Countrywide Mortality Surveillance for Action (COMSA) Sierra Leone
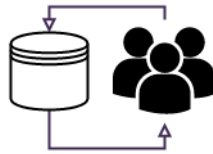
Data Access and IT Plan

Submitted to the
Bill & Melinda Gates Foundation
Grant # OPP1159622
September 30, 2019

Ministry of Health and Sanitation
Sierra Leone

Centre for Global Health
Research

Njala University

NOTE: This draft data access plan and IT plan will be further refined after consultations with Government of Sierra Leone and key partners.

Table of Contents

COMSA - Draft DAP and IT plan

COMSA - Draft DAP and IT plan

Abbreviations

BMGF	Bill & Melinda Gates Foundation
C4	COMSA Central Coordinating Committee
CDGT	COMSA Data Governance Team
CGHR	Centre for Global Health Research
CHAMPS	Child Health and Mortality Prevention Surveillance Network
COD	Cause of Death
COMSA	Countrywide Mortality Surveillance for Action
CRVS	Civil Registration and Vital Statistics
DDI	Data Documentation Initiative
eVA	Electronic Verbal Autopsy
GOSL	Government of Sierra Leone
ICD-10	International Statistical Classification of Diseases and Related Health Problems (10th version)
IDSR	Integrated Disease Surveillance and Response
IT	Information Technology
LMIC	Low and Middle Income Country
MCCD	Medical Certification of Cause of Death
MDS	Million Death Study
MITS	Minimally Invasive Tissue Sampling
MOHS	Ministry of Health and Sanitation
NCRA	National Civil Registration Authority
PSU	Primary Sampling Unit
SL-SRS	Sierra Leone Sample Registration System
SRS	Sample Registration Systems
VA	Verbal Autopsy
WHO	World Health Organization

Document Revision History

Date	Description	Version
Sep. 30, 2019	Initial version submitted to BMGF	1.0
Oct. 15, 2019	Revision asked by BMGF	1.1
Nov. 7, 2019	Includes feedback from Sierra Leone partners	1.2

COMSA SL - Background

Most deaths in low and middle income countries (LMICs) like Sierra Leone occur at home and without medical attention, resulting in Cause of Death (COD) information being mostly unknown (Jha 2014). For example, only 3% of the world's children who died in 2010 had a complete medical certificate at death (Black et al. 2010). Sub-Saharan African countries account for half of global child deaths, and disproportionately lack routine, reliable, nationally-representative COD information. This lack of information greatly limits evidence-based resource allocation, misses identifying new diseases or outbreaks of existing diseases, reduces accountability over health expenditures and hampers population-based research on risk factors. Practical solutions to address this first problem already exist, namely random samples of the COD using Sample Registration Systems (SRS). The Indian Million Death Study (MDS), led by CGHR in collaboration with the Registrar General of India, is the largest of such solutions and has successfully demonstrated "proof of concept" for low-cost, high-impact mortality surveillance. Linkage of mortality through electronic verbal autopsy (eVA) and clinical-pathological confirmation such as Minimally Invasive Tissue Sampling (MITS) could substantially improve global confidence in these COD data. Finally, there is an urgent need to create shared learning experiences and have tools and methods that are widely replicable in other settings.

Study Objectives

1. **Establish the Sierra Leone-SRS.** The COMSA investment will support the government of Sierra Leone (GOSL) to develop and implement a sample registration system of births and deaths with COD in the total population using verbal autopsy. The SL-SRS will draw a statistically representative sample of about 5% of the total population and 5% of all deaths and stillbirths (about 3500 deaths) and run from 2019-28 (COMSA will fund the 2019-22 start-up period). The SL-SRS will track pregnancies, stillbirths, births, deaths by cause at all ages, and in-/out-migration.

2. **Support integrated MITS clinical-pathological testing in Bo District.** Biological confirmation of child deaths is a novel addition to understanding COD, but poses major logistic and organizational challenges. COMSA and CHAMPS will collaborate to support MITS data collection of ~150 - 200 child deaths, drawn from facilities and communities. Bo District (population 575,000) will be the starting site, chosen because it is a peri-urban district with rural and urban populations. This site will supplement the Makeni District already begun by CHAMPS.

3. **Strengthen routine death reporting, mortality system integration and use of data.** This will support GOSL's overall CRVS plan, including linkage with the 117 phone reporting system.

COMSA - Draft DAP and IT plan

COMSA SL - Data Access Plan

Introduction

The Countrywide Mortality Surveillance for Action in Sierra Leone (COMSA SL) data access plan describes how data will be accessed, shared, and disseminated. Guided by the Bill & Melinda Gates Foundation (BMGF) data access principles, COMSA SL has developed a set of policies and procedures for accessing the data collected and produced in this project. The objective of this document is to delineate an approach for providing efficient user access to data while protecting the privacy of study participants, and preserving data security and integrity. This data access plan (DAP) serves the immediate needs of COMSA but it may be modified in the future based on feedback from data users.

Collaborating Partners

The COMSA SL involves the following participating institutions:

- Ministry of Health and Sanitation Sierra Leone (MoHS)
- Njala University (NU)
- Centre for Global Health Research (CGHR)

All participating institutions will act as data stewards and will advocate for the proper use of COMSA SL's data, including the adoption of high standards of data privacy and security. All data collected during COMSA SL will be owned by the Government of Sierra Leone but access to the data at various levels (aggregated and de-identified individual-level data described below) will be provided to the general public, researchers, and partner institutions.

CGHR will build capacity and transfer key skills to MoHS so that all relevant activities continue once this project ends. Similarly, the MoHS will be involved in the implementation of this plan and in any future transition plans.

Purpose of the Data Access Plan

This document defines the COMSA Data Access approach to promoting the transparent collection, processing, analysis, and sharing of COMSA surveillance data, while ensuring privacy, availability, security, and governance. CGHR will implement, in collaboration with Sierra Leone in-country partners, a comprehensive data access framework of practice that:

- Outlines data governance guiding principles;
- Details data governance structure;
- Defines the type and quality of data and information to be disseminated;
- Defines the contexts, conditions, and criteria for accessing data; and
- Outlines the process for requesting, monitoring, and safeguarding data.

While COMSA focuses on providing transparent management, rapid dissemination, and clearly stated rules for access and use of COMSA data, national and international laws governing data use and sharing must be adhered to and enforced. As a result, this data access plan also addresses our approach to understanding the circumstances under which adherence to regulations on data use and management will apply for COMSA in reference to applicable laws, regulations, and policies. Finally, this document serves as a guide to data governance, recognizing that strategies and standards will continue to evolve over time.

Data Overview

The life-cycle of COMSA SL's data begins with data collection in the field from a selected number of sampling areas. Households within these areas are surveyed for recent mortality events (Fig. 1, eVA app). Deceased household members are eventually assigned an ICD-10 cause of death code (Fig. 1, CME app). After the initial survey, a re-sampling of 3% of households is conducted to ensure the quality of data collected (Fig.1, Re-sampling app). Structured field data are ingested by a relational database, whereas unstructured data (e.g. image and audio files) are uploaded directly to the server. Both the relational database and the file-based storage system reside in a secure cloud-based data server. COMSA SL's data warehouse will be comprised of these data alongside a data cataloguing application. This data warehouse in turn powers a web platform of several graphical user interfaces (GUIs) including websites, dashboards, widgets, and a data catalogue. End users will have access to data based on their access rights and approval from the COMSA Central Coordinating Committee (C4). Figure 1 shows a simplified view of the overall architecture and data flow of COMSA SL.
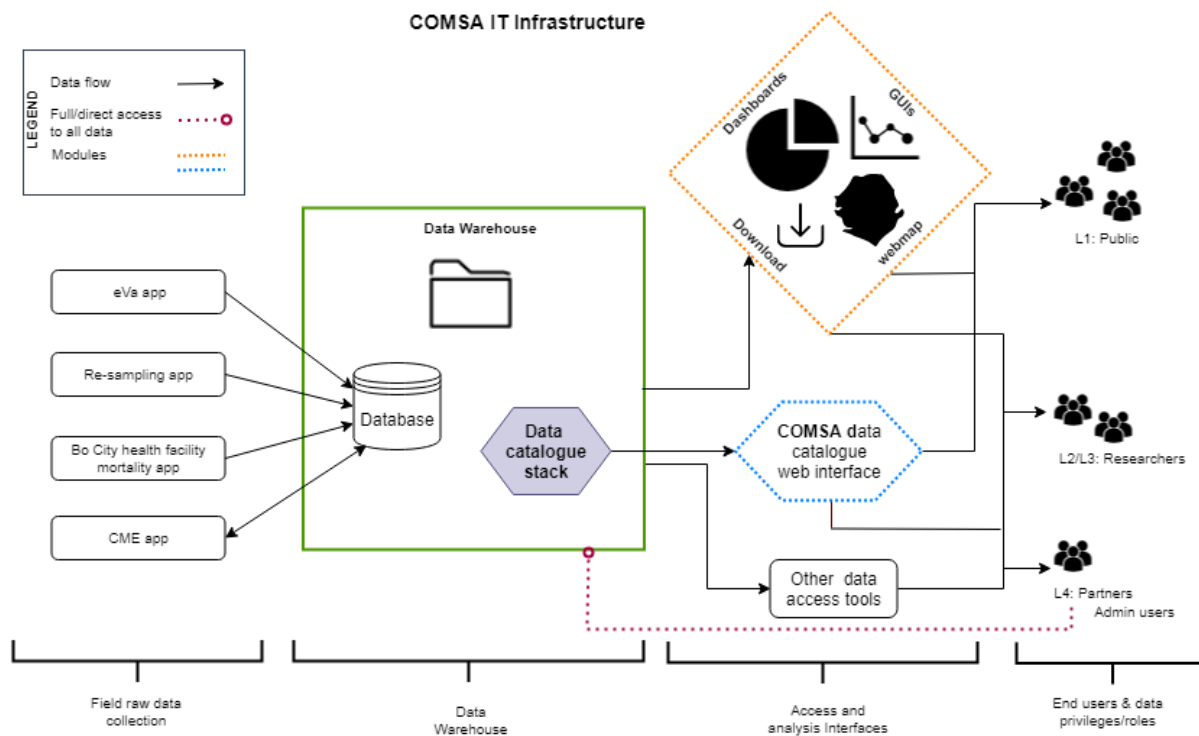
Figure 1: A high-level view of COMSA's IT infrastructure

Data Collection

COMSA SL field data collection will take place in hundreds of census enumeration areas (i.e. primary sampling units, PSUs), across 190 chiefdoms in 16 districts and 54 regions. Two main data collection exercises will be carried out, one at the community level and a second one at the health facility level.

Community Level Data Collection

Each PSU contains about 100 households and roughly 550 people. All current household members and deaths that occurred in the household within the previous year will be enumerated. For all recorded deaths an electronic verbal autopsy (eVA) will be carried out to establish a cause of death and assign an ICD-10 code. To ensure the collection of high quality data, a re-sampling of 3% of death events will be carried out.

