Ethics Protocol*

Articulating epidemiological data harmonization initiatives as practical and collaborative experiences

2025-04-29

Project Title: Articulating epidemiological data harmonization initiatives as practical and collaborative experiences
Submitted Materials: zackbatist.info/CITF-Postdoc/irb-docs.pdf
Principal Investigator: Zachary Batist
Protocol: 25-01-057
Submitted: 2025-01-30
Approved: 2025-03-03
Amended: 2025-05-xx

Recruitment and consent

Will this study involve recruitment of human study participants?

- ☒ Yes
- ☐ No

How are potential study participants identified and/or recruited to the study? Explain how potential participants are identified or introduced to the study, and who will recruit participants. Will the investigator/s require any special permissions or access to the target population e.g. clinic access, patient registries or records, mailing lists, community access?

The study draws from semi-structured interviews with 10-15 individuals from a single case: the Covid-19 Immunity Task Force (CITF) Databank. The CITF was an initiative whose mandate was to catalyze, support, fund and harmonize knowledge on SARS-CoV-2 immunity for federal, provincial, and territorial decision-makers in their efforts to protect Canadians and minimize the impact of the COVID-19 pandemic. The CITF Databank provides continued support to

*This is an automatically generated PDF. Refer to the project website for continuous updates at https://zackbatist.info/CITF-Postdoc.

the research community by centralizing research data provided by CITF-funded studies and by making the data accessible for extended research. This study draws from interviews with individuals who lead, support or participate in the CITF Databank's operations, including professional researchers, research trainees and administrative and technical support staff.

The case was selected partially out of convenience, but this does not discount the prospective analytical value that it affords. As a community-led initiative, the CITF Databank presents an opportunity to investigate how epidemiologists balance the values deriving from their own epistemic culture with the challenges of coordinating collective efforts; it has established an explicit governance structure, which provides me with terms and concepts to examine through a critical lens; it comprises people working in a multitude of roles, many of whom are locally available; and it is a venue where multiple relevant epistemic conflicts and challenges manifest themselves, which are rich sources from which a qualitative researcher may ascertain competing and complementary value regimes. Interviews will be oriented by the study's goal to document processes of reconciling different stakeholders' interests as they converge in the formation of a common data resource.

The study participants are individuals who are affiliated with the CITF Databank. Participants are identified through consultation with project leaders, which helps determine the value that an interview will bring and to help make preparations prior to the interview. The principal investigator will be clear with project leaders that they should not pressure prospective participants to participate in the study, and that their participation should be treated as separate from their regular duties. The principal investigator will then approach the recommended individuals to introduce the study and its objectives and to invite them to participate as research subjects. If these individuals express interest in participating in the study, the principal investigator will schedule a time to sit for an interview. Some interviews may be conducted remotely using internet-based video conferencing software, depending on participants' availability.

Describe the consent process. If alternate processes for seeking consent are planned (e.g. verbal, online, waiver), please provide a rationale and outline the procedure of obtaining and documenting consent and/or assent, where applicable.

Once individuals express their interest in participating, participants will be provided with an informed consent document that outlines in more detail the goals of the study, the roles of the participant, how they will be recorded, how data pertaining to them will be retained, and the potential risks and benefits pertaining to their involvement. This document will also describe how participants' personally identifiable information will be managed and used. Participants will be asked to read and sign the document in order to obtain written informed consent. For interviews that will be held remotely using internet-based video conferencing software, participants will asked to send their signed informed consent documents in PDF format to the principal investigator. At the start of each interview the researcher will reiterate participants' rights and ask them to orally reaffirm their consent before proceeding.

Is there a relationship between the study participants and the person obtaining consent and/or the principal investigator/s?

☒ Yes
☐ No

If yes, please explain the nature of the relationship, and outline the steps that will be taken to avoid the perception of undue influence.

The project that serves as the case in this research is the Covid-19 Immunity Task Force (CITF) Databank, which the principal investigator currently serves as a postdoctoral researcher. The interviews will remain focused and limited in scope, and will not touch on matters relating to other aspects of their work. Moreover, prior to and throughout their involvement as research participants, frank and open discussion will be encouraged regarding collective expectations and to articulate the boundaries between participants' relationships with the principal investigator as colleagues and as research subjects.

Risk-benefit assessment

Describe the foreseeable risks to study participants. What risks are attributable to the research, including cumulative risks? Which risks are participants normally exposed to in the course of their clinical care or in their daily activities as they relate to the research questions/objectives?

Participation in this study does not involve any physical, psychological or legal risks. However, the principal investigator will be asking participants to share detailed information about their work practices and work relationships, and there is minimal risk that their responses may disrupt or complicate their professional reputations. Since this study is based on a single case whose identity has already been publicly established, it is not feasible to fully obfuscate the identities of prospective study participants.

To mitigate against the minimal risk of reputational harm, the principal investigator will provide participants with opportunities to shield their identities in public presentations of the findings when the risk is more clearly apparent. When the principal investigator senses that there may be heightened risk to reputational harm, he will approach the potentially affected participants and invite their input on how to proceed. Alternatively, the principal investigator may take action to shield the individuals' identities independently of their added input. This aligns with this study's emphasis on co-constituting knowledge about epidemiological research practices and value regimes through collaboration with the participants whose work will be examined.

What procedures are in place to monitor and assess participant safety for the duration of the study?

