



PROPOSAL OF A RESEARCH PROJECT

GENERAL INFORMATION

Project title:

Leveraging Artificial Intelligence and Data Science Techniques in Harmonizing, Accessing and Analysing SARS-COV-2/COVID-19 Data in Rwanda

Short title: Leveraging Artificial Intelligence in SARS COVID 19 Data Analysis in Rwanda (LAISDAR Project).

Project duration (in months): **24**

Countries in which research will take place: **Rwanda**

PROJECT LEADER

Name: **Dr Francine Birungi MD, MPH, PhD**

Mailing address:

Building, Suite (optional): **School of Public Health Buildings, Kicukiro**

Number and Street or PO Box: **KG 11 Ave (Street),**

District, City: **Gasabo, Kigali**

Province/State: **Kigali Town**

Country: **Rwanda**

Postal Code: **PO Box: 3286, Kigali**

Telephone number (with country and area codes): **+250788 462 574**

Email address: fbirungi@nursph.org or birufrance@yahoo.fr

Mailing address:

PROPOSING INSTITUTION

The **proposing institution** normally receives and administers the funds. If this institution has not received IDRC funds before, the Institutional Profile Questionnaire must be completed. If the funds are to be administered by a third party, please contact IDRC to discuss the arrangements.

Legal Name: **University of Rwanda**

Mailing address:

Building, Suite (optional): **Remera Campus Buildings**

Number and Street or PO Box: **KK 737 Street, Gikondo**

District, City: **Kigali**

Province/State:

Country: **RWANDA**

Postal Code: **P.O. Box 4285**

Telephone number (include country and area codes): **+250 789001218**

Facsimile number (include country and area codes):

E-mail address: **vc@ur.ac.rw**

COLLABORATING INSTITUTIONS (if applicable)

Collaborating institutions are those that are jointly proposing research to IDRC in collaboration with the proposing institution and **will each receive funds** directly from IDRC. In this case, each collaborating institution must complete a separate grant application (Section C need only be completed by the proposing institution that will coordinate the project). In addition, the institution coordinating the project must complete a consolidated budget form (see page 17).

Not Applicable

PARTICIPATING INSTITUTIONS (if applicable)

Participating institutions are those that will participate directly in the conduct of the research but **will not receive funds** directly from IDRC. In this case, the proposing institution is responsible for disbursement of funds and for ensuring that all institutions abide by the standard terms and conditions that apply to the IDRC grant.

1. University of Rwanda, Rwanda:
 - 1.1 College of Medicine and Health Sciences (CMHS), Rwanda
 - 1.2 African Centre of Excellence for Data Science (ACE-DS), College of Business and Economics, Rwanda
 - 1.3 Centre of Excellence in Biomedical Engineering and e-Health (CEBE), College of Science and Technology, Rwanda
 - 1.4 Centre of Gender Studies, College of Arts and Social Sciences (CASS) at UR
2. Rwanda Biomedical Center (RBC), Ministry of Health, Rwanda
3. Regional Alliance for Sustainable Development (RASD), Rwanda
4. Ghent University (UGent), Belgium

PARALLEL FUNDS (Summary)

Parallel funds are contributions that will be made by other **international funding agencies (or foundations)** to the project. Details must be provided in Section D, page 15.

Donor	Amount	Currency
-------	--------	----------

Not applicable

ABSTRACT OF RESEARCH PROJECT (should not exceed 250 words)

This proposal aims to leverage Artificial Intelligence and Data Science Techniques to create a scalable framework for gathering, harmonizing, and accessing available SARS-CoV-2 / COVID-19 data in Rwanda, with a common interface for data querying, analysis and reporting on top of the data hub. The project will gather all existing fragmented data and will collect new enriched longitudinal data. We propose making optimal use of existing (open source) solutions, standards and tools offered by the Observational Health Data Sciences and Informatics (OHDSI) community especially the OHDSI Common Data Model (CDM) through the Observational Medical Outcomes Partnership (OMOP) initiative. This largely used OMOP CDM will help in pooling data from multiple existing cohorts and to standardize data elements creating a ready-to-use dataset. The supported data harmonization framework will help making the inventory and mapping of needed data elements/variables from each data (static and longitudinal) to OMOP CDM. After harmonizing, this project will create a common data portal interface for access, querying and analytics including OHDSI tools and machine learning techniques on top of the data hub portal. Within this framework all data will be securely accessed in federated manner, meaning that no data will leave its owner location; which will comply with data privacy regulations. This project will ultimately generate data to help in the identification of key factors influencing the susceptibility to infection and clinical manifestation, to assess the optimized therapeutic and clinical management options, and provide evidence-based recommendations for health policies in preventive strategies, protective actions, and disease management.

DETAILS OF RESEARCH PROJECT

Please attach your proposal, which must be based on the following guidelines. Proposals should not exceed 20 single-spaced pages (excluding appendices).

RESEARCH PROBLEM AND JUSTIFICATION: archive.idrc.ca

Since March 14th, 2020 when the first case of COVID-19 was confirmed in Rwanda, there have been an increasing cases numbers all over the country, in first phase due to trans-borders populations and currently spreading into the community. As of September 8th 2020, 447603 people have been tested in the country confirming 4439 positive cases and the total deaths are currently 20 on a total Rwandan population of 11 Million . However, as the tests are made only in high risk groups(known contacts with positive cases), the real number of positives cases (mild to asymptomatic) is assumed to be higher. The government institutions, policymakers and researcher are encouraging all initiatives to increase data, and they are pushing for more data analysis to know more about the real COVID situations and the trends. There is need for accurate data on the prevalence, incidence and evolution of the disease. Strategies to prevent and control COVID-19 and mitigate its effects have been initiated in at different levels in Rwanda. However, these strategies need continuous adjustment as the characteristics and the dynamics of the virus are progressively discovered. This project comes to add new knowledge on the dynamics of COVID-19 in Rwanda by highlighting the trends and its characteristics.

This proposal is targeting the COVID-AI call specifically on theme 4 “Strengthening data systems and information sharing about COVID-19”. The SARS-COV-2/COVID-19 data has the potential to transform our disease understanding and advance science but also to understand outcomes which enable efficient preventive or treatment measure. However, in Rwanda like in other countries this data is currently fragmented, incomplete and scattered across multiple institutions including hospitals, clinics and testing sites that have captured vast amounts of data on the disease. Analysing those fragmented COVID-19, datasets brings poor evidence. Pooling all those datasets together in one single dataset is challenging as they have different data structure and data owner may fear break in data privacy. Therefore, we need an innovative approach to analysed all data together, and this is the added value of this proposal.

Moreover, some tested patients were referred to testing sites from healthcare facilities and therefore they have there a usual patients’ healthcare records or Electronic Health Record (EHRs), that can help to understand underlying comorbidities and risk factors. In Rwanda all those data are stored in separate databases and different formats, and this drastically limits its usefulness both for research and benefit patients and the community. The current project is proposing to leverage Artificial Intelligence (AI) and other Data Science (DS) techniques to create a scalable framework for inventorying, harmonizing and federating the accumulated data from COVID-19 patients and converting it to a standardized data format so that it can be used as part of wider studies on the disease. Each dataset will be mapped to a common data model, already in use for other observational studies thanks to the Observational Health Data Sciences and Informatics (OHDSI) community especially the OHDSI Common Data Model (CDM) through the Observational Medical Outcomes Partnership (OMOP) initiative. The data will remain under complete control of the original data owner, thereby ensuring ethical and local data privacy rules are respected. The harmonized data will include not only COVID-19 diagnosed/serotyped but also non-infected individuals as they will come from normal hospitals electronic health records (EHRs) or testing databases with positive and negative results. In second stage, this project will collect new data in a longitudinal way. Those new longitudinal data will be enriched with patient reported outcome (PROs) and will be in the same standardized model by design. The surveys will be conducted through mobile application questionnaires completed by direct phone calls and face-to-face surveys. anonymous geofencing data will be collected as well. The project outcome will be leveraging all federated data with Machine Learning (ML) and other mathematical methods to drive evidences. Those evidence fulfil the government of Rwanda priorities and need in predicting and monitoring the burden of COVID-19 in the Rwandan community, on hospital admissions and overall infection rates and monitor the impact of various public health measures on the pandemic evolution in the country. Finally, the proposed approach is scalable by extending the list of new datasets or updating the existing one and all data will remain available for future usage. The same approach is also applicable for other diseases and pandemics like Ebola virus, Influenza, and others.

OBJECTIVES:

The general objective

Leveraging Artificial Intelligence and Data Science Techniques in Harmonizing, Accessing and Analysing SARS-COV-2/COVID-19 Data in Rwanda

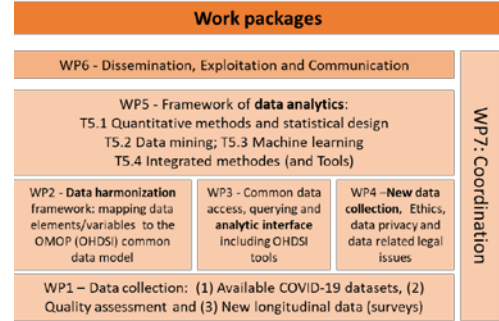
The specific objectives

1. To inventory all existing collected data on COVID-19 in Rwanda, assess its quality, periodicity, and readiness towards the common data model.
2. To create the framework for data harmonization to the OMOP Common data model (OMOP CDM): We will start by mapping full hospital patients records, focusing on 8 hospitals located in regions with high number of COVID-19 patients and completing with other isolated datasets.
3. To collect OMOP CDM based prospective enriched data on COVID-19 from the community through mobile surveys applications, face-to-face validation survey, and potentially other sources if available like geofencing data.
4. To build a common data query interface with analytical tools including OHDSI open source tools and ML.
5. To leverage both traditional mathematical modelling techniques, statistical methods and machine learning methods for prediction models for the burden of COVID-19 in the community but also the potential impact on hospital admissions or overall infection rates and the impact of various public health measures on 1)the pandemic evolution in the country; 2)on social-economic situation, and on 3)mental health (stratified by gender and other vulnerable groups).

METHODOLOGY:

Conceptual and theoretical framework: Work plan and work packages
 The project work plan will be organized into 7 different work packages (WPs):

Figure 1. Work packages of the project



Conceptual and theoretical framework: Technical aspects

The current proposal methodology will have 4 steps:

Step 1: Data gathering /collection (applying existing methodology)

That includes the inventory of the existing data from the first 6 months of the pandemic in Rwanda (first case was identified in march 2020) and the 1-year data collection (through Mobile Applications surveys, telephone calls and face-to-face).

