MORAMANGA DSS

SITE VISIT REPORT: MURAMANGA/IPM-MADAGASCAR

Version No	Visit date	Prepared by
1.0	29 th September to 05 th October 2019	 Amek Nyaguara Michael Kinuthia

Contents

Acronyms
1.0 Background
2.0 Overview of the Moramanga HDSS
3.0 Comparison between Moramanga DSS and KHDSS 4
4.0 Database architecture
4.01: Initial structure
4.02: New structure after IPM team visit to Kilifi
4.03: Reviewed structure – after Kilifi team visit to IPM8
4.03: Specific recommendations
5.0 Moramanga EDCS
5.1. Technology
5.2 Specific recommendations10
6.0 EDCS data security measures
7.0 Data cleaning organization
7.0.1 Data manager
7.0.1.1 Current Setup
7.0.1.2 Recommended improvements
7.0.2 Data supervisor
7.0.2.1 Current Setup
7.0.2.2 Recommended improvements
7.0.3 Field supervisor13
7.0.3 Field supervisor
7.0.3 Field supervisor 13 7.0.3.1 Current Setup 13 7.0.3.2 Recommended improvements 13
7.0.3 Field supervisor137.0.3.1 Current Setup137.0.3.2 Recommended improvements138.0 IPM DSS future plans13
7.0.3 Field supervisor137.0.3.1 Current Setup137.0.3.2 Recommended improvements138.0 IPM DSS future plans138.0 Appendix 1: Timetable14

Acronyms

- KHDSS- Kilifi Health and Demographic Surveillance System
- EDCS- Electronic Data Capture System
- IPM- Institut Pasteur de Madagascar
- QC- Quality Control
- VA- Verbal Autopsy
- **GPS** Global Positioning System
- PID- Person Identifier
- HH Household
- FW Field worker
- DM Data manager
- DS Data supervisor
- FS Field supervisor

1.0 Background

Following a request from Institut Pasteur de Madagascar (IPM) though MHURAM project to have Kilifi HDSS technical staffs offer an advisory exchange visit to Moramanga HDSS run by IPM. Two members of the KHDSS team travelled to IPM in, Madagascar for 5 days (29th September to 5th October 2019) for collaborative visit and the main objectives the visit were to:

- 1. Review Moramanga HDSS electronic database design:
 - Database architecture
 - Database security
- 2. Review Moramanga database application
 - Data capture system process flow
 - System validations for data collection application
- 3. Data collection tools
 - questionnaires and definitions
 - Hardware
- 4. Data synchronization
- 5. Staffing and responsibility

- Roles of the data manager
- Roles of the data supervisor
- Roles of the field supervisor
- 6. Data capture security measures
 - Tablet encryption
 - Application security

2.0 Overview of the Moramanga HDSS

The Moramanga HDSS was established by IPM in the year 2012 and it is located in Moramanga district which is 110 kilometers (2hrs, 30minutes drive) away from IPM head office. The HDSS covers 3 divisions (Communes) constituting 30 villages with a total population of 78,000 residents. Table 1 below shows some comparison between Kilifi and Moramanga HDSS.

	KILIFI HDSS	MHURAM
Population coverage	290 000	78 000
Farthest distance within	1 hour using motorcycle	2 hours using
HDSS		motorcycle
Round	3 per year	1 per year or after 2
		years (subject to funds)
Unit support	Logistic (Drivers)	Logistic (Driver)
	Community Liaison	
	Officer Health & Safety	
Job list	Field workers (36)	Field workers (10)
	Field supervisors (5)	Field supervisor (3)
	Data supervisors (2)	Data supervisor (2)
	Cartographer (1)	Data manager (1)
	Field manager (1)	Field manager HDSS (1)
	ICT Developer (1)	
	Data manager (1)	
	Data Base administrator	
	(1)	

3.0 Comparison between Moramanga DSS and KHDSS

	Clinician VA (1)	
Time allocated to HDSS	Majority 100%	100% for FW and Field Supervisors, the rest of the staffs work on other studies as well
Funding	Core funding	Overheads from Projects
Relation with community	Community engagement	Projects specific

4.0 Database architecture

4.01: Initial structure

IPM database was initially designed to handle data collection for specific round that is data collected each round were stored separately in its own database.