Health Facility Level Data Collection

Once community (household level) data are collected, COMSA will collect mortality data at selected health facilities in Bo District within PSUs. The main purpose of this is to strengthen routine death reporting and analysis. Medical information and other relevant data from deaths that occurred at health facilities will be collected and uploaded to COMSA's IT platform. All facility deaths will be assigned a cause of death code if not already assigned by the health facility.

COMSA - Draft DAP and IT plan

For more detailed information about the study and data collected please refer to the COMSA SL Protocol.

Data Access Guiding Principles

COMSA SL's data access principles reflect those of the BMGF:

1. **Governance**: All of COMSA's partners will jointly provide policies and procedures for data governance within Sierra Leone. In-country collaborators will be encouraged to imbue governance procedures with local knowledge, values, and customs to facilitate data management, access, and use by in-country researchers and institutions.
2. **Confidentiality**: Respect will be given to matters of identity, privacy, and confidentiality as they pertain to the individuals and communities from or about whom data are collected. Respect will also be given to matters of attribution as they pertain to researchers, evaluators, and their collaborators.
3. **Attribution**: Proper attribution will be given where it is due. We recognize that COMSA's work in Sierra Leone is only possible due to the many collaborators. It is important for COMSA to recognize the work of field data collectors, researchers, and collaborators.
4. **Accountability**: All processes and procedures for data access will be transparent, clear, and consistent with data management standards that ensure quality data, appropriate security, and equitable access.
5. **Innovation**: Open data access encourages diversity of analysis and opinion. Detailed documentation will provide consistent understanding of the context and nature of data collection as well as their scope and limitations. COMSA SL will facilitate the evaluation of alternative hypotheses and permit meta-analyses and other analytic innovations.
6. **Efficiency**: Practices to enable efficient access to the data will be adopted to promote the sharing, analysis, and publication of data in a timely manner. All data and related publications will be made available to all collaborators openly and promptly.
7. **Collaboration**: COMSA has been planned so that collaboration between researchers can be fostered. Collaboration is crucial as it catalyzes efficient multidisciplinary research.

Data Governance

Members of the COMSA Central Coordinating Committee (C4) will form the COMSA Data Governance Team (CDGT). The CDGT will be comprised of the Director of Disease Prevention and Control; the Director of Reproductive and Child Health; the Director of Health Security and Emergencies, MOHS; the Director of Policy, Planning and Information, MOHS; the District Medical Officer, Bo District; the Dean Elect of the School of Community Health Sciences, Department of Community Medicine and Clinical Studies, Bo Campus, Njala University; and the Data Manager of the Centre for Global Health Research. The main function of the CDGT will be to develop and revise data sharing and management policies and procedures (Fig. 2). Although most data requests will follow standard data agreement protocols such as agreeing to an End User License Agreement, there will be some requests that will require approval by

COMSA - Draft DAP and IT plan

the CDGT. As the work expands, the team will also include members of the collaborating institutions and appointed experts in epidemiology, statistics, informatics, and data management. The CDGT will convene on a consistent basis to discuss data access and sharing needs. The CDGT will appoint one of its members as the go to person for approval of standard L3 data requests.
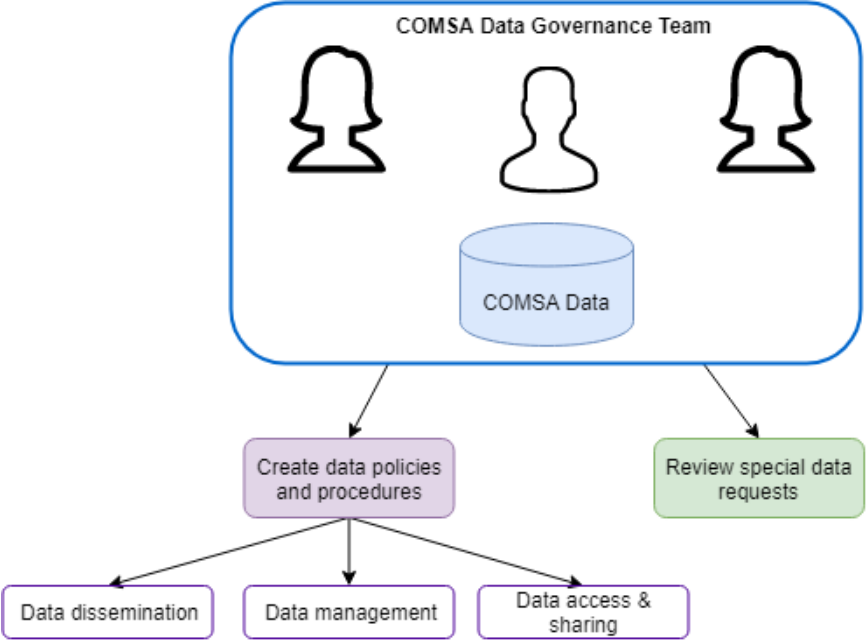
Figure 2: COMSA's data governance structure and functions

As the project progresses the CDGT will update current data policies and procedures and create new ones. Within the CDGT's scope of work, it will need to discuss issues such as,

- Is COMSA meeting its data security and privacy benchmarks?
- What additional data cuts or subsets should be available to the public?
- How can COMSA scale up the dissemination and use of its data?
- Is there a need for new data levels and data access levels?
- Is there room to improve COMSA's data access methods?

The CDGT will encourage active participation of all COMSA's current and future collaborators. The vision for the CDGT is to be the place where future data challenges are solved, and where data policies and procedures are developed. The CDGT will foster research collaboration and keep COMSA's data openly accessible to Sierra Leone and global users.

Data Access

Data Access Levels

The COMSA SL data governance team will be responsible for sharing and monitoring of data at all levels of access. The turnaround time for fulfillment of data access requests and approvals will be dependent on the level of the dataset requested and the method of dissemination. Data access levels, methods, and other relevant policies will continuously improve as the CDGT gets feedback from different users.

Level 1: Publicly Available Datasets
Access to COMSA SL Level 1 (L1) datasets will be unrestricted and open to the public through COMSA's platform. COMSA SL will request the appropriate citation of any open data or materials utilized for presentations or publications. L1 datasets will be in aggregate format, and therefore will not contain identifiable information.

COMSA SL will utilize common web monitoring tools (e.g. Google Analytics) to track industry standard metrics regarding page-views, frequency of downloads, and static content interaction. Reports will be made accessible to the appropriate collaborators and stakeholders per standard operating procedures defined by the CDGT.

Level 2: De-identified Individual-level Datasets
Following the practices of the INDEPTH Network, access to Level 2 (L2) datasets will require online registration and agreement to COMSA SL's Data Use Agreement (DUA). The DUA will state, among other points, that datasets and materials that use COMSA's data should be made available on COMSA's platform within two weeks of publication. All data requests and downloads by users of L2 datasets will be logged and monitored. COMSA SL will request the appropriate citation of any open data or materials utilized for presentations or publication. Similar to L1 data, L2 datasets and materials that use COMSA's data should be made available on COMSA's platform within two weeks of publication.

All identifiers of individuals will be removed in such a way that the data could not be used alone or in combination with other information to identify individuals. L2 data will contain district names and codes only, no smaller geographic subdivision will be present. See Appendix 1 for additional details.

Level 3: Identifiable Limited Individual-level Datasets
Level 3 (L3) data are individual-level records without names of people or contact information. Access to L3 datasets will require registration, agreement to the DUA and the submission of a 250-word study abstract/protocol. Applications will be reviewed and either denied or approved by the CDGT within a 1-3 days. Upon approval the applicant is required to complete a Non-Disclosure Agreement. Only under exceptional circumstances will requests to use L3 data be denied. Special requests for Level 3 datasets such as those that require geographic data smaller than chiefdoms will be subject to IRB oversight and may require IRB approval. Note that L3 data will never contain names or other personal contact information. See Appendix 2 for additional details.

In addition to requesting the appropriate citation of any open data or materials utilized for presentations or publication, COMSA SL will request the opportunity for the pre-publication review of any publications derived from Level 3 COMSA SL data. All data requests and downloads by users of L3 datasets will be logged and monitored. Datasets and materials that use COMSA's data should be made available on COMSA's platform within two weeks of publication.

Unstructured data such as images and audio files will fall under the L3 data classification. The CDGT will govern the access to these data when the appropriate requests are made. However, de-identifying of these data will take place before any data sharing and/or analysis takes place.

Level 4: Identifiable Individual-level Datasets
Identifiable individual-level data that include names and contact information will not be shared with users under normal circumstances. If access to Level 4 (L4) data is requested , the CDGT and an IRB will review such request and either approve or denied it. The default procedure will be that the analysis of L4 data be carried out by COMSA's team members and the output shared with the requester. If this is not possible, then the data will be shared with approved researchers under strict security to protect privacy of study participants.

L4 data will mostly be used for internal data linking and for running quality assurance and quality checks. Only COMSA team members with superuser roles will have direct access to these data. The high restriction on L4 data may change in the future as the Data Governance Team gets feedback and requests from users. All changes in access of L4 data will be deliberated and approved by the CDGT and the IRB.

In summary, L1 and L2 data sets will be immediately accessible on COMSA's platform. L3 data request applications will be evaluated within one week by the CDGT. Fulfilment of special requests of L3 data may vary depending on the complexity of the request and/or the terms of the data use agreement. L4 data won't be shared with users as they are meant for internal usage only. In the future, the CDGT may decide to alter current data levels and may create new ones as needed.

COMSA - Draft DAP and IT plan

Table 1: COMSA's data levels, access features, and governance matrix

Level	Data Features	Access Method	Required Agreements & Documents	Governance	Frequency of Data Refresh	Restrictions
L1	Summarized, tabulated, or aggregated datasets and webmaps. All data are de-identified	-Publicly available -Direct access through web interfaces	-None	Monitor usage: downloads, page-views, map-views, citations	Available for the length of the project and three years after. Updated every 120 days* for the first two years and every six months afterwards.	-**Unrestricted**: Open to the general public -**Limitations:** To ensure statistical de-identification, obfuscation methods and/or date shifting methods may be employed to obscure case counts or event timing.
L2	De-identified individual-level dataset	-Registration online -Access via API or scripts	-DUA	Monitor usage: downloads, citations	Available for the length of the projects and three years after. Updated every 120 days* for the first two years and every six months after that.	-**Minimally Restricted:** Access to Case-level de-Identified data will be limited to users who register online and agree to terms and conditions of the DUA. -**Limitations:** Depending on the context of use, statistical obfuscation methods and/or date shifting methods may be employed to obscure case counts or event timing to ensure statistical de-identification.
L3	Identifiable limited individual-level data	-Registration online -Access via API or scripts	-DUA -NDA -250-word abstract	Data request review and approval by CDGT and monitor usage: downloads, citations	Defined by application request and terms of DUA	-**Moderately Restricted:** Only available under special requests -**Limitation(s):** Depending on the context of use, statistical obfuscation methods and/or date shifting methods may be employed to obscure case counts or event timing to ensure statistical de-identification. Data have no names or contact info.
L4	Identifiable individual-level data	-Registration online -Access via scripts, secured FTP	-EULA -DUA -NDA -Accredited IRB -Study protocol	-Data request review and approval by CDGT -IRB oversight -Policies and procedures set by CDGT	-Defined by application request and terms of DUA -Updated daily for internal admin use only	-**Highly Restricted:** Normally only available to project administrators for internal linking and QA/QC. Approved researchers will need to provide a study protocol among other requirements to access these data. These data will be available through a login-based FTP server.

