Advancing Epidemiological Research in Africa: Federated Infrastructure, Data Harmonization, and Knowledge Transfer for Scalable Public Health Insights – Technical contribution of the BRIDGE NETWORK project.

Shirah Cashriel[1], Lars Halvorsen[1], Anne Li[1] , Claude Mambo Muvunyi[2] , Marc Twagirumukiza[3], Pascal Coorevits[4] On behalf of BRIDGE NETWORK Project researchers.

[1]edenceHealth, Kontich, Belgium

[2]Rwanda Biomedical Center, Kigali, Rwanda

[3]Ghent University, Faculty of Medicine and Health Sciences, Department of Basic and Applied Medical Sciences, Ghent, Belgium

[4]Ghent University, Faculty of Medicine and Health Sciences, Department of Public Health and Primary Care, Unit of Medical Informatics and Statistics, Ghent, Belgium

Background

Epidemiological research relies heavily on high-quality, standardized data to analyse disease patterns and inform public health policies and interventions. However, variability in health information systems and data formats hinders interoperability and large-scale (inter)national studies. To address these challenges, the BRIDGE NETWORK project[1] and more specifically work package 3 focuses on data harmonization and federated research infrastructure as well as knowledge transfer regarding these processes across institutions and countries in Africa. Also included in the BRIDGE NETWORK project[1] is the creation and roll out of different educational programs at the African universities to ensure the network can be built and maintained by local personnel. Through this work, we hope to contribute to a more robust and accessible evidence-based data network for public health and epidemiology.

Methods

Our approach utilizes the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)[2] of the Observational Health Data Sciences and Informatics (OHDSI) community as a standard to integrate heterogeneous data sources, including District Health Information Software (DHIS-2) modules. The aim is to collaborate with PhD students and postdoctoral students trained within BRIDGE NETWORK project on developing mapping strategies as well as the Extract, Transform, Load (ETLs) implementations to align clinical and epidemiological health data with OMOP CDM. In parallel, we will implement a federated research infrastructure at both national and international levels, allowing secure, decentralized data analysis without the need for data sharing. This project builds on previous work done in the LAISDAR project[3] where two widely used electronic medical records systems (EMR) used within Rwanda were mapped towards the OMOP CDM. The existing infrastructure from this project will be updated and used as a starting point to deploy ETLs in the other countries involved in the BRIDGE NETWORK project. The project team expects that the PhD candidates and postdoctoral fellows leading the research, along with their thesis or research topics, could influence the project's technological and implementation choices.

Results

In the first phase, it is important to understand the current OMOP CDM landscape within the participating countries, as well as to get a view on the existing technical skills of our partners. This way we can build on the foundation of the knowledge at each of the institutions as well as efficiently utilise the available resources. Since Rwanda already has an existing network of OMOP instances from a previous project[3], this country will act as a proof of concept for the roll out of the available software suite. All Rwandan sites have been provided with either a Mac Mini (MacOS) or a Mini PC (Windows OS). To ensure the heterogeneity of the sites, all machines have been converted to Asahi Ubuntu 24.04.01 Desktop LTS. An installation guide and video were provided to instruct the users at each site in the conversion process. An installation guide was also circulated to walk the users through the OMOP set-up and deployment. The ETLs available at this stage only support the transformation to the OMOP CDM of

two different EMR systems, openMRS[4] and openClinic GA[5], although the aim of the project is to expand the number of ETLs to also support other EMR systems, DHIS-2 modules and Patient Reported Outcome Measures (PROMs) data. We aim to leverage commonly used vocabularies such as ICD-10 and SNOMED CT, in combination with the different relationships represented in the OMOP vocabulary to increase the efficiency of mappings. After running the ETLs, several OHDSI tools will be run as part of the pipeline such as Ares, Atlas and Data Quality Dashboard (DQD). We are also exploring ways to connect and integrate different open standards such as FHIR and openEHR into our data pipelines.

The next step in the process is to set-up a federated national and international network to support large-scale studies. Vantage6[6] will be deployed as it is an open-source privacy-preserving federated learning framework. Each participating site will run a vantage6 node that operates locally on harmonized OMOP CDM data, executing distributed computations without exposing patient-level information. The network will utilise secure containers to process queries, ensuring that only aggregated results or model parameters are shared between sites. Within this setup, institutions can collaboratively perform epidemiological studies while maintaining full data sovereignty. Additionally, this national infrastructure is designed to scale by linking to an international vantage6 network, allowing cross-border research while enforcing local governance policies. A portal using Keycloak for authentication will be served at each national central server. This portal will give the user access to extended dashboarding tools such as Superset and provide the admin users with metrics using Prometheus and Grafana.

