

Report: Second TWG Meeting on the Establishment of a Global Repository of Epidemiological Parameters

June 20, 2023

WHO Hub for Pandemic and Epidemic Intelligence

More than 50 participants from diverse backgrounds in public health intelligence and related fields convened for the second Technical Working Group (TWG) meeting on June 20, 2023. The focus of this meeting was to advance the establishment of a standardized data model designed to capture and present critical epidemiological parameters effectively. This initiative represents a crucial step forward in enhancing the consistency and accessibility of key epidemiological parameters for public health professionals, infectious disease modelers and researchers globally.

Report written by: Patricia Ndumbi, Megan Evans and Rebecca Lais
Formatting and Publishing: John Fass

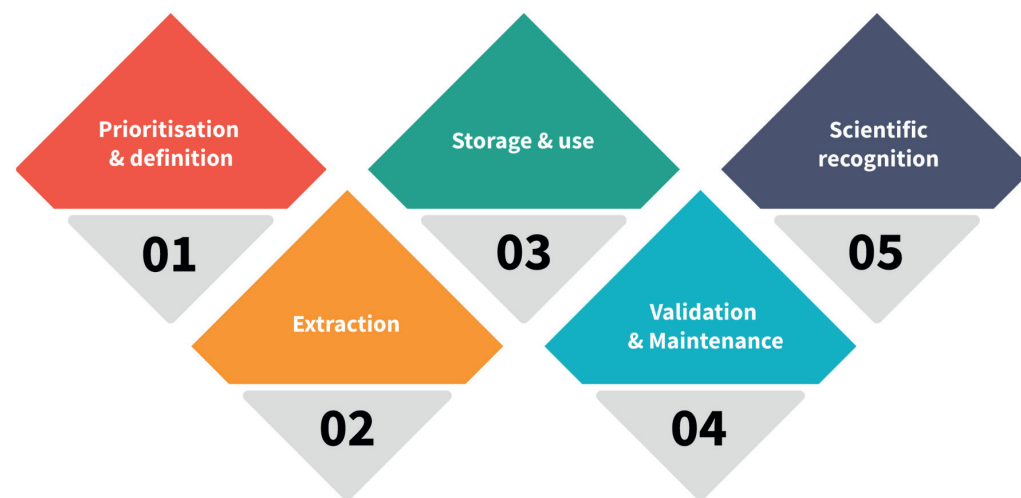
BACKGROUND

The epidemiological parameters community consists of a global collaborative working group coordinated by WHO, which aims to develop a global repository of parameters. Epidemiological parameters are used in mathematical models to help better understand the transmission dynamics of pathogens and to determine the potential impact of outbreaks in terms of morbidity, mortality, and geographical spread over time. By enabling faster and more transparent insight generation at the beginning and during an outbreak, the global epidemiological parameter repository will serve as an essential global public tool to inform and guide public health interventions designed to mitigate the spread of diseases and reduce their impact on affected populations.

The epidemiological parameters project consist of five work streams, 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 the current workshop was on the first work stream.

CONSULTATION AIM

The primary objective of this workshop was to seek consensus on a standardized data model for the extraction and representation of essential epidemiological parameters.



More details on each workstream can be found at the project website:
<https://worldhealthorganization.github.io/collaboratory-epiparameter-community/#/?id=our-workstreams>

CONSULTATION PROCEDURE

The workshop opened with an introduction on key objectives from Patricia Ndumbi (WHO) and Julia Fitzner (WHO), followed by a summary of the key outputs from the first technical working group meeting, delivered by Megan Evans (WHO).

Lisa Waddell (PHAC) then presented examples of different ways disease parameters are reported in the literature, emphasizing the need for accommodating diverse information and providing sufficient context in the data model. The presentation included examples of case reports and studies with different types of observations, prompting discussions on the minimum data requirements for inclusion in the database.

Participants then engaged in Mentimeter activity facilitated by Carmen Tamayo Cuartero (Epiverse-TRACE) to gather information on participants’ areas of expertise and the public health objectives they believed would benefit from accessing the epidemiological parameter database.