Data Use Agreement, DUA; Non-Disclosure Agreement form, NDA; Application Programing Interface, API; Quality assurance and quality control, QA/QC.
*Refers to period from capture of death record in the field to posting on COMSA's platform

Data Requests

L1 data will be available to the public directly without the need for registration or submission of a data request. L2 data will be available with an online registration and agreement to COMSA's Data Use Agreement (DUA). Access to L2 data will be immediately available after registration. L3 data will require online registration and a data access request application to start. L3 data requests will be reviewed by the CDGT and accepted or rejected accordingly. If accepted, requesters will also have to sign an NDA form, subject to IRB. Figure 3 shows a data access request flow chart. Similar to L3, L4 data will be available through a data access request application. COMSA SL team supervisors (and/or superusers) will have access to all microdata in order to carry out data quality assurance and quality checks.)
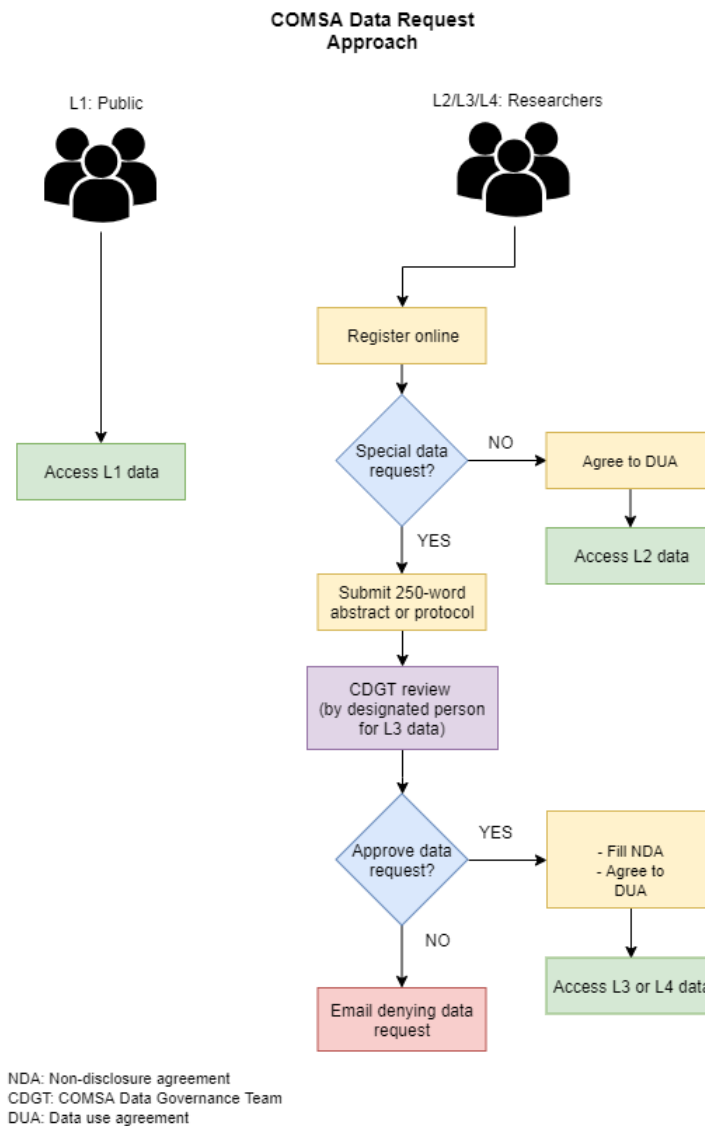

Figure 3: COMSA SL's data request approach for L2, L3, and L4 users and data. L1 data are accessed without the need of any agreements. In addition to a DUA and NDA, L3 and L4 data access need IRB approval as well.

COMSA - Draft DAP and IT plan

Data Dissemination

The dissemination of COMSA SL's data is a critical part of the project. Robust procedures together with an extensive but accessible web platform comprised of various technologies will facilitate the efficient dissemination of COMSA SL's data. Initially, CGHR will be in charge of disseminating data through COMSA SL's IT platform. While doing so, CGHR will provide capacity building and skills transfer to MoHS so that they continue the dissemination of data at the end of the project.

Initial Phase-in and Availability Thereafter

The COMSA approach to quickly disseminating data may be challenging when balanced with a need for accuracy and reliability in the newly accumulated surveillance information.

In an effort to reach a reasonable accord, we propose an initial "phase-in" period to allow for preliminary analysis and interpretation of the data. The phase-in period is effective for up to 6 months from the first date of the completion of data collection in a phase I province. The phase-in period is not meant to apply to each individual province. During the phase-in period, access to the curated (L2-L4) surveillance data from each case is only available to CGHR, MOHS, and Njala University. Summary level information (L1 data) is reportable within the context of case counts or other aggregate-level metrics.

After the phase-in period has elapsed, all levels of curated case-data from the initially deferred cases will be accessible. Subsequent curated case data would then be made available, in their entirety, within 120 days of data collection.

At any time, including the phase-in period, data and information supporting the reporting of notifiable conditions will be made immediately available through appropriate reporting mechanisms to the MOHS and Njala University. Release of information to support case reporting requirements for a specific case could precede final autopsy determinations to allow for public health action prior to release of fully curated COMSA case data.

As changes in methodological approaches and techniques occur over time, it is anticipated that COMSA, after consideration of potential impact of the advances/changes, will reopen applicable prior cases and perform additional analysis and/or generate refined cause of death determinations. The timeline for release of these new findings will depend on the operational and analytic approaches required to generate findings for individual cases and aggregate data sets. All data dissemination from COMSA is subject to constraints defined within Material and Data Transfer Agreements, Protocols, and Informed Consents.
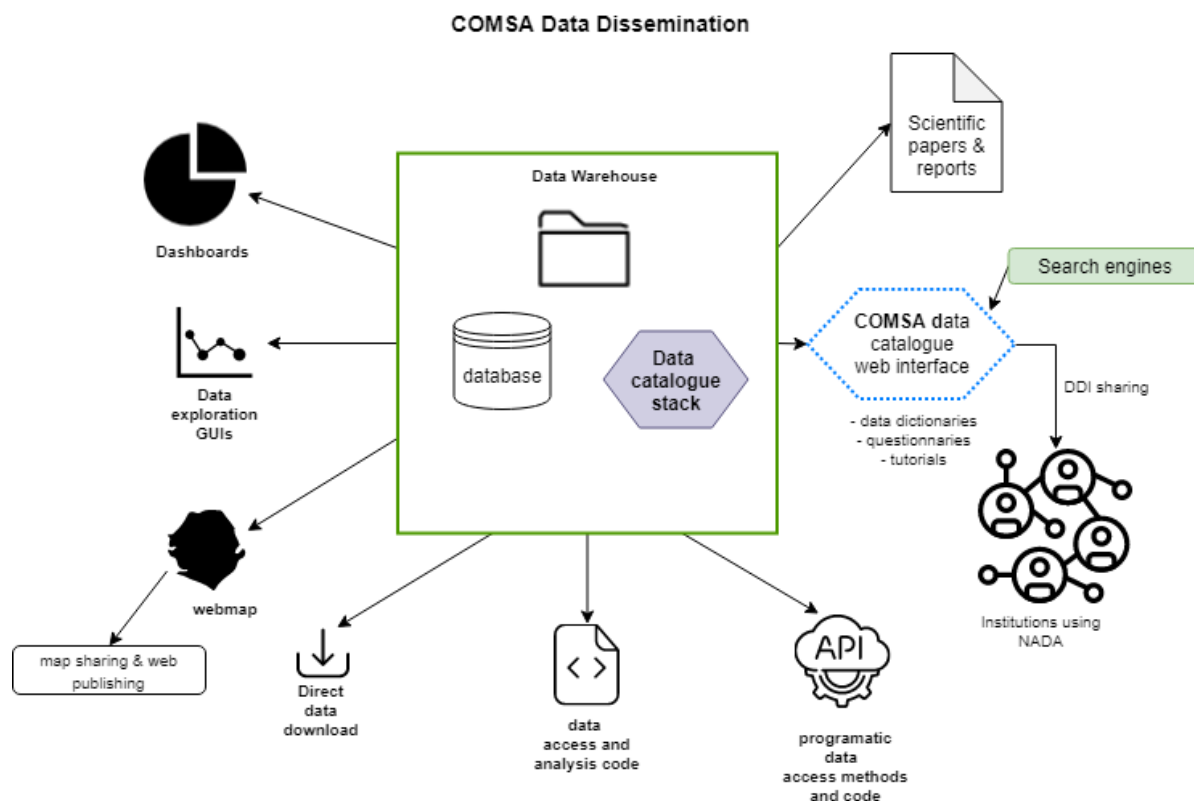
COMSA Data Dissemination

Figure 4: COMSA's data dissemination approach. In terms of user access, L1 data users will have access to data through all the ways depicted above except through data access and analysis code (scripts) and APIs which are only available to L2, L3 and L4 registered users.

Technologies

As shown in Figure 4, COMSA SL will offer users a range of Graphical User Interfaces (GUIs) to interact with its data and metadata. Project administrators can utilize dashboards to grant access rights to users and monitor data collection, page-views, and data usage by various levels of access. Also, users can explore the data and download datasets via web dashboards and other data visualization widgets. A webmap interface will allow users to understand the geographical context of COMSA's data. L1 data will be available for direct download via GUIs. Researchers (L2 and L3 users) with specific data needs will be able to access the data through custom programming scripts and APIs. Moreover, we will also use other data widgets to provide additional data exploration and visualization tools to COMSA SL's users. Finally, COMSA will also provide a data cataloguing application allowing users to easily find datasets and their corresponding metadata.

Data Discovery

COMSA SL's main website (public access front end) will include appropriate text and HTML tags for adequate search engine indexing and optimization. Standard tagging methods such as those of Schema.org will be used so that third-party search engines (Google, Yahoo, Bing, and Yandex) recognize data structures embedded within web pages and index them appropriately. Similarly, a data dashboard

COMSA - Draft DAP and IT plan

and data cataloguing application will provide adequate meta-tags for within-website search and discovery through keywords. Moreover, we will make sure that all publications that use COMSA's data in any shape or form reference back to the main website and other relevant resources. Additionally, the website will contain tutorials about how to discover, access, and use COMSA SL's resources.

