Safer Births Bundle of Care 2 (SBBC II) Data Management Plan

1. Introduction

a. Data collection sites: This study includes 142 health facilities from 5 regions in Tanzania (Manyara, Geita, Tabora, Shinyanga, and Mwanza). Different data as identified in the SBBC study protocol are collected.

b. Data collection methods:
Data is collected in several ways, both electronically (using tablets/mobile phones), and semi-automatic (training equipment).

c. Databases and Servers; We utilize different databases and servers to store data depending on the data type and tools used for data collection.

- The Open Data Kit (ODK) central database is stored on the hospital server, Debian 10 server, which is owned and maintained by Haydom Lutheran Hospital with a high level of security. This database consists of
    i. Retrospective clinical baseline data (i.e., events/actions during intrapartum care and patient outcome data), collected electronically (using tablets) from the delivery registries, and/ or inpatients registries, death registries, or referral books in all the study health facilities.
    ii. Prospective clinical implementation data (i.e., events/actions during intrapartum care and patient outcome data) that are collected electronically (using tablets) from the delivery registries, and/ or inpatients registries, death registries, or referral books in all study health facilities, using the "SBBC II Mtuha data form".
    iii. The use of clinical SBBC tools (Moyo, NeoBeat, and Upright bag mask) and extra clinical information are collected from patient's file (i.e., partograph/Labor care guide) in all study health facilities using the "SBBC II Extra form".
    iv. Health facility readiness and service availability assessment collected to assess the service readiness for intrapartum care at the start of the project, using the "SBBC II Facility readiness assessment tool".
    v. Mentorship assessments checklist is completed to describe the facility on availability/accessibility of the services of intrapartum care and identify any area that needs improvement/emphasize, using the "SBBC II Mentorship checklist tool".
- The NeoNatalie Live database is stored on a Microsoft Azure server, Haydom domain owned, controlled and maintained by Laerdal Global Health in which there is limited access for only identified people. Haydom Lutheran Hospital has a user agreement with Laerdal. This database consists of simulation training data from health care workers (HCW) that are collected semi-automatically by the NeoNatalie Live simulator and electronically uploaded by the operator.
- The Learning Improvement and Facilitation Tool (LIFT) system database is stored on the Microsoft Azure server, Haydom domain owned, controlled and maintained by Laerdal Global Health in which there is

limited access for only identified people. It consists of streamlined training program assessments through its innovative digital solution, conducting facilitator-led assessments and collecting participant assessments digitally.

d. **Data accessibility and use:** Data can only be accessed by those identified in the Data Transfer Agreement (DTA) and delegated by the principal investigator as per Good Clinical Practice. Since data are owned by Haydom Lutheran Hospital, any investigators/teams (external and internal) before involvement and use of data are required to request permission and approval from the Haydom Lutheran Hospital for his/her involvement and use of data. External investigators/teams must have the permission from Tanzania Commission for Science and Technology (COSTECH) and National Institute for Medical Research (NIMR) to access and used the data.

2. Data management

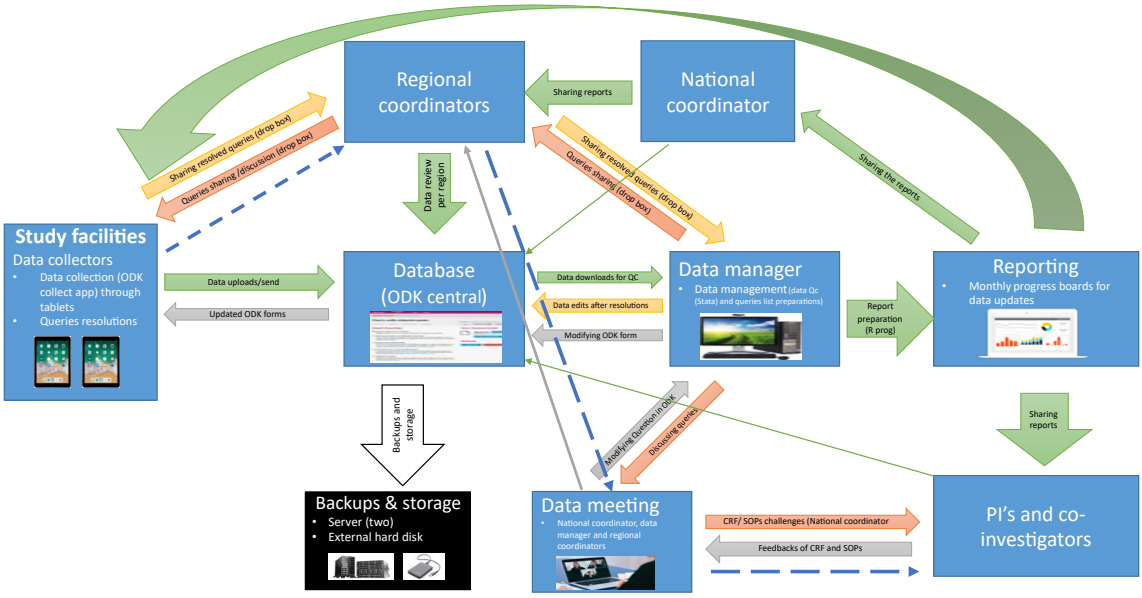
A. SBBC II Mtuha data collection form and Extra form

