

Concept Note

Strategic Recommendations for Future Cohort Research to Support Epidemic and Pandemic Response

VERSION 1, 30 April 2025

Authoring institution: Amsterdam Institute for Global Health & Development

On behalf of the END-VOC Consortium

Confidentiality status: Public

1. Introduction

Background and Rationale

This concept note concerns cohort research, understood as observational epidemiological studies of population groups who often share a common exposure, outcome, and/or other relevant characteristic. This concept note addresses both prospective and retrospective cohort studies of all sizes.

Evaluating the global circulation and impact of infectious and environmental diseases through multinational cohort studies is essential for public health protection, yet also challenging. Common challenges include disparities in cross-national and cross-institutional healthcare infrastructure and genomic sequencing capabilities, and cohort design and harmonization. Furthermore, ethical, legal, and administrative challenges around data sharing across borders hinder timely integration and analysis of cohort data.

These complexities of cohort research are highly relevant to pandemic preparedness and response. In the context of Covid-19, real-time evidence on variant behaviour—including transmissibility, immune evasion, and treatment resistance – was critical for responsive public health decision-making.

Evaluating variants of concern through cohort studies globally enables both anticipatory and reactive measures. Therefore, the END VOC consortium members seek to **formulate strategic recommendations for future cohort research** to inform future epidemics and pandemics.

Purpose and Intended Audience

The purpose of this concept note is to collate the main strengths and weaknesses reported by members of the END-VOC consortium when conducting cohort research during the Covid-19 pandemic. This Concept Note also offers key future opportunities and threats identified by END-VOC consortium members.

Based on these insights, this concept note formulates key discussion questions that serve as a departure point for further exploration with the END-VOC consortium members to supplement and refine, and later for discussion with the wider stakeholder network of cohort research (e.g. WHO, CEPI, ECDC, EMA, and other cohort consortia such as ReCoDID and ORCHESTRA).



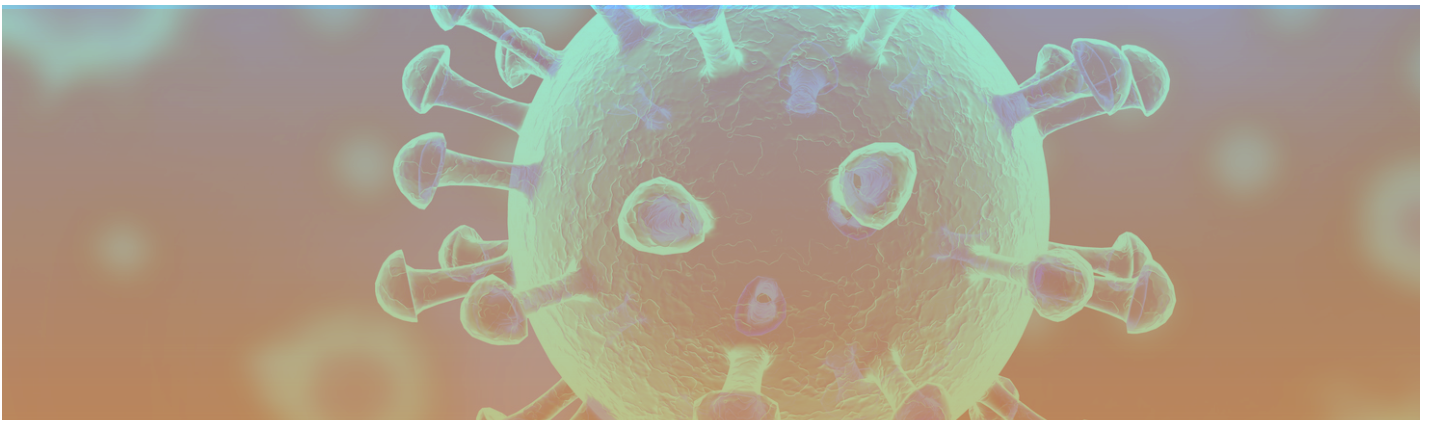
2. Methodology

Stakeholder views were gathered through **seven semi-structured interviews** with researchers actively involved in cohort-based COVID-19 research within the END-VOC Consortium. Each interview, conducted in March 2024, lasted between 20 and 40 minutes and included one to three respondents from each of seven national teams – two from low- and middle-income countries (LMICs) and five from high-income countries.

