

Research Data Management Plan

**Establishing a European database**

**of patients on dialysis or living with a kidney transplant**

**that have COVID-19**

**ERACODA**

(The **ERA**-EDTA **CO**VID-19 **Dat**abase for KRT patients)

Research Register number: 2020-00208

|  |  |
| --- | --- |
| **Date** | **Version** |
| 20-04-2020 | 2.5 |
| **Completed by** |
| Role in the study | Database manager | Signature |
| Full name | Dr. L.M. Kieneker |
| Department | Dept. NephrologyUniversity Medical Center GroningenP.O. Box 30.0019700 RB GroningenThe Netherlands | Signature date |
| **Authorized by**  |
| Role in the study | Principal investigator UMCG | Signature |
| Full name | Prof. dr. R.T. Gansevoort |
| Department | Dept. NephrologyUniversity Medical Center GroningenP.O. Box 30.0019700 RB GroningenThe Netherlands | Signature date |

Table of Contents

[1. Details about the research 3](#_Toc37757393)

[2. Description of the data collection 5](#_Toc37757394)

[3. Data storage and archiving 6](#_Toc37757395)

[4. Data documentation and metadata 7](#_Toc37757396)

[5. Data availability, access and re-use 8](#_Toc37757397)

# Details about the research

|  |
| --- |
| **SUMMARY** |
| Project full title: | Establishing a European database of patients on dialysis or living with a kidney transplant that have COVID-19 |
| Project acronym: | ERACODA |
| UMCG Research Register nr: | 2020-00208 |
| Subsiding party:  | NA |
| Principal investigator: | Prof. dr. R.T. Gansevoort |
| Partner organisations: | European Renal Association-European Dialysis and Transplant Association (ERA-EDTA), Milano, ItalyInformation Management Research (IMT), UMCG |
| Project duration: | Start: 23-03-2020 | End: Not known yet |
| Objectives / aims of the research project: |
| Primary aim: | To assess the extent of risk for mortality and to identify risk factors for morbidity and mortality of COVID-19 patients on chronic kidney replacement therapy |
| **ROLES AND RESPONSIBILITIES DATAMANAGEMENT** |
| Data collection: | All ERACODA participants |
| Data quality control: | Dr. M. PenaDept. Clinical Pharmacy and PharmacologyUniversity Medical Center GroningenDr. L.M. KienekerDept. NephrologyUniversity Medical Center GroningenDrs. H. de VriesDept. NephrologyUniversity Medical Center Groningen |
| Data processing: | Dr. M. PenaDept. Clinical Pharmacy and PharmacologyUniversity Medical Center GroningenDr. L.M. KienekerDept. NephrologyUniversity Medical Center GroningenDrs. H. de VriesDept. NephrologyUniversity Medical Center Groningen |
| Data analysis: | Dr. M. PenaDept. Clinical Pharmacy and PharmacologyUniversity Medical Center GroningenDr. L.M. KienekerDept. NephrologyUniversity Medical Center Groningen |
| Data archiving: | Dr. L.M. KienekerDept. NephrologyUniversity Medical Center GroningenDrs. H. de VriesDept. NephrologyUniversity Medical Center Groningen |

##

# Description of the data collection

Data on COVID-19 positive patients on chronic kidney replacement therapy will be entered in the ERACODA database from patients’ electronic medical records. These data consists of patient characteristics at time of presentation in the hospital, out-patient clinic or dialysis center, COVID-19 related characteristics, follow-up of the patient and patient outcomes. For a detailed description of all items collected, please see the ERACODA REDCap codebook. Only data that have been obtained during routine clinical care (i.e. there will be no additional interventions) and that are required to answer the research questions will be collected in the database. These source data will remain and be archived in the local institution as usual for clinical care.

For the collection of the data, a study specific electronic Case Report Form (eCRF) has been designed in [REDCa](https://DocPortal.umcg.nl/management/hyperlinkloader.aspx?hyperlinkid=539252d3-6af1-4c56-8a67-115f73dc8301)p, a commercially available, state-of-the-art database program, that is hosted on the secured servers of the UMC Groningen within the European Union. Participants will have to register as user by a two factor authentication process. After participants log into REDCap, they have to choose one of the authentication options: 1) Google authenticator or 2) e-mail. In the eCRF the data will be entered by participants. The infrastructure of REDCap has access management, audit trail and an automated back-up in place. The use of this infrastructure is in line with current University Medical Center Groningen (UMCG) policy and (inter)national standards.