Data collection: We use a data collection form (SBBC II Mtuha data form) designed for SBBC II purposes to collect de-identified clinical events/actions, and outcomes level data during intrapartum care in each health facility. Information is extracted from labor and delivery registers, and/or inpatient registries, death registries, or referral books. The SBBC II Extra form is designed to collect de-identified information on the use of clinical SBBC II tools (i.e., Moyo, NeoBeat, and Upright bag mask) from the patient's file, plus information about patient source of admission, gestation age, and fetal heart rate on admission, collected from the patient's file (i.e., partograph/Labor care guide) These data collection forms were converted to an electronic ODK format that is used for data collection (entry) and transferred to the ODK central database by dedicated data collectors (n=70). Each data collector covers approximately 2 sites/facilities, entering data once or twice a week.

Data type: These are clinical de-identified data from labor and delivery registers and/or inpatient registries, death registries, or referral books, and patient's files (partograph/labor care guide) are collected and entered in the ODK forms (SBBC II Mtuha data form and SBBC II Extra form. It includes all facility deliveries from all the included facilities in the project.

Data flow: The flow and data quality control processes of de-identified clinical data from the SBBC II Mtuha data form and SBBC II Extra form are explained in Figure 1.

Figure 1. SBBC II Mtuha data and Extra form data management flows



Data process and quality checking

- The data collectors manage daily quality control procedures and safe transfer of data to the central research servers at Haydom on a weekly basis. A central data manager at Haydom performed a second quality control and resolves the remaining issues with the data collectors, before storing the data.

- The data quality control process started during ODK form creations and modification by ensuring customization of the questions on the ODK form. Continuous second data quality checking is performed monthly, by the data manager at Haydom, once the data are in the database to check for completeness and consistency of the data collected, by using data analysis software in STATA. A list of codes is prepared to check the data completeness and consistency.

- To help resolve errors, incompleteness and inconsistencies, a list of queries is prepared and shared with data collectors at each site/facility. They are required to resolve each query within a week after identification.

- Key clinical indicators and training data are processed monthly and presented in standardized reports of each facility and region. These reports are shared with the specific region and all facilities monthly, serving as a local feedback and benchmarking mechanism to drive local discussions and evaluations of own work as part of the continuous quality improvement component of the project.

Data accessibility

The clinical de-identified data collected via SBBC II Mtuha data form and Extra form, can be accessible from the ODK central database to specific people identified on different levels depending on their role in the study.

i. Data collectors have no access to the database, their role is to collect data and upload, once they have sent/uploaded the data, they cannot change anything
ii. Regional coordinators have access to the database but can only review the data of his/her region to monitor the progress of work. They do not have access to other regions and cannot edit/change anything on the data collected.
iii. One Data manager, also the administrator of the database, is the one who has access to editing any data submitted in the database after having proof from the source (data collector) that needs to be edited/changed. The proof used is the resolved query list from data collector after doing verification from the source (i.e., labor and deliveries registries, and/ patient sources). One IT administrator oversees the maintenance of the server and updates of the database.
iv. Principal Investigators (PIs) and co-investigators who have been identified can access the data only by viewing the data submitted in the database, but they cannot edit/change data that has been collected directly to the database. They can download and work with the dataset in their own office password protected computer, for those who are outside the country (Tanzania) will be given access through virtual private Network (VPN).

Data storage and security

- Clinical de-identified data are stored in multiple ways; automatic back-up system on multiple servers in different locations, password protected external hard disk with double authorization as well as one password protected computer used for data management with single authorization which is managed by the data manager. All electronic devices (tablets, computers, and savers) with data, and those used for data collection, are password protected for security purposes. The access to the database is restricted within the country, those who are outside and identified by the PIs and have the permission from Tanzania Commission for Science and Technology (COSTECH) and National Institute for Medical Research (NIMR), will be given access through virtual protected network (VPN).