The interview consisted of open-ended questions exploring:

- **expectations and experiences** related to data collection, sharing, access, collaboration and interoperability, including main challenges and barriers;
- **the adaptability of local regulations**, institutional procedures, and research practices to these challenges and barriers;
- **reflections on improving cohort research design** and conduct for future pandemics.

The interviews were transcribed, and pseudo anonymised, then analysed for strengths and weaknesses experienced in Covid-19-related cohort research, as well as opportunities and threats to future cohort research for pandemic preparedness.



3. Analysis

Strengths

- **Established large-scale national cohorts** offer a ready infrastructure for pandemic research, enabling rapid data collection without new recruitment. Such cohorts may also have rich longitudinal health and lifestyle data, enabling longitudinal analyses of variables that would otherwise be absent from recently established cohorts.
- **National health registry law and an existing model and online platform** for international data sharing and research collaboration (including ethical approval requirements) provided immediate clarity about the requirements for (cross-border) data sharing, as well as a secure, remote online platform for granting data access to researchers abroad.
- **Maintenance of 'in house' core infrastructure** for surveillance and genomic sequencing.
- **Availability of a 'trusted research environment'** for secure access to cohort data. 'Trusted research environments', such as UCL's Data Safe Haven, offers both the software/hardware and governance framework for pooling, harmonising and sharing data across different research projects. A TRE allows for secure remote access to sensitive cohort data, promoting high standards for data privacy and trust between the people, projects and governance procedures involved.
- **Development of best practices and templates** with legal experts for aligning research practices and agreements for data sharing with GDPR.



Weaknesses

- **Ethical approval delays.** Ethical approval requests were handled quicker during the Covid-19 pandemic when more assessments were conducted with a faster turn-around time compared to standard practice by ethical committees after the pandemic. However, some countries also require additional ethical approval by a national authority, which further delays the research.
- **Existing national cohorts may be designed for other objectives** or target populations, omitting important sub-populations or variables of relevance to variants of concern. These omissions might necessitate creating separate (smaller) cohorts.
- **General dearth of funding for cohort studies.** Less funding is available for non-interventional cohort studies (in particular, those sponsored by public authorities) compared to studies testing pharmaceutical interventions (and funded by a manufacturer). Lower funding for the former constrains the cohort size. Limited funding also restricted more advanced data collection (e.g., through neurological exams for long COVID), preventing richer data capture.
- **Heterogeneous data governance laws and standards,** including heterogeneous interpretations of the recently adopted EU GDPR, across different countries and institutions (e.g. hospitals) for data sharing, management and use pose challenges for integrated global cohort research.
- **Lack of complete, harmonised, and good quality data** when research was initiated for the first time during an emergency that left little time for preparation.
- **Different capacities of the institutions involved** (e.g. financing, facilities) fragmented the research designs and methods that could be harmonised across the consortium.
- **Limited infrastructure in and transportation to remote research sites:** Periodic internet outages and unreliable connectivity disrupted access to online platforms needed for research collaboration. Many research materials needed to be imported due to limited local availability and managing (international) transport logistics to a remote site caused further delays.
- **Data access challenges.** Some institutions required on-site data access only (e.g. researchers needing to travel to a foreign country), limiting cross-border analysis. Moreover, the inability to download or export data from TREs slows research.
- **Digital data collection tools may not be acceptable to respondents,** excluding harder-to-reach populations. Elderly populations and technologically underserved groups were harder to include due to their difficulty using digital survey tools.
- **Time investment to tailor standardised data collection tools.** Standardised international questionnaires (i.e. designed by a foreign partner institution) need tailoring to make it comprehensible in the local language and culture, consuming time and effort.
- **Standardised data collection tools may not yield in-depth understanding** of complex conditions. Reliance on questionnaires limited insight depth. Lack of qualitative interviews or clinical evaluations hindered understanding of complex conditions such as long COVID.