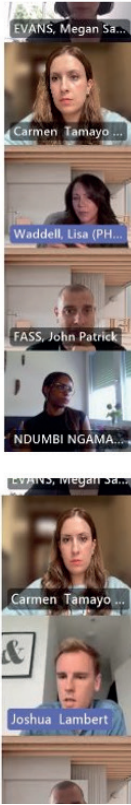
Joshua Lambert (Epiverse-TRACE) then gave a presentation discussing the use cases of epidemiological parameters within different public health objectives. This was followed by a Miro board exercise where participants were asked to identify required parameters, contextual information, and limitations related to capturing parameters for an example epidemiological use case selected to help prompt reflection and discussion.

Finally, Finlay Campbell (WHO) presented the results of the survey conducted to gather feedback on the proposed data model. The survey results were then discussed section by section with the participants.

Mpox	Source study	Incubation period	
Study 1	Case report	Malaise 3 days/Rash 3 days	Summary estimates can be provided in many ways. e.g. mean, median with different measures of variability. Some parameters may need additional information to specify the definition of the parameter.
Study 2	Summary of surveillance data collected at country level during early outbreak	exposure to first symptom onset: <ul style="list-style-type: none">mean 7.6 days (95% CrI: 6.2–9.7)median 6.4 (95% CrI: 5.1 – 7.9) and a standard deviation of 1.8 days (95% CrI: 1.6–2.2).95th percentile 17.1 days (95% CrI: 12.7–24.3) *Same data provided for exposure to rash onset	US data up to June 6, n=22 +18 cases from the Netherlands. Analysis included constructing a doubly censored dataset with assumed log-normal distribution of the incubation period and use of the Metropolis-Hastings Markov chain Monte Carlo algorithm for calibration.
Study 3	Summary of an outbreak from single city.	<ul style="list-style-type: none">mean 8.2 days (SD = 4.7).Only confirmed cases: mean 7.6 days (SD = 4.1) and range 2 to 20 days	368 cases Cologne Germany up to Sept 17 2022.
Study 4	Retrospective study of 16 historical outbreaks in Congo	exposure and symptom onset: <ul style="list-style-type: none">median 7 (range 0–17; IQR 1–13) days.	16 outbreaks, 2001-2021 with 327 persons investigated. Only 29 people had incubation period data.

Exercise example

- Example use-case
Understand transmission in the first two weeks of an outbreak in region x
- What parameters are required?
- Mean and range for a generation time or serial interval
 - Distribution parameters
- What other contextual information is required/desired?
- When and where was the data collected

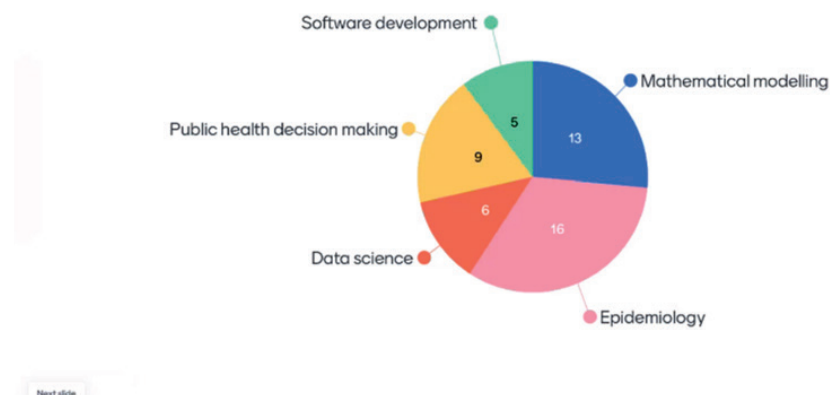


KEY OUTPUT AND DISCUSSION POINTS

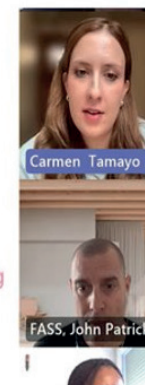
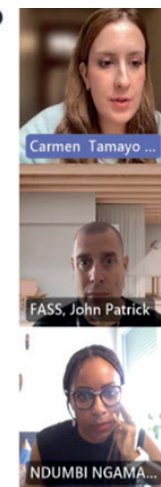
Mentimeter

The results of the Mentimeter survey illustrate the varied expertise of the TWG, spanning fields such as epidemiology; mathematical modelling; public health decision-making; data science; and software development. A wide range of public health objectives were listed by participants, with the most cited being 'evaluating the impact of interventions and control measures' and 'modelling disease transmission dynamics and estimating reproduction numbers'. Anticipated use of the parameters repository primarily revolved around essential functions like epidemic forecasting, risk assessment, contact tracing, and the formulation of public health strategies.