Expected data sources will be in different formats. Ranging from Covid-19 related data registered in Excel documents, via data sources containing Minimum Clinical Data (MCD) in DHIS2¹ and other systems, to more granular Electronic Medical Record (EMR) data in Open Clinic², OpenMRS³ and other EMR systems. We will start by mapping full hospital patients records, focusing on 15 hospitals located in regions with high number of COVID-19 patients and completing with other isolated datasets. The list of hospital will be defined at the start of the project as data on COVID-19 are still increasing. In total the study will target use of 1Terabytes (1000 Gigabytes) size data which size is big enough for such AI modelling.

The new data collection will follow validated guidelines/principles⁴ in terms of data collection and will be done by a longitudinal approach using mobile App questionnaires: A minimum of 214 people per administrative district (6.420 persons throughout Rwanda totalizing 154.080 survey-entries over 24 weeks) will be required for mobile App responses weekly for 6 months (24 weeks). A minimum sample of 10 persons per district will be reached out by the data collector (2 times: at the beginning of the study and the end) via validation phone call or face-to-face questionnaire if the COVID-19 situations in Rwanda allows. A sub-group of patients cured from COVID-19 will be specifically followed. If a followed subject has a medical file in participating hospitals, the two datasets will be linked with possibilities of linkage data request in future.

The respondents will be randomly sampled from national population registry of each district, thanks to National Institute of Statistics (NISR) Authority. The sample will proportionally include males and females based on number of inhabitants. There's a risk of having not sufficient numbers of respondents and/or they don't report regularly, that's why a team of data collectors will call subjects once a week to complete the missing data and to enhance the response rate. Each participant will receive mobile fee connection and internet bundle each week to allow data collection. To mitigate the expected gap of the gender digital divide but also of selected persons without a mobile phone anymore, the consortium establish mitigation measures including-but not limited to, leveraging the community healthcare workers (CHWs). Each village in Rwanda have a CHW who is participating into various ministry of health (MoH) programs and they have all received the mobile phones from MoH. If we select a respondent without a mobile phone we will liaise with nearer CHW to reach out to him. We included into the budget the service pay to connect the involved CHWs. The other measures will be specified and tested into the sampling plans and practical data collection plans which will be developed at the beginning of the project.

The questionnaires (which will be translated in 3 languages, Kinyarwanda, English and French in Mobile application) include 10 modules (at least 8 of them has to be fulfilled by the project): 1)Demographics; 2)Face mask use; 3)Hand hygiene; 4) Respect of social distancing measures and risk minimization measures; 5) Recent risk situations exposures and COVID-19 measures. On the outcome side, the collected data will include 6) Coronavirus like-Signs and symptoms; 7)Mental health indicators (based on General anxiety disorder-GAD); 8) Social economic impact (based on loss of income, or categories); 9) Covid-19 test results; and 10) if available the geofencing data (no personal data to be collected): Only the Ethical committee approved anonymous phone tracking enabled at individual device on voluntary basis. The sampling plans and practical data collection plans will be developed at the beginning of the project. The sampling and data collection plans will help to overcome biases especially integrate gender dimension to deal with gender digital divide gap, known worldwide but also in Rwanda.



Figure 2-The Modules of new data to be longitudinally collected

Step 2: Infrastructure for data harmonizing (developing novel techniques)

For data harmonization the custom designed ETL scripts will be developed per data source to extract, transform and load the source data to an OMOP CDM database instance. In early stages when the hospital EHRs are not yet harmonized, we will also use synthetic data approaches to help automate harmonization processes.

¹ <https://www.dhis2.org/> (Accessed on June 5th, 2020)

² <http://openclinic.sourceforge.net/>(Accessed on June 5th, 2020)

³ <https://openmrs.org/how-openmrs-can-help-countries-flatten-the-covid-19-curve/>(Accessed on June 5th, 2020)

⁴ <https://www.reachresourcecentre.info/toolkit/data-collection-processing/> (Accessed on June 5th, 2020)

METHODOLOGY CONT'D:

The data owner-side infrastructure will include the OMOP CDM database instance, the Arachne client, the OHDSI Atlas¹ analytical tool, R Studio², and Jupyter¹. The data harmonization process converts the observational data from the format of the source data system to the OMOP Common Data Model (OMOP CDM), the CDM supported by the Observational Health Data Sciences and Informatics (OHDSI) organization. This project will benefit from consortium members (lead by the UGent with *Edence Health NV* company experts support) in the steps involved in the data harmonization process, typically:

- **Mapping workshop:** this a face-to-face (in person or via video conference) workshop, usually a full day, where the initial mapping from source data to OMOP CDM is discussed in detail.
- **Structure mapping + final mapping doc:** Based on the mapping workshop, documentation and notes, the structure mapping is finalized and documented in the mapping document. This forms the basis of the ETL design.
- **Code mapping:** depending on which source terminologies are used in the data source, mapping the local codes to the standard vocabularies used in OMOP CDM (LOINC, SNOMED, RxNorm, etc.) can be either a short, easy process or a long, involved one with multiple iterations.
- **ETL² implementation:** the ETL script(s) to transform the source data into the OMOP CDM database instance; normally done in Python.
- **ETL testing:** the ETL scripts are tested both on development data, and ideally also on the data source's test data.
- **ETL deployment:** once the ETL scripts will be tested successfully, and packed and deployed using GitHub and Docker.

The data harmonization process will differ quite substantially for different data sources.

In terms of architecture design, we propose the following conceptual framework (Figure 3):

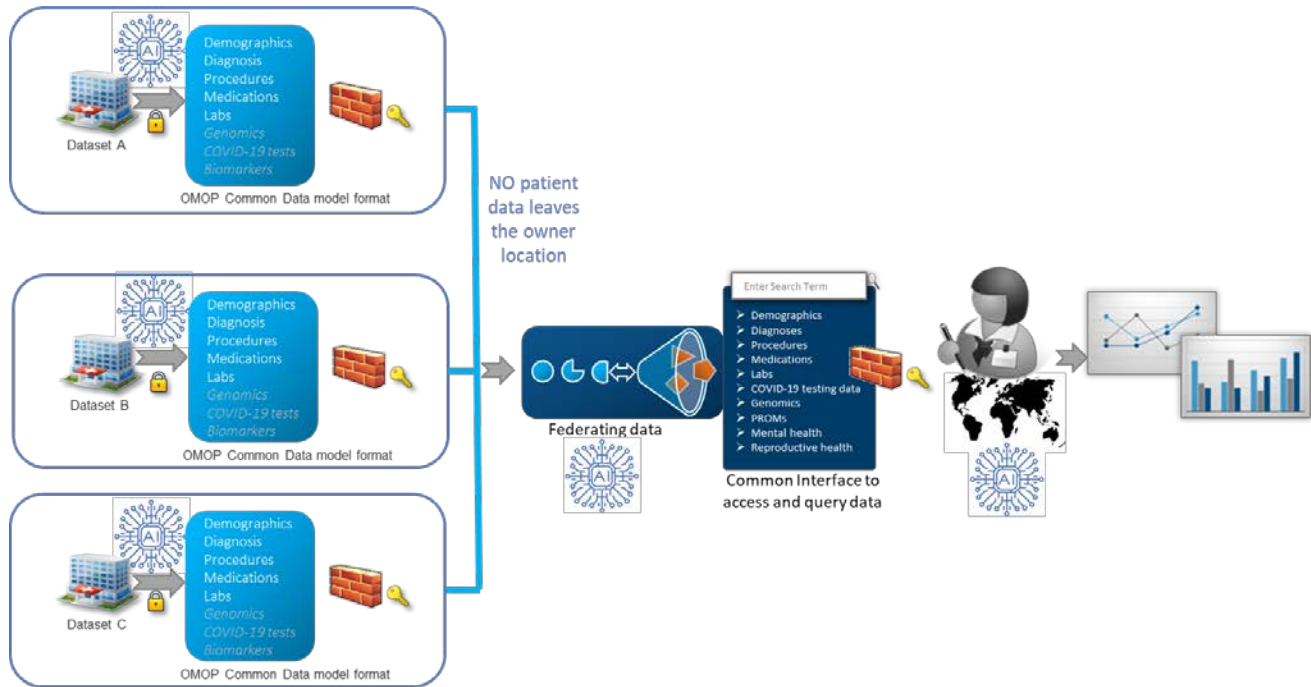


Figure 3. High-level conceptual framework of the project

Step 3: Infrastructure for data access, query, and data analysis (Mixing existing methods and innovative techniques)

The central platform data access, query, and data analysis, or central site setup, will manage and coordinate the studies that will be performed across the participating data sources. The central site should at a minimum consist of a database, and Atlas instance, a catalogue of data sources, an R Studio instance, and possibly also a Jupyter server instance.

Depending on the network infrastructure chosen (see above), there may also be an installation of the Arachne central server. The database, for example a PostgreSQL database, will include an OMOP CDM schema, as well as additional schema(s) to support a central data catalogue and study coordination.

- There are new techniques with regards to the creation of synthetic data and using data to help automate harmonization processes and training models: This approaches will be also used in our project from early beginning when the harmonized data from hospitals EHRs are not yet available, specially leveraging the OHDSI community available mock up data (like Synthea³ – etc) to train different algorithms /models, before we use them on real data.

¹ <https://www.ohdsi.org/atlas-a-unified-interface-for-the-ohdsi-tools/> (Accessed on June 5th, June 2020)

² <https://rstudio.com/> (Accessed on June 5th, June 2020)

- The OMOP CDM schema will have the same OMOP CDM vocabulary version as the participating sites and will allow studies to be prepared and tested. If needed, a synthetic data set (e.g. Synthea) or available local data set can be loaded.
- There will also be result schemas that will be able to hold the Achilles output per data source site – this will allow a central view on the descriptive statistics for each site.
- The database will also be the place to gather aggregated results from the data source sites as part of defined studies.

The OHDSI Atlas instance is integrated with the PostgreSQL database (in use, open source). The central Atlas instance will, as mentioned above, allow cohort definitions and studies to be prepared, and to view descriptive statistics for each participating site. The R Studio and Jupyter instances will allow development and testing of R scripts as part of a study design, or to analyze data collected from data source sites as part of studies. The Arachne central server setup will allow central management of network studies, with tight integration with the OHDSI tools such as Atlas.