4.02: New structure after IPM team visit to Kilifi

IPM team redesigned the database to allow continuous update of DSS database in each enumeration round.

In-migrants and trans-migrants are searched from the IPM DSS register and if found, they are linked to their new place of residence.

When registering new person in the field, that person is assigned a temporary DSS person ID, which is replaced, after acquiring a real DSS PID when field data is synchronized to the production database.

Field and production database have same design. Data is collected and stored in individual tablet databases (Fragments) in the field (In Moramanga) and synchronized daily into production database at the IPM which is 110 Kilometers from the field.

Below is the IPM database structure:



4.03: Reviewed structure - after Kilifi team visit to IPM



4.03: Specific recommendations

- Synchronization algorithm had been designed to overwrite the temporary DSS ID for new persons after propagating back the keys to the dependent tables.
 - The recommendation is to **maintain the temporary DSS PID** in addition to the real DSS PID, this is to enable reconstructing back of the transformed data incase need be. It could also be very useful; in case you want to check and confirm that after consolidating the data from different tablets that data integrity across related tables (relationships) is maintained and for

troubleshooting purposes in case of issues with synchronizations that affects propagating of keys across related tables.

- Vaccine records were stored in tbl_enumeration_person having a column for each vaccine and a column for each vaccine date of vaccination. This meant to introduce new vaccine, two new columns had to be added for each vaccine introduced.
 - The recommendation is to normalize the table further and have an additional table holding vaccine information with columns (Name of vaccine, date of vaccination and foreign keys field). This will enhance scalability, if a new vaccine is introduced, there will be no need to change database structure.
- Newborn demographic details (Names, data of birth and sex) were being recorded both in persons and delivery results table. This is to take care of children born in DSS, but later moved out of DSS.
 - The recommendation is to maintain all demographic information in one table – Persons table, then have a flag to mark the person within the DSS and outside the DSS. This will eliminate the need to have demographic details stored in different tables (delivery, visitors etc.) and will help in future, in case of linkage with the hospital, the PIDs will be derived from one table – persons table. n
- Changes to records in the database are updated such that they overwrite the existing values without maintaining a log.
 - The recommendation is to have an **audit trail** that logs changes made to existing records in the database. It can have attributes – Change type, record identifier (primary key in the corresponding table), field being changed, old value, new value, staff initials (person making change), timestamp (date and time of change) and source of information.
- Database referential integrity is validated/checked at the application level only.
 - The recommendation is that, in addition, add **database constraints** to enforce referential integrity at the database level. This will ensure accuracy and consistency of data within relationships is maintained.
- Backups. The plan was to have a backup in IPM created at midnight daily.
 - We recommend to also have daily backup for each tablet in I the field office at Moramanga. This will enable recovery of daily work in case of tablet malfunction Also at the end of the round, it is recommended to consolidate the data in all the backups and compare counts of consolidated data to counts of the records in the production database. Ideally, if synchronization completed without flaws, the counts should match.

5.0 Moramanga EDCS

5.1. Technology

Moramanga DSS EDCS is a custom-made application developed using PHP/MySQL technologies and the VA questionnaire is designed using ODK.