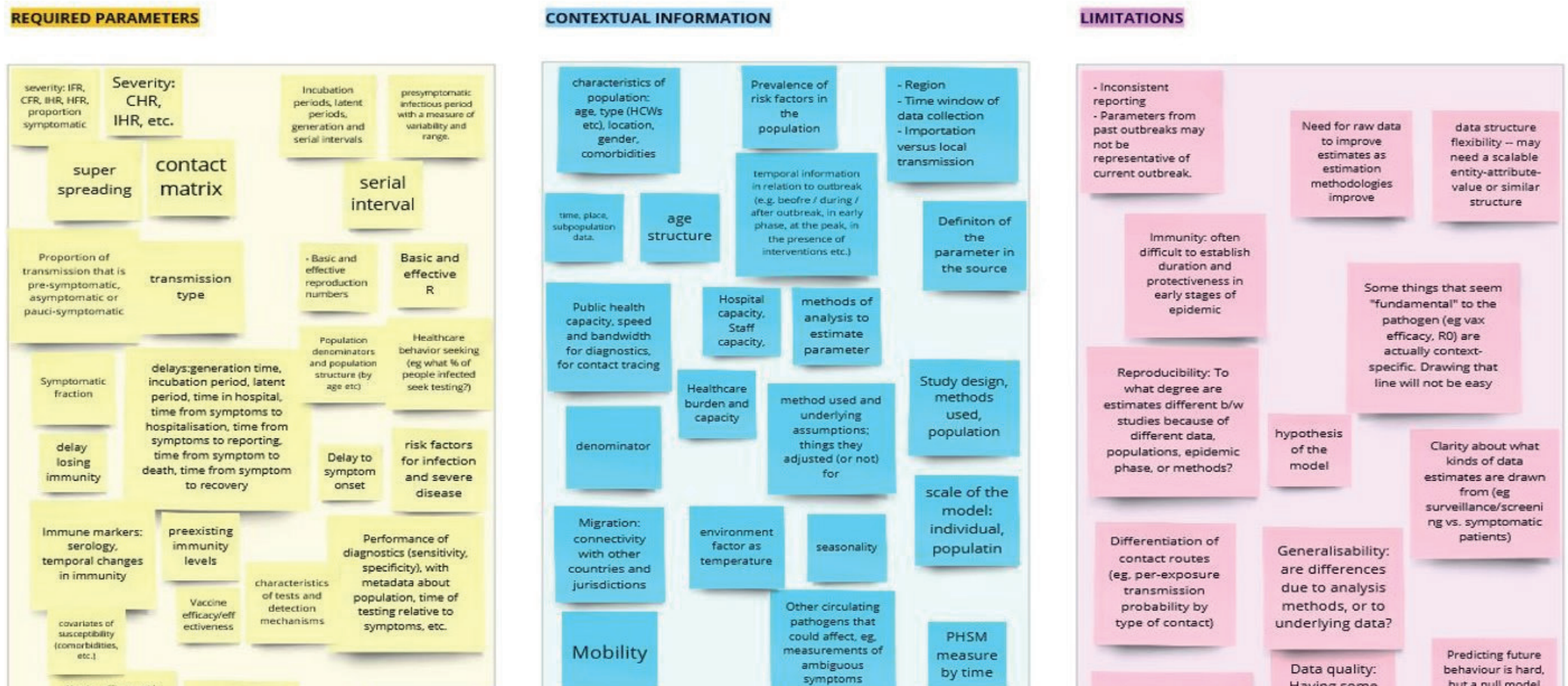
What are your main area(s) of expertise?