Step 4: Data analysis and interpretation (Mixing existing methods and innovative techniques)

The federated datasets are challenging to analyse with traditional statistical methods, because they are, like other real-world-data (RWD), 1) collected without any intention for being used in research; 2) incomplete and not cleaned and 3) collected in sporadic way, not pure longitudinal approach so no way to derive cohort-like data from them.

The current project will leverage the AI techniques including is Machine learning techniques and data mining that bring an added value in discovering hidden patterns or relationships between data points.

The Machine learning model consists of two modules: GRU-ODE⁴, responsible for learning the continuous dynamics of the latent process that generates the observations and GRU-Bayes, responsible for dealing with incoming observations and update the conditional current estimate of the latent process. Those two steps and modules are similar in essence to the propagation and update steps of a Kalman filter. With GRU-ODE, we are able to project in time the hidden process $h(t)$ and hence indirectly future observations. GRU-Bayes perform the update of the hidden state conditioned on new observations. Yet, unlike a Kalman filter, this approach allows to learn very complex dynamics for the latent process.

The Figure 3 below show the overall architecture that we propose to support this project. The design incorporates the following parts: Central platform: includes a data catalogue describing the different data sources, the Arachne⁵ central hub, a central OHDSI Atlas instance, a central database, as well as R Studio and Jupyter.

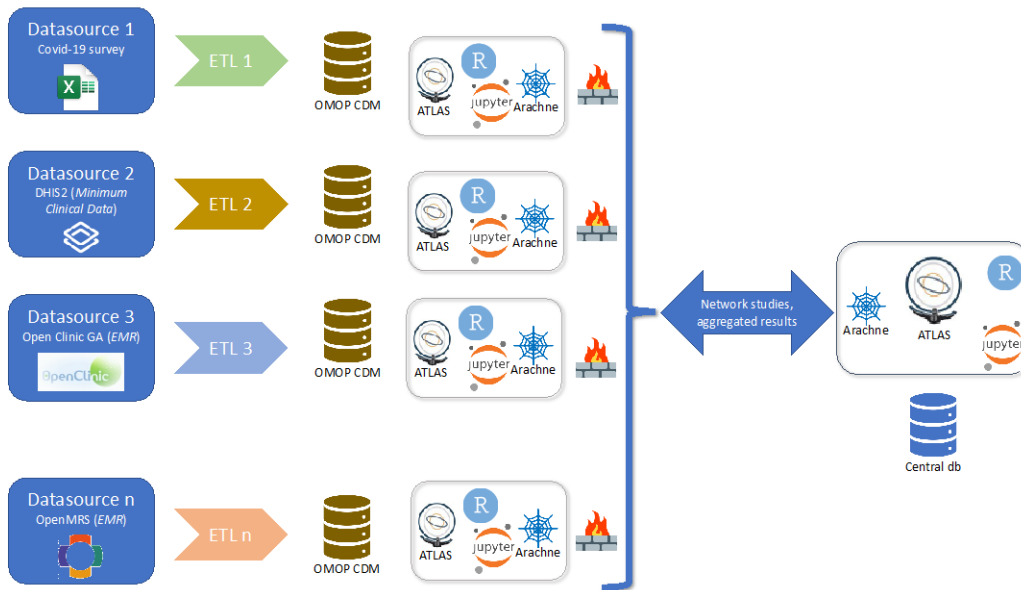


Figure 4. High-level view of the architecture supporting the data analysis interface

¹ <https://jupyter.org/> (Accessed on June 5th, June 2020)

² **ETL** is short for extract, transform, load, three database functions that are combined into one tool to pull data out of one database and place it into another database. Extract is the process of reading data from a database.

³ <https://synthetichealth.github.io/synthea/>

⁴ <https://arxiv.org/abs/1905.12374>

⁵ <http://www.dotaceu.cz/getmedia/4c8c06e9-53e5-47c0-b18d-2c41c682e42f/Arachne-User->(Accessed on June 5th, June 2020)

User participation

The project will fall under University Research policy where all research projects are monitored by Directorate of Research & Innovation at College level, and reporting to Deputy Vice-chancellor in charge of Academic Affairs and Research (UR/DVC AAR). For this purpose no other public officials nor advisory board is foreseen.

Gender considerations.

The virus does not discriminate. In order to respond effectively to the crisis, we need a whole-society approach to understand its differential impact on women and men. Supporting gender analysis and sex-disaggregated data is an integral part of this project. The gender COVID-19 related data are still scarce and a little is known in Rwanda on the topic. Therefore collecting related data will be a key activity to bridge the gap and contribute to best gender driven policies locally and in the region. Specifically we articulated the gender considerations in this project in a multi-lever approach including:

- Fostering gender balance throughout in research teams, in order to close the gaps in the participation of women (Figure 5). Our research consortium has a core principle to respect the minimum of 30% female participation, and is also achieved in this proposal submitting team. The consortium brings together a team of senior investigators and managers with a significant participation of female experts in their teams. Thanks to Rwandan government policy on gender (known worldwide), our project will integrate all measures to integrate gender at all research stages.

Furthermore, this project will follow the UR Gender Policy and its vision “to create a University that promotes a diverse staff, student and stakeholder community, in which all people are valued, respected and treated equally and equitably in terms of gender”. We will follow the seven guiding principles and objectives; and the complementing strategies to implement our message of gender equity through our interactions with UR-CMHS students and staff, health staff and community members. Concretely, this project will keep gender equity as a target by: 1) Using the University of Rwanda Tool Kit and checklist on Gender Equality during the whole project¹; and 2) Striving to achieve gender equity in project team, program committee, participants (included in indicators)

- Integrating the gender dimension in research and innovation (R&I) content that will help improve the scientific quality and societal relevance of the produced knowledge, technology and/or innovation. Gender has been integrated as a transversal theme and not a vertical aspects. Gender facets are found in all COVID-19 consequences including morbidity in general and mental health problems in particular and socio-economic outcome.

Social and cultural factors related to gender such as specific considerations for some collected data elements will be addressed as well, eg. reproductive health data, the usage of gender-sensitive research questions and gender-impartial language.

Moreover, the sampling will pay special emphasis to gender proportional balance while collecting new COVID-19 data and gender key output/aspects will be driven from data analysis. Last-but not the least, this project having EU partners, will also integrate guidelines from the "Toolkit Gender in EU-Funded Research" of the European Commission as general guideline for research (Figure 5).

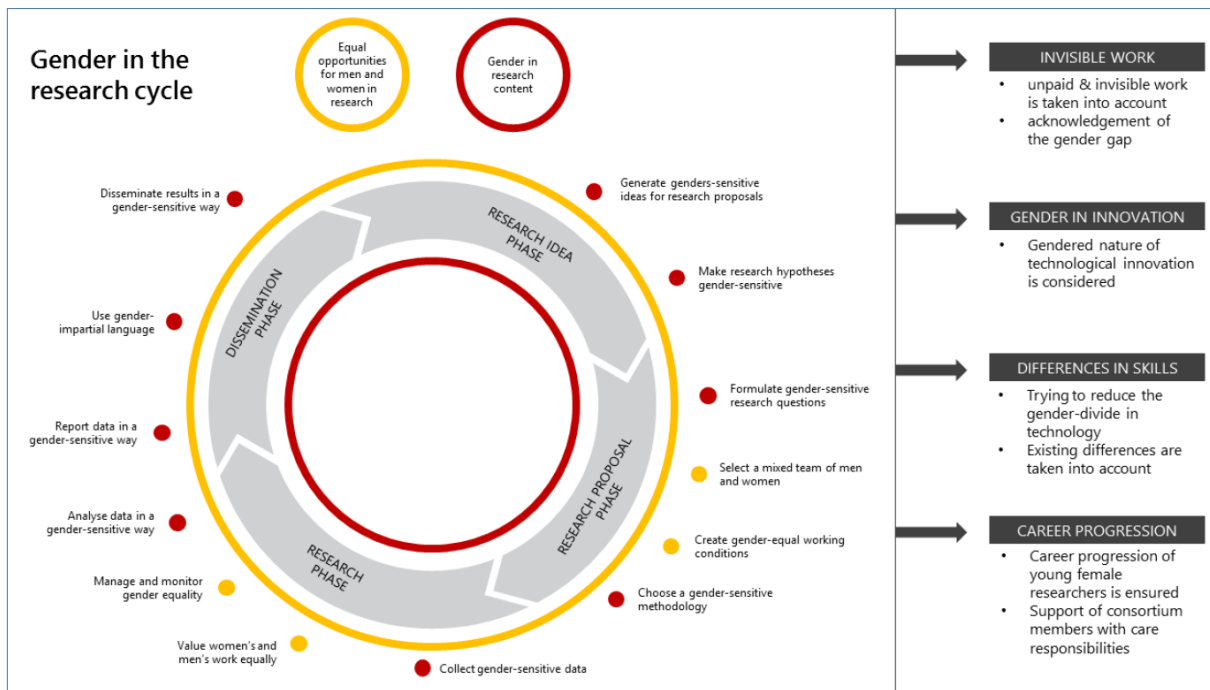


Figure 5. Gender Analysis in this research project¹

¹ <https://ur.ac.rw/documents/UR%20Gender%20Policy.compressed.pdf> (Accessed on September 8th 2020)

Concretely, this project will keep gender equity as a target by 1) Using the University of Rwanda Tool Kit and checklist on Gender Equality during the whole project and 2) Striving to achieve gender equity in project team, program committee, participants (enrolled subjects) Finally, at a general scope, the consortium agrees to tackle gender issues both on a content and on a structural level of the project. Within the research content, e.g. through the usage of gender-sensitive research questions and gender-impartial language, and in terms of opportunities for men and women, e.g. in creating gender-equal working conditions and valuing women's and men's work equally.