Data Formats

Data will be available in widely used formats such as Excel, CSV, and JSON (or GeoJSON). Also, data will be available in database-ready formats such as SQL dumps. Similarly, users will be provided with APIs for downloading data directly to their analysis environments, such as Stata, R, or Python. The available data formats may change or expand as COMSA SL incorporates user feedback. Summarized data reports will also be made available in HTML or PDF format.

Metadata

In order for COMSA's data to become widely accessible it must first be searchable. Metadata standards are important as they readily provide structured key information about data and allow it to be searchable on the web. The Data Documentation Initiative (https://ddialliance.org/), DDI, is a widely adopted international metadata standard. COMSA will adopt the DDI standards to properly describe all data collected in the field as well as other derived data products. DDI files together with field questionnaires, data dictionaries, ICD codes, reports, publications, and other ancillary information will be available to the public in COMSA's online data catalogue. Metadata on the COMSA SL website will be accessible by an open-source cataloguing platform powered by DDI files such as Dataverse. This data catalogue will allow users to search and access COMSA SL metadata. DDI files can be easily shared among institutions that use such metadata, which furthers data dissemination.

As part of COMSA SL's data dissemination strategy, we will also create tutorial and sample code (R and Stata) for researchers to quickly become familiar with the data and how to use them properly. These materials will also be available in COMSA's online data catalogue.

In short, COMSA will leverage the power of DDI to build a data catalogue for efficient discovery and sharing on the web. COMSA will not only strive to enable the free flow of data from the field to analytical environments but also produce and share relevant meta-data. The ready access to data and meta-data will catalyze open, reproducible, and transparent research.

Data Storage and Retention

All data will be captured electronically in the field through laptops and tablets. Structured data will then be uploaded to a relational database residing on a server. Similarly, unstructured data (images and audio files) will be uploaded to folders in the same server. The MOHS will be involved from the start in the planning and execution of data backups and storage. The database and folders with data will be backed up on the cloud and on off-line drives on a daily basis during the length of the project. Similarly, backups will also be done on the MOHS ICT Unit cyber-infrastructure for sustainability and capacity building of

COMSA - Draft DAP and IT plan

the existing health management information system. After the period of performance of the project, the same backup scheme will take place on a weekly basis for one year and on a monthly basis for two more years. Data will be available in the cloud at all times. Once the project is completed the Government of Sierra Leone will continue to make available the data and IT products through cloud infrastructure. The Government of Sierra Leone (specifically MOHS) will be the final custodian of these data as well as the IT products.

Future Data Storage and Retention

At the end of the COMSA project, all data and IT products will be handed over to MOHS for future permanent storage and retention. Moreover, the overall stewardship of the data together with the data will be handled to the government of Sierra Leone (MOHS) and the BMGF in accordance with applicable agreements. Furthermore, at project completion, COMSA SL data access and dissemination will be subjected to terms and policies as defined by the BMGF (or their designee) based on existing MOUs. These terms may be subject to the discretion of the in-country partners or the ministries of health.

Data Security and Privacy

All procedures for data collection and storage will ensure high standards of data security and privacy including best practices from other organizations such as the United Kingdom's Biobank and follow similar procedures mentioned in the IRB submission to MOHS. For instance, field data will be stored on very secure relational databases on servers with strict client authentication controls. Similarly, all client-to-server communications will be encrypted with Secure Sockets Layer (SSL) and the database itself will be setup to require SSL connection by the client by default. Moreover, remote access to backend databases will only be allowed through SSH tunneling.

L1 users (general public) will only have access to aggregated data while L2 and L3 users will have access to micro-data (individual-level data). Individual-level data will be de-identified by removing all fields with personal information, such as names and dates of birth. (Appendix 1 shows additional characteristics of de-identified data.) L1, L2, and L3 structured data will be stored on a single relational database. L1, L2, and L3 users will only be given the minimum privileges needed to access L1, L2, and L3 data respectively. Further, all L3 logins will have set expiry dates to ensure strict user management. L2 and L3 data will have a random unique record identifier.

Identifiable data (L4 data) will be stored on a separate relational database that will only be accessible to L4 users (project administrator users and selected approved researchers) based on specific logins and IP addresses. The random unique record identifier available in L2 and L3 data will also be stored in this database. The default database access port will not be used; instead a new port will be opened and shared only with L4 users. Only one designated person within the team at a time will have superuser access to all data from the backend. The designation of the superuser will be determined and approved by the CDGT. In the case of a breach of security or privacy, the CDGT will be informed immediately and corrective actions will be implemented.

In summary, a clear data user and role management strategy will be devised for COMSA SL to allow data access in an efficient manner that does not compromise data security or privacy. We have past experience from the Million Death Study, which had 900,000 records and not even a single breach of data security and privacy reported. (Note: further data security points are mentioned in the IT Plan.)

COMSA - Draft DAP and IT plan

COMSA SL - Information Technology (IT) Plan

Introduction

The COMSA SL IT Plan describes the IT systems and procedures that are being used to enable collection, storage, and dissemination of data. COMSA SL is using state of the art IT technologies and platforms while also taking into account the ground realities of operating in a resource constrained country where power and internet access are unreliable and intermittent. COMSA SL is leveraging open source technologies and platforms wherever possible to not only maximize the return on investment but also to ensure interoperability. Custom software has been developed for the core aspects of the system when no suitable open source software exists. COMSA SL is also investing in knowledge transfer to the personnel in Sierra Leone to ensure sustainability of the system over the long term.

COMSA SL's information technology infrastructure has three major systems:

1) SL-SRS: The Sierra Leone Sample Registration System will cover identification and mapping of 600 primary sampling units drawn from the 2015/16 census (with each unit having about 100 households and 550 people), household enumeration, verbal autopsy and dual physician coding, support systems like progress monitoring, and payment processing with the final output being levels of births, deaths, and causes of deaths at the national and district level.

2) Facility-Based COD Reporting: COMSA SL will add facility-based COD reporting (using the WHO standard death certificate) starting with Bo District Hospital. This involves integration within existing systems, including IT systems, so as to avoid duplication of efforts and data collection at the local level. We will work within the open-source District Health Information System Software (DHIS2) promoted by WHO and currently being piloted by the MOHS.

3) Data Access and Integration Modules: This system comprises the components used to store the collected data and enable its sharing between different parts of the different systems and modules to provide controlled access to the data according to the specifications laid out in the data access plan.

The SL-SRS System

The SL-SRS system is composed of two systems. The first is for the startup baseline enumeration (in 2019/20) covering household membership, prevalence of some key morbidity and risk factors, and dual physician coded causes of death for all deaths occurring in 2016 to 2018.

The second is the annual netting of events and cause of death.
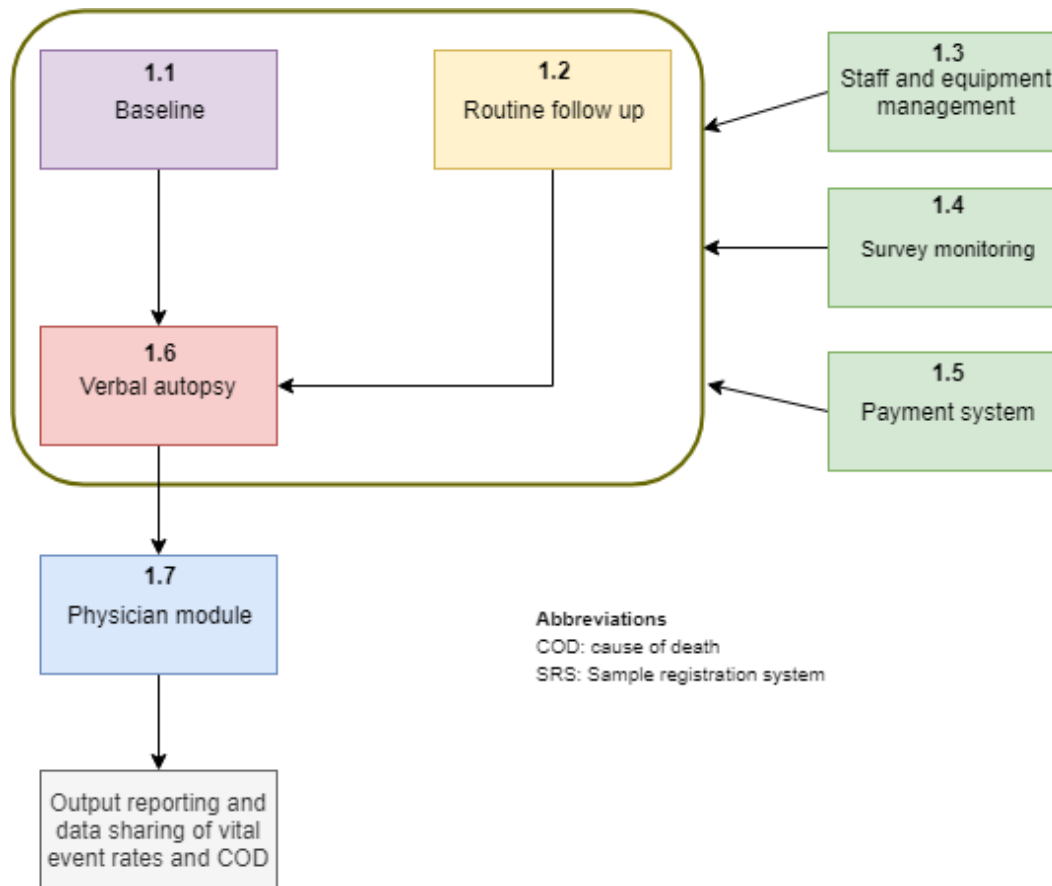
Figure 5: COMSA will require seven modules.

Baseline system (Module 1.1)

The baseline capture system involves the SL-SRS staff identifying sample units within the districts, mapping the boundaries for the study, identifying households, listing the members in the household, and identifying the deaths that occurred in the household in the last three years, and then conducting a verbal autopsy on those identified deaths.

The verbal autopsy data collected will be coded by trained physicians in a process which involves double coding, reconciliation, and adjudication with rules built in to optimize the process. This system will include modules for quality assurance, cleaning, and resampling.