For what specific use cases will you access the epiparameter database? (e.g., contact tracing, hospital bed capacity, estimating severity)
30 answers



KEY OUTPUT AND DISCUSSION POINTS



Miro board illustrating key discussion points relating to required parameters, contextual information, and limitations.

Example use case: Quantifying the spread of a disease in a specific population or geographic area and assessing the impact of a public health interventions.

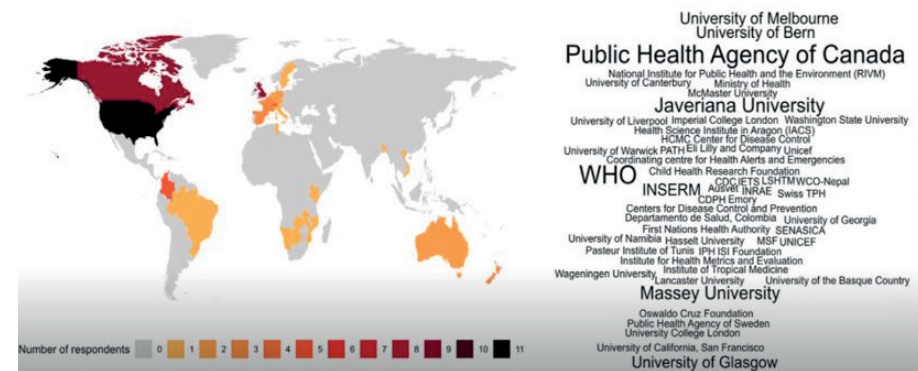
KEY OUTPUT AND DISCUSSION POINTS

Contextual information: The group emphasized the added value of incorporating contextual information alongside data, particularly underlining its significance during the COVID-19 pandemic. Although the type and extent of contextual information to include will require ongoing discussion, it was collectively agreed that temporal, geographical and socio-demographic factors were particularly relevant.

Data source: Significant deliberation centered around how to capture information about the source of extracted data. Participants stressed the importance of providing sufficient information for users to track and access the original data source. This might necessitate the inclusion of a Digital Object Identifier (DOI), a URL, and/or the name of the institution, organization, or research group responsible for the data.

Repository format: The discussion also revolved around the repository's structure and design to ensure optimal usability, including streamlined and standardized data entry processes.

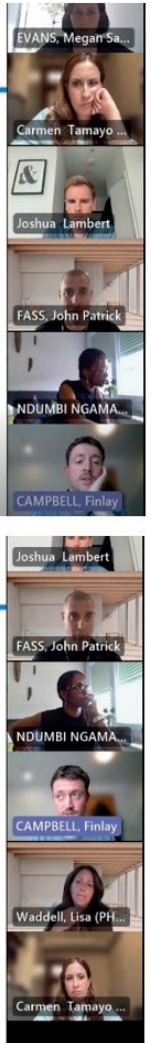
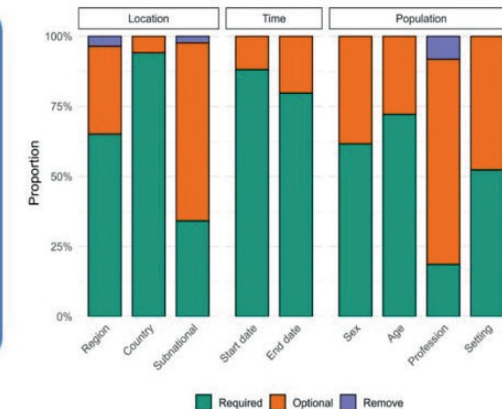
86 respondents from 25 countries and 53 Institutions



Context

Results & Feedback

- Basic location and time indicators considered necessary by most respondents
- Context is important for some parameter types (e.g. transmission rates highly context specific)
- Population data needed for disaggregated parameters
- Large number of optional population covariates that **vary between parameters and studies** and cannot be specified clearly ahead of time



KEY OUTPUT AND DISCUSSION POINTS

Pathogen, disease and vector designations: There was agreement that a predefined list of pathogens could help streamline and standardize the data entry process. Many participants echoed the importance of maintaining flexibility in the disease and vector designation to avoid undue constraints and accommodate diverse needs.