Conclusion

By harmonizing and integrating clinical and epidemiological data, alongside implementing a federated infrastructure, this work aims to enhance the efficiency and reproducibility of public health research in Africa by establishing a robust research framework. The integration of OMOP CDM with federated analysis enables scalable, privacy-preserving studies, fostering international collaboration and accelerating evidence generation. These findings will provide valuable insights to the broader OHDSI community and demonstrate effective knowledge transfer within a large-scale international consortium.

References

1. BRIDGE NETWORK project. Scaling Up Research in SSA Countries in Infectious Diseases: African-European Training Network in Integrated Health Informatics and Data Sciences. https://cordis.europa.eu/project/id/101145734 Accessed on 27 March 2025.
2. OHDSI/OMOP CDM. Welcome to OHDSI. https://www.ohdsi.org/data-standardization/ Accessed on 27 March 2025.
3. Nishimwe, A., Ruranga, C., Musanabaganwa, C. et al. Leveraging artificial intelligence and data science techniques in harmonizing, sharing, accessing and analyzing SARS-COV-2/COVID-19 data in Rwanda (LAISDAR Project): study design and rationale. BMC Med Inform Decis Mak 22, 214 (2022).
4. OpenMRS: medical record system. 2025. https://openmrs.org/ Accessed on 27 March 2025.
5. Verbeke F. OpenClinic GA: open-source integrated hospital information management system. 2025. https://sourceforge.net/projects/open-clinic/ Accessed on 27 March 2025.
6. Moncada-Torres, A., Martin, F., Sieswerda, M., Van Soest, J., & Geleijnse, G.. VANTAGE6: an open source priVAcy preserviNg federaTed leArninG infrastructurE for Secure Insight eXchange. AMIA Annual Symposium proceedings. AMIA Symposium 2020, 870–877.

BRIDGE NETWORK Project Consortium list of researchers as of 30 March 2025 (alphabetical order):

Andrew Boulle, Angela Koech, Alphonse Kpozehouen, Alex Riolexus Ario, Alfred Mubangizi, Anne Li, Charles Sossa, Claude Mambo Muvunyi, Colette Azandjeme, Edgard Marius Ouendo, Edwin Akugizibwe,

Elenizoumpoulia Papadopoulou, Emmanuel Kimera, Everlyne Mathu, Eva Langat, Fede Dewulf, Felix Ocom, Freija Descamps, Gemechu Tadesse, Geert Byttebier, Geremew Tasew, Getachew Tollera Eticha, Ghislain Sopoh, Grace Wairimu, Gregory Tweheyo, Hillary Koros, Hillary Otema Ario, James Orwa, Jan Frederik Frøen, Jean Claude Semuto Ngabonziza, Jean Marie Vianney Kwizera Nkurunziza, Jessica Senekal, Johanna Lucia Thorbjornsrud Nader, John Rubaihayo, Jonas Gbaguidi, Joshua Wesana, Juliet Namagulu, Katherine Johnston, Kelly Tara Dolgner, Kjetilelias Telle, Kirsty Garson, Kwizera JMV, Lars Halvorsen, Liezel Mitchell, Linda Gambakwe, Luc Belmans, Mack Kigada, Mahima Venkateswaran, Marc Lamont, Marc Twagirumukiza, Marit Hasle, Marleen Temmerman, Masresha Tessema, Meg Osler, Minja Veraangenyreynold, Mulugeta Geremew Geleso, Ncumisa Honono, Negga Asamene Abera, Nicola Mulder, Noël Moussiliou Paraïso, Nontobeko Tena-Coki, Pacifique Nizeyimana, Pascal Coorevits, Pius Achanga, Rahim Khawaja, Roar Samuelsen, Samuelsen Roar, Shirah Cashriel, Veronica Robi, Yolaine Glele Ahanhanzo, and Lyndon Zass.

Acknowledgments

A special thank you to Emma Gesquiere for her extensive efforts and invaluable input on this project.