

Grand Challenges ICODA webinar series

Privacy Enhancing Technologies (PETs): The use of Trusted Research Environments in global health research

Neil Postlethwaite Technical Director, ICODA

Joseph Ouma Senior Programmer, periCOVID study

Dr Carl Marincowitz Principal Investigator, PRIEST Study



BILL& MELINDA GATES foundation









Overview of Trusted Research Environments

Neil Postlethwaite

Technical Director International COVID-19 Data Alliance (ICODA) / HDR Global

- Overview of Privacy-Enhancing Technologies (PETs)
- Background on ICODA
- Introduction to the Five Safes
- The ICODA Workbench as our Trusted Research Environment

Trusted Research Environments in practice

Joseph Ouma

Senior Programmer for the periCOVID study, Makarere University Johns Hopkins University (MUJHU), Uganda

Dr Carl Marincowitz

Principal Investigator of the PRIEST Study for Low and Middle-Income Countries, Sheffield University, UK

- Brief overview of research project
- Experience using TREs in project

Q/A



Privacy-Enhancing Technologies



Traditional PETs:

Encryption in transit and at rest

De-identification techniques

Emerging PETs:

Trusted Research Environments

Federated analytics

Differential privacy

Multi-party computation

Homomorphic encryption

Reference: https://cdeiuk.github.io/pets-adoption-guide/what-are-pets

Enables researchers to **collaborate**, **link data**, **run analyses**, and **share code and results** within a secure environment.



Researchers

Other names:

- Trusted Execution Environments (TEE)
- Virtual Research Environment (VRE)
- Digital Research Environments (DRE)
- Secure Research Environment (SRE)
- Secure Data Environment (SDE)



Benefits of Trusted Research Environments





Researchers



Secure airlocked environment: Users are authenticated with access based on roles and permissions. Outputs are quarantined before release



Configurable tools and services: Analysis tools can be added and configured. Federated queries offer connection to other TREs



Enhanced collaboration: Shared workspaces allow users to work collaboratively on research projects



Single, secure data storage: Custodians can store and manage data within the TRE



Secure remote access: Accessible to accredited users remotely, allowing project teams to work globally



Scalable compute power: Processing power is provided by TRE servers and not individual user's machines, and can be scaled based on demand



Enhanced Trustworthiness: Data is stored securely, providing reassurance to public that their sensitive data is managed correctly



Data custodians



Public, patient, practitioners



The International COVID-19 Data Alliance (ICODA) was convened by HDR UK in mid 2020 to focus on a rapid response the COVID-19 pandemic, using data science approaches and leave a lasting infrastructure, processes and learnings

ICODA is trying to add distinctive value to other data platforms in three key ways:

- 1. Improving discoverability: working with multiple data providers to make "meta data" available for researchers to search and find datasets
- 2. Providing a secure analysis platform: relevant data from multiple sources can be brought together on the ICODA Workbench, a secure Trusted Research Environment for researchers to address specific research questions.
- **3.** Facilitating federated access: where it is not possible to transfer data, providing a federated approach to allow researchers to send analysis to the data (rather than moving data to the researcher) in one or more different locations.

We prioritise the importance of demonstrating **trustworthy data stewardship** and making data FAIR – findable, accessible, interoperable, reusable

PET

PET

Introduction to Five Safes



Safe People

Researchers are trained and authorised to use data safely

Safe Projects

Research projects are approved by data owners and for the public good

Safe Settings

A Trusted Research Environment provides a secure environment within which to perform research

Safe Data Data is curated to protect any confidentiality concerns

Safe Outputs

Screened and approved outputs that are non-disclosive

The ICODA Workbench

Allows researchers to access data in a trusted research environment with the opportunity to work on important research questions

Data Access

- Invited researchers register
- Login via 2 Factor Authentication
- Permits global access .