Parameter estimates: The participants engaged in a productive discussion about providing a more comprehensive and nuanced data representation. One suggestion involved introducing an “uncertainty” field for estimates and an “other” category for additional contextual data. Additionally, the idea of incorporating a binary flag to distinguish between Frequentist and Bayesian approaches was discussed, with the understanding that interested parties could access more detailed information as needed.

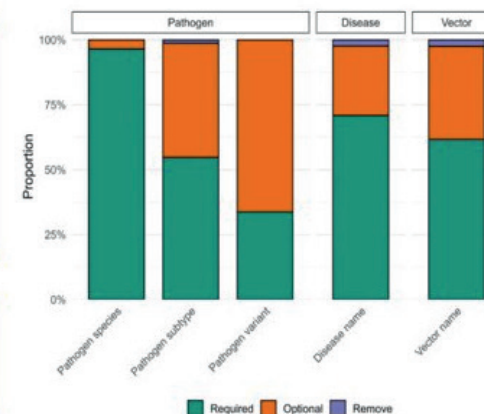
Pathogen, disease and vector designations

Results & Feedback

- Importance of subtypes, variants depends on pathogen (e.g. influenza vs measles)
- Higher-level classifications like pathogen family could be useful (e.g. flavivirus, alphavirus)
- Automated field-completion if pathogen species is provided?

Discussion

- Should pathogen species be selected from a list or free-text?



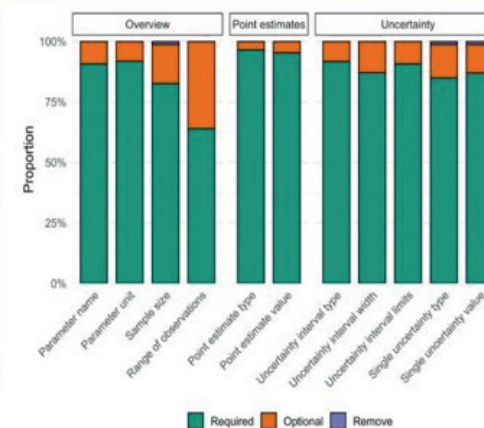
Parameter estimates

Results & Feedback

- Clear parameter definitions needed
- Uncertainty listed as a “required” field by >90% of participants

Discussion

- Should a measure of uncertainty be necessary for submission?
- Can/should parameter names be selected from pre-specified list?



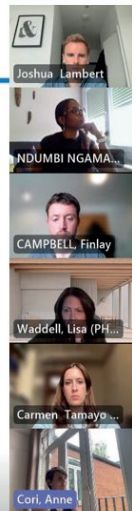
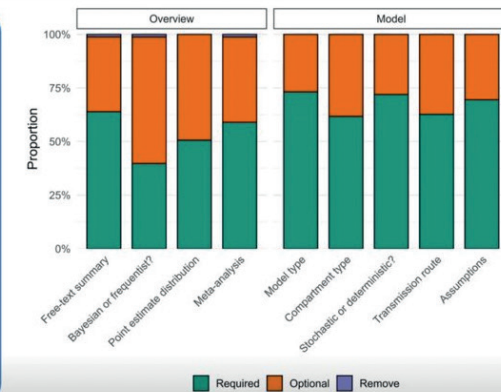
KEY OUTPUT AND DISCUSSION POINTS

Methods: It was recognized that multiple modeling and statistical approaches could be used to estimate parameters. The participants suggested the inclusion of a checkbox to differentiate between parameters estimated from a model vs. parameters estimated from observed data. Concerns were raised about how to effectively extract a parameter that is estimated multiple ways in a single study. Effective documentation and appropriate capture of the estimates and measures of variability were considered very important.

Methods

Results & Feedback

- Specify study type (e.g. descriptive, observational study)
- "Models have unique features that cannot be specified according to predefined categories: simple categorization may be misleading."
- "The ultimate reference must be the original paper. Unclear how useful it would be to collect these details in free-text fields, unless you can decide upon a short list of valid values."
- "Meta-analysis requires a lot more information to be extracted. I am not sure meta-analyses have a place in this database."



