## 

## **Recency Data Manager II**

**Report to:** Health Informatics Systems Analyst

**Location:**  Uganda

**Period: April 1, 2020 – September 30, 2021** *(renewal subject to availability of funds)*

**The Company**

The University of California, San Francisco (UCSF) through its affiliate international NGO Global Programs for Research and Training (GP) is engaged in projects in Uganda, Kenya, Burundi, Cambodia, Malawi, Namibia, Vietnam, Thailand, and Zambia (to be referred to as Recency Countries) funded by the U.S. Centers for Disease Control and Prevention (CDC). UCSF’s role is to provide technical assistance to Ministries of Health and its partners, including multiple comprehensive IPs, to roll out the use of a rapid test for recent infection (RTRI) for HIV surveillance and public health response. As part of UCSF’s and Recency Countries’ surveillance and survey activities, a lot of data is collected, managed, stored, and analyzed at various levels. This quality of the data and trust in results is heavily dependent on fidelity to research protocols and procedures through Continuous Quality Improvement (CQI) activities. Quarterly reviews of recency testing sites are performed to confirm that there is adequate staffing, materials, and posted procedures to ensure that recency is carried out methodically and purposefully. These site visits produce a lot of data which itself needs to be cleaned, processed into an analytical format, and ultimately visualized on PowerBI. Consequently, there is need to hire a CQI Recency Data Manager across all Recency Countries to coordinate and solve data quality, consistency, and validity issues that they actively manage, analyze and visualize.

**The Position**

The Recency Data Manager will primarily be responsible for maintaining several in-country SQL-based databases for CQI datasets from recency studies, as well as related datasets (public/programmatic/survey), for purposes of ensuring fidelity to implementation and study protocols. He/She will be responsible for correcting data using approved methods, extracting/transforming/loading using SQL Server (ETL) of data into an analysis-ready format, and ultimately using PowerBI to visualize data. Additionally, he/she will work closely with the UCSF CQI team to maintain, clean, and code Excel-based data containing root cause analyses and action plans. He/She will also be responsible for providing custom datasets upon request in a custom format (e.g. as an Excel file). The Data Manager is expected to work closely with Informatics Developers, Programmatic/MOH staff, and UCSF employees. The Data Manager will also coordinate with CQI Officers from various countries to manage CQI data from supervision visits to facilities to ensure, for example, that the RTRIs are being administered properly, informed consent is obtained, and that all documentation is completed correctly.

Specifically, data management encompasses the following dimensions of which the candidate applying should have adequate knowledge and demonstrated application:

* **Analysis and visualization:** Approved methods for cleaning and transforming data to enable analysis, visualization, and use at various levels (MOH, stakeholders, etc.)
* **Data management:** Collection, management, integration, analysis, visualizing and reporting, as well as supporting data use and interpretation of data in line with proper data management procedures and best practices using the SQL language.
* **Appropriate use:** Ensuring privacy and confidentiality through the construction of minimum use data sets and guidance on appropriate statistical practices that achieves the objectives of an internal/external data user
* **Data security:** Verifying data users as well as the implementation and documentation of appropriate administrative, technical, and physical safeguards for all data (along with appropriate penalties)
* **Data quality:** Ensuring data standardization, verification, and validation procedures are implemented and documented to ensure data accuracy, availability, timeliness, completeness, credibility, and solid understanding of its limitations
* **Public disclosure:** Documenting all public data requests and ensuring appropriate communication between data stewards and data users for arranging data disclosure prioritization and procedures

**Roles and Responsibilities:**

Analysis and Visualization

* Cleaning and ETL of data using SQL scripts (stored procedures, functions, etc.) periodically
* Qualitative coding of datasets from various countries’ CQI activities
* Transforming data into data warehousing schema to enable analysis and visualizations
* Use of PowerBI to display data on graphs, maps, and line-lists

Documentation

* Creation of documentation and a solid understanding of all data flows (from source to analysis and use) for all CQI-related data
* Maintenance and understanding of all SQL-based documentation which details data cleaning, standardization, validation, and analytical procedures
* Maintenance and ongoing analysis of data quality as well as tracking of all 'bugs'
* Documenting and cataloging all data elements (data dictionaries and metadata) for the various datasets as well as the relationships between them (e.g. ERD's for data linkages)
* Documenting all changes to variables and data structures (i.e. change management)
* Development of all policies and procedures for requesting and approving data access for internal and external data use partners
* Active maintenance of logs which contain information on all users which have access to certain data sets as elements, as well as intended use (e.g. analysis, publications, etc.)
* Documenting minimal use dataset creation criteria and parameters

Communication and Coordination

* Ensuring that policies and procedures for data access are adhered to by internal staff and external partners
* Coordinating between appropriate project staff and data stewards to determine data request prioritization and minimal dataset development timelines
* Coordinating with the informatics team to ensure proper prioritization of minimal dataset development and fixing data quality issues

Communicate Performance Management

* Communicate appropriate vs. inappropriate uses of datasets
* Preparing publications on data

Technical Assistance

* Ensuring appropriate use through guiding analytical approaches and ensuring that correct conclusions are drawn from source data
* Creation of minimal use datasets to meet external data user needs, requiring an understanding of SQL and procedures/functions
* Developing and evaluating electronic survey tools (e.g. ODK) to ensure data quality and consistency

Technical Documentation

* Ensuring that technical guidelines for anonymization of datasets are followed for external data use while still maintaining data linkages across appropriate datasets

**The Person**

**Minimum requirements**

* A Bachelor’s degree or Master’s degree in Data Science, Data Management, Computer Science, Software Engineering or a related field with sufficient training in database related disciplines
* At least five years of working experience managing active large study/surveillance datasets
* Demonstrable experience with at least one of the following: Oracle, SQL Server, MySql, PostgreSQL
* Fluent in complex data manipulation using SQL, especially more advanced scenarios
* Demonstrable advanced expertise in writing complex SQL queries and statements.
* Demonstrable experience with data visualization using at least one of the following platforms: PowerBI, Tableau, Qlickview
* Experience working in the health sector, public health, and/or PEPFAR programs
* Demonstrated good communication and interpersonal skills, especially working remotely
* Ability to coordinate with multiple partners in various countries to carry out duties as assigned, especially working remotely

**Considered an Advantage**

* Certification in one or more popular database management systems (DBA in Oracle, SQL Server or similar)
* Experience with quality improvement activities such as PDSA cycles, root cause analyses, and monitoring and evaluation of programs *highly* preferred
* Background in statistics and/or data analysis
* Experience with at least one statistical analysis software package: R, SAS, STATA, SPSS
* Experience with international health/research/data projects

**Interested?**

We offer an attractive package commensurate with qualifications and experience. The closing date is **31 March 2021**. Email your cv and application letter to **info.uganda@ucglobalprograms.org**. **Only electronic CV’s will be accepted**.