

COVID-19 Patient Data Consortium for Research

COVID-19 Patient Registry (CPR) on REDCap

Data Governance

I. DOCUMENT SUMMARY

This section summarizes the key points and provisions of this Data Governance document.

OVERVIEW

The COVID-19 Patient Data Consortium for Research or COVID Patient Registry (CPR) is a centralized repository for COVID-19 patient data that is a research resource intended for the study of COVID-19 disease and its sequelae. The registry allows data aggregation and interoperability across research institutions and health facilities to promote research collaboration, data sharing, and accelerate discovery and technological advances in COVID-19 healthcare.

The repository is stewarded by the Center for Informatics at the University of San Agustin, which is in-charge of the development and implementation of the informatics infrastructures requirements and enforcement of data governance.

KEY PROVISIONS

- Data in the repository is governed by this data governance document, which has been discussed and agreed upon by consortium members, to guarantee usability (i.e., quality and consistency), as well as data privacy, confidentiality, and security (protection and accessibility).
- Data contributed to the repository by consortium members (hospitals) are intended for research use only; any use that is not for research is prohibited and sanctionable and prosecutable under the full weight of Philippine law.
- Data to be used for research will be de-identified to protect patient privacy.

II. MISSION, VISION, AND GOALS

Our mission is to establish the COVID-19 Patient Data Consortium for Research, or the COVID-19 Patient Registry (CPR), to respond to the growing need for high-quality research-ready patient data on COVID-19 and to utilize the data in research on disease phenotypes, clinical course, risk factors, patient care, treatment/drug response, diagnostics, vaccines, health outcomes, AI/ML, and many other topics.

With the CPR, we envision creating value for patients, researchers, hospitals, physicians, and public health using high-quality research-ready data on COVID-19.

III. DECISIONS

This data governance document shall be the guiding policy for the planning and implementation of all data-related activities of the consortium.

IV. CPR TEAMS/DUTIES AND RESPONSIBILITIES

The CPR consortium is currently composed of the following institutions and health facilities: University of San Agustin - Center for Informatics (USA-CFI), National Center for Mental Health (NCMH), Makati Medical Center (MMC), Lung Center of the Philippines (LCP), The Medical City (TMC) Group of Hospitals, and Philippine General Hospital (PGH). Future consortium members from Visayas and Mindanao are also being recruited. All the hospitals in the consortium are responsible for contributing their full dataset on COVID-19 patients.

The **Data Governance Committee (DGC)** is composed of at least one representative from each participating institution. The DGC will continue to develop and update the data governance of the consortium. The DGC will also serve as the **Research Program Committee (RPC)** which will set the research agenda. The data steward is the CFI, which will enforce the provisions of this data governance.

The current members of the DGC are:

1. Romulo de Castro, PhD
2. Salvador Eugenio Caoili, MD, PhD
3. Fresthel Monica Climacosa, MD, PhD
4. Lyre Murao, MD
5. Roland Gilbert Remenyi, PhD
6. Jesus Emmanuel Sevilleja, MD, MPhil
7. Marissa Alejandria, MD, MSc
8. Janice Caoili, MD
9. Sullian Naval, MD

V. COMMUNITY

The stakeholders of the COVID-19 Patient Registry are the patients, hospitals, clinical researchers and scientists, and physicians. The users of the CPR are the researchers, data contributors (from each consortium member institution), data curators, and the data stewards/database administrator (CFI). The CPR aims to benefit public health through research and information targeted to health policymakers, consortium members, and their respective communities.

VI. DATA OWNERSHIP, ACCESS, AUDIT, AND DELETION

- The data contributed to the patient registry remains the contributors' property, but they are contributing it to research.
- The consortium aggregates, curates, and makes the data available for research through its Research Program.
- Access to data by authorized users is 24/7 unless the database server is under maintenance or during extraordinary circumstances (e.g., disasters and other emergencies).
- To maintain and secure the data, the data steward may access the data; however, the steward cannot use or modify it unless the data owners grant permission.
- Institutions that leave the consortium can have their data deleted from the registry.

VII. COVID-19 PATIENT REGISTRY USE AND DATA USE POLICIES

The use of the patient registry will be in accord with the Research Program's objectives and aims defined by the consortium. Requests for the use of the data will be subject to the approval of the RPC. Data accessed from the patient registry should not be used for research and activities which are explicitly prohibited by the laws of the Republic of the Philippines.

