# Refining the Global Repository of Epidemiological Parameters data schema
Collaboratory workshop

Meeting report
WHO Hub for Pandemic and Epidemic Intelligence
Berlin, Germany
May 14 - 16, 2024

This event was a 3-day in-person workshop dedicated to refining the Global Repository of Epidemiological Parameters (GREP) data schema. A total of 28 people from a range of interdisciplinary, multisectoral and geographic backgrounds joined this workshop, which included a series of presentations and hands-on group activities. This workshop fostered strong momentum within the community to move towards the development of GREP as an essential public good to inform and guide public health interventions worldwide.
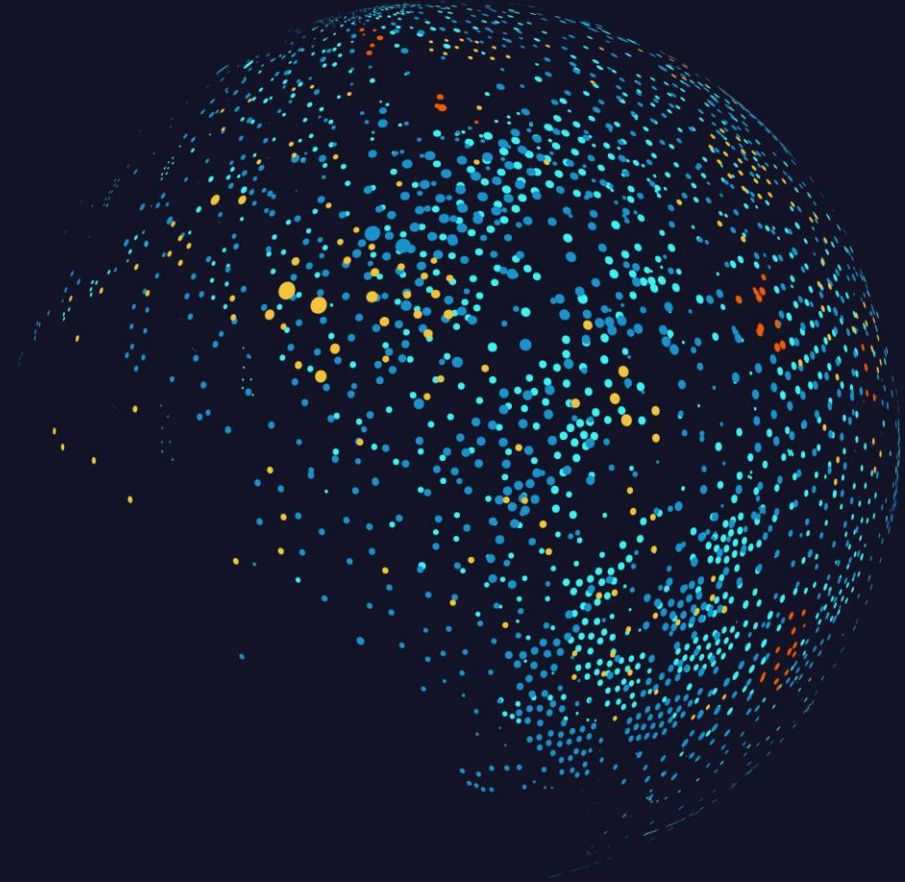
Collaboratory

World Health Organization | HUB

BACKGROUND

The EpiParameter community is a global collaborative working group coordinated by the World Health Organization (WHO), with the aim of developing a Global Repository of Epidemiological Parameters (GREP). This repository is intended to be accessible by modelers, epidemiologists, subject matter experts and decision makers, to inform mathematical models and other epidemiological analyses and, by extension, public health preparedness and response.

This project consists of five workstreams, including definition of a standardized data model for key parameters, development of extraction protocols, storage and use of parameters, mechanisms for maintenance and quality assurance, and frameworks for recognition and contribution incentives. The focus of this workshop was on the first three workstreams.