Tooling

- Collaborative workspaces with open source tooling (R, Python, SQL)
- Bring your own data, code and 3rd party software or tools
- Specialised clinical research tools • available from ICODA partners

Compute

- Web based tools or clustered virtual machines (Windows, Linux, GPU)
- Scalable cloud computing



Global Portfolio of Driver Projects

All working within ICODA's Workbench



ICODA

DP-REHCORD

Harvard T.H. Chan School of Public Health Impact of COVID-19 on Health Service Delivery and Institutional Mortality: A Multi Country Consortium

DP-ISARIC

Universidad de La Sabana in Colombia Data Descriptor, Reference Coding, and Characterization of the Systemic Complications of Critical Care Patients Included in the ISARIC COVID-19 Dataset

DP-IROC

MRC/UVRI & LSHTM Uganda Research Unit and Makarere University

Incidence and risk factors for COVID-19 amongst pregnant and lactating women and their infants in Uganda.

DP-ACCORD

Western Cape Government Health Department and the University of Cape Town

Addressing Critical Covid-19 Questions Through Research Using Linked Population Data

DP-PIH-CovCo

Partners in Health of Haiti, Malawi, Mexico, and Rwanda

The Impact of COVID-19 on Chronic Care Patients' Health Care Utilization and Health Outcomes in Haiti, Malawi, Mexico and Rwanda

DP1-EFCT Duke University, Certara, Cytel, BMGF, NMD Group Evaluating the efficacy and` safety of COVID-19 treatments

DP2-iPOP

University of Edinburgh, University of Manitoba, Murdoch Children's Research Institute, University of New South Wales + iPOP Consortium International Perinatal Outcomes in Pandemic (iPOP)

DP-PRIEST

University of Sheffield & Cape Town

Pandemic Respiratory Infection Emergency System Triage (PRIEST) Study for Low and Middle-Income Countries

DP-EFFECT

DP-IDS-COVID 19

Middle Income Country

Fiocruz, Brazil

Fiocruz, Brazil Effectiveness of COVID-19 Vaccination in Brazil Using Mobile Data

Evaluating Effects of Social Inequalities on

the COVID-19 Pandemic in a Low- and



https://icoda-research.org/research/driver-projects/

DP-RASUP

Fiocruz, Brazil

Routine Assessment of Infections, Prevention, and Control of SARS-CoV-2 on Unequal Populations

DP-CHAIN

City University of Hong Kong

Characterizing COVID-19 Transmission Chains for Precision Mitigation Using Epidemiological Survey Data



Trusted Research Environments in practice

Joseph Ouma

Senior Programmer, periCOVID study Makarere University Johns Hopkins University (MUJHU), Uganda

Dr Carl Marincowitz

Principal Investigator, PRIEST Study for Low and Middle-Income Countries Sheffield University, UK



BILL& MELINDA GATES foundation





IROC: Incidence and Risk factors for COVID-19 among pregnant and lactating women and their infants in Uganda



Joseph Ouma, IROC team, Uganda

HDR-UK Ref Number: 2021.0096

Project period: 5th July 2021 to 4th October 2022



Outline

- Project team
- IROC, Goals and Objectives
- Data, Progress and Achievements
- Experiences using the workbench
 - Benefits
 - Challenges



Core team



- Prof Kirsty Le Daore, Principle Investigator. MRC/UVRI & LSHTM Uganda Research Unit and Makerere University John's Hopkins University Research Collaboration (MUJHU), Uganda.
- Joseph Ouma, Senior Data Manager/Biostatistician, Makerere University John's Hopkins University Research Collaboration (MUJHU), Uganda.
- Dr Lauren Hookham, Program manager peri-COVID study, in Uganda
- Gordon Rokundo, Data Manager, Makerere University John's Hopkins University Research Collaboration (MUJHU), Uganda.
- Mary Kyohere, Program manager, Makerere University John's Hopkins University Research Collaboration (MUJHU), Uganda.
- Prof Philippa Musoke, Executive Director, Makerere University John's Hopkins University Research Collaboration (MUJHU), Uganda



Collaborations and partnerships



- Makerere University John's Hopkins University Research Collaboration (MUJHU), Uganda.
- MRC/UVRI & LSHTM Uganda Research Unit
- Kawempe National Referral Hospital, Uganda
- Division of Health information, MoH, Uganda
- Department of MNCH, MoH, Uganda
- The International COVID-19 Data Alliance (ICODA)



The gap



- Real time data collection, analysis, reporting and USE to rapidly support changes to guidelines and care are critical
- Electronic Health Records (EHR) address a number of challenges, providing a platform that supports clinical care, real-time data collection, data analysis, reporting and disease surveillance.

