

Establishing a data hub for COVID-19 data in Kenya and Malawi

An outbreak of COVID-19 and its variant has been spreading rapidly across the world since it first emerged in December 2019. Currently, the pandemic presents a serious threat to many people, but especially those in low and middle-income countries (LMICs). Over the past year, the outbreak has caused the ill-health and loss of millions of lives, and significant disruptions in global socio-economic activity. Even though a number of vaccines have been developed and are under distribution across the globe, dissemination has been slow, particularly in the developing world.

The pandemic has put governments on the spot, compelling them into action within a limited policy space. The length and severity of the impacts of the virus depend on the projected length and location(s) of each wave, as well as whether there is a concerted, fast-tracked response. In the Global South, health systems that were already weak have been negatively affected as health facilities and healthcare workers cater to the ill with limited resources.

Owing to the uncertainty surrounding COVID-19, it is imperative that governments and non-state actors work together to develop systems that provide for proper epidemiological planning and management especially where health systems are weakest, and hence populations most vulnerable.

It is for this reason that the African Population and Health Research Center (APHRC), in collaboration with partners from Malawi (Malawi Polytechnic), UK (The ALPHA network of the London School of Hygiene & Tropical Medicine - LSHTM), South Africa (South African Population Research Infrastructure Network - SAPRIN), and France (Committee on Data of the International Science Council - CODATA), have come together to establish a data hub for COVID-19 data in Kenya and Malawi.

Objectives

The main objective of this project is to develop a scalable, coordinated COVID-19 data ecosystem in Kenya and Malawi, which can be up-scaled to the regional level. The approaches will be informed by artificial intelligence (AI) and data science (DS), using multidisciplinary research for an inclusive COVID-19 response. The project will strengthen the capacity to share data and information, and to inform policy implementation in Kenya and Malawi, two low- and middle-income countries (LMICs).

The study's specific objectives include:

- To develop a cloud-based OMOP ([Observational Medical Outcomes Partnership](#)) common data model to collect and collate diverse types of disaggregated COVID-19 data from multiple data sources for sharing with the scientific community for use and re-use.
- To mine and predict the evolution of the COVID-19 pandemic using AI and data science tools and algorithms.
- To strengthen capacity of key stakeholders to enhance data use and analytical capacity of technical staff involved in the data production pipeline (supply, production and use).
- To improve policy engagement and learning to support advocacy and decision support tools for technical stakeholders in government policy making, political leadership and health policy planners.

Approach

Building the structure for bringing together open, FAIR data from African health systems and African health researchers requires addressing both technological and cultural issues. Dubbed 'the INSPIRE Platform for Evaluation and Analysis of COVID-19 Harmonized data (PEACH)', the study will harness a

robust suite of data standards and technologies, diverse data integration methodologies, the power of AI and data science and a trusted governance and policy environment to produce a comprehensive data hub for COVID-19 data. This will follow a multi-pronged approach as follows:

1. *Development of a governance framework for data integration*
2. *Use of a Common Data Model (CDM)* based on epidemiological questions from core COVID-19 questionnaires and clinical data reported from health facilities, this model will be used to store harmonized COVID-19 data for this project. The COVID-19 data hub will be informed by the needs of the primary end-users (policymakers and researchers) and the availability of the data from producers.
3. *Development of data management protocols* for bringing the partner data into the CDM and for automatic data integration of the relevant epidemiological surveillance and clinical data.
4. *Use of analysis workbenches for COVID-19 data*, one that incorporates the use of artificial intelligence and machine learning to identify vulnerable populations, inequalities in access to treatment and the severity of the impact of COVID-19 within Kenya and Malawi. The second workbench will use causal analysis to provide a set of interrelated recommendations to policy makers on public health, food security and the economy over time.

At a Glance

Partners

- The African Population and Health Research Center (APHRC)
- Malawi Polytechnic
- The ALPHA network of the London School of Hygiene & Tropical Medicine
- South African Population Research Infrastructure Network (SAPRIN)
- Committee on Data of the International Science Council (CODATA)

Donors

International Development Research Centre (IDRC)

Duration

Two years

Target beneficiaries

- Direct beneficiaries: policymakers from the relevant ministries, departments and agencies
- Indirect beneficiaries: media, healthcare providers, the citizens of Malawi and Kenya

Geographical focus

Kenya and Malawi