Ethical considerations, a responsible approach (Rights-based, Inclusive, Ethical and Sustainable) to research and implementation

This consortium is include senior researchers seasoned on matters of rights-based, inclusive, ethical and sustainable research thanks to experience from previous research cooperation financed by Sweden where standards procedures were implemented but also from ongoing Sida cooperation with University of Rwanda. The consortium, through series of consortium meetings and work packages leader, will ensure the values and principles underpinning the UN human rights conventions; non-discrimination, participation, accountability and transparency. All measures addressing power imbalances within and between institutions and programs, ensuring transparent, inclusive and ethical research processes, enhancing good governance of research institutions will be included in 2 documents which will be drafted and signed by all research team at the beginning of research implementation: The "Standards of procedures (SoP)" and "Best practices procedure (BPP)". Moreover the research teams will discuss along the research implementation the data security and ethical implications which the consortium will address explicitly, robustly, ethically, in full compliance with Rwandan legislation. On top this project will follow also guidance from "Specific Ethics Considerations during Health Emergencies" notes for applicants to the IDRC call for proposals. We have dedicated the whole WP4 to "Ethics, data privacy and data related legal issues". Within this work package, and at very early stages (Month 1) the consortium will interact with relevant Ethical committee or Institutional Review Board (IRB) related to each dataset to assess the requirement of a secondary use of the data in an anonymized and secured way. The WP4 will also use the data collection to assess the acceptability of the using AI in Rwanda (through participants satisfactory questions), and the analysis will contribute to a broader understanding of responsible AI research and implementation in Rwanda. As no guidelines exist in Rwanda, the international guidelines will be used as basis for this WP4 including the Recommendation on Artificial Intelligence (AI) – the first intergovernmental standard on AI – was adopted by the OECD Council at Ministerial level on 22 May 2019 on the proposal of the Committee on Digital Economy Policy (CDEP)².

Moreover on this, there be a final retrospective or analysis summarizing key findings and recommendations for similar approaches in the future in Rwanda but also in African regions with similar settings.

All data will be anonymized, and no data will leave its original location. For new data to be collected within this project, Ethical clearance will be applied. This project will support construction of a state-of-the-art technology platform, interfaces and tools to enable queries to be conducted on a network of COVID-19 cohort data sources in order to derive such patient numbers. These data sources are then supposed to already exist, and each have their own consent or other legal basis for holding data on patients. The responsibility of this project is to verify that its intended processing of the data is compatible with each data source's legal basis, terms of reference and other local policies including local ethics approvals if needed. This project will therefore collect a set of data access agreements with these data sources after verifying this legal and ethical compatibility, and then enable a remote connection to them.

Regarding the participants in the prospective data collection part, they will be requested to sign the consent form before enrolling in the study during the very first call by research surveyors.

The participant will be informed that the participation is voluntary and that they can withdraw from the study at any time during the study period. The benefits to participate in the study include but not limited to: By continuing to respond to survey questions, a participant may acquire more knowledge about COVID-19 preventive measures, and they can be role models in their implementation. If they have any depression symptoms, they may be relieved to see that scientists care and even by talking to the phone with data collectors maybe a kind of therapy.

There are no documented risks to participate in this study but long telephone usage every week with a relatively big survey questionnaire pose a minimal risk worth mentioning. There are no incentives planned for participating in this study but due to the fact that the participants will need internet connection to respond to survey questions, the participants will be given 500 RWF to purchase internet bundles.

The data governance ecosystem

Given the focus on sensitive clinical data, it is important to govern data adequately and ensure appropriate management of the data. Therefore the consortium will specifically address the data governance matters, from the sources of data, their integration and use ensuring suitable privacy protection and information governance.

No patient data will be shared even anonymized. As the IDRC embraces the principle of sharing research data and encourages researchers to make their data openly available, the researchers will be able to access data where they get only the aggregated data (no data download). Each individual will need to register and request access to the whole or a part of data available from the common analytical interface. He or She will sign data access agreement, limited by the research project duration. The same procedures methodologies are in use at similar datasets access request such as the Clinical Practice Research Datalink (CPRD)³ in UK or the European Surveillance System (TESSy) - ECDC⁴. Research findings will also be made accessible to research participants through a login credentials, but also findings will be disseminated to participants through various media.

Capacity building through trainings.

This project will contribute to research but also to capacity building by training staff training before and during the project both at UR and at participating hospitals. The training foreseen (see also the budget lines) include: 1) data mapping infrastructure; 2) Training on Surveys instruments and 3) Training on Sensitive patient data handling, Data harmonization, interoperability and medical terminology: A team from Ghent University (Belgium) will train the Rwandan locals on OHDSI OMOP CDM mappings including terminology and coding.

¹ "Toolkit Gender in EU-Funded Research" of the European Commission, Directorate General for Research and Innovation together with Yellow Window Management Consultants, Chapter 2.1 "How to Make Research Gender-Sensitive", 2011.

² <https://oecd.ai/ai-principles> (Accessed on September 8th 2020)

³ <https://www.cprd.com/Data-access> (accessed 25 June 2020)

⁴ <https://www.ecdc.europa.eu/en/publications-data/european-surveillance-system-tessy> (accessed 11 June 2020)

Organizational matters:

This proposal gathers 6 partners together into a collaborative research where each one brings specific expertise. The uniqueness of the consortium stands in gathering together both academics experts in data sciences and e-Health, academic centres of excellence, government agency (Ministry of Health/ Rwanda biomedical centre), hospitals, healthcare professionals, private SMEs (RASD Rwanda) active in innovation in data management/sharing and community outreach.

The project will adopt a governance model that will promote the active participation of all involved stakeholders, including both academia and health facilities, national COVID-19 testing sites and other interested in using data to generate evidences. The consortium will be a challenging project to manage due to its size, its ambition, the variety of activities and their interdependencies. For this purpose, we foresee setting up a dedicated project management office (PMO), and a permanent research officer.

The consortium will establish a Work Package Leaders (WPLs) monitoring meetings to evaluate the project progress and to ensure alignment, scientific coherence and coordination across WPs. This will have the decision power on all high level design, budget re-allocation and other major issues. It's led by the Project Promoter Prof Marc Twagirumukiza, who also remain the overall scientific coordinator.

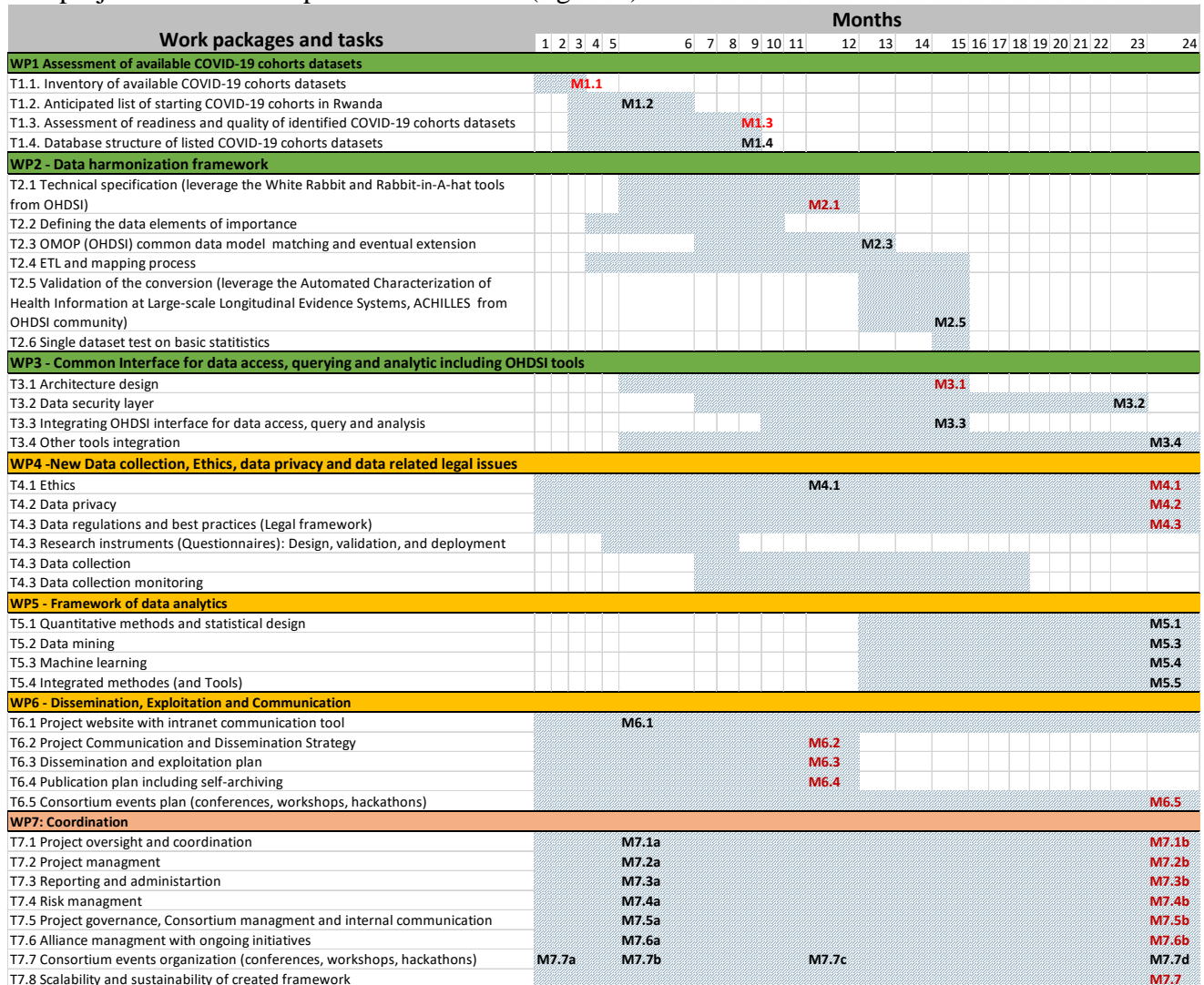
Table 1. Project WPs Leaders

#WP	Work Package Title	Lead
1	Data Collection	MoH/RBC
2	Data harmonization framework	UGent* KUL
3	Common Interface for data access, querying and analytic including AI techniques and OHDSI tools	UGent*
4	Ethics, data privacy and data related legal issues	UR/CEBE
5	Framework of data analytics	UR/ACE-DS
6	Dissemination, Exploitation and Communication	RASD Rwanda
7	Coordination	UR/CMHS

With sub-grant to **Edence Health N.V (Belgium) company as IT and technical infrastructures expert*

PROJECT SCHEDULE:

The project activities are planned as follows (figure 4):