The progress of the survey will be monitored continuously. The performance of the survey staff and physicians will also be monitored using metrics generated by the various modules. Payments which are due are automatically computed and scheduled for payment.
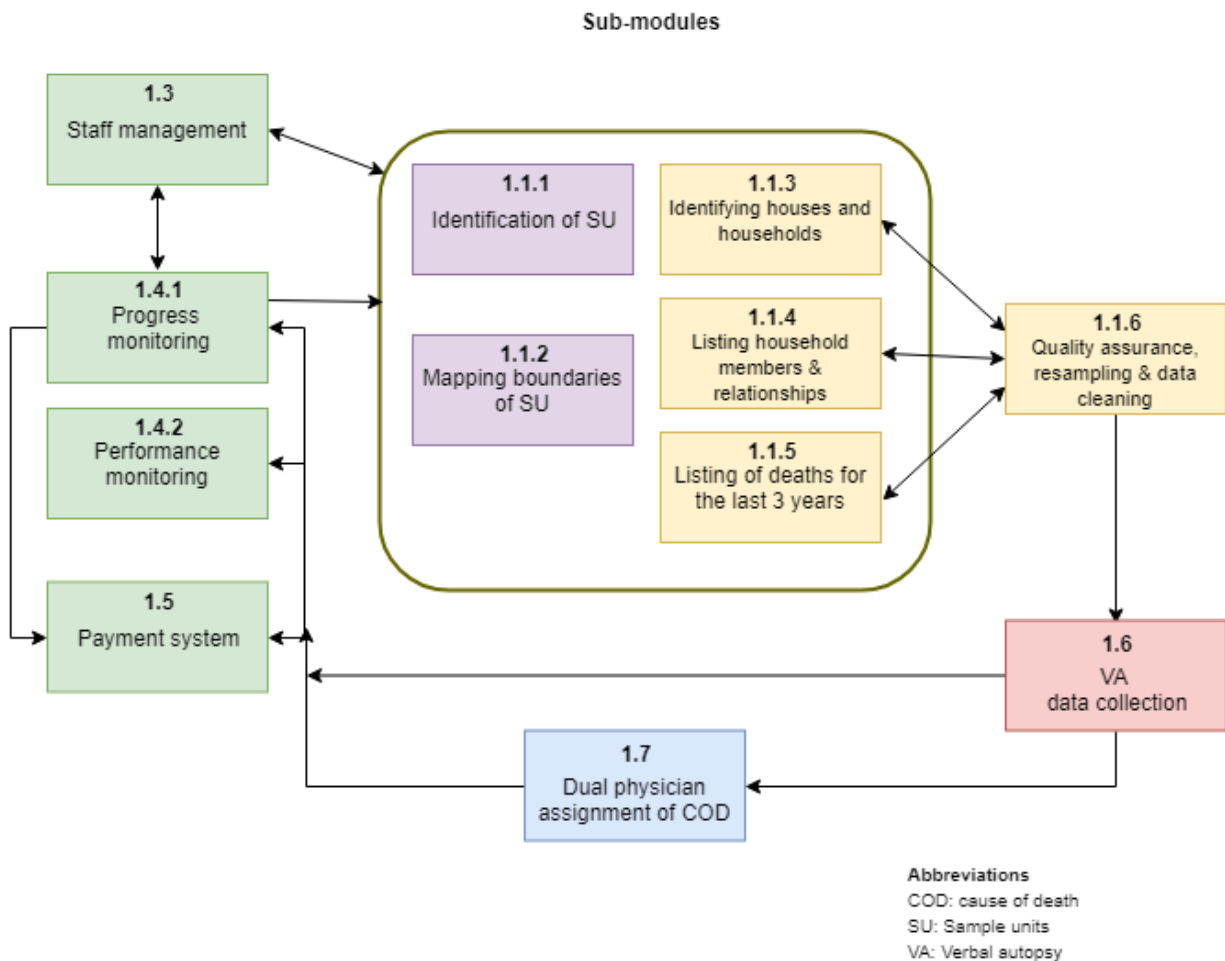
Figure 6: Module 1.1 sub-modules

Sub-modules

1.1.1 Identification of Sample units
This sub-module covers initial identification of areas using maps from the census and based on sampling methods used.

1.1.2 Mapping boundaries of Sample units
This sub-module covers mapping of identified areas using Open Street Maps to guide the Enumeration and VA stages.

1.1.3 Identifying Houses/Households
This sub-module covers the listing of the houses and households using appropriate forms. This includes getting a GPS reading.

COMSA - Draft DAP and IT plan

1.1.4 Listing household members and relationships

This sub-module covers the collection of details of the household members.

1.1.5 Listing of deaths

This sub-module covers forms which let the surveyor collect the details of deaths for the last 3 years, which forms part of the baseline.

1.1.. Quality assurance, resampling, and data cleaning

This sub-module includes resampling for a randomly selected subset of the data collected by surveyors. This sub-module performs checks on the data being collected and flow to enable timely corrections and updates to processes. Data cleaning functions will help clean up errors that occurred in the fieldwork.

The aggregated data from this sub-module is then passed on to the following modules and sub-modules:

- 1.4 Survey monitoring
    - o 1.4.1 Progress monitoring
    - o 1.4.2 Performance monitoring
- 1.5 Payment system, where the work done by personnel can be used to make payments

Routine follow-up system (Module 1.2)

The routine follow-up system will involve re-interview in year 2 and 3 to record changes (household membership, in- and out-migration, and births and deaths) in households in the areas assigned to them. (Note: sample frame is households WITHOUT replacement.)

Once a year, the SL-SRS survey staff will visit the areas and update the baseline data. A verbal autopsy will be conducted on the new deaths that have occurred in the last year and put through a process of quality checks and cleaning.

The verbal autopsy data collected will be coded by trained physicians/CHOs/senior nurses in a process which involves double coding, reconciliation, and adjudication with strict rules built in to optimize the process. The progress of the survey will be monitored continuously; payments which are due are automatically computed and scheduled for payment.

The performance of the survey staff and physicians will also be monitored from metrics generated by the various modules.

Figure 7: Module 1.2 sub-modules (yellow boxes)

Sub-modules

1.2.1 Yearly update
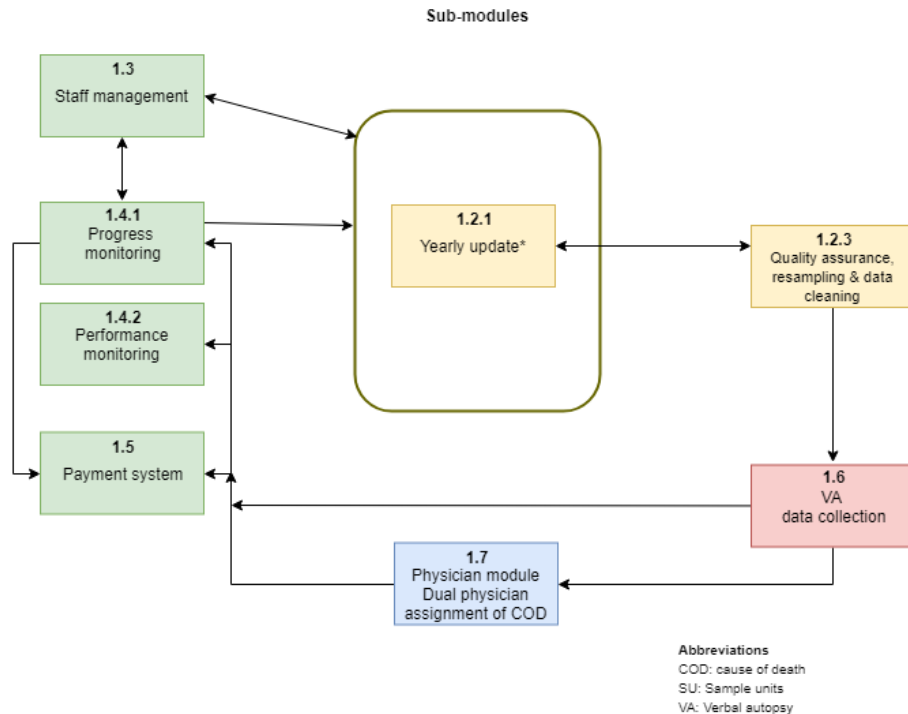
This is for the survey staff to do an annual update of the household data. This facilitates identifying changes/events (births, deaths, pregnancies, etc.) in the households within the Sample Units.

1.2.2 Quality assurance, resampling, and data cleaning

Similar to the module in the baseline enumeration, this sub-module covers the quality checks, assignment of resampling, and data cleaning of erroneous data.

Staff and equipment management (Module 1.3)

To sustain the study, the availability of trained staff at the right time is very important. These modules will help plan, execute, and monitor staff requirements effectively while keeping quality as the primary focus. It will facilitate timely action to ensure the continuity of the study.

The staff will require equipment (especially IT equipment) to do their job. An effective system of inventory management and tracking of these will be part of this sub-module.

Figure 8: Module 1.3 sub-modules

Sub-module

1.3.1 Recruitment needs
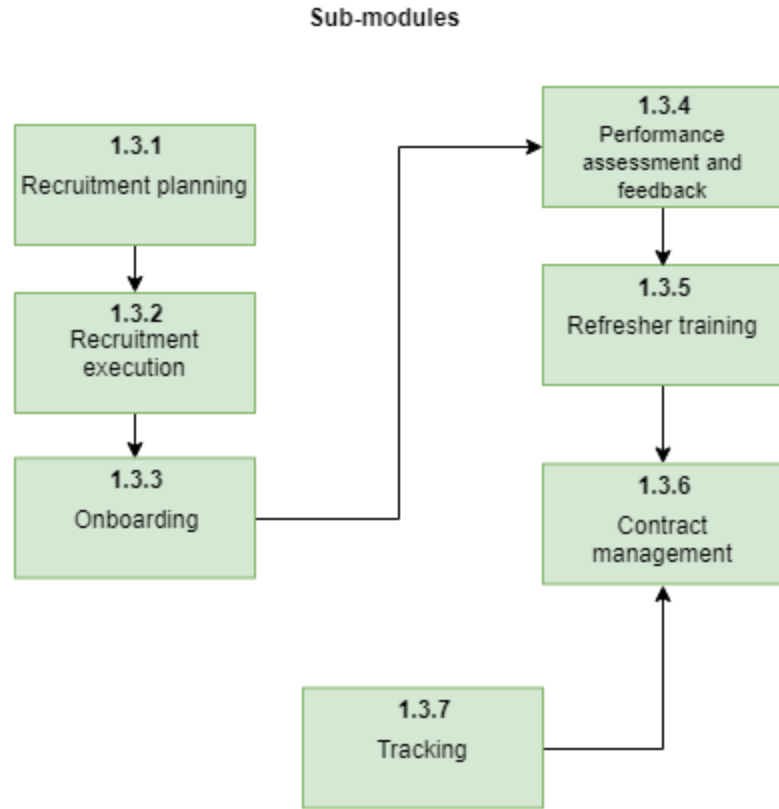
This sub-module is intended to provide data and reports for monitoring required/open/closed positions for key personnel like health surveyors, physicians, and supervisors for the survey based on location.

1.3.2 Recruitment

This sub-module covers the tracking of the status of recruitment of the planned positions.

1.3.3 Onboarding

This sub-module covers the training of the personnel for their job. It covers automated training content with teaching videos along with skill and knowledge tasks with automated scoring and acceptance/rejection.

1.3.4 Performance Assessment and feedback

COMSA - Draft DAP and IT plan

These sub-modules contain the ongoing review of the performance of the personnel, assessment of their performance and qualities, providing feedback, and taking corrective action.

1.3.5 Refresher/Corrective training

This sub-module facilitates the scheduling and conduct of offline and online training for existing personnel who may require these.