Prior to each interview, and as part of the procedure for obtaining informed consent, participants will be informed about their rights. The principal investigator will invite participants to reflect on their responses and express their desires to obfuscate their identities when reporting specific responses.

The principal investigator has established a consistent habit of maintaining reflexive notes regarding his encounters with research subjects, which include notes about signs of apprehension expressed by study participants. These notes inform his sensitivity to the degree of risk and support his decisions regarding whether to approach participants regarding potentially sensitive responses.

Describe the potential benefits of the study for: (1) the study participants; (2) the population under investigation, and (3) the field of research.

This study contributes to the development of better epidemiological data-sharing infrastructures by articulating the social, collaborative and discursive aspects of a well-established and exemplary data-sharing initiative. In particular, the study demonstrates how these factors relate to, overlap with or conflict with technical, institutional and epistemic factors. By explicitly framing data sharing as a social and collaborative activity, the study informs the design of more effective data-sharing infrastructures that better support the contextualization of data and enhance data's value in contexts of reuse. This work therefore documents and reflects on the ways in which epidemiologists mobilize distributed records, and helps develop practical solutions that enable more effective collaborative workflows. Additionally, this study may directly benefit participants by framing the experiences they address during interviews in ways that they might not have otherwise considered, thereby encouraging greater reflexivity in their own work.

Privacy and confidentiality

Please describe the measures in place for meeting confidentiality obligations. How is information and data safeguarded for the full cycle of the study: i.e. during its collection, use, dissemination, retention, and/or disposal?

Commitment to maintaining full confidentiality is not warranted given the minimal risk profile associated with this study, and since the vast majority of recorded data will not contain sensitive information. At the same time, the principal investigator is committed to working with study participants to ensure that information that may carry elevated risk is handled with care and in accordance with participants' wishes.

In situations whereby a participant desires to obscure their association with information they provided, the principal investigator will have access to unobfuscated versions of those responses containing information that may identify the participant. These unobfuscated records, which may include audio and video records of interview sessions, as well as unedited transcripts and textual notes containing information that may reveal the participants' identities, will be kept

in secure and encrypted media, and destroyed within five years of concluding the study, which provides sufficient time to revisit the data and produce additional research outputs. However, transcripts edited in a manner that reduces potential of elevated risk may be kept, published and archived.

The study is committed to adhering to fundamental data security practices, including those specified in McGill University's Cloud Directive, which regulates the curation of sensitive research data. Physical records will be kept in a locked drawer in secure workspaces, either at McGill University's School of Population and Global Health or at the principal researcher's home office. Digital records will be stored on encrypted and password-protected drives and on secure servers approved or managed by McGill University under the Cloud Directive.

Recordings of remote interviews conducted using internet-based video conferencing software will be made using the software's built-in recording tools. Only video conferencing software approved by the Cloud Directive will be used.

If a contracted cloud/storage service provider or online survey tool is used, provide information on the service provider's security and privacy policy, location of its servers, data ownership, and what happens to the stored data after the contract is terminated. For more information, please consult the University's directive.

The study uses file-sharing software hosted by the CITF Databank at McGill University's School of Population and Global Health to backup all files maintained for this study. These backups include unobfuscated versions of files containing potentially sensitive information. The software used to manage these backups is managed by McGill University and has been approved for storing sensitive research data by the Cloud Directive.

The study uses the secure GitLab instance hosted by the Surveillance Lab within the Clinical and Health Informatics Research Group at McGill University to store and track changes to sensitive research data. This software is managed by McGill University and has been approved for storing sensitive research data by the Cloud Directive.

The study maintains a website where the principal investigator shares documentation that supports the study and reflects on the work as it progresses. This is hosted using GitHub Pages and is backed up using Dropbox. No sensitive research data passes through these services.

Recordings of remote interviews conducted using internet-based video conferencing software will be made using the software's built-in recording tools. Only video conferencing software approved by the Cloud Directive will be used.

Please explain any reasonable and foreseeable disclosure requirements (e.g. disclosure to third parties such as government agencies or departments, community partners in research, personnel from an agency that monitors research, research sponsor, the REB/IRB, or regulatory agencies).

No disclosure requirements are foreseen.

If there are plans for retaining participant and/or study data for future use, please describe the context for its use, requirements for potentially re-contacting study participants and consent, and how the data will be stored and maintained for the long term.

Data will be published in compliance with ethical standards for sharing open social science research data. In particular, care will be taken to disassociate the identities of participants with statements they may have made that carry elevated risk.

The principal investigator may invite select participants to collaborate on papers presenting the findings or advocating for actions based on the project's findings.

Secondary use of data studies: if the study involves data linkage, please describe the data that will be linked and the likelihood that identifiable information will be created through the linkage.

This project does not rely on data deriving from other studies. The data may be reused in related work being undertaken under the same grant and by those who access the openly accessible data after they are published.

Managing conflicts of interest

Conflicts of interest do not imply wrong-doing. It is the responsibility of the investigator to determine if any conflicts apply to any person/s involved in the design and/or conduct of the research study or any member of their immediate family. Disclose all contracts and any conflicts of interest (real, perceived, or potential) relating to this research project. Conflict of interest may also arise with regard to the disclosure of personal health information.

- ☒ Not applicable. There are no conflicts of interest to disclose.
- ☐ Yes, there are conflicts of interest to disclose.

If yes, please describe the conflicts of interest (real, potential, and perceived), and the procedures for managing declared conflicts. Not applicable.