**Only main milestones are displayed for image clarity proposes.*

Figure 6. Gantt diagram of project activity timelines

Table 2. List of milestones

Milestone number	Milestone name	Related work package(s) (See the description in the timelines table)	Due date (in month)	Means of monitoring and verification
MS1	1 st MPIs bi-monthly meeting (Virtual)	WP1	1	List of invited and confirmed members to the 1st Meeting and Meeting's Minutes
MS2	1 st Consortium meeting (Virtual)	WP1, WP2, WP3, WP4, WP5, WP6, WP7	1	List of invited and confirmed members to the 1st Meeting and Meeting's Minutes
MS3	Establishment of PMO in UR	WP7	3	PMO staff and their contracts + Physical verifications (photos)
MS4	1 st version of available COVID-19 datasets	WP1	6	Report
MS5	1 st hackathon on Database structure of listed COVID-19 datasets	WP1	6	Report
MS6	Validated Survey questionnaires	WP1	6	Report
MS7	2 nd Consortium meeting	WP1, WP2, WP3, WP4, WP5, WP6, WP7	6	List of invited and confirmed members to the 1st Meeting and Meeting's Minutes
MS8	1 st round of data collection launched	WP3	7	List of participants
MS9	1 st hackathon leveraging the White Rabbit and Rabbit-in-A-hat tools from OHDSI on initial datasets	WP2, WP3, WP4	8	List of participants
MS10	1 st hackathon defining the data elements of importance	WP2, WP3, WP4	9	List of participants
MS11	1 st hackathon defining OMOP (OHDSI) common data model-matching and eventual extension	WP2, WP3, WP4	9	List of participants
MS12	1 st hackathon ETL and mapping process	WP2, WP3, WP4	9	List of participants
MS13	Full ETL and mapping process finished	WP1, WP2, WP3, WP4, WP5, WP6, WP7	12	Agreement
MS14	Single dataset test on basic statistics	WP2, WP3, WP4	12	Report
MS15	6 round of new data collection closed	WP4	12	Report
MS16	3 rd Consortium meeting	WP1, WP2, WP3, WP4, WP5, WP6, WP7	12	List of invited and confirmed members to the 1st Meeting and Meeting's Minutes
MS17	4 th Consortium meeting	WP1, WP2, WP3, WP4, WP5, WP6, WP7	16	List of invited and confirmed members to the 1st Meeting and Meeting's Minutes
MS18	Project sustainability plan	WP1, WP2, WP3, WP4, WP5, WP6, WP7	20	Report
MS19	Thematic framework of validations analytics: Paper submitted (Immunity assessment and COVID-19)	WP5, WP6, WP7	24	Manuscript
MS20	Project closing meeting	WP1, WP2, WP3, WP4, WP5, WP6, WP7	24	List of invited and confirmed members to the 1st Meeting and Meeting's Minutes

Expected outputs and outcomes of the research

The output of this project is double: 1) A harmonized, shared, accessible, and ready for analysis data, federated within a start-of-the-art framework. 2) Data driven evidences leveraging machine learning and data mining techniques. This project outcome will be **an impact on science, on healthcare delivery and on government policy regarding COVID-19 in Rwanda**. The gathered data can be, in scalable and sustainable way, be analyzed and 1) provide evidence-based recommendations for health policies in preventive strategies, protective actions, and disease management and 2) provide a live monitoring dashboard for the burden of COVID-19 in the community but also the potential impact on hospital admissions and overall infection rates. Finally, this work will contribute to predicting the impact of various public health measures on the pandemic evolution in the country. A **knowledge update, outcome use and impact plan** will be developed along the project implementation within WP6. On top of this timeline activities, the consortium set up 20 milestones for the project overall progress monitoring (Table 2)

In terms of quantitative anticipated outputs: 1) Eight hospitals data will be mapped to OMOP CDM then harmonized; 2) Eleven COVID-19 datasets will be federated; 3) one new dataset from new longitudinal data collection will be available; 4) A central analytical interface will be developed where all harmonized dataset and hospital EMRs can be accessed; 5) At least 6 thematic publications will be produced

Scaling

The proposed approach is scalable by extending the list of new datasets or updating the existing one and all data will remain available for future usage. Currently this proposal will start with 8 hospitals (selected in area where there's a high number of COVID-19 patients) and 12 separate datasets. Adding an additional dataset connection into the system will require a minimum work as the infrastructure will be in place already. The same is true if in the near future an additional hospital wants to join the platform of data sharing. The same approach is also applicable for other diseases and epidemics like Ebola virus, Influenza, and others. This is a crucial contribution to the healthcare and science in our region plagued by emerging epidemics (Ebola recently, Covid-19 currently, etc.) The main challenge here will be the skills and capacity building at different institutions. Also this is a new innovative area and culturally we expect some misunderstanding, fear of loosing ownership on their data, etc. To mitigate this the capacity building activities and key stakeholder education programs at major target institutions will ensure use, uptake and sustainability. Finally, the federated database and the whole infrastructure will stay in place and maintained by UR SPH team. Moreover the consortium will continue to leverage the infrastructures for other indications with different grants -currently working on a pilot project in non-communicable diseases.

Study limitations

With the current COVID19 data, our study has one limitation of not including high risks groups that were severely affected by the pandemic in Rwanda. These clusters are prisons, schools, and transit centers. The AI data analysis will always take this into consideration. To avoid selection bias, there is no acceptable way to include pre-identified clusters even though the risk is known. We are aware that by asking the questions on mental health, some participants may want to seek care after the answering the questions and have the general score on depression score. The study has added a key information at the end of the questionnaire to inform the participant that "in case of need of medical help they will call 114, an MoH toll free number to orient the patient for the closest health facility". Another limitation derive to the risk of having not sufficient numbers of respondents and/or they don't report regularly, that's why a team of data collectors will call subjects once a week to complete the missing data and to enhance the response rate. Each participant will receive mobile fee connection and internet bundle each week to allow data collection. To mitigate the expected gap of the gender digital divide but also of selected persons without a mobile phone anymore, the consortium establish mitigation measures including-but not limited to, leveraging the community healthcare workers (CHWs).

RESULTS AND DISSEMINATION:

The dissemination and exploitation activities are crucial here and the whole work package is dedicated for this (WPs 6). A detailed dissemination strategy, will be prepared in WP6, focusing on the promotion of project results in order to attract the target audience, raise their awareness and engage them in the project activities. The aim of the exploitation will be to ensure the optimal use of this project's results after its completion, and therefore speed up the potential of their uptake by health practitioners, researchers, regulators and policy makers. This research project consortium includes Rwanda Biomedical Centre, a technical body of ministry of health in Rwanda, which will ensure that the project outcome contribute to the government of Rwanda priorities and needs.

Collaboration with other networks and initiatives is considered of utmost importance. Specific activities will be carried out within WP6 to ensure permanent interaction and coordination with external initiatives. Networking activities with other projects and initiatives to promote synergies of efforts and mutual leverage of results will be carried out. The strategy for exploitation of results is at the very core of our proposal.

This will be initiated in WP4 on 3 outcome analysis and will leverage similar efforts happening in OHDSI community¹. An example of this project outcome exploitation and use would be a revisit of the previous analysis done on fragmented data on hydroxychloroquine². To further promote alignment with emerging cohorts Several workshops will be organised by the project, open to external stakeholders, initiatives etc. To ensure that dissemination and exploitation of the project results align throughout the lifecycle of the project, activities will be divided into three overarching strategic areas, which, as outlined below, are designed to accord with more detailed phased-activity for both dissemination and exploitation respectively.

In terms of ML and AI analytic techniques, the project be adapting existing approaches and testing novel approaches suitable for such short periodicity data. If a novel approach or methodology is found robust it will be disseminated through a publication as an open source. Other outputs (to be elaborated as dissemination plan) including the lessons learned reports on data governance (legal, ethical and social), on gender considerations will be share to facilitate future project and smooth scaling and uptake.

¹ An international characterisation of patients hospitalised with COVID-19 and a comparison with those previously hospitalised with influenza. Source: www.medrxiv.org/content/10.1101/2020.04.22.20074336v1 (accessed 25-June-2020)

² Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study. Source: <https://www.medrxiv.org/content/10.1101/2020.04.08.20054551v2> (accessed 25-June-2020)

INSTITUTIONS AND PERSONNEL:
PREVIOUS RELEVANT PROJECTS AND ACTIVITIES INLINE TO THIS PROJECT

1. SmartHealthcare (SI-2020-01-50, 01.01.2020 – 31.12.2021), A pilot project on digitalizing the hospital patient trajectory and secondary use of Electronic Health Records (EHRs) in Rwanda (**Running** -Participant: UGent & UR/CEBE & RASD Rwanda)
2. Africa Rapid RT-RPA-Based Test with PCR Detection For The Severe Acute Respiratory Syndrome Coronavirus2 (Africa RRT-RPATSARS Cov2) (**Running** -Participant: RBC)
3. Within the University of Rwanda-Sweden Programme for Research, Higher Education and Institutional Advancement, funded by Sida (Swedish International Development Cooperation Agency) Prof. Jean Bosco Gahutu was team leader for the sub-programme of "Medicine and Health Sciences" in the 2008-2013 and 2013-2019 agreements; Prof Jean Bosco Gahutu is the team leader for the sub-programme of "Infectious Diseases" in the 2019-2024 agreement; output 2008-2020: 11 PhD graduates. (**Running** -Participant: UR)
4. NIH D43 grant (D43TW010335) under the PI/PD Mutimura Eugene, whose main aim is to leverage previous partnerships with academic and public health experts to create and strengthen scientific leadership and expertise in Rwanda for clinical research in human immunodeficiency virus (HIV)- and antiretroviral therapy (ART)- associated cardiovascular and metabolic disease. (**Running** -Participant: RASD Rwanda)
5. NIH funded U24 grant (U24HL136990-02) with an objective to build in-country capacity to conduct late-stage (T4) translation research (TR) for hypertension (HTN) control in Rwanda. (**Running** -Participant: RASD Rwanda)
6. eStandards (H2020 643889, 01.05.2015 - 30.04.2017), eHealth Standards & Profiles in Action for Europe and beyond (**Closed**-Participant: UGent)
7. VALUeHEALTH (H2020 643847, 01.04.2015 - 31.03.2017), Business model for sustainable eHealth services in Europe (**Closed**- Participant: UGent)
8. SemanticHealthNet (FP7 288408; 01.12.2011 - 31.05.2015), Semantic Interoperability for Health Network (**Closed**- Participant: UGent)
9. SALUS (FP7 287800; 01.02.2012 – 30.04.2015), Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies (**Closed**- Participant: Prof Marc Twagirimukiza)
10. EHR4CR (IMI-JU 115189; 01.03.2011 – 28.02.2015), Electronic Health Records for Clinical Research (**Closed**- Participant: UGent)
11. EMIF (IMI-JU 115372; 01.01.2013 – 31.12.2018), European Medical Informatics Framework (**Closed**- Participant: UGent)
12. EHR4CR (IMI-JU 115189; 01.03.2011 – 28.02.2015), Electronic Health Records for Clinical Research (**Closed**- Participant: UGent)

PARTNERS
1. University of Rwanda (UR)
Description of the Partner

UR (<https://ur.ac.rw>) is the sole public Higher learning Institution in the Rwanda that was established by the government of Rwanda through the law no 71/2013 of 10th September 2013. The University of Rwanda resulted from merging the former seven (7) Higher Learning Institutions namely; National University of Rwanda (NUR), Kigali Institute of Science and Technology (KIST), Kigali Health Institute (KHI), Institute of Agriculture and Animal Husbandry (ISAE), Umutara University (UP), School of Finance and Banking (SFB) and Kigali Institute of Education (KIE).