5.2 Specific recommendations

- Household Completeness check currently there no check in place to validate the completeness of a HH.
 - Recommendation is to have button pressed by the field worker to mark completion of a HH. On click of this button, validations check algorithm is executed to check the completeness of the HH data and in case of an issue, an alert pops up notifying the FW of the issue – this enables the field worker to sort out data quality issues before leaving the HH. HH is marked as complete only if all validations check passes.
- Field time management currently there is no a mechanism to track field time management to effectively manage FW time usage.
 - Recommendation to have start and end timestamps log per HH. This can be executed in the backend, where, when a FW worker starts enumerating a HH, the application captures and logs current timestamp as start time and when complete button (In bullet above) is pressed and validations check pass, current timestamps is logged as end time. This would be very helpful to the DSS coordinator to understand FW time usage in the field.
- Users management. Current users are DSS round specific and not tracked across different rounds.
 - The recommendation is to have a permanent users management module that shall always generate a unique FW code over time. That means even if a FW leaves IPM, the code is not reassigned to a new FW, rather the new FW get a totally new and unique code. If a FW rejoins, his old account is reactivated and continues using the old FW code. This helps mapping of data to the FW who collected it over time.
- Versioning. Currently the application is not under version control.
 - The recommendation is to set up a configuration management component that will assist in revision management. This helps in management of changes in the application, and incase the changed version is unstable, and has bugs, it facilitates reversing back to a previous stable version. It also enables more than one developer to work on the same project simultaneously by managing codes merging and conflict resolution. Explore, Git or SVN.
- Application Testing. Currently the application was being tested against the user's requirement.
 - Recommendation is to have in addition a parallel data collection during the testing phase paper-based data collection and EDCS then compare the data from the two approaches. For two weeks pilot had been done on paper, parallel data collection can be simulated by re-entering sample questionnaires in EDCS using same values as in paper questionnaire covering all possible scenarios and carry out data comparison. When this test passes, the data from the testing tablets is synced to the main/production in IPM and second check is performed between the data in the tablets and data in the production database. This is to verify that synchronization process works as expected and propagating of keys across

related tables works fine – this is important since if it is not working as expected, it can lead to data corruption.

- Field Visit. During the field visit, below issues were flagged out
 - On portrait mode, the tablet could not show color changes on enumerated residents. Recommendation is to train FW to use landscape mode.
 - Some residents' names did not match actual names given by the respondents. Recommendation to investigate whether during work allocation, there were any system flaws, or it is just the previous FW who did not capture the names correctly.
 - When creating a new HH, there is a prompt to key in the count of HH members, and a form with similar number of input fields is generated. Incase mid interview you get that the count was wrong, there no provision to add new fields. This needed to be sorted before rollout.
 - While searching in-migrants, the provision to confirm related members to the immigrants, was to re-search using HH number. This would need a FW's initiative to do a search again which is not guaranteed. The recommendation is to add these details on the person view dashboard for easy verification.
- Audit trail incorporate audit trail (Change log) in the application to capture all changes made by the FW.

6.0 EDCS data security measures

Below are some of security measures discussed and recommended:

- Role based access control -Each user's account to have predefined roles and privileges. This enables specific component of the system to be restricted to authenticated and authorized users only.
- **Application security.** -System design to detect and prevent common application hacking methods. Below are some of the design consideration recommended:
 - SQL injections Design of data execution models in a manner that attackers cannot access and execute malicious SQL statements in the database. Also, all passwords fields to be fully encrypted.
 - Session attacks- The system to automatically kill all sessions, if the application is idle for a moment Explore the actual time.
 - Hiding code Server-side code to be hidden from client users to minimize chances of attackers studying code algorithms with intent of breaking into the system.
 - Error reporting Error reporting to be structured in a manner that it cannot suggest cord execution mechanism. Debugging to be disabled at production to ensure attackers do not feed the system with improper data with the aim of studying error messages.

- Limited users' access to system files and database folder- Access right to system files and database folders to be limited to administrators only. This is to minimize chances of a field worker colluding with an attacker to transfer both system and database files to another unauthorized machines.
- **System recovery** -Daily backups to be performed per tablet in Moramanga, to enabled recovery in case of crash.
- Weak system user account. -The database system account limited to CRUD (Create, Read, Update and Delete) functionality only. This is to ensure, even in case of leakage, the database system account cannot be used to do large data transfers, replications or dropping of the IPM EDCS database.
- **Tablet encryption**-Full disk encryption feature. It is designed to protect data by providing encryption for entire volumes. Procedure shared in appendix 2.
- Limited installed application-Only applications that support IPM EDCS application to be installed on the tablets and only administrator's account which are not available to other users that has rights to install more applications. Thus, Data access and manipulation will be limited to the IPM EDCS application only.
- **Data query logs/Audit trails** Logging and tracking of all data queries by authorized users through the IPM EDCS system to minimize chances of users querying information that is not related to the tasks they are supposed to be performing.