CONSULTATION AIM

Previous meetings had invited community members to prioritize essential epidemiological tasks which the GREP data schema should be able to support. However, GREP was yet to be evaluated against a wider range of use cases, hence the requirement for this workshop.
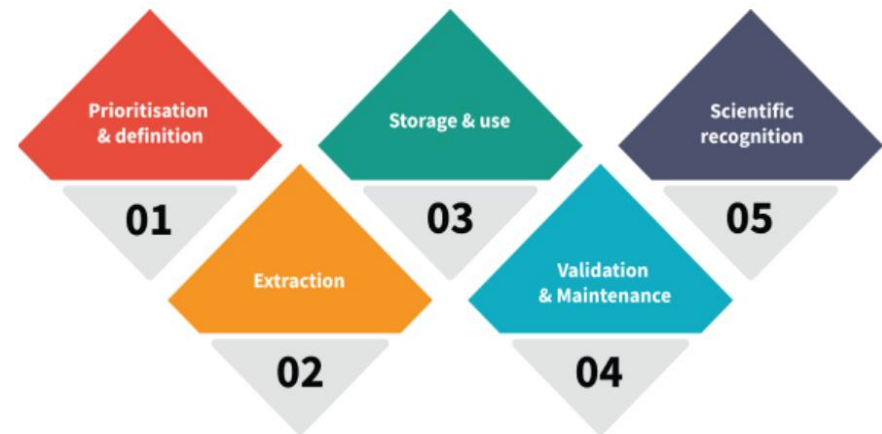
HIGH LEVEL SUMMARY OF PRESENTATIONS

There were several speakers and presentations over the course of the workshop, summarized below.

Day 1 – Prioritization and definition of parameters
- Opening remarks – Oliver Morgan (WHO)
- Introduction to the Collaboratory – Julia Fitzner (WHO)
- Summary of community progress and overarching workshop objectives – Patricia Ndumbi (WHO)
- Introduction to the GREP data schema – Finlay Campbell (WHO)

Day 2 – Extraction of parameters
- Introduction to the xMart platform – Julia Fitzner (WHO)
- Challenges of extracting parameters from various sources – Lisa Waddell (Public Health Agency of Canada, PHAC)
- Results from systematic reviews on WHO blueprint diseases and an overview of the {epireview} R package – Anne Cori (Imperial College London, ICL)

Day 3 – Storage and use of parameters
- Introduction to the Epiverse-TRACE initiative – Adam Kucharski (London School of Hygiene and Tropical Medicine, LSHTM)
- Exploration of the {epiparameter} R package – Joshua Lambert (LSHTM)
- Concluding remarks – Patricia Ndumbi (WHO)