The mission of UR is to deliver quality education and develop innovative teaching and research meant to address the problems of the population, the students, the nation, the region and the world. University of Rwanda has substantial expertise in assessing and monitoring data. Gathering, harmonizing, and accessing all available SARS-CoV-2/ COVID-19 cohorts are among the topmost challenging problems. UR with collaboration from other partners will work to contribute to develop skills and methods to more accurately estimate the burden of these problems and track their course. From UR, 3 institutions will participate in this project: The CMHS, ACE-DS and CEBE;

a) UR/College of medicine and health sciences (CMHS)

The CMHS The College of Medicine and Health Sciences is part of the University of Rwanda; whose formation was put in law on October 15th 2013. The College, also known as UR-CMHS has a central role to play in the social and economic development of the nation through training of medical doctors and health professionals. In support of this, the College engages in capacity building relevant to this cause by spearheading health training initiatives and capacity building, through robust and relevant research programs

Main tasks

The UR/CMHS is the leading coordinator of the project but also in charge of 1) Strategic design of the research project (i.e. designing the project so that it leads to useful findings in the health sector); and 2) Clinical and public health implications of research findings (i.e. CMHS schools team will support the interpretation of research findings and contribute to elucidating their usefulness in COVID-19 prevention and control).

Key personnel

Dr. Francine Birungi, MD, MPH, PhD (F) Lecturer, Postgraduate Coordinator and Head of the Department of Epidemiology and Biostatistics at School of Public Health. Her research area focuses on Infection diseases, especially in on Childhood TB and TB-Non-Communicable Diseases co-morbidities. Also, she is interested in research focus on Maternal and Child Health. She has been involved in teaching, research as principal investigator or co-investigator and grants writing since 2009.

Dr Stefan Janssen, PhD (M) is the Deputy Director Ag., Center for Mental Health, University of Rwanda and his research focus on Global Mental Health, with a particular interest for translating successful local practices and solutions into an academic framework. He also performs research in the area of Knowledge-Attitude-Behaviour Change through interventions in road safety, radio dramas and more. In addition, he is active as a neuroscientist, setting up the first fMRI neuroscience experiments to be conducted in Rwanda. He will be assisting the project coordinator.

Dr. Annette Uwineza, MD, PhD(F) Senior Lecturer of Medical Genetics, Head of the Department of Clinical Biology, School of Medicine and Pharmacy. Her research largely focuses on identification of genetic aetiologies of intellectual disability in Rwandan. She acquired knowledge on bioinformatics allowing analyzing data of next generation sequencing; designing genome wide associated studies. Her team seeks to develop new techniques in Human Genetics in Rwanda. She is currently member of the National team of COVID-19 task force and she is involved in supervision of laboratory analysis of SARS-CoV-2 Testing at the National Reference Laboratory

Ms Cecile Ingabire, RN MPH (F) is an Assistant lecturer at University of Rwanda, College of Medicine and Health Sciences, School of Nursing and Midwifery. She is leading the module of Health measurement and Research and coordinating the research activities at Huye campus. She has a strong interest in Research methodology and Biostatistics. She is fluent in a wide variety of software and data management systems. Her area of research is Non Communicable diseases and Epidemiology methods.

b) UR/African Centre of Excellence in Data Science (ACE-DS)

The ACE-DS is one of 24 Eastern and Southern Africa Higher Education Centres of Excellence in the World Bank's ACE II Project. The main objective of the project is to strengthen selected Eastern and Southern African higher education institutions to deliver quality post-graduate education and build collaborative research capacity in the regional priority areas. African Centre of Excellence in Data Science (ACE-DS) has been established at the University of in October 2016 with financial support of the World Bank of 4.5million USD for five years. ACE-DS is a regional centre with combined expertise in statistics, economics, business, computer science, and engineering to use big data and data analytics to solve development challenges. ACE-DS works with many partners at national, regional and international levels. Biostatistics is one of programs available in ACE-DS. ACE-DS is member of the ARISE Network and has hosted ARISE Fellows workshop in January 2019. ACE-DS has a strong and experienced team to contribute in this project on networking of existing EU and international cohorts of relevance to COVID-19.

Key personnel

Charles Ruranga, BSc, MSc, PhD, (M) is Senior Lecturer in Department of Applied Statistics at the University of Rwanda. He has participated in management of many projects (e.g. NUFFIC-NPT, SIDA, World Bank). His primary research interests lie in the domain of econometrics, data management and Data analysis. Charles has participated as PI in the development of proposal for establishment of the African Center of Excellence in Data Science (ACE-DS) and University of Rwanda won a grant of 4.5million USD to establish ACE-DS. Charles is the Center Leader OF ACE-DS and ACE-DS is regional center working with many partners.

Joseph Nzabanita, MSc, PhD, (M) is a Senior Lecturer in the Department of Mathematics, School of Science, College of Science and Technology, University of Rwanda, Kigali, Rwanda. His research interest is multivariate statistics, focusing on growth curve modelling. He has considered the classical and extended Growth Curve Model (bilinear regression) with a linearly structured covariance matrix, e.g., banded, Toeplitz, special structure with zeros or some mix. Also, he introduced and studied the 2-fold Growth Curve Model, which is an extension of the Growth Curve Model as a tensor version. Currently, he is interested in developing and applying multivariate models in machine learning. Joseph Nzabanita is serving as Head of Research in the African Centre of Excellence in Data Science (ACE-DS), one of the 24 centres of excellence funded by the World Bank through ACE II project.

Kabano Ignace, MSc, PhD fellow (M), is a Senior Lecturer of Demography and Statistic and Head of Training, African Centre of Excellence in Data Science/University of Rwanda. Senior Project Impact Monitoring and Evaluation Consultant, InterSocial/WB Certificate in Land Acquisition, Resettlement and Social Sustainability/Senior Social Safeguard Analyst/Co-Founder and Senior Consultant, Centre for Integrated Development Research and Action (CIDRA Ltd/www.cidragroup.com).

Dr Uwimana Annie MSC, PhD in Mathematical Statistics. Lecturer at UR/Center of Excellence in Data Sciences (ACE-DS)

Centre of Excellence in Biomedical Engineering and eHealth (UR/CEBE)

The CEBE is specialized center at, the University of Rwanda, dedicated to Research and capacity building in Biomedical engineering and eHealth. It aims to develop a critical mass of labour-intensive support personnel for the sector through the creation of an enabling environment for hands-on skills development and R&D capability in strategic areas such a biomedical engineering and e-Health. CEBE is already participating in a various eHealth projects including the project with UGent to digitalize and harmonize the patient medical data from 4 hospitals + 2 health centers in Rwanda, using OHDSI OMOP common data model. The UR/CEBE is key player in this consortium in training activities.

Celestin Twizere, PhD (M) is a PhD holder in Engineering of Biomedical Electronics, Electromagnetism and Telecommunication from the University of Roma Tre, Italy in 2011. In 2002, he joined the Department of Physics in the Faculty of Education and in 2006, the Department of Electrical and electronics Engineering, National University of Rwanda as an Assistant Lecturer.

Center for Gender Studies at the University of Rwanda (UR/CGS)

The Center for Gender Studies (CGS) is under the College of Arts and Social Studies (CASS) at the University of Rwanda (UR). The Center engages in teaching, research and community outreach activities. The CGS has developed partnerships and collaborations with international and regional academic and non-academic institutions and carried out joint research.

Kagaba Mediatrice, PhD (F), lecturer in the Department of Development Studies and a researcher in the Center for Gender Studies at the University of Rwanda. Her area of expertise includes gender, peace, identify, conflict and development. She received her PhD from Gothenburg University, Sweden in 2016. Her current research focuses on Masculinity Studies, Gendered Conflicts and Identities, Peace and Development Studies, Gender and Health, Women and Human rights, and Gender and Lived experiences.

His research interests include Ground Penetrating Radar, Wireless Communication and Networks, Electromagnetic Radiation and Propagation and Smart Antenna for mobile Communication. Actually, Celestin is the Centre Director of the Regional Centre of Excellence for Biomedical Engineering and e-Health (CEBE), an AFDB funding of around 19 million USD. The key strategy of CEBE is to position itself to operate in a knowledge creation and transfer, industry partnership, and technical service delivery ecosystem to achieve high-impact results in support of several key health sector initiatives around medical equipment and devices, rehabilitation and mobility devices, and e-health system services.

2. Rwanda Biomedical Center (RBC)

Description of the Partner

RBC (www.rbc.gov.rw) Rwanda Biomedical Centre (RBC) is the nation's central health implementation agency. RBC was established in 2011 through a merger of fourteen key health institutions. RBC strives to improve the health of the Rwandan population by providing high quality, affordable and sustainable health care services. This is accomplished through the implementation of preventative, rehabilitative, and curative health interventions. RBC conducts scientific research, provides diagnostics services, and implements innovative health interventions to protect the nation against diseases and other health threats.

Main tasks

RBC is leading the data collection on COVID-19 in Rwanda. On this project it will facilitate the data gathering, providing guidance on new data collection and on data analysis, to meet the government needs. For this key contribution, RBC have an MPI position on this proposal. RBC will play a vital role in project outcome/ generated knowledge uptake, use and impact.