MU-JHU CARE LTD

Goal and objectives of IROC



To determine risk factors for adverse pregnancy and neonatal outcomes during the COVID 19 pandemic using the EHR;

- 1. Develop a risk factor model for indirect and direct effects of the pandemic on MNCH services in Uganda;
- 2. Identify at risk populations who would benefit from vaccination catch up programmes;
- **3. Formulate guidance on infection prevention and control** during pregnancy and in the postpartum period;
- 4. Develop a Ministry of Health COVID19 Dashboard for MNCH data for this and future pandemics



Setting and data



- Kawempe National Referral Hospital (KNRH), provides MNCH services to >30,000 women annually
- UgandaEMR, an electronic medical records system is used to capture patient level service delivery data based on MoH HMIS tools
- Data on women accessing Antenatal, Delivery and postnatal services at KNRH for Jan 2020 – Oct 2021 used for the study.
 - Outcomes death (maternal, neonatal, infant<6 months), prematurity, stillbirth, miscarriage, gestational diabetes, antepartum & postpartum hemorrhage
 - Other variables Maternal demographics (age, level of education, marital status, area of residence –rural/urban) and parity, gravidity, gestational age.
- Data on Government initiated restrictions- showing lockdown timelines from the CoVID19 Dashboard, Ministry of Health

Progress and achievements



Stakeholder engagement meetings with MoH and KNRH staff were held:

- Guided dashboard development,
- selected indicators for dashboard,
- built excitement and acceptability and
- initiated use case as well as ongoing commitment
- Risk factor model and at risk population identified
 - manuscript to be submitted for publication by end of September 2022
- Abstract on data visualization accepted for poster presentation at the conference in Cape town South Africa





Visualization dashboard complete shared with KNRH



器 EMR / EMR Dashboard ① Last 30 days ~ 3 ÷ ŝ \square Data ANC Flags obs Maternity Data Cleanups **Patient Flags** OBS **ANC Details** MCH:006 Q 88 L&D ANC OutComes Labor And Delivery 12 **ANC Details** Delivery Methods ~ ANC1 ANC2+ Admissions Deliveries Discharges 2056 2299 2159 1625 901 SVD Breech Other 985 552 97 4 4 **Discharge Condition** ANC1 **Breathing Normal Referral IN** 901 PreTerm Term 989 513 Transfer Fistula 1726 Run Awa 433 ANC2+ 1347 2 4 0 2957 ී 30 Mother 49 126 23 23 20 16 10 Child 0 02/07 01/13 01/18 01/23 01/28 02/02 5 APH - PPH Malaria Hypertensive 0 0 0 💳 Obstructed 💳 Sepsis 💳 Anaemia 💳 Reptured 3 Ab... APH PPH Ma., Hig., Ob., Se., An., Re., Ect., Oth., - Ectopic - Other



Workbench onboarding and preparations

- Training and signing up on Synapse
 - Team were introduced to Synapse environment, available tools to share, organize and promote research work and outputs.
- Data Processing Agreement (DPA) for data sharing and processing were completed
- Training on mandatory ICODA workbench courses and modules completed.
- User accreditation and granting access to study, IROC, workspaces on the workbench.





IROC data in the workbench



- Large dataset >49,700 antenatal and >51,680 Labour and delivery attendance beyond the workbench data table editor specifications of 100x10,000.
- Windows VM was deployed to the workspace
 - Interactive sessions held with the workbench team to inform expansion of workbench space and utility
- R Markdown modules made available via the virtual machine
 - through interactive sessions, IROC team were shown how to use them for data analysis



IROC data in the workbench

- Data curation support by MMS Holdings Inc. USA.
 ICODA curation partner
 - Curation done outside the workbench
 - Another data processing agreement signed
 - Supported Improvement of the study meta-data
- Statistical support/advise from ICODA's Statistical Expert Group.
- Further interactive meetings/training with ICODA/Workbench team on use of the workbench tools including export and use of the analysis modules on virtual machines.