1.3.6 Contract management

This sub-module covers tracking of contracts; for example, contracts for personnel may include exits, renewals, equipment-related, vendor contracts, etc.

1.3.7 Tracking assets

This sub-module provides for tracking all the assets being used for the surveys and managing inventories, asset assignment, repairs, and replacement.

Survey monitoring (Module 1.4)

A set of user friendly dashboards will let users get a quick glimpse of progress of the survey and each individual activity, in terms of data collected, data quality (errors), etc. These dashboards will also allow viewing of the performance of individual staff on rate and quality. The individual worker will also be able to view their own performance data.
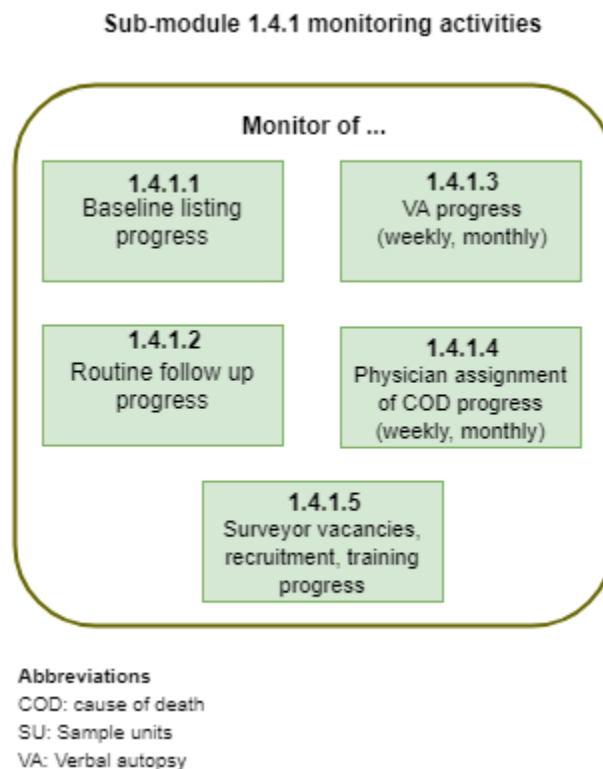
Figure 9: Sub-module 1.4.1 monitoring activities

COMSA - Draft DAP and IT plan

Sub-modules

1.4.1 Overall Progress monitoring

This covers features for the management personnel to monitor the overall progress of the various survey activities.

1.4.1.1 Monitor Baseline listing progress
This sub-module provides various charts and reports for tracking the progress of the baseline survey to the management personnel. It covers performance, progress, and quality related information.

1.4.1.2 Monitor Routine follow up progress
This sub-module provides various charts and reports for tracking the progress of the monthly and yearly updates, whenever they are scheduled, to the management-related personnel. It covers performance, progress, and quality related information.

1.4.1.3 Monitor Verbal autopsy progress overall, weekly, monthly
This sub-module provides various charts and reports for tracking the progress of the verbal autopsy survey to the management personnel. It covers performance, progress, and quality related information.

1.4.1.4 Monitor Physician assignment of COD progress overall, weekly, monthly
This sub-module provides various charts and reports for tracking the progress of the cause of death assignment by the physicians to the management personnel. It covers performance, progress, and quality related information.

1.4.1.5 Monitor Health surveyor vacancies/recruitment/training progress
This sub-module provides an overview of the progress in staffing activities like recruitment and training.

Sub-module 1.4.2 monitoring activities

Monitor of ...

1.4.2.1
Surveyor performance*
(weekly, monthly)

1.4.2.2
Physician
performance*
(weekly, monthly)
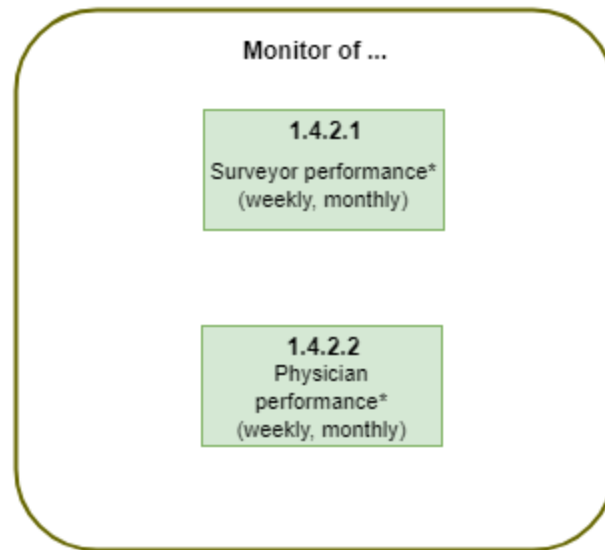
Notes
-Individual performance will be monitored with metrics such as speed and quality. Overall, weekly, and monthly performance will be monitored.

Figure 10: Sub-module 1.4.2 monitoring activities

1.4.2 Monitor staff performance (on the job)

These cover features to monitor the performance and quality of individual personnel on their assigned tasks.

1.4.2.1 Monitor health surveyor's performance
This sub-module contains charts/reports for the health surveyor on speed and quality, reflecting overall, weekly, and monthly performance.

1.4.2.2 Monitor Physician performance
This sub-module contains charts/reports for the physician on speed and quality, reflecting overall, weekly, and monthly performance.

Payment system (Module 1.5)

Timely and accurate payments are one of the keys to a successful implementation. Modules to track the variable payments of the field staff and physicians will ease the overhead involved in this task. Payments will be automatically calculated; appropriate documents will be created and put through an appropriate approval process, with a log of history of changes and corrections.
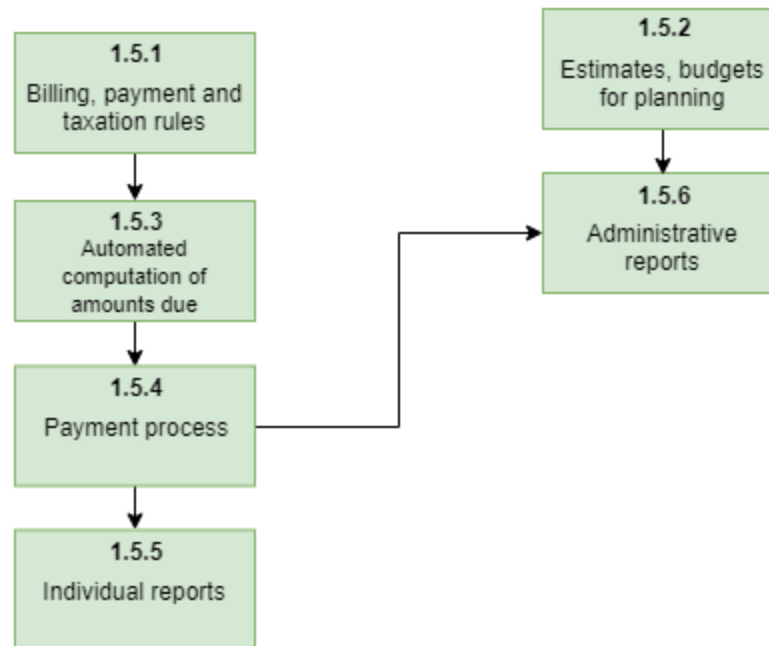
Notes
- 1.5.3: Amounts due refers to bills, taxes, and payments due on a schedule
- 1.3.4: This includes statements, approvals and supporting documents
- 1.3.5: Individual reports contain bills, taxes and payment information
- 1.3.6: Admin report contain expenditures, taxes, planned vs. actual expenses, etc.

Figure 11: Module 1.5 sub-modules

Sub-modules

1.5.1 Configuring Rules
For billing, payment and taxation rules

1.5.2 Estimates, budgets

1.5.3 Automated computation
Of bills, taxes, and payments due on a schedule

1.5.4 Payments
Statements, approvals, supporting reports

1.5.5 Individual reports
Bills, taxes and payment information

1.5.6 Administrative reports
Expenditures, taxes, planned vs. actual, etc.

COMSA - Draft DAP and IT plan

Verbal autopsy data collection (Module 1.6)

Verbal autopsy involves a health surveyor going to households where a death has occurred and collecting data in a structured questionnaire and then writing down a narrative of the events as described by the respondent in a chronological order. Adequate tools for prompting for symptoms and capturing the narrative information will be required.

Sub-modules

Notes
- 1.6.1: Age specific forms (neonatal, child, and adult) with relevant quality control and flow optimization rules
- 1.6.2: Narratives will be captured using context sensitive information, spell checker, and other tools to aid interviewer
- 1.6.3: Quality assurance consists of VA-specific re-sampling and data cleaning
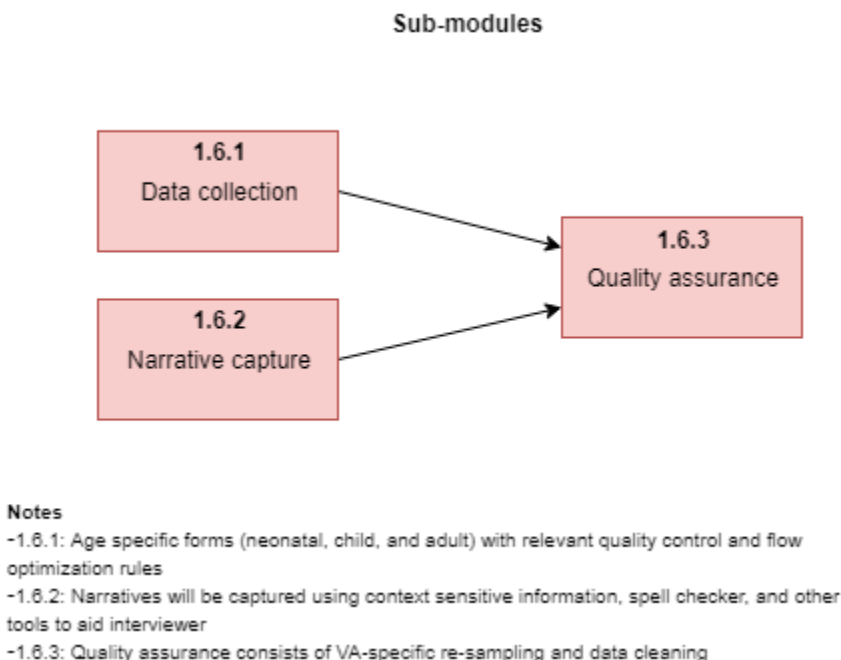
Figure 12: Module 1.6 sub-modules

Sub-modules

1.6.1 Data collection
Through appropriate forms (neonatal, child, adult) based on age using all the relevant quality control and flow optimization rules

1.6.2 Narrative capture
Using appropriate context sensitive information like positive symptoms, treatment history guidelines, and other tools such as spell check, etc.