Key Personnel

Dr Nsanzimana Sabin, MD, PhD, (M) is the Director General of Rwanda Biomedical Centre. He has extensive experience in HIV program design, strategic planning, implementation, and operational research with a focus on global care and treatment of people living with HIV. Dr. Sabin holds a Medical Degree (MD) and a master's degree in Clinical Epidemiology from the University of Rwanda; a PhD in Epidemiology from the Basel Institute of Clinical Epidemiology and Biostatistics at the Swiss Tropical and Public Health Institute, University of Basel, Switzerland. Dr. Sabin serves as a Fellow at the African Scientific Institute (ASI), a Fellow at Royal College of Physicians of Edinburgh (FRCP) and serves as an adjunct Assistant Professor of Global Health Delivery at the University of Global Health Equity (UGHE) in Kigali and teaches Epidemiology and Research Methodology at the University of Rwanda.

Ms Clarisse Musanabaganwa, MSc (F), is a PhD Candidate (Epigenetics & Bioinformatics) at UR and USF /USA . She holds a Master's degree in Public Health field from Manchester Metropolitan University-UK, Masters of science in Epidemiology from University of Rwanda, she holds a postgraduate diploma in clinical research and pharmacovigilance and a Bachelor degree of science in Bio-Chemistry, Biotechnology and genetic engineering from Bangalore University-India. She has been working in medical research Center /Rwanda Biomedical Center Since 2012 to date Research Interest: Molecular Epidemiology (Epigenetics and Epigenomics) Drug Discovery and Clinical Trials Health System Strengthening Research Impact Evaluation studies

Mohammed Semakula (M), BSc, MSc is a National Strategic plan/Operational Plan advisor (NSP/OP advisor) at Rwanda Biomedical Center (RBC). He has a strong knowledge in statistics methods and data analysis. He has been Senior Bio-statistician, Director of research statistics in Medical Research Center and member of National Health Research committee. He has been member and chair of steering committee in charge of establishing council of Rwandan statisticians. He has been Deputy President of Young African Statisticians and demographers. He is active member of civil registration and vital statistics experts' group in Africa and member of Royal statistics society with Gradstat title. He is familiar with data modelling and management of big datasets, He worked in different technical expertise areas of statistics such as study design, Protocol development, sampling, data analysis, monitoring and evaluations, Demography and spatio-modelling. Semakula has excellent transferable skills in data management, analysis and modelling as well as strategic planning and capacity building. Recently, he is supporting Rwanda Biomedical Center, HIV national control program to elaborate and operationalize national strategic plan for Rwanda as well as supporting HIV researches. Mr. Semakula holds a degree in Applied Statistics from University of Rwanda and a Master's degree of Statistics, Biostatistics from University of Hasselt, Belgium. Mr. Semakula is advanced quantitative analyst and visiting lectures to Universities in Rwanda.

3. Ghent University (UGent)

Description of the Partner

UGent (www.ugent.be) is a top 100 university and one of the major Belgian universities, founded in 1817, and currently counting over 46,000 students and 15,000 employees. Located in Flanders, the Dutch-speaking part of Belgium and the cultural and economical heart of Europe, Ghent University is an active partner in national and international educational, scientific and industrial cooperation. With a view to cooperation in research and scientific service, numerous research groups, centres and institutes have been founded over the years. We distinguish ourselves as a socially committed and pluralistic university in a broad international perspective. UGent has 11 faculties (counting in total 86 faculty departments) and 9 administrative departments. Financially UGent has an annual revenue of 410 million euro, 58% of the revenue is used for the personnel costs, 35% for operations and equipment, the remaining 7% goes to the depreciation of the university investments.

Main tasks

UGent will primarily bringing AI mixed methods expertise needed for this project. That include the Mobile application questionnaire design, Virtual surveys and, Machine learning techniques. UGent is coordinating input from KU Leuven and other international relevant COVID-19 data to this project. For this key contribution, RBC have an MPI position on this proposal.

Key Personnel

Prof dr Marc Twagirumukiza, MD, PhD, (M) is Professor of Medicine, Clinical Pharmacology and Therapeutics at the Faculty of Medicine and Health Sciences of the Ghent University in Belgium. His primary research interests lie in pharmacological treatment outcome and medicines access in Low-and-middle income countries, focusing on Africa. He is leading efforts to harmonize prospective cohorts on COVID-19 internationally (in collaboration with European commission actors through H2020 program) and has a networking with Easter-Africa countries (Rwanda, Uganda, Kenya, Tanzania, Burundi) actors active in COVID-19 projects with data collection component.

Prof dr Pascal Coorevits, BICT, MSc, PhD, (M) is Professor of Medical Informatics and Statistics at the Faculty of Medicine and Health Sciences of the Ghent University in Belgium. He has participated in many EU, national – and transatlantic eHealth research projects (e.g. ASSIST, EHR-Implement, EHR-QTN, ARGOS, HITCH, EHR4CR, SemanticHealthNet, EURECA, ...). His primary research interests lie in the domain of Electronic Health Records (EHRs) and are oriented towards various aspects of quality labelling and certification of EHRs (EHR functional descriptive criteria, EU and US quality labelling and certification models, tools/methodologies for EHR evaluation and conformance testing). Pascal leads the data quality assessment Task Force of i-HD.

De Brouwer Edward, PhD Fellow, (M) is a PhD fellow in Artificial intelligence and Machine Learning for Healthcare at Katholieke Universiteit Leuven (KU Leuven), Belgium. His PhD is centered on “Machine learning modeling of time-dependent patient trajectories (supervisor Prof. Yves Moreau)”. His research aims at developing robust machine learning models to retrieve most relevant information from this data. This relies on the modelling of hidden variables that represent an unobserved disease state. One is matrix-factorisation with dynamic latent vectors, the other is the use of recurrent neural networks (LSTMs). More than extending the current state of the art implementations, they provide a new perspective on the problem. To make the most of these, he aims also provide a scheme to share data and their models across distinct entities without having to disclose any information. In this way, he addresses the pressing issue of data sharing and centralization.

4. Regional Alliance for Sustainable Development (RASD), Rwanda

Description of the Partner

Regional Alliance for Sustainable Development- Rwanda, (RASD Rwanda), a research institution with enormous expertise and experience in clinical and epidemiological research as well as in project management. RASD Rwanda is a national non-governmental organization whose primary goals are to leverage national institutional research, training and healthcare programs: to implement innovative clinical, epidemiological and translational/late stage studies that provide evidence-based findings; build scientific, clinical, laboratory and translational research capacity for Rwandans and improve access to healthcare services for the most vulnerable populations

RASD-Rwanda's strategic pillars include: To develop and strengthen regional partnerships in East African countries and beyond; To develop and strengthen programs to tackle ongoing epidemiological shifts including NCDs (T2D, CVD, heart failure, cancer); To promote operational research, scientific capacity building through community grassroots based intervention; To strengthen programs aimed at programs in HIV infected persons, and other vulnerable individuals; To promote and strengthen population changes and dynamics such as population density, reproductive health, sexually transmitted infections through research and healthcare, education to promote technology- and competence-based solid foundation for development; To develop and strengthen local, regional and international partnerships with institutions and Universities in clinical and operational research and healthcare programs advancing independent capacity for clinicians and scientists to perform high-impact research for the improvement of health outcomes for Rwandans.

The RASD-Rwanda is very well set up to address and develop critical needs in the policy and implementation of the integration of critical care in Rwanda because in its position as a long term partner with the Ministry of Health and the Rwanda Biomedical Centre, It has been in the past few years working on supporting the development of the dissemination and implementation science with several stakeholders in health care research in Rwanda.

Main Tasks

RASD will lead the WP6 focusing on Dissemination, Exploitation and Communication.

Key personnel

Dr. Eugene Mutimura PhD: (M) is a Senior Investigator and trainer with substantial research and training experience in implementing research and trainings programs in Rwanda and central Africa. He is the head of RASD-Rwanda with an International research collaboration Washington University in St. Louis, he also heads the National Council for Science and Technology, a government think tank involved in setting of priorities for national science and technology innovation, research and development agenda.

Dr. MUTABAZI Vincent MD MSc: (M) is medical doctor, with a Masters in Epidemiology. He has been a co-investigator on several NIH funded research projects that mainly investigated novel ways of HIV prevention, HIV related cardiometabolic studies and implementation science projects. He has also been involved in working with scientists to set priorities for building research infrastructure in Rwandan institutions that include the institutionalization of research capacity

Aurore NISHIMWE (F) is a Lecturer in the school of health sciences, College of Medicine and Health Sciences, The University of Rwanda. She holds a Master of Science in Health Informatics. Mrs. Aurore is also pursuing her Ph.D. in public health with a focus on e-Health implementation research at University of the Witwatersrand, South Africa and a One Health fellowship through Tufts University (USA). The One Health Fellowship is a collaborative project between Tufts University, the University of Global Health Equity and the University of Rwanda and is funded by the Cummings Foundation. Mrs. Aurore' areas of research interests include Implementation Research, Telemedicine and e-Health, Health informatics and Public Health especially non – communicable diseases. Her interests are in eHealth interventions. She teaches Information Technology in Healthcare, Knowledge Management in Healthcare Delivery and Software based clinical decision making and support systems. She also teaches Essentials of emergency and critical care, pediatric anesthesia and patients monitors, Research methodology.



SECTION E: QUALIFICATIONS AND EXPERIENCE OF RESEARCH TEAM

PROJECT LEADER

Family name: **Francine**

Given names: **Birungi**

Job title: **Head of Department of Epidemiology and statistics, School of Public Health, University of Rwanda**

Mailing address:

Building, Suite (optional): **School of Public Health Buildings, Kicukiro**

Number and Street or PO Box: **KG 11 Ave (Street),**

District, City: **Gasabo, Kigali**

Province/State: **Kigali Town**

Country: **Rwanda**

Postal Code: **PO Box: 3286, Kigali**

Telephone number (with country and area codes): **+250788 462 574**

Email address: fbirungi@nursph.org or birufrance@yahoo.fr

Facsimile number:

LANGUAGE SKILLS: Indicate competence in all languages used by the research team.
E = excellent; G = good; F = fair.

Language	Read	Write	Speak	Understand
English	Yes	Yes	Yes	Yes
French	Yes	Yes	Yes	Yes
Other (Swahili, Kinyarwanda)	Yes	Yes	Yes	Yes

CURRICULUM VITAE: Please include an up-to-date curriculum vitae that includes the following information.

Work experience: List all positions held, beginning with your current position.