2	Reason_for_Admission	N
з	1 FRESH SCAR IN LATENT LABOUR	
в	1 P.SCAR AND OBSTRUTED	
в	1 P/S	
в	1 p/scar	
Э	1 FRESH P/S	
в	1 FRESH P/S, OBLIQUE AT TERM	
5	1 FRESH P/S, UTI	
в	1 FRESH P/SCAR	
1	1 FRESH PREVIOUS SCAR IN LABOR	
Э	1 FRESH SCAR	
в	1 FRESH SCAR CONTRACTING	
5	1 FRESH SCAR IN LABOUR	
в	1 FRESH SCAR, SIPI	
в	1 FRESH SCARCONTRACTING	
в	1 FRSEH SCAR	
D	1 LOSE NUCHAL CORD, OLIGOHYDROMIOS	
3	1 OSC + CONTRACTED PELVIS + BIG BABY	
в	1 P / S	
в	1 P SCAR WITH TWIN PREG	
в	1 P/ S	
в	1 P/ SCAR	
в	1 P/ SCAR ,IN LBR	
2	1 P/, PROM	
в	1 P/S	
в	1 P/S .IN LABOUR	
в	1 P/S IN ACTIVE LABOR	
в	1 P/S IN LABOUR	
в	1 P/S IN LATENT LABOR	
в	1 P/S & 1VBAC IN LATENT, POOR PROGRESS LABOUR	
з	1 P/S + CORD AROUND THE NECK	
в	1 P/S + OBLIQUE LIE	
2	1 P/S . 1 VBAC . PET	
3	1 P/S . 1VBAC	



Benefits using the TRE



- The Aridhia service desk was very supportive in guiding the team: on uploading data onto the workspace; using the exported modules and in resolving user challenges.
- Single secure data storage accessible remotely by only accredited study team members.
- Secure file storage and sharing for large files removed the burden of sharing data files and related meta data with team members.
- Analytics and tools to model data more efficiently were available for use.



Benefits using the TRE



- While using the VM, there is restricted access to external resources such as the web, reassuring on data security.
- Using the airlock provision, team members can only download datasets after authorization from the workspace administrator.
- Improved collaboration as team members contributed directly on the data analysis codes and other inputs as the analysis progressed

Challenges

- Used tools/modules outside of the environment The team was not able to optimally use them.
- Using data from routine service delivery, containing text data fields-The team worked collaborating with MMs to complete the data curation.
- Using the R Markdown modules in the VM environment, required familiarity with R program
- Working in the environment needing large bandwidth, hence patience to run and re-run scripts

Stakeholder engagement session







Acknowledgements

- ICODA workbench team/engineers
- ICODA SEG,
- MMS, ICODA data curation partner
- Ministry of Health, Uganda
- Kawempe National Referral Hospital
- Study participants







The PRIEST (Pandemic Respiratory Infection Emergency System Triage) Study for Low and Middle-Income Countries

The role of the Trusted Research Environment and Work Bench



RESEARCH ARTICLE

Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: The PRIEST observational cohort study

Steve Goodacre¹*, Ben Thomas¹, Laura Sutton¹, Matthew Burnsall¹, Ellen Lee¹, Mike Bradburn¹, Amanda Loban¹, Simon Waterhouse¹, Richard Simmonds⁰, Katie Biggs¹, Carl Marincowitz¹, Jose Schutter¹, Sarah Connelly¹, Elena Sheldon¹, Jamie Hall¹, Emma Young⁰, Andrew Bentley², Kirsty Challen³, Chris Fitzsimmons⁴, Tim Harris⁵, Fiona Lecky⁰, Andrew Lee¹, Ian Maconochie⁵, Darren Walter⁷

School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, United Kingdom,
 Intensive Care, Manchester University NHS Foundation Trust, Wythenshawe Hospital, Manchester, United Kingdom,
 Emergency Department, Lancashire Teaching Hospitals NHS Foundation Trust, Preston, United Kingdom,
 Emergency Department, Sheffield Children's NHS Foundation Trust, Sheffield, United Kingdom,
 Emergency Department, Barts Health NHS Trust, London, United Kingdom,
 Emergency Department, Barts Health NHS Trust, London, United Kingdom,
 Emergency Department, Manchester University NHS Foundation Trust, Wythenshawe Hospital, Manchester, United Kingdom

* s.goodacre@sheffield.ac.uk

Abstract

Objectives

We aimed to derive and validate a triage tool, based on clinical assessment alone, for predicting adverse outcome in acutely ill adults with suspected COVID-19 infection.