Opportunities

- **Leverage established cohorts.** Collaboration with institutions that already have cohorts can reduce setup time for research projects.
- **Standardise and pre-approve international data sharing agreements.** Pre-approved data sharing templates between institutions in the EU and (at least) partner countries eligible to receive EU framework programme funding for research would reduce bureaucratic delays. These data sharing agreements should adhere to the applicable laws for data protection and privacy (e.g. GDPR). It could be explored whether the European Commission could facilitate these template data sharing agreements.
- **Develop a harmonised database** to guide basic anonymised data collection that can be shared internationally. This database can contain few variables but still allow for simple modelling, which would kickstart data sharing and collaborative research in a large international consortium while data protection issues are worked out for the more complex or sensitive data.
- **Establish an international base for hardware and online platform for data sharing** where existing anonymised cohort data can be pooled and stored, independent of (and ideally before) any global health emergency. In the event of an emergency, this data can be easily accessed for analysis.
- **Scaling of existing TRE Models through training and dissemination.** Successful platforms like UCL's Data Safe Haven can serve as scalable templates for other countries or projects. Workshops and toolkits on TRE setup, data sharing policy, and GDPR compliance could empower institutions and reduce friction.
- **Provide targeted legal support for researchers on data protection.** Making legal experts (e.g., on data protection) available to research teams could help expedite the establishment of agreements for data pooling, storing and sharing among projects involved, and overall streamline (cross-border) data operations. Templates pre- approved by legal experts and (where possible) institutions for data sharing and ethical agreements could reduce inter-project negotiation time dramatically.
- **Explore the potential of computer / AI generated data** specific to different countries' genetic profiles that could be used for modelling and analysis without having to share real patient data.

Threats

- **Difficulties securing qualified human resources** for cohort research, particularly when projects are of a short-term nature.
- **Unpredictable changes to ethical clearance procedures** and requirements during pandemic response and recovery.
- **Unforeseen disruptions in international trade and transport** to source research materials can impact project timelines and lab testing.



4. Questions for Discussion

The following questions are based on the insights reported above and serve as guiding questions for further discussion with the END-VOC consortium members and, later, external stakeholders (where relevant).

1. What support structures and tools (e.g., pre-approved legal templates) can help harmonize data governance standards across institutions and countries to enable timely collaborative research? Where are those tools available or who should develop them/how should they be developed?
2. What support structures, tools, and practices can help address divergent (and lengthy) ethical approval procedures across countries and institutions?
3. How do 'trusted research environments' (TREs) enhance or hinder data sharing, and how can TREs be optimised (e.g. through training) to support international, collaborative science while preserving data security?
4. What contingencies can be planned at the stage of project design or onset to mitigate future threats from trade disruptions, technological outages, or fragmented institutional capacities that delay cross-national cohort-based research efforts (particularly during health emergencies)?
5. How should research teams address the trade-offs between standardized international data collection tools and the need for cultural localization and richer qualitative insights into complex conditions?
6. In what ways can existing cohort infrastructures be adapted or supplemented to include underrepresented populations and emerging variables relevant to new health threats like variants of concern or long COVID?
7. Is the potential to use AI-generated or synthetic data for international cohort analysis desirable from a scientific, ethical, and societal perspective? Should this approach be integrated into cohort research projects, and if so, how?
8. Given the underfunding of non-interventional cohort studies, how can the value of longitudinal, observational data be better communicated to funders to secure more sustainable investment?

5. Next Steps

This concept note summarises the **key strengths and weaknesses** experienced by END-VOC consortium members during COVID-19 cohort research, while also highlighting **future opportunities and threats** identified for improving such efforts.

The result is a non-exhaustive list of discussion questions that will serve as the basis for an internal discussion with consortium members at the END-VOC 2025 General Meeting. At that time, the procedure for **formulating strategic recommendations** will be decided.