1.6.3 Quality assurance, resampling, and data cleaning for VA data

Dual physician assignment of COD (Module 1.7)

Physician sub-modules

Abbreviations
COD: cause of death
ICD-10: international classification of disease 10th revision
VA: Verbal autopsy
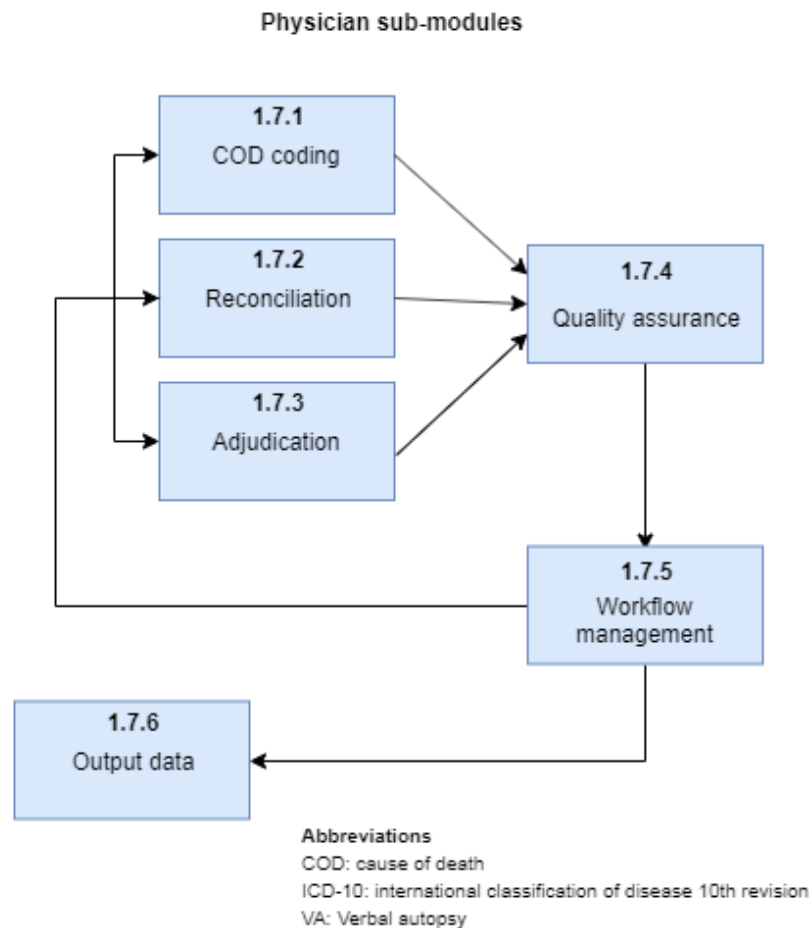
Notes
-1.7.1: Provides summarized VA information & narrative, tools for entering keywords, and context sensitive information to aid physician assign COD (ICD-10 codes)
-1.7.2: Provides COD and keywords from one physician to another to aid reconciliation of COD. Also, provides tools and information from 1.7.1
-1.7.3: Provides COD and keywords from one physician to another to aid reconciliation of COD. Also, provides tools and information from 1.7.1 & 1.7.2
-1.7.4: Provides tools and processes for data quality assurance and exception/error handling
-1.7.5: Provides matching with configurable rules such as ICD-10 equivalence codes

Figure 13: Module 1.7 sub-modules. All these modules integrate to provide the different levels of data (and user) access shown in figure 3 above.

Sub-modules

1.7.1 Coding

This sub-module provides summarized death information, narrative, tools for entering keywords, and context sensitive information for the physician to come out with the cause of death.

1.7.2 Reconciliation

This sub-module provides summarized death information, both physicians' assigned COD and keywords, and context sensitive information for the physician to determine the correct cause of death (ICD-10).

COMSA - Draft DAP and IT plan

1.7.3 Adjudication

This sub-module provides summarized death information, both physicians' assigned COD and keywords at each of the previous stages, and context sensitive information for the physician to determine the correct cause of death (ICD-10).

1.7.4 Data quality assurance and exception handling

This sub-module handles any errors in the ICD codes or those which have carried over from VA. Also, it provides the feature of cancellation of records which have errors.

1.7.5 Workflow management

This sub-module manages the workflow (coding->reconciliation->adjudication) of the coding process including matching of 2 physician codes with configurable rules like equivalence rules for ICD codes.

1.7.6 Output data

This sub-module merges all the data from the various reports and provides consolidated output data for analysis.

Facility-based Deaths System

The facility-based deaths system is a pilot study that will take place at Bo Hospital in Bo District. The facility system will enhance the existing IDSR (and eIDSR) and CRVS systems by adding death-related information starting with a pilot in Bo District focused on higher level facilities only (hospital). If successful, and using funding from outside this project, the system may be extended to all facility-based deaths nationally.

It will involve capturing information on facility and community deaths, assigning cause of death to each. This captured information with the assigned cause of death will serve the purposes of reporting on a regional level of this data via DHIS2 as well as registration and certification of these deaths through the CRVS system.

The data will be validated using the national estimates from the SRS system as well as actual data for the SRS units themselves to identify gaps.

Physician sub-modules

| 2.3 Surveillance reporting | 2.1 Health facility deaths & COD capture | 2.5 Comparison with SL SRS |
| 2.4 CRVS systems | 2.2 Community deaths & COD capture | |

Abbreviations
COD: cause of death
CRVS: Civil Registration and Vital Statistics
IDRS: Integrated Disease Surveillance and Response
MCCD: Medical Certification of Cause of Death
SL SRS: Sierra Leone Sample Registration System

Notes
-2.1: Health facility deaths and COD are from MCCD. Maternal and child extensions are from IDSR
-2.2: CBS and 117 reported cases will be included
-2.3: There will be integration with DHIS2
-2.4: Registration and certification possibly with existing or upcoming systems from NCRA
-2.5: Comparison will be done with SRS SL estimates to identify potential gaps

Figure 14: Bo City health facility deaths & COD capture modules

Modules

2.1 Facility deaths and COD (MCCD) capture

This module will facilitate the capture of deaths at a facility (from facility records, IDSR, and maternal and child extensions of IDSR).

2.2 Community deaths and cause of death

This module will facilitate the capture of actual reported deaths that occurred in the community from health personnel, CRVS, and the 117 reporting system which also covers death reporting.

2.3 Surveillance reporting

This module covers a potential addition to the existing DHIS2 mortality reports (to be explored).

2.4 CRVS systems

This module was originally identified as a means to integrate with CRVS systems for registration and certifications, from the National Civil Registration Authority (NCRA).

2.5 Compare with SL-SRS estimates and identify gaps

This module intends to provide a comparison of death estimates for the community of the SL-SRS system with the data of actual deaths and provide information on the gaps in death reporting (details to be worked out).

COMSA - Draft DAP and IT plan

Integration and Data Reporting System

This sub-module contains common administrative modules which would be required for the system as well as module/external system integration and reports for use by various agencies and the public.

Module integration and data reporting

3.1 User, role management

3.3 Internal integration

3.5 Data access portal

3.2 Authentication & authorization

3.4 External integration

Abbreviations
API: application programming interface
COD: cause of death
DHIS2: District Health Information Software version 2
eIDRS: Integrated Disease Surveillance and Response
NCRA: National Civil and Registration Authority

Notes
-3.1: This also includes linking users to staff
-3.2: Activity logs, single sign on and other key access metrics will be monitored
-3.3: This includes data transfer modules and integration with APIs of other internal modules and sub-modules
-3.4: This integrates with external systems such as DHIS2, eIDSR and NCRA
-3.5: Data & reports are mainly for external users such as non-partner agencies within Sierra Leone, donor agencies, health agencies, etc.

Figure 15: Integration and data reporting modules

Modules

3.1 User/role/ACL management

3.2 Authentication, authorization

Includes activity logs, single sign on, etc.

3.3 Integration features

This module covers features for data transfer modules with the ability to check data, transform data from one format (e.g. from eVA to physician coding or as a WHO format output) to another, and integrate with API of various modules or sub-modules listed in other sections.

3.4 Integration with external systems

COMSA - Draft DAP and IT plan

This module covers both importing and exporting data in suitable formats (e.g. WHO VA format, the related eIDSR, CRVS, 117, etc.) A generic import/export feature will be the default fallback in case of limited automation capabilities.

3.5 Data Access Portal

This module is responsible for making the data accessible to the outside world in the form of aggregated data, i.e. dashboards or datasets (L1 to L4 data of the data access plan). This module will provide facilities to host datasets, and allow registration, request approval, dataset download, and appropriate security and access control.

The users could include the public, researchers, various agencies within Sierra Leona, and various external agencies like donors, health agencies, etc. The availability of data in these reports shall follow a well-defined policy/protocol for sharing data (details in the data access plan).

Data Security Policies

The following data security policies will be applicable to all modules:

1) All server access will be controlled by system administrators.
2) All access to all modules will be authenticated/authorized and be role-based (adhering to standard security policies).
3) Database access is restricted to specific users only (i.e. database administrators).
4) Field data (on laptops) will be cleaned as soon as the collection is completed for a phase.
5) Web-based access will be done on SSL and all standard web-based protection (e.g. cross origin controls, cross site scripting, etc.)
6) Multimedia data (e.g. audio and images) will be stored on secure cloud-based platforms with access granted explicitly to eligible users.
7) Proper backup and restore procedures will be in place in order to protect the data.
8) Development and testing environments will be clearly delineated from the production environment.

References

1) Black, R. E., Cousens, S., Johnson, H. L., Lawn, J. E., Rudan, I., Bassani, D. G., . . . Mathers, C. (2010). Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet, 375(9730), 1969-1987. doi:10.1016/s0140-6736(10)60549-1
2) Jha, P. (2014). Reliable direct measurement of causes of death in low- and middle-income countries. BMC Medicine, 12(1), 19. doi:10.1186/1741-7015-12-19
3) Setel, P. W., Sankoh, O., Rao, C., Velkoff, V. A., Mathers, C., Gonghuan, Y., . . . Lopez, A. D. (2005). Sample registration of vital events with verbal autopsy: a renewed commitment to measuring and monitoring vital statistics. Bull World Health Organ, 83(8), 611-617. doi:/s0042-96862005000800015

Appendices

Appendix 1: Definition of De-identified data (L2)

De-identified data are records that have the following fields removed:

- Names
- All geographic subdivisions smaller than a district (or comparable country-specific subdivisions), including street address, city, county, and their equivalent geocodes
- All elements of dates (except year) for dates directly related to an individual, including birth date and date of death; age categories for individuals are not considered identifying information if such categorization does not violate small cell size restrictions for potential re-identification of individuals
- Telephone numbers
- National/Demographic ID
- Medical record numbers
- Any other unique identifying number, characteristic, or code

Appendix 2: Definition of Identifiable limited individual-level data (L3)

Identifiable limited data are records that have the following fields removed:

- Names
- All geographic subdivisions smaller than a county (or comparable country-specific subdivisions, chiefdoms in Sierra Leone), such as street address, and their equivalent geocodes. (Special requests of L3 data may contain smaller geographic references such city names and postal code areas if approved by CDGT)
- All elements of dates (except year) for dates directly related to an individual, including birth date and date of death; age categories for individuals are not considered identifying information if such categorization does not violate small cell size restrictions for potential re-identification of individuals
- Telephone numbers
- National/Demographic ID
- Medical record numbers
- Any other unique identifying number, characteristic, or code

COMSA - Draft DAP and IT plan

Appendix 3: Data Use Agreements

(Note: these agreements are modified versions of one provided by the INDEPTH Network found here: http://indepth-network.org/data_sharing_policy/annex%203.pdf)

COMSA-SL Data Use Agreement for L2 data

1. Only registered users are allowed to use the data found in COMSA-SL data repositories
2. Data and other material provided by COMSA-SL will not be redistributed or sold by the data user to other individuals, institutions or organisations without COMSA-SL's written permission.
3. The data user will not attempt to re-identify people found and/or mentioned in the data, and she/he will not make use of the identity of any person or establishment discovered inadvertently in the data. Any such discovery will be reported immediately to COMSA-SL.
4. Any books, articles, conference papers, theses, dissertations, reports or other publications employing data obtained from COMSA-SL will cite the source as per the citation requirement provided with the data.
5. An electronic copy of all publications that use the requested data will be sent to COMSA-SL within two weeks after publication.
6. The original collector of the data, COMSA-SL, and the relevant funding agencies bear no responsibility for the data's use or interpretation or inferences based upon it.
7. Means of Acceptance: By checking on the 'I accept' box and 'Submit' button below, the data user acknowledges and accepts the terms of this agreement.