HIGH LEVEL SUMMARY OF ACTIVITIES

The workshop involved multiple hands-on activities and discussions for participants.
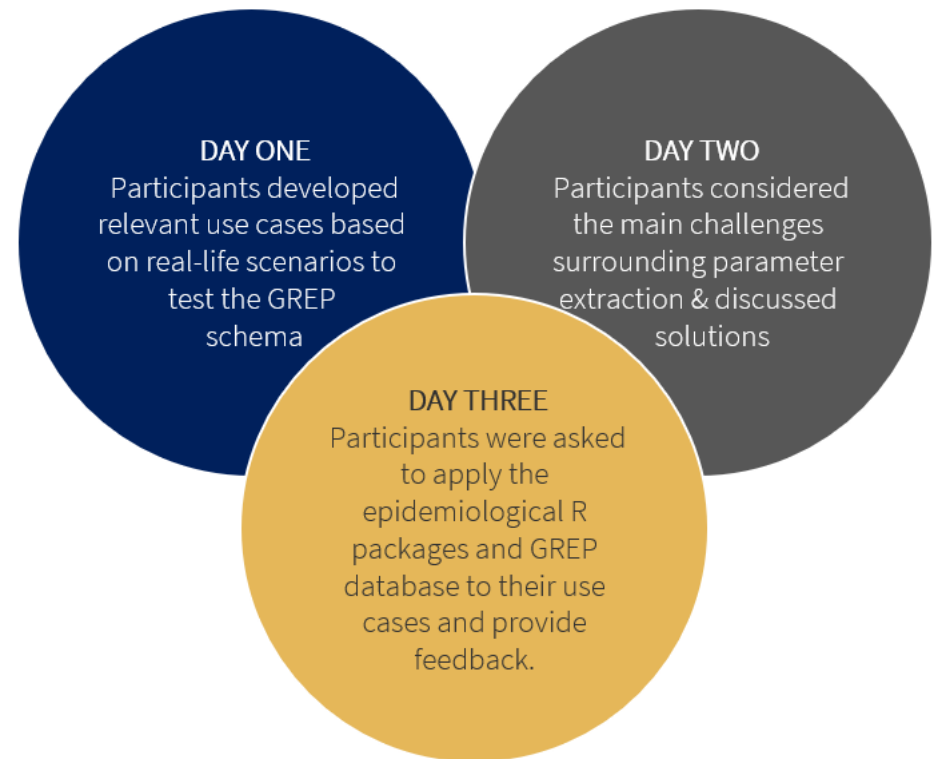
Day 1 – Prioritization and definition of parameters
- Develop real-life epidemiological use cases to test the GREP data schema – led by Carmen Tamayo Cuartero (LSHTM)

Day 2 – Extraction of parameters
- Identification of key challenges in parameter extraction, suggesting bottlenecks, solutions and tools to improve GREP's efficiency and utility – led by Lisa Waddell (PHAC)

Day 3 – Storage and use of parameters
- Testing the {epireview} and {epiparameter} packages with the use cases developed on day 1 to assess their applicability and interoperability, and to provide code contributions – led by Joshua Lambert (LSHTM).

DAY ONE
Participants developed relevant use cases based on real-life scenarios to test the GREP schema

DAY TWO
Participants considered the main challenges surrounding parameter extraction & discussed solutions

DAY THREE
Participants were asked to apply the epidemiological R packages and GREP database to their use cases and provide feedback.

DAY ONE OVERVIEW

On the first day of the workshop, participants were tasked with developing relevant use cases based on real-life epidemiological scenarios to assess the suitability of the Global Repository of Epidemiological Parameters (GREP) data schema.

Once participants had tested their own use cases, they then swapped with other groups to assess the GREP from a different perspective. Stemming from this exercise, four key areas of feedback were generated for improving the GREP. These improvements included:

- Upgrades to data schema functionality, including the requirement for the GREP to be stored in a relational format with the ability to easily interact with and filter data.
- Inclusion of additional variables not currently featured in GREP, for example, vaccine efficacy, intervention effects, setting, population and clade.
- Provision of a data dictionary for each of the variables included in GREP.
- Guidance on the process to input new data and whether new entries should include systematic reviews and grey literature (not currently included).

Next steps from this day of the workshop include the establishment of a GREP technical working group to work on the new technical requirements for the data schema in collaboration with the broader community.

Collaboratory

DAY TWO OVERVIEW

On the second day of the workshop, participants were invited to consider the main challenges surrounding parameter extraction. Small groups were tasked with identifying key bottlenecks and associated solutions. Groups were also invited to share methods or tools that could be used to improve the efficiency of maintaining GREP considering these challenges.

Each group presented their identified challenges, highlighting the key priorities of interest. The four topics decided upon for further development included:
1.   Disaggregation and context
2.   Human resources and incentives
3.   Quality and consistency
4.   Reporting guidelines

Participants were invited to join groups for the above priorities based on their personal choice. Groups were then invited to present their progress and plans to continue the work in the future.

The activities from this day were designed to identify areas of collaboration for community members to work on in the future. Participants from each group will be invited to engage in topic-specific follow up activities, such as virtual workshops, webinars and paper drafting.