B. **Individual skills simulation training data (NeoNatalie Live data)**

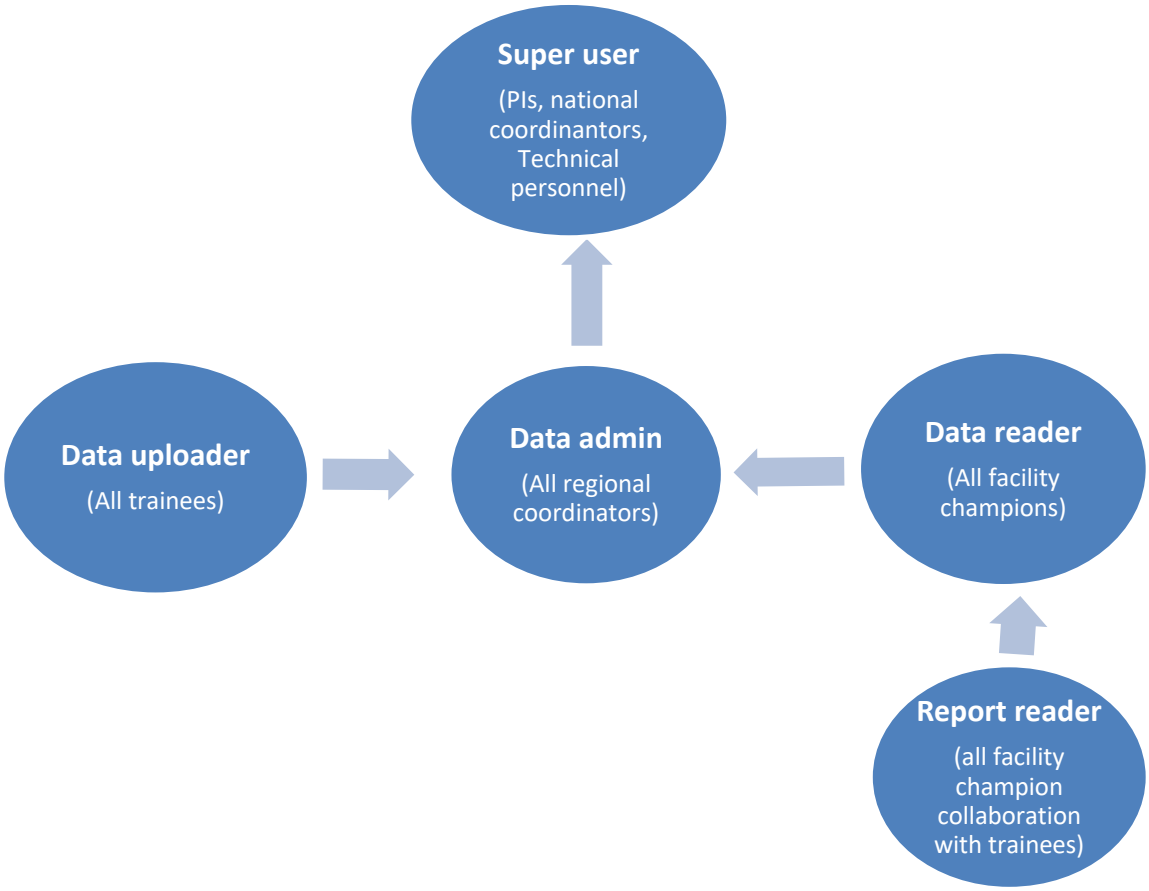
Data collection tool: The NeoNatalie live app is used to collect data from simulation training; both individual skills and team training progress over time. The app is installed in study devices (tablets)

Data types: It collects simulation trainings data, which is captured semi-automatically by the innovative training simulator, i.e. NeoNatalie Live. It collects the frequency of training, both for individual and group scenarios. The simulator

system provides feedback after the end of each training session, and this data is also stored.

Data flow: NeoNatalie Live Simulation training data access and flow are explained in figure 2.

Figure 2. NeoNatalie Live and LIFT data flow and access



Data process: NeoNatalie Live data is collected semi-automatically during every training on the study devices (tablet). Data is uploaded regularly to the Microsoft Azure cloud server, named Liveborn, by the data collectors. Haydom has a user agreement with Laerdal.

Data accessibility: NeoNatalie Live data is accessible for different levels of access, there are super user, data admin, data reader and data uploader.

- **i.** Super users can manage the user, view dashboard, view weblog and upload data.
- **ii.** Data admin can view dashboards, view weblogs and upload data, Data readers can view dashboard and weblog,
- **iii.** Report readers can view dashboard and
- **iv.** Data uploaders can only upload data.

Data storage and security; Simulation training NeoNatalie Live data are stored in the Liveborn, Microsoft Azure cloud server, Haydom domain owned, controlled and maintained by Laerdal Global Health in which there is limited access for only identified people. Since it collected identified information using the names of learners, access was given to those delegated by the principal investigator. At the end of the study, data will be downloaded, and each learner will be assigned a unique ID that will be used to de-identify them. When these data need to be used before the end of the study, the de-identification must be done by assigning each learner a unique identification and the names need to be removed from the dataset to observe the privacy and confidentiality of the participants.

