

Data Documentation Template

1 Study Description

1.1 Identification

1.1.1 Title

Contains the full authoritative title of the data collection. A full title should indicate the geographic scope of the data collection as well as the time period covered.

m-Africa: Building mobile phone-connected diagnostics and online care pathways for optimal delivery of population HIV testing, prevention and care in decentralised settings

1.2 Overview

1.2.1 Abstract

An unformatted summary describing the purpose, nature, and scope of the data collection, special characteristics of its contents, major subject areas covered, and what questions the PIs attempted to answer when they conducted the study. A listing of major variables in the study is important here.

In this study we are using mixed methods studies to evaluate the feasibility and acceptability of implementing mobile-phone connected HIV diagnostic tests and online clinical care pathways within current healthcare system in KwaZulu-Natal, South Africa. In parallel we are developing and piloting new mobile phone connected diagnostic tests for use in community settings. These tests use an application linked to a mobile camera to interpret the results. The performance of the tests will be benchmarked to gold-standard diagnostic tests (Enzyme Linked ImmunoSorbent Assays (ELISA) and Polymerase Chain Reaction (PCR)) using patient samples (n=100).

1.2.2 Kind of Data

The type of data included in the file: survey data, census/enumeration data, aggregate data, clinical data, event/transaction data, program source code, machine-readable text, administrative records data, experimental data, psychological test, textual data, coded textual, coded documents, time budget diaries, observation data/ratings, process-produced data, etc. No description just a single phrase, e.g. Genetic sequences

Images of POC HIV tests to train and evaluate a smartphone application to automate interpretation of commercial HIV tests and biobarcode tests.

1.2.3 Unit of Analysis

Basic unit(s) of analysis or observation that the study describes: For PANGEA is each record a sequence from a specimen, or are there multiple records for a single specimen or study participant

Photographs of individual test strip responses.

1.3 Scope

1.3.1 Topics Classification

MeSH subject headings

[HIV-1](#)

[Incidence](#)

[Phylogeny](#)

[Epidemics](#)

[Population Surveillance](#)

[Rural Population](#)

[HIV Infections](#)

[Africa](#)

1.3.2 Keywords

Keywords summarize the content or subject matter of the survey. As topic classifications, these are used to facilitate referencing and searches in electronic survey catalogues.

HIV rapid test, Point-of-care test, Lateral Flow test

1.4 Coverage

1.4.1 Country

Indicates the country or countries covered in the file

South Africa

1.4.2 Geographic Coverage

Information on the geographic coverage of the data. Include the total geographic scope of the data, and any additional levels of geographic coding provided in the variables.

Demographic surveillance area of the Africa Health Research Institute.

1.4.3 Universe

A description of the population covered by the data in the file; the group of persons or other elements that are the object of the study and to which the study results refer. Age, nationality, and residence commonly help to delineate a given universe, but any of a number of factors may be involved, such as age limits, sex, marital status, race, ethnic group, etc. The universe may consist of elements other than persons, such as specimen, sample or isolate. In general, it should be possible to tell from the description of the universe whether a given individual or element (hypothetical or real) is a member of the population under study. Also known as universe of interest, population of interest, and target population.

??

1.5 Producers and Sponsors

1.5.1 Investigators

The person, corporate body, or agency responsible for the data collection's substantive and intellectual content. Repeat the element for each author and use the affiliation attribute if available. Invert first and last name and use commas.

Remarks: The author in this element should be the individual(s) or organization(s) directly responsible for the intellectual content of the data collection.

Name	Affiliation
Herbst, Kobus McKendry, Rachel Turbe, Valerian	AHRI University College London University College London

1.5.2 Funding

The source(s) of funds for production of the data collection. If different funding agencies sponsored different stages of the production process, use the role attribute to distinguish them. Also includes a field for the grant/contract number of the project that sponsored the data collection effort.

Agency	Abbreviation	Grant number	Role
Medical Research Council (UK)	MRC	M-AFRICA MR/P024378/1	

--	--	--	--

1.5.3 Acknowledgements

Statements of responsibility not recorded in the title and statement of responsibility areas. Indicate here the persons or bodies connected with the work, or significant persons or bodies connected with previous editions and not already named in the description. For example, the name of the person who cleaned the data collection might be cited here, using the role and affiliation attributes. Does not include funders.

Name	Affiliation	Role

1.6 Sampling

1.6.1 Sampling Procedure

The type of sample and sample design used to select the survey respondents to represent the population. May include reference to the target sample size and the sampling fraction

Survey data collection will be integrated into the routine AHRI population data collection platform, and will not require additional household visits. Qualitative data will be collected by an experienced qualitative research support team.

1.7 Data Collection

1.7.1 Dates of Collection

Contains the date(s) when the data were collected/produced.

August 2017 to present.

1.7.2 Notes on Data Collection/Production

Used to describe noteworthy aspects of the data collection/production.

Images of lateral flow tests were collected by AHRI fieldworkers at participant's home, upon participant's consent, using the AHRI work tablets (Samsung Galaxy Tab A10).

1.8 Data Processing

1.8.1 Other Processing

Used to indicate additional information about the methodology and processing involved in the production of the dataset.

The raw images have been stripped of the GPS coordinates attached in the metadata, and renamed to pseudo-anonymise the dataset.

1.9 Data Access

We include data access statements in our publications, addressing interested researchers to contact our group for access to data sharing privileges. The study investigators will identify a data custodian that will approve data requests on behalf of the study or refer requests to the study steering committee if the nature of the request is such that sensitive data need to be shared. Interested researchers will be required to submit a proposal to use any data for independent analyses. Applicants will be asked to state their analysis objectives, to demonstrate they have the capacity to conduct the intended investigations, and to show that their team has sufficient understanding of the local study context to be able to interpret the results appropriately. Previously we have made raw experimental data available upon request.

1.10 Contacts

1.10.1 Contact persons

Individuals listed as contact persons will be used as resource persons regarding problems or questions raised by the user community. The URI attribute should be used to indicate a URN or URL for the homepage of the contact individual. The email attribute is used to indicate an email address for the contact individual.

Name	Affiliation	Email	URI

2 File Description

2.1 Data Files

2.1.1 Contents

Abstract or description of the file. A summary describing the purpose, nature, and scope of the data file, special characteristics of its contents, major subject areas covered, and what questions the PIs attempted to answer when they created the file. No need to repeat information already contained in the abstract in paragraph 1.2.1.

jpeg pictures of lateral flow tests, captured at the time of interpretation.

3 Variable Description

A code book of the variables in the data file.

Name	Definition	Data Type and Codes