OUTCOME AND FEEDBACK

Outcome

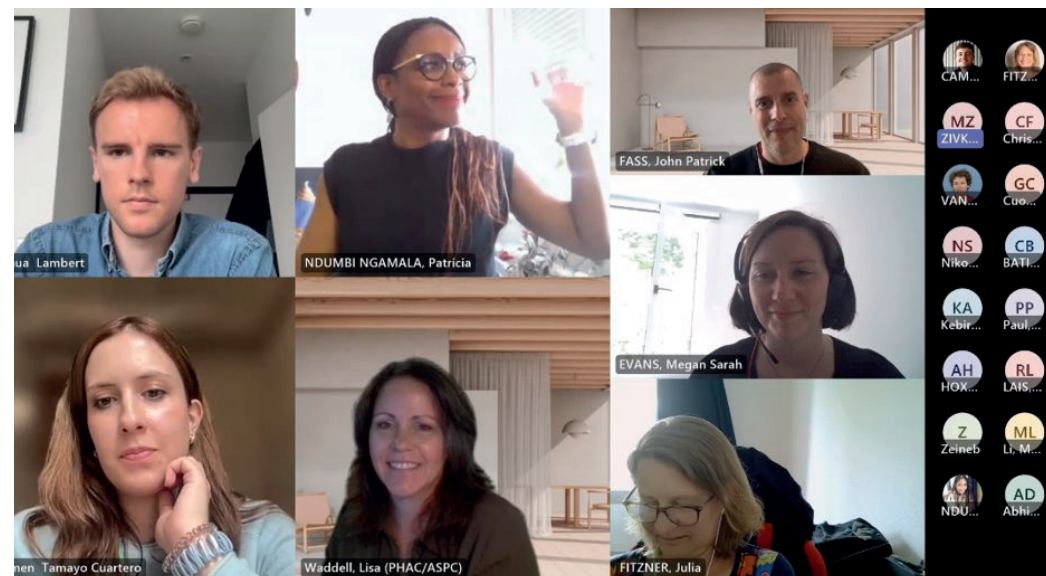
The meeting's key outcome was a shared commitment to advancing public health intelligence by establishing a standardized data model, enriching it with contextual information, ensuring transparent data sourcing, and structuring the repository format for optimal usability. These collaborative efforts underscore the TWG's dedication to improving public health decision-making, forecasting, and risk assessment.

Feedback

The meeting received positive feedback from participants, who applauded the forward-thinking approach and recommended incorporating a testing phase for the initial versions of the database.

As one participant aptly put it: "Great to see all this up-front thinking and solicitation, and I hope that there will be an iterative/agile process of feedback and improvement as the first versions of this database are rolled out".

Moreover, participants commended the interactive nature of the sessions, recognizing that they effectively addressed the most critical discussion points throughout the workshop. Some participants suggested experimenting with a more concise format for future sessions to maintain focused engagement from start to finish.



PARTICIPANTS

Imperial College London

Gina Cuomo-Dannenburg
Anne Cori

Institut Pasteur de Tunis

Ben Miled Slimane

Javeriana University

Candida Rosa Diaz Brochero
Zulma Milena Cucunubá Pérez

London School

of Hygiene & Tropical Medicine

Sebastian Funk

Global.health

Abhishek Dasgupta

University of Capetown

Sheetal Silal

Public Health Agency of Canada

Tricia Corrin

University of Melbourne

James McCaw

WHO HQ

Cedric Stephane Bationo
Ana Hoxha
Henry Laurenson-Schafer
Bernadette Mirembé
Karl Schenkel
So Yoon Sim
Katelijn Vandemaele

Pandemic Sciences Institute, University of Oxford

Christoph Fraser

WHO Pandemic Hub

Julia Fitzner
Rebecca Lais
Oliver Morgan
Prabasaj Paul
Rithika Sangameshwaran
Mia Zivkoviz

ORGANIZERS

WHO Pandemic Hub

Finlay Campbell
Megan Evans
John Fass
Patricia Ndumbi

Epiverse-TRACE initiative at London School of Hygiene & Tropical Medicine

Adam Kucharski
Joshua Lambert
Carmen Tamayo Cuartero

Public Health Agency of Canada

Lisa Wadell