7.0 Data cleaning organization

7.0.1 Data manager

7.0.1.1 Current Setup

DM creates data cleaning scripts, executes them in the production database hosted in IPM and generates data query reports. The reports are transmitted via mail to Moramanga for action by data supervisor.

The DM supervises the two data supervisors – to achieve this, the plan was to have DM carry out 2 weeks supervisory visits to Moramanga each month.

7.0.1.2 Recommended improvements

Considering the database is hosted at IPM and not accessible from Moramanga, it was clear that the two weeks that the data manager will be in Moramanga per month– will be not productive. It is for this fact that it is recommended that, it shall be efficient for DM only to visit Moramanga when there are specific issues to address e.g. during kick off of a round, a case of major data issue or rollout of new feature in DSS.

7.0.2 Data supervisor

7.0.2.1 Current Setup

The two data supervisors sit in Moramanga. Receive data queries from the DM and resolve the ones that do not require field verifications and forward the queries that need verification to the field supervisor.

They compile feedback from query report and forward back to DM for subsequent cleaning.

7.0.2.2 Recommended improvements

Considering the database is hosted at IPM and not accessible from Moramanga, it was clear that the two data supervisors would not be utilized fully. This is because they will fully depend on DM to perform data quality check and only receive compiled query report.

It was also clear that there was a risk of the data supervisor acting just like a messenger – just receiving data queries from DM and passing to field team, and forwarding the feedback back to DM.

From this fact, it was recommended that:

- 1.0 One data supervisor to sit in IPM and work with the DM to perform data quality checks and compile data query reports in liaison with the DM.
- 2.0 The second data supervisor to remain in Moramanga and assist in data queries in field, maintain daily data backups in the field and receive data queries from IPM.

7.0.3 Field supervisor

7.0.3.1 Current Setup

Three data supervisors supervising ten field workers.

7.0.3.2 Recommended improvements

It was observed that the ratio of 3 field supervisor for 10 field workers is very high. Considering there was strain on the current fieldworker and need to fully utilize the three field supervisors, it was recommended that:

- 1.0 The field supervisors to assist in field data collection during morning sessions and carry out supervisory duties in the afternoon.
- 2.0 The field supervisors to carry out field verifications that would be needed instead of re-assigning the task to a field worker.
- 3.0 To train one field supervisor to carry out VA interviews

8.0 IPM DSS future plans

- Come up with a hospital database
- link HDSS database with health facilities database in real time?
- Come up with CSMI database Immunization of the children

9.0 Appendix 1: Timetable

Visit program of the KHDSS team to MHURAM project, Moramanga Madagascar September 22th – 28th 2019 Version of August 27th, 2019

Institutions: Institut Pasteur de Madagascar (IPM), Kenya Medical Research Institute (KEMRI) Wellcome Trust

September 22th: KHDSS Team arrival (Kenya airways)

Contacts:

Rila Ratovoson	rila@pasteur.mg	+261 34 09 128 55
Reziky Mangahasimbola	mreziky@pasteur.mg /	+261 32 65 040 71
	treziky@gmail.com	
Amek Nyaguara	ANyaguara@kemri-wellcome.org	
Michael Kinuthia	MKinuthia@kemri-wellcome.org	
Miora Rasoloarison –	pm_epidemio@pasteur.mg	
Mialy Rakotoniaina –		
Anna Rakotoarison		

Day 1: Sunday 22 th September 2019			
-	Arrival of KHDSS Team at Ivato airport – Transfer to the hotel	Driver of IPM	
	Day 2: Monday 23 th		
08h30 –	Presentation with IPM Direction; Head	Hotel pick-up at 08h00	
12h00	of Epi-RC Unit	(KHDSS Team) - IPM	
	Presentation of MHURAM (background,	Driver	
	procedures, teams, organization of the	EPI-RC unit	
	field team)	EPI-RC meeting room,	
	Participants: Rila – Reziky – Mano –	Dodin building	
	Maheninasy – Maria – Landru – KHDSS		
	Теат		
12h00	Lunch		
14h –	Presentation of the architecture of the	EPI-RC meeting room,	
17h30	MHURAM database and collection	Dodin building	
	tools.		
		Back to the hotel	
		(KHDSS Team) - IPM	
		Driver	