RESEARCH PROGRAM

Consortium members will receive first priority to do research using the CPR's COVID-19 patient data. The second priority will go to unaffiliated researchers whose topics contribute to the country's COVID-19 health response.

The topics identified by the consortium as priority research areas are:

- Clinical course and disease severity
- Risk factors
- Comorbidities and co-infections
- Hematologic characteristics
- Immunologic status and vaccination history
- Imaging characteristics (X-ray, CT, Ultrasound, etc.)
- Neurologic
- Psychiatric and other complications
- Various treatment modalities such as hemoperfusion/dialysis
- High flow oxygen
- Convalescent plasma
- Herbal Medicine use

VIII. SECURITY PROGRAM (CPR - Security Program)

Data Storage

Covid-19 Patient data is aggregated and stored in a clinical data warehouse which is the patient registry developed in the REDCap (Research Electronic Data Capture) platform that has been implemented by the [REDCap Users' Consortium](#) (RUC), of which the Center for Informatics (CFI) is a founding member and maintains the CFI instance. The patient registry is also covered by the [data governance of the RUC](#). Also, the CFI's abiding [principles](#) of data stewardship apply. In developing the patient registry, we incorporated data audit trails (data created by, data created on, data last updated by, data last updated on) as an additional security measure. REDCap's data security protects the data with [CPR - User Passwords](#) password protection, user management & rights control, and compliance with HIPAA (the United States' Health Insurance Portability and Accountability Act - there is a log of user activity).

The server where REDCap is installed also provides several security features to protect data within, including password protection, advanced DDoS mitigation, secure socket layer, advanced performance analysis, firewall rule configuration, server hardening, database backup, security audit, intrusion prevention, content recovery, hosting restoration, email spam review, malware & blacklist removal, database only recovery, setup local backup.

A mirror database using the Knack platform is utilized to combine all the records in the REDCap database in order to deduplicate records, i.e., identify if some records from different hospitals are from the same patient. This database is secured with the following features: password protection, password encryption, IP blocking, roles & permissions, advanced logins, data encryption, record level security, and version tracking (<https://www.knack.com/tour/security>).

Lastly, the CFI will perform regular management (password/credentials, network, practices) of services and end-point devices, and regular user training in person and/or online.

User Training (CPR - User Training Program)

Users are trained on ethical data practices ([CPR - Good Data Practice and SOPs](#)), use of unique and strong password ([CPR - User Passwords](#)), use of secure wi-fi networks, use of verified website address/link (with HTTPS), awareness on possible phishing emails and other social engineering attacks, and the use of REDCap.

User Responsibilities (CPR - User Roles)

There are three user roles: (1) **data contributor/curator**, responsible for ensuring the submission of high-quality COVID-19 patient data from their hospital to the

COVID-19 Patient Registry. The curator is also responsible for checking the integrity and quality of data submitted. In cases where there are only scanned copies of paper medical records of COVID-19 patients, the curator is responsible for digitalizing the scanned record to integrate into the database; (2) **database administrator/data steward**, which is the Center for Informatics and its authorized staff, responsible for administering, managing, and maintaining the databases. The data steward is responsible for data storage, security, data deduplication, availability, and accessibility. Only the authorized database administrators can see the full dataset with identifiers but is bound by the data governance rules; (3) **researcher** is the primary user of the aggregate data and can query the database to obtain information on the cohort for their research. Researchers limit the use of the data according to what is stipulated in their approved research proposal. The CPR user responsibilities include agreement and compliance with the CPR data governance, reporting security and privacy incidents.

User Rights Management ([CPR - User Rights Management](#))

The users of the patient registry are provided limited rights that stipulate their interactions with the database. The **data contributor/curator** has the ability to upload records, create records, view own records, edit records from their own hospital. The **database administrator/data steward** has full access and can manage users, create records, edit records, disambiguate/join & delete records, and view/download full datasets with identifiers. The **researcher** has the ability to use the query interface to find cohorts and request de-identified patient records.