During the execution of the project data analyses will take place weekly on the data collected that far using the statistical software programme SAS version 9.4. This data does not contain (in)directly identifiable data. All the syntaxes describing data processing and analyses will contain or be accompanied by comments explaining the code and the decisions made. The syntaxes will be checked by the data quality controllers/epidemiologists.

At completion of the data collection phase, all data collected in the eCRF in REDCap will be locked (i.e. database lock) for editing. Before database lock the data controllerswill perform data cleaning checks on missing or invalid data. After data cleaning and lock, datasets will be extracted from the eCRF in .csv format and stored within the UMCG network environment (G:-drive).

We will minimise the collection of possibly identifiable information, i.e. names and addresses will not be collected in the eCRF and consequently in the data files for analysis. Year of birth, sex and ethnic background are necessary for the research and will be collected. All participants will receive a research ID number in the datasets, consisting of site ID, record number of the site, country, and the hospital code of the patient. The hospital code of the patient will only be used for easy patient retrieval during the study and will help prevent the entry of duplicate records (e.g. because several doctors add the same patient to the database). This code will be deleted when a patient record is completed and closed and will NOT be saved in the final database as described in the document “Managing Patient Record IDs: Standard Operating Procedure <version number and date>”.

The data will be pseudonomized since the patients’ identifier of the national or regional registry (for dialysis patients) or Eurotransplant identifier (for kidney transplant patients) is requested in the database. The reason for pseudonymisation instead of anonymisation is that this will enable record linkage with these registries to investigate follow-up status of patients included in the ERA-EDTA COVID-19 KRT database. This registry number will be stored in a pass-word protected file separated from the other data collected in the study.

# Data storage and archiving

Paper source data and study files will remain and be archived in the local institution as usual for clinical care.

The digital data, both raw and processed/analyzed data, will be temporarily stored at a secured study specific folder within the UMCG network (G:-drive). This study specific folder can only be accessed by authorized personnel, who are involved in this study. Access to the study specific folder is arranged and maintained by the Functional Administrator of the department.

After the research project has been completed (i.e. data collection, data analysis and submission of research article completed) all the digital data will be transferred to a study specific folder on the UMCG Research Drive for long-term storage. The Research Drive can only be accessed by the Principal Investigator and by authorized personnel after approval of the Principal Investigator. The UMCG Research Drive is in line with current UMCG policy and fits the requirements of the Corporate Information Security Officer (CISO) and (inter)national standards.

This data storage environment allows for careful access management and is only accessible with username and password and is compliant with UMCG policies. Data storage is backed-up automatically every day on the servers of the UMCG. Management of the data and access permissions are managed by the research project’s data manager. The raw data containing identifiable information will be kept strictly separate from the processed data.

The digital data will not exceed 15 GB. The research data will be stored for 15 years after the data collection has been completed (i.e. last research assessment for the last patient has been performed).

# Data documentation and metadata

A description of the research project is registered in the UMCG under the number 2020-00208. This register is only accessible for UMCG researchers.

The eCRF specifically designed for this study in REDCap also provides a codebook containing descriptive information on all data variables and if applicable its units of measurement. This data dictionary or code book is extracted from REDCap in CSV file format.

For further analyses, extractions of the cleaned raw data will be made from this information system that will be explored using statistical and data science tools (SAS, version 9.4). All syntaxes describing data processing and analysis will contain or be accompanied by comments and descriptions explaining the code and the decisions made. The syntaxes will be checked by data quality controllers/epidemiologists.

# Data availability, access and re-use

Access permissions are controlled by the research project’s data manager and Principal Investigator. The raw data containing identifiable information will be kept strictly separate from the processed data and can only be accessed by an independent data manager of the department of Information Management Research (IMT).

Data will only be used for the present project, and not be made available for re-use.