Methods

We undertook a mixed prospective and retrospective observational cohort study in 70 emergency departments across the United Kingdom (UK). We collected presenting data from 22445 people attending with suspected COVID-19 between 26 March 2020 and 28 May 2020. The primary outcome was death or organ support (respiratory, cardiovascular, or renal) by record review at 30 days. We split the cohort into derivation and validation sets, developed a clinical score based on the coefficients from multivariable analysis using the derivation set, and the estimated discriminant performance using the validation set.





OPEN ACCESS

Citation: Goodacre S, Thomas B, Sutton L, Burnsall M, Lee E, Bradburn M, et al. (2021) Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: The PRIEST observational cohort study. PLoS ONE 16(1): e0245840. https://doi.org/10.1371/journal. pone.0245840.

Editor: Itamar Ashkenazi, Technion - Israel Institute of Technology, ISRAEL

Received: November 6, 2020

Accepted: January 9, 2021

Published: January 22, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0245840 **Original research**

BMJ Open Potential solutions for screening, triage, and severity scoring of suspected COVID-19 positive patients in lowresource settings: a scoping review

Sarah Hirner,¹ Jennifer Lee Pigoga ⁽¹⁾,² Antoinette Vanessa Naidoo,² Emilie J Calvello Hynes ⁽²⁾,³ Yasein O Omer,^{2,4} Lee A Wallis,² Corey B Bills ⁽²⁾

ABSTRACT

Objectives Purposefully designed and validated

screening, triage, and severity scoring tools are needed

(LRS). This review aimed to identify currently proposed

severity scoring of patients with suspected COVID-19

on initial presentation to the healthcare system and to

Design A scoping review was conducted to identify

studies describing acute screening, triage, and severity

scoring of patients with suspected COVID-19 published

between 12 December 2019 and 1 April 2021. Extracted

information included clinical features, use of laboratory

and imaging studies, and relevant tool validation data.

Participant The initial search strategy yielded 15232

Results Most studies were from China (n=41, 33.1%) or

of 37 studies provided validation data: 4 prospective and

33 retrospective, with none from low-income and lower

Conclusions This study identified a number of screening,

proposed for patients with suspected COVID-19. No tools

specific to resource limited contexts is crucial to reducing

were specifically designed and validated in LRS. Tools

triage, and severity scoring tools implemented and

the United States (n=23, 18.5%). In total, 57 screening, 23 triage, and 54 severity scoring tools were described. A total of 51 tools–31 screening, 5 triage, and 15 severity scoring—were identified as feasible for use in LRS. A total

evaluate the utility of these tools in LRS.

articles: 124 met inclusion criteria.

middle-income countries.

mortality in the current pandemic.

and/or implemented methods of screening, triaging, and

to reduce mortality of COVID-19 in low-resource settings

Naidoo AV, et al. Potential solutions for screening, triage, and severity scoring of suspected COVID-19 positive patients in lowresource settings: a scoping review. BMJ Open 2021;11:e046130. doi:10.1136/ bmjopen-2020-046130

To cite: Himer S, Pigoga JL,

 Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/ bmjopen-2020-046130).

Received 21 October 2020 Accepted 25 June 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹University of Colorado Denver

Strengths and limitations of this study

- We provide the first review of COVID-19 screening, triage, and severity scoring tools both proposed and implemented among initial patient presentations to the healthcare system.
- Many screening, triage, and severity scoring tools have been proposed and implemented, but none are specific to low-resource settings (LRS).
- We identified 51 tools—31 screening, 5 triage, and 15 severity scoring—that have variables feasible for collection in LRS.
- Feasibility, however, does not predict that a tool will be accurate or effective, and no tools from this review were validated in LRS.
- It is likely that many tools being used in healthcare systems worldwide are not published and thus cannot be described in this review.

healthcare systems worldwide by increasing demand for personal protective equipment (PPE), diagnostics, oxygen and mechanical ventilators.⁶ Low-resource settings (LRS) have limited access to these resources and remain disproportionately challenged during the COVID-19 pandemic.^{7 B} Even in regions where viral transmission remains low, patients with suspected COVID-19 require precautions, and confirmed cases require



Aims

Primary objective: develop clinical risk stratification tools to predict need for inpatient admission from the ED for patients with suspected COVID-19 infection applicable to LMIC settings.