DAY THREE OVERVIEW

On the third day of the workshop, the discussion focused on workstream 3: storage and use of parameters. The sessions on day 3 were built on the use cases developed on day 1.

Participants were asked to apply the epidemiological R packages ({epiparameter} and {epireview}) and GREP to their use cases. Participants were also asked to provide feedback on how easy the R packages were to understand, how easy it was to load the data, and how interoperable the R packages with epidemiological parameters were with existing epidemiological/outbreak analysis tools (e.g. {EpiEstim}, {EpiNow2}).

Key learnings highlighted were that participants found the packages easy to install but noted the need for more documentation and a data dictionary; recommended clearer functionality and consistent labelling for parameter identification; suggested a future dashboard with basic metrics and subset capabilities; and suggested visualizations and synthesis methods to facilitate parameter estimate selection.

∵ Collaboratory
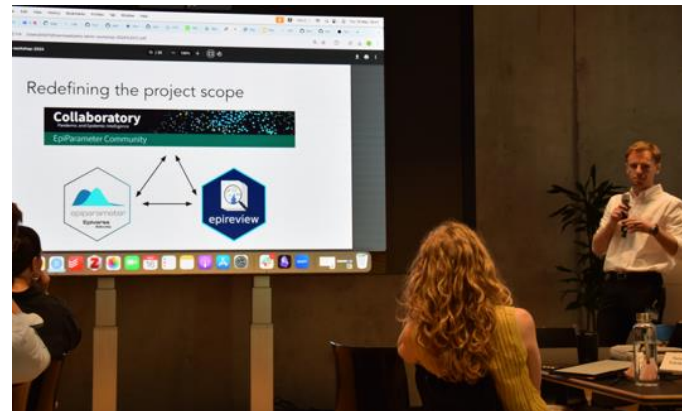
LOOKING AHEAD

At the time of writing (July 2024), the GREP minimum viable product (MVP) is in development, with planned socialization and presentation to the broader community for late 2024. Based on community feedback, the GREP will continue to be refined. The MVP is planned to include:
- Relational data schema format for a minimum of two diseases
- a simple data entry interface
- a data visualization tool for exploring and disaggregating the data
- tutorial material for connecting to the API for using the GREP in R packages.

The MVP will be made available to all community members by December 2024. Beyond this, the Collaboratory plans to productionize the GREP between January and June 2025 by onboarding developers to refine the data entry and validation process. A public-facing version of the GREP is then anticipated to be available by December 2025.

Longer term, ongoing work will seek to:
- ensure that the GREP is maintained to a high standard and is continuously adding in new priority diseases.
- Streamline GREP processes so they are replicable and scalable
- Implement AI solutions where appropriate

GREP MVP
Aug – Dec 2024

GREP Production
Jan – Jun 2025

GREP Deploy
Jun – Dec 2025

GREP Expansion
Ongoing

Collaboratory

PARTICIPANTS

Data.org
Chris Hartgerink

Epicentre
Anton Camacho
Paul Campbell

Epiverse-TRACE initiative at London School of Hygiene and Tropical Medicine
Adam Kucharski
Carmen Tamayo Cuartero
Dongxuan Chen
Joshua Lambert

European Centre for Disease Control
Bastian Prasse

Imperial College London
Anne Cori
Christian Morgenstern
Kelly McCain
Patrick Doohan
Ruth McCabe

John Hopkins University
Kate Grabowski

Oxford University
Abhishek Dasgupta

Public Health Agency of Canada
Irena Papst
Lisa Waddell
Michael Li

Telethon Kids Institute
Saras Windecker

UNC Gillings School of Global Public Health
Justin Lessler

University of Cape Town
Jared Norman

WHO HQ
Olivier Le Polain
Patricia Ndumbi

WHO Pandemic Hub
Chloe Rice
Elisa Di Maggio
Finlay Campbell
Julia Fitzner
Prabasaj Paul

Collaboratory