	Participants: Rila – Reziky – Mano –	
	Maheninasy – Maria – Landru – KHDSS	
	Team	
	Dinner	hotel)
	Day 3: Tuesday 24 th	
08h00 –	Working meeting on the database	Hotel pick-up at 07h30
12h00	(objective: check the organization of	(KHDSS Team)
	the new database and propose areas	IPM Driver
	for improvement)	FPI-RC meeting room
	Participants: Reziky –Maria – Landru – KHDSS team	Dodin building
12h00	Lunch	
14h00 –	Working meeting on the database	EPI-RC meeting room.
17h00	(objective: check the organization of	Dodin building
	the new database and propose areas	
	for improvement)	Back to the hotel
		(KHDSS Team) - IPM
	Participants: Reziky –Maria – Landru –	Driver
	KHDSS team	
	Day 4: Wednesday 25 th	<u> </u>
08h00 –	Working meeting on data collection	Hotel pick-up at 07h30
12h00	tools and sending (Initial census,	(KHDSS Team)
	follow-up survey) Link between data	EPI-RC meeting room,
	and GPS	Dodin building
	Participants: Peziky - Maria - Landry -	
	KHDSS team	
12h00	Lunch	
14h00 –	Working meeting on data collection	EPI-RC meeting room,
17h00	tools and sending (Initial census,	Dodin building
	follow-up survey) Link between data	Pack to the hotal
	and GPS	(KHDSS Team) - IPM
	Particinants: Reziky –Maria – Landry –	Driver
	KHDSS Developer & Data manager	
19h00	Dinner	KHDSS Team (At their
1		
		hotel)

07h00- 12h00 12h00 12h00 13h30- 14h30	Visit of Moramanga: HDSS area (departure at 07h00), hospital and CSMI (how to link the data there?) Participants: Rila –Mano/Mahenina - Reziky – Maria – Landru – KHDSS Team + MHURAM Team Moramanga Lunch Data cleaning organization at Moramanga	Hotel pick-up at 06h30 (KHDSS Team) - IPM Driver Coq d'Or Moramanga	
17h30	Arrival at Antananarivo		
19h00	Dinner	KHDSS Team (At their hotel)	
Day 6: Friday 27 th			
08h00 – 12h00	Data cleaning organization at IPM Antananarivo (role's identification, tools, data flow) Participants: Rila –Mano - Reziky – Maria – Landru – KHDSS Team	Hotel pick-up at 06h30 (KHDSS Team) - IPM Driver EPI-RC meeting room, Dodin building	
12h00	Lunch		
14h00 – 16h00	Data cleaning organization at IPM Antananarivo (role's identification, tools, data flow) Participants: Rila –Mano - Reziky – Maria – Landru – KHDSS Team	EPI-RC meeting room, Dodin building	
16h00 –	Wrap up	EPI-RC meeting room,	
17h00	Participants: Laurence B./ Rindra – Rila –Reziky – KHDSS Team	Dodin building	
19h00	Dinner	KHDSS Team (At their hotel)	
	Day 7: Saturday 28 th		
-	Departure of KHDSS Team to Ivato airport	IPM Driver	

Kilifi team explaining to moramanga HDSS team on how to carryout parallel test on application using pre-filled questionnaire at the field office

10.0 Appendix 2: Encrypting Android device

Below find the procedure to encrypt if you are running on Android.

Before encrypting ensure:

- 1. that your data is backed up, in case of any failure during the process, that would make your data no longer accessible.
- 2. Also ensure you have more than 80% of battery power, and that your battery is functioning optimally.

Step 1 Select 'Settings'

Step 2.

On the settings screen select 'Lock screen and security'

Step 3

On the Lock Screen and Security screen select 'Encrypt device'

Step 4.

The device will ask you to setup a PIN that will be used after the process to access your data. *Note: Kindly try with one device first and test to ensure all your processes are not affected.*