Secondary objectives:

- consult with patient and clinical stakeholders to develop clinically usable and contextually appropriate risk stratification tools.
- externally validate new and previously developed risk-stratification tools in different income settings and waves of the pandemic
- identify the optimal cut-off points for the risk stratification tools and compare their performance with existing triage methods (e.g. National Early Warning Scores and the WHO hospitalisation algorithm for pneumonia).

Research team and collaboration





- Clinical researchers (Myself, Dr Fuller and Professor Goodacre)
- Statistical and Machine Learning Expertise (Information School, Professor Bath, Dr Hassan and Dr Sbaffi)
- Clinical and data expertise (Professor Wallis, Professor Hodkinson and Dr McAlpine)
- Collaboration Sudan (led by Dr Omer)

Data sets

UK, PRIEST study

N = 20,698 (adults)

26th March and 28th May 2020

Western Cape, South Africa

N = 446,084 (adults)

27th August 2020 to 11th March 2022

Sudanese Multi-centre study

N = 2583 (adults)

3rd January 2020 to 17th December 2021

Outputs (Achieved and Projected)

- Validation of existing (including PRIEST) clinical risk-stratification tools in Western Cape (including Omicron sub-group)
- Development new modified version of PRIEST tool applicable to LMICs using Western Cape dataset
- Validation of new and existing tools in Sudanese population
- Comparison of Machine Learning tools to traditionally developed clinical risk-stratification tools

Role of the TRES/Workbench



Advantages for PI



Advantages for analysis



Access for our team anywhere in the world (in theory!) TRES/Workbench

Suite of simple descriptive applications

Good for linux and python

Useful for collaboration across teams for analysis

Caveats

At the start of the project and for a while after, the opened terminals did not last for long and disappeared very quickly, even in less than 2 minutes upon moving to another tab or window. This problem was resolved later.

Some of the frequent and normal keyboard shortcuts (such as ctl+c, ctl+v) can't be used when working on the terminals or writing scripts, which causes lots of delays in the work.

Needs very large internet bandwidth to maintain connection.

Hard to access/install external libraries. This also applies to famous providers like IBM.

Not having the permission for copying results from the workspace to some files outside the workspace introduces more manual work. Trying to print every single result to a file then asking for permission to airlock is not an optimal solution, especially if the results will come from different scripts.

Conclusions

- TRES/Workbench- key for pooling and managing our datasets
- Due to restrictions on use of some datasets (Sudanese)-analysis had to occur on TRES
- Probably possible to conduct all analyses on TRES but due practical limitations some analysis conducted in UoS secure environment

Additional resources



Privacy Enhancing Technologies (PETs). Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance - Information Commissioner's Office (ICO) https://ico.org.uk/media/about-the-ico/consultations/4021464/chapter-5-anonymisation-pets.pdf

Repository of PET Use Cases - Centre for Data Ethics and Innovation https://cdeiuk.github.io/pets-adoption-guide/repository/

Building Trusted Research Environments - Principles and Best Practices; Towards TRE ecosystems (paper) - UK Health Data Research Alliance, NHSX

https://zenodo.org/record/5767586#.YyM 2uzML0o

An Introduction to Privacy Enhancing Technologies (online course)

- Health Data Research UK Futures https://hdruklearning.csod.com/ui/lms-learning-details/app/course/cec9fd53-df0b-4821-aaec-ba73019f75ab





Thank you, any questions?

Neil Postlethwaite

Technical Director, ICODA

Joseph Ouma Senior Programmer, periCOVID study

Dr Carl Marincowitz Principal Investigator, PRIEST Study



BILL& MELINDA GATES foundation



Convened by HDRUK Health Data Research UK