Academic qualifications: List chronologically (starting with most recent) all degrees, diplomas, or certificates held or expected.

Credentials: List up to six awards, distinctions, of professional designations you have received and feel are pertinent to this application.

Funded research: List up to eight grants or contracts that you have received from IDRC or other sources. List them in chronological order starting with the most recent.

Publications: List your relevant professional publications.

Other information: Include any other information you feel is relevant to this application.



KEY MEMBERS OF RESEARCH TEAM

(for each key member of the team attach a brief curriculum vitae)

Family name: Twagirumukiza	Given names: Marc
Institution: Ghent University (Belgium) & University of Rwanda	
Job title: Professor of Medicine, Clinical Pharmacology and Therapeutics	
Project role/responsibility: Project Overall Research Coordinator, MPI (Multiple Principal Investigator model)	
Family name: Ruranga	Given names: Charles
Institution: University of Rwanda (UR)/Center of Excellence in Data Sciences (ACE-DS)	
Job title: Senior Lecturer	
Project role/responsibility: MPI (Multiple Principal Investigator model)	
Family name: Twizere	Given names: Celestin
Institution: University of Rwanda (UR)/Center of Excellence in Biomedical Engineering and eHealth (CEBE)	
Job title: Senior lecturer, Director of the centre	
Project role/responsibility: MPI (Multiple Principal Investigator model)	
Family name: Musanabaganwa	Given names: Clarisse
Institution: Ministry of Health (MoH) / Rwanda Biomedical Center (RBC)	
Job title: Director of Medical Research Center (MRC)	
Project role/responsibility: MPI (Multiple Principal Investigator model)	
Family name: Nsanzimana	Given names: Sabin
Institution: Ministry of Health (MoH) / Rwanda Biomedical Center (RBC)	
Job title: Director General	
Project role/responsibility: MPI (Multiple Principal Investigator model)	
Family name: Jansen	Given names: Stefan
Institution: University of Rwanda (UR)/College of Medicine and Health Sciences (CMHS)	
Job title: Research Coordinator, Directorate of Research and Innovation	
Project role/responsibility: Consortium manager, MPI (Multiple Principal Investigator model)	
Family name: Mutumura	Given names: Eugene
Institution: Regional Alliance for Sustainable Development- Rwanda, (RASD Rwanda)	
Job title: Senior Investigator & Lecturer at University of Rwanda	
Project role/responsibility: MPI (Multiple Principal Investigator model)	



KEY MEMBERS OF RESEARCH TEAM

(for each key member of the team attach a brief curriculum vitae)

Family name: Uwineza	Given names: Annette
Institution: University of Rwanda (UR)/College of Medicine and Health Sciences (CMHS)	
Job title: Senior Lecturer in Human Genetics / Director of Allied Health Services at University Teaching Hospital Kigali (CHU K)	
Project role/responsibility: Investigators Team Member	
Family name: Ingabire	Given names: Cecile
Institution: University of Rwanda (UR)/School of Nursing (SoN)	
Job title: Assistant Lecturer (RN, MPH)	
Project role/responsibility: Investigators Team Member	
Family name: Semakula	Given names: Mohammed
Institution: Ministry of Health (MoH) / Rwanda Biomedical Center (RBC)	
Job title: Director of Biostatistics	
Project role/responsibility: Investigators Team Member	
Family name: Mutabazi	Given names: Vincent
Institution: Regional Alliance for Sustainable Development- Rwanda, (RASD Rwanda)	
Job title: Director of Research Center	
Project role/responsibility: Research Officer, Investigators Team Member	
Family name: Nzabanita	Given names: Joseph
Institution: University of Rwanda (UR)/Center of Excellence in Data Sciences (ACE-DS)	
Job title: Senior lecturer	
Project role/responsibility: Investigators Team Member	
Family name: Kabano	Given names: Ignace
Institution: University of Rwanda (UR)/Center of Excellence in Data Sciences (ACE-DS)	
Job title: Senior Lecturer	
Project role/responsibility: Investigators Team Member	
Family name: Nishimwe	Given names: Aurore
Institution: Regional Alliance for Sustainable Development- Rwanda, (RASD Rwanda)	
Job title: Investigator & PhD Student	
Project role/responsibility: Investigators Team Member	



KEY MEMBERS OF RESEARCH TEAM

(for each key member of the team attach a brief curriculum vitae)

Family name: Uwimana	Given names: Annie
Institution: University of Rwanda (UR)/Center of Excellence in Data Sciences (ACE-DS)	
Job title: Lecturer	
Project role/responsibility: Investigators Team Member	
Family name: De Brouwer	Given names: Edward
Institution: Katholieke Universiteit Leuven (Belgium) on behalf of PhD work via Ghent University	
Job title: PhD Student	
Project role/responsibility: Investigators Team Member	
Family name: Coorevits	Given names: Pascal
Institution: Ghent university (Belgium)	
Job title: Professor, Head of Medical Informatics and Statistics Dpt	
Project role/responsibility: Investigators Team Member	
Family name: Kagaba	Given names: Mediatrice
Institution: Centre for Gender Studies (CGS), one of the Academic Centres within the College of Arts and Social Sciences (CASS) at UR	
Job title: Lecturer in Development Studies Department and Center for Gender Studies at UR, and Vice President of Graduate Women International (Rwanda Chapter).	
Project role/responsibility: Investigators Team Member (Specialized on Gender Transversal Theme)	
Family name: Kabakambira	Given names: Jean Damascene
Institution: University Teaching Hospital (CHU K), Kigali, Rwanda	
Job title: Lecturer in Head of Dpt of Medicine	
Project role/responsibility: Investigators Team Member	

Appendices

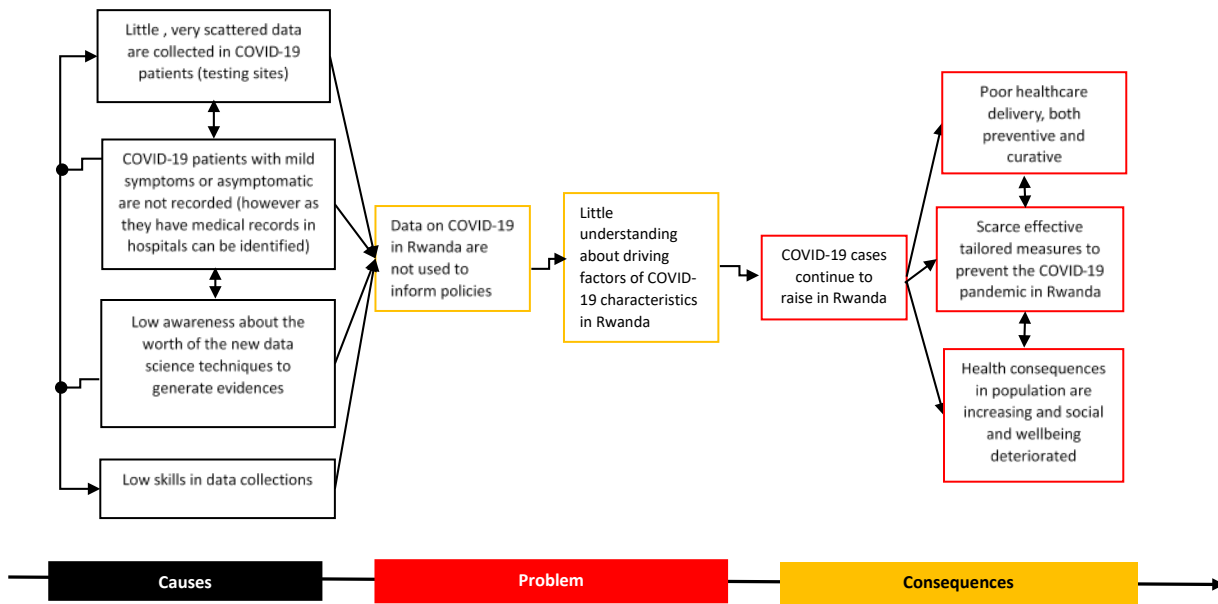


Figure 7. Problem tree

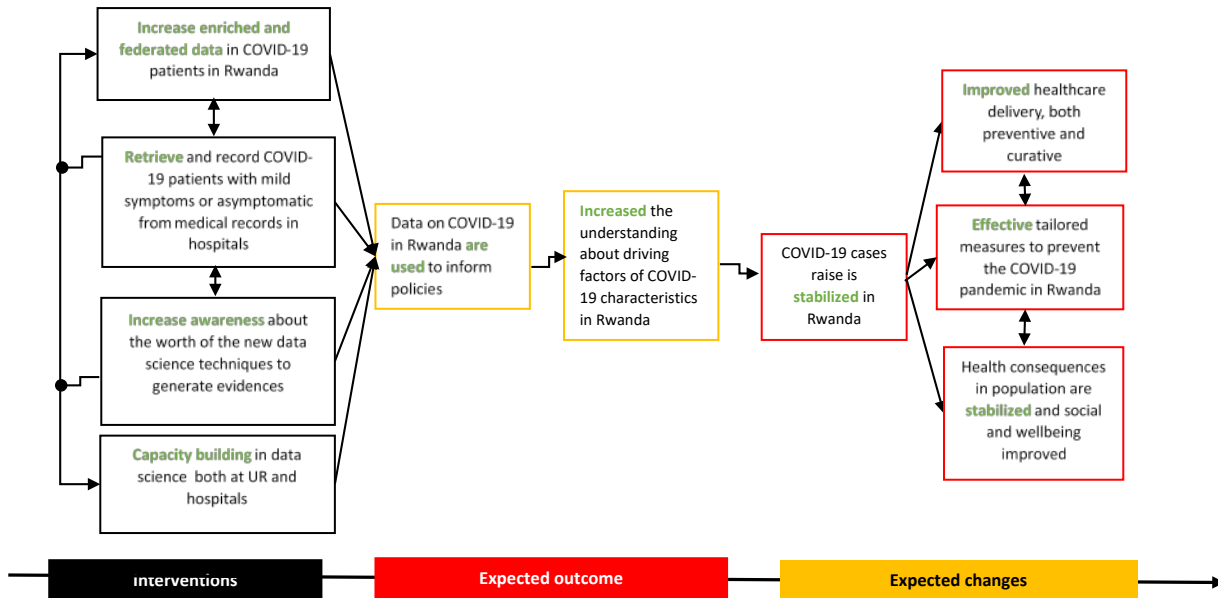


Figure 8. Theory of change