COMSA-SL Data Use Agreement for L3-L4 data

1. Only registered users are allowed to use the data found in COMSA-SL data repositories
2. Data and other material provided by COMSA-SL will not be redistributed or sold by the data user to other individuals, institutions or organisations without COMSA-SL's written permission.
3. The data user will not attempt to make use of the identity of any person or institution found in the data for non-scientific purposes.
4. Any books, articles, conference papers, theses, dissertations, reports or other publications employing data obtained from COMSA-SL will cite the source as per the citation requirement provided with the data.
5. An electronic copy of all publications that use the requested data will be sent to COMSA-SL within two weeks after publication.
6. The original collector of the data, COMSA-SL, and the relevant funding agencies bear no responsibility for the data's use or interpretation or inferences based upon it.
7. Means of Acceptance: By checking on the 'I accept' box and 'Submit' button below, the data user acknowledges and accepts the terms of this agreement.

Appendix 4: COMSA SL Staff and IT support

Staffing and IT support needs (FTEs)

Support would be of two types: helpdesk and technical support to end users. There would also be troubleshooting support for software installed on the machines as well as the server. Software maintenance support is essential until the module stabilizes and for change management thereon. Appropriate contracting and information is essential.

Appendix 5: COMSA SL IT deployment

Description

The surveyors will collect data on their laptops in the field and upload data to the VA server. They will use mobile phones to capture GPS and transmit it to the eVA application using Traccar.

The integration module for VA to data warehouse uploads the collected data to the data warehouse server (after transformation).

The integration module for VA to CME reads the data collected, and transforms and stores them on the CME server. CME server assigns the records to physicians who log in to the CME server and code the records.

Once the records are coded, the integration module for CME to data warehouse uploads the cause of death data to the data warehouse server (after transformation).

The dashboard for monitoring uses data from the eVA and CME systems to generate the performance/quality charts/reports and is accessed by a limited set of users including managers/QA personnel and other entities monitoring progress.

The output data for the data portal is generated from the warehouse and uploaded after anonymization and adding relevant metadata. Researchers and others who need the data will register on the data portal and be able to access the data.
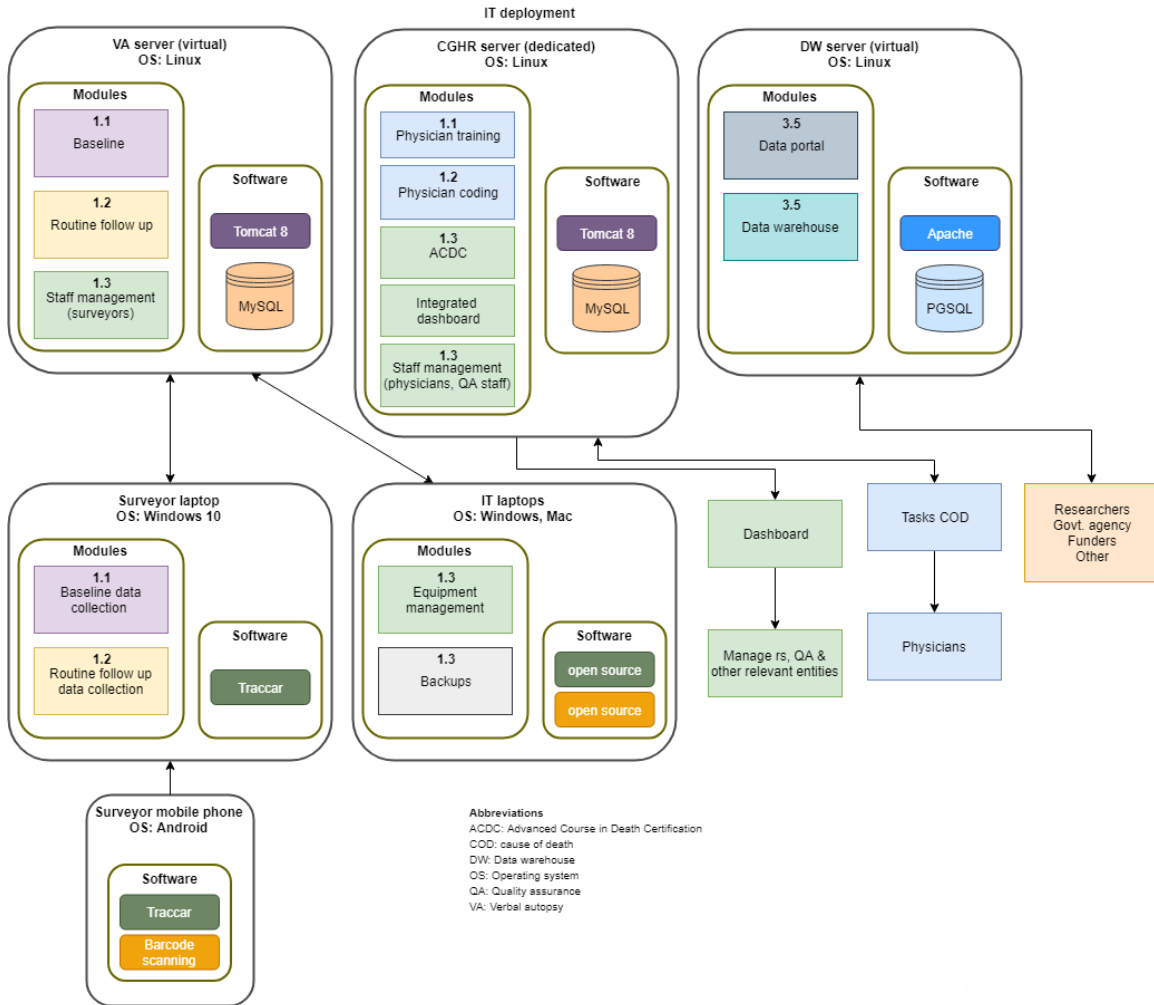
COMSA - Draft DAP and IT plan

Figure 16: COMSA IT deployment

Appendix 6: COMSA IT module status & draft milestones (subject to revision)

Table 2: COMSA's list of modules and sub-modules with milestone reference dates

ID	System	Module/sub-module	Sub-module/Activities	Comments	Status	Date
1	SL-SRS					
1.1		Baseline			Deployed	15-May-19
1.2		Routine follow up		Yearly	Pending	
1.3		Staff and equipment				
1.3.1		Positions		covered in 1.4	Pending	
1.3.2		Recruitment			Pending	
1.3.3		Onboarding			Deployed	15-May-19
			Physician training system		Deployed	
			eVA training	Manual process		
			MCCD training		Deployed	
1.3.4		Performance Assessment and feedback	Narrative review		Deployed	26-Sep-19
			Surveyor performance dashboard	Dashboards for the monitoring staff deployed. Individual dashboard pending	Deployed	
			Physician performance dashboard	Individual physician dashboard deployed. Comparative reports available	Deployed	
1.3.5		Refresher/Corrective training Planning and conducting			Pending	
1.3.6		Contract management (exits, renewals, equipment, vendor)			In process	
1.3.7		Tracking assets			Pending	
			Inventory		In process	
			Assignment		In process	
			Repair		Pending	
			Replacement		Pending	
1.4		Survey monitoring	COMSA Dashboards	Part of dashboards	Deployed	

COMSA - Draft DAP and IT plan

1.4.1.1		Monitor Baseline listing progress			Deployed	15-May-19
1.4.1.2		Monitor Routine follow up progress			Pending	
1.4.1.3		Monitor Verbal autopsy progress			Deployed	15-Jun-19
1.4.1.4		Monitor Physician assignment of COD progress	Will have to be ported to the new integrated dashboard		Deployed	
1.4.1.5		Monitor Positions				
			Health surveyor vacancies/recruitment/Training status	Current positions status deployed Future Requirements pending	Deployed	15-Jul-19
			Physician vacancies , training status		Deployed	15-May-19
1.4.2		Monitor staff performance				
			Health surveyor		Deployed	
			Physician	To be upgraded to be in included in the integrated dashboard	Deployed	
1.5		Payment system		Will have to be customised	Deployed	
1.6		Verbal autopsy (VA)			Deployed	
1.7		Physician module		Dependent on integration of VA to CME (2.6) for deployment	Deployed	
2.	Facility-based deaths					
2.1		MCCD coding			Pending	
2.2		117 community death		General csv inputs, direct integration (with warehouse/) Status uncertain	Pending	
2.3		Integration with DHIS2 (Surveillance)		General csv outputs, from data collected from MCCD and 2.2) status uncertain	Pending	
2.4		CRVS integration		General csv outputs /inputs, direct integration with warehouse. NCRA CRVS system status uncertain	Pending	

COMSA - Draft DAP and IT plan

2.5		Compare with SRS SL estimates and identify gaps (details to be worked out)		Uncertain. To go through some analysis to the data warehouse and should be reported in a dashboard	Pending	
3.	Integration & data reporting					
3.1		User/role/ACL management and linking users to staff			Deployed	15-May-19
3.2		Authentication, authorization, activity logs, single sign on etc.			Deployed	
3.3		Internal integration between modules (ETL)			In process	15-Nov-19
			VA to CME		In process	
			VA to data warehouse/data portal		In process	
			CME to data warehouse/data portal		In process	
			VA to Machine learning		In process	
			Machine learning outputs to data dashboards/portal		In process	
3.4		External integration module	csv export/import module	ETL (to support 2.2, 2.3, 2.4)	In process	15-Nov-19
3.5		Data access portal			Pending	15-Nov-19

--END --

COMSA - Draft DAP and IT plan