Risks and Mitigation

SECURITY RISKS	MITIGATION MEASURES
Hacking	REDCap utilizes SHA-256 and AES-256 during data transit and storage, which are both 256-bit encryption algorithms that provide high levels of security.
Phishing Attack / Login compromise	If an attacker manages to login using a user's credentials and causes damage, the compromised account will be immediately deactivated and investigated to prevent further damage.
Catastrophic Data Loss	Data can be exported as CSV, TXT, or JSON files anytime by the users for their own backup. The data stewards perform manual backups regularly to mitigate the risk of further data loss.

Unauthorized Database / App Manipulation	REDCap has built-in audit trails: (1) activity logging and (2) data field logging. The former is accessible to the data stewards, while the latter is available to all users. Once an end-user reports an unauthorized manipulation of their respective data, the audit trail can be used to provide details on the unauthorized data manipulation. The users will also be advised to observe specific precautionary measures to prevent such unauthorized transactions from happening in the future.
--	---

COVID-19 Patient Registry User Management and Technical Issues

CONCERN	COURSE OF ACTION
REDCap account request	The REDCap account creation usually takes one day for approval if all the necessary documents are submitted.
Technical concerns (Bug reports/accessibility issues)	Issues are reported first to the Data Stewards to analyze or resolve, then elevated to the REDCap administrator.

Data Destruction

CONCERN	COURSE OF ACTION
A hospital requests for termination of collaboration	The collaborating hospital shall communicate to the consortium about the termination of collaboration. Relevant data must be exported before the agreed termination date. After this date, all data and records from the said hospital will be disposed of appropriately or destroyed.
A hospital requests for termination of a user account	The collaborating hospital shall communicate to the consortium about the termination of a user account.

IX. PRIVACY PROGRAM (CPR - Privacy Program)

The database steward aggregates, disambiguates, and de-identifies the data that is made available for research.

The consortium also collects contact information of end-users (scientists and researchers) who request access to the CPR for the sole purpose of contacting them, validating, and updating end-user information. It will never share any patient data to entities that are neither part of the consortium nor identified scientists and researchers approved for access. It reserves the right, at its sole discretion, to modify its data governance by posting the updated version to inform all users.

The revision date shall be included to allow users (researchers) to monitor changes easily. End-users will also be sent a soft copy of updated policies every time revisions are made. After any such changes, the continued use of CPR constitutes an end-user's acceptance of the new data governance. Should an end-user disagree with any of the applicable terms of the latest revision, he/she must communicate the issue to the consortium via email. These issues shall be discussed by the consortium for resolution.

X. CONTINUITY

The contributions of the members of the CPR currently fund its operation. The members will seek other means to ensure continuous operations, including government and international agency grants, collaborative financing, etc.

In the event that the CPR can no longer be maintained, the consortium will determine the action to be taken on the data.

XI. CONFIDENTIALITY

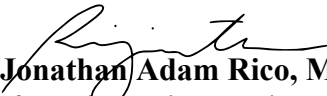
All users granted access to protected health information (PHI), research data, and results are bound by implicit confidentiality and non-disclosure under penalty of law.

In signing this document, the end-user agrees to the aforementioned provisions:

Signature over printed name: _____

Date: _____

This data governance document was prepared by the Center for Informatics:



Jonathan Adam Rico, MSc
Center for Informatics - University of San
Agustin



Pia Regina Fatima Zamora, MD, PhD
Center for Informatics - University of San
Agustin




Romulo de Castro, PhD
Center for Informatics - University of San Agustin


With inputs from consortium members:



Salvador Eugenio Caoili, MD, PhD
College of Medicine, UP Manila




Fresthel Monica Climacosa, MD, PhD
College of Public Health, UP Manila




Eyre Murao, PhD
UP Mindanao

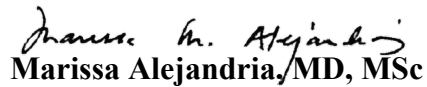
Signed off on by the representatives of the collaborating hospitals:



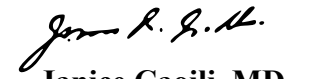
Roland Gilbert Remenyi, PhD
The Medical City




Jesus Emmanuel Sevilleja, MD, MPhil
National Center for Mental Health



Marissa Alejandria, MD, MSc
Philippine General Hospital



Janice Caoili, MD
Makati Medical Center



Sullian Naval, MD
Lung Center of the Philippines