C. Learning Improvement and Facilitation Tool (LIFT)

Data collection tools: The LIFT scenario app is used to collect in-situ simulation training data. The app guides facilitator-led assessments and collects participant assessments digitallyThe app streamlines training program assessments through its innovative digital solution to provide immediate feedback on participant performance through automated scoring and offers data-driven insights for progress tracking and informed decision-making.

Data type: It collects semi-automatic simulation training data from the simulator, and facilitator-led assessments, and participant assessments digitally.

Data flows: The data flow of LIFT scenarios are similar to the NeoNatalie Live data, explained in Figure 2.

Data process: LIFT scenarios data is collected semi-automatically during every training on the study devices (tablet). Data is uploaded regularly to the Microsoft Azure cloud server, Haydom domine, named Liveborn, by the data collectors. Haydom has a user agreement with Laerdal.

Data accessibility: LIFT scenarios data is accessible for different levels of access, there are super user, data admin, data reader, and data uploader.

- **i.** Super users can manage the user, view the dashboard, view the weblog and upload data.
- **ii.** Data admin can view dashboards, view weblogs and upload data, Data readers can view dashboards and weblogs,
- **iii.** Report readers can view the dashboard and
- **iv.** Data uploaders can only upload data.

Data storage and security: LIFT scenarios simulation training data are stored in Liveborn, Microsoft Azure cloud server, Haydom domain owned, controlled, and maintained by Laerdal Global Health in which there is limited access for only identified people. Since it collected identified information using the names of learners, access was given to those delegated by the principal investigator. At the end of the study, data will be downloaded, and each learner will be assigned a unique ID that will be used to de-identify them. When these data need to be used before the end of the study, the de-identification must be done by assigning each learner a unique identification and the names need to be removed from the dataset to observe the privacy and confidentiality of the participants.

D. SBBC II facility readiness assessment tool and SBBC II Mentorship checklist tool data

Data collection tools: SBBC II facility assessment tool is used to collect data during assessment of the health facility readiness to assess the service readiness for intrapartum care at the start of the project. SBBC II mentorship checklist tool is used to describe the facility on availability/accessibility of the services of intrapartum care and identify any area that needs improvement/emphasize during each supportive mentorship activity conducted in all study health facilities. These data collection tools are converted to ODK.

Data types: Health facility readiness data are collected to assess the service readiness for intrapartum care at the start of the project, and mentorship data are collected during each mentorship visit describing information about the facility on availability/accessibility of the services of intrapartum care.

Data flow: During the readiness facility assessment and every time supportive mentorship is conducted, the national and regional facilitators (who have undertaken mentorship) must fill in the SBBC II Mentorship checklist tool, ODK form for each facility. After they complete filling in the checklist form, they upload the form to the ODK central database.

Data process: Health facility readiness and mentorship data collected are uploaded in the ODK central database, they will be stored and downloaded for compilation of facility readiness report and mentorship reports and other project purposes.

Data accessibility: For supportive mentorship/supervision data can be accessible from the ODK central database to specific people identified on different levels depending on their role in the study.

- **i.** Regional coordinators have access to the database but can only review and download the data of his/her region to monitor the findings observed during the mentorship for their respectful region.
- **ii.** One Data manager, also the administrator of the database, is the one who has access to editing any data submitted in the database after having proof from the one who completed that form (national and regional facilitators) that needs to be edited/changed. The proof used is the resolved query identified by national and regional facilitators after doing verification from the source (i.e. health facility). One IT administrator oversees the maintenance of the server and updates of the database.
- **iii.** Principal Investigators (PIs) and co-investigators who have been identified can access the data only by viewing the data submitted in the database, but they cannot edit/change data that has been collected directly to the database. They can download and work with the dataset in their own office password protected computer, for those who are outside the country (Tanzania) will be given access through virtual private Network (VPN).

Data storage and security; Supervision data are stored in multiple ways; automatic back-up system on multiple servers in different locations, password protected external hard disk with double authorization as well as one password-

protected computer used for data management with single authorization which is managed by the data manager. All electronic devices (tablets, computers, and savers) with data, and those used for data collection, are password-protected for security purposes. Access to the database is restricted within the country, those who are outside and identified by the PIs and have the permission from Tanzania Commission for Science and Technology (COSTECH) and National Institute for Medical Research (NIMR), will be given access through virtual protected network (VPN).

3. **Ethical and legal requirements:** All the investigators involved in the study must be trained in Good Clinical Practice and Research Ethics as part of an effort to be obliged in both ethical requirements and scientific merit. Before submitting any manuscript to a journal, the authors must submit the manuscript to the Institution's review board (NIMR) to be reviewed and get approval for publication submission to any journal.