks the eCRF for completeness and accuracy based on the paper CRF
- Completes the "Form Status" item

MTN-042b QC Review Step #2

- Compares data entered on eCRF with patient chart
- Follows-up on any QC trends/issues, as described in Site Specific SOPs
- After consultation and agreement with original abstractor, makes updates, as needed, to eCRF
- Updates the “Form Status” item
- Indicates that QC#2 was completed by initialing and dating the QC2 columns on the PTID Linkage Log

3.5. Quality Assurance Activities/Management Team Review

After study activation, each site will typically undergo 8 weeks of data collection as well as a formal pause of approximately 1 week to complete quality assurance (QA) activities. Formally pausing to complete QA activities is recommended by the management team, however, per site discretion, the 1-week QA pause may be omitted (e.g. if pausing has the potential to cause issues with IRBs and/or relationships with identified healthcare facilities). In this case, data collection will continue during QA activities.

QA activities with pausing will proceed as follows:

- Sites to abstract data for a period of 1 week (7 days, e.g. Monday-Monday) as soon as possible after site activation, starting on a date agreed upon by the management team
- Formally pause data abstraction activities for approximately 1 week to allow time for reports to be created, reviewed, and discussed on a call between MTN-042B management team members and the site.
 - *NOTE: during the pause, no charts will be abstracted, and sites would not be expected to ‘go back’ and review these charts retrospectively, i.e. any deliveries during these dates would be not included in the 042B dataset*
- Data will be reviewed to check for patterns of missing or unexpected data. Any problematic patterns will be discussed with the site and additional training will be provided, if necessary. The study team will make any needed adjustments to the data collection procedures to address any data quality findings or logistical challenges.
- On a date agreed upon with the management team, data abstraction activities will be resumed for a period of 7 weeks

For sites who do not pause, the procedures above will be followed with the exception that data collection will continue while reports are created/reviewed/discussed with the management team.

Ongoing reviews of data tables will be performed throughout the data abstraction period, as needed.

Due to the nature of chart review studies, it is possible that there will be occasional missing or unexpected data. SCHARP will not query the site about individual records with missing or unexpected data.