KU LEUVEN

DEPARTMENT OF PUBLIC HEALTH
& PRIMARY CARE

VICAR

Data Management Plan

Version 1 : December 29, 2024

Responsibility of Prof. dr. Marc Sabbe (KU Leuven)

Disclaimer:

This Data Management Plan (DMP) is a living document. We are committed to continuously updating it to ensure that all aspects concerning data management, including collection, storage, processing, sharing, and disposal, are thoroughly covered. As new phases of the VICAR project emerge, and our understanding of the corresponding data improves, we will revise this document accordingly. This ensures that our DMP remains current and comprehensive. The version number can be found at the end of the document.

Contents

Abbreviations

- AI : Artificial Intelligence
- DMP : Data management Plan
- DTA : Data Tansfer Agreement
- CTC : Clinical Trial Center
- GDPR : General Data Protection Regulation
- LRD : KU Leuven Research and Development
- MDR : Medical Device Regulation
- MTA : Material Transfer Agreement
- PI : Principal Investigator
- RDR : Research Data Repository
- VICAR : Ventilation In Cardiac Arrest

1 Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project.

Dataset name/ID	Description	New or reuse	Data type	File format	Data volume
Demcon Raw	Raw data measured with setup using the Demcon Turbine Blower	New	Numerical	.txt-file	< 1 TB
Demcon Processed	Processed files derived from the "Demcon Raw" dataset measured with the setup using the Demcon Turbine Blower	New	Numerical	.csv-file	< 1 TB
Dräger Raw	Raw data measured with the setup using the Dräger Ventilator	New	Numerical	.txt-file	< 1 TB
Dräger Processed	Processed files derived from the "Dräger Raw" dataset measured with the setup using the Dräger ventilator	New	Numerical	.csv-file	< 1 TB
ZOLL Raw	Raw data measured with the setup using the Demcon Turbine Blower	New	Numerical	.txt-file	< 1 TB
ZOLL Processed	Raw data measured with the setup using the ZOLL 731 Series portable ventilator	New	Numerical	.txt-file	< 1 TB
VICAR-matrix	Database containing all settings and derived parameters for the three ventilators	New	Numerical	.csv-file	< 100 GB
Models	The models and algorithms used during the machine learning/AI phase of the project	New	Model	N/A	< 1 TB
Human study data	Clinical data from clinical trial (e.g. vital signs, lab results, intervention details, outcomes)	New	Numerical and clinical data	.csv-file	< 500 GB
Animal study data	In vivo animal study data (e.g. physiological measurements, outcomes, biomarkers)	New	Numerical and experimental data	.csv-file	< 100 GB
MDR	Deliverables according to MDR regulations	New	Textual	.doc-file	< 100 GB

Note: The data related to human clinical studies and animal studies will be further detailed and updated accordingly as soon as more information becomes available.

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL, etc.) per dataset or data type:

No data will be reused from earlier projects.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

Yes, dual use (approval number pending)

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

Yes, this project will involve processing personal data, including potentially special category data such as health or biometric data, in compliance with GDPR regulations. Details and any required approvals will be updated in the DMP at a later stage, once available.

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)? If so, please comment per dataset or data type where appropriate.

Yes, the VICAR project has potential for commercial valorization, including through non-exclusive licensing, technology transfer, and collaborations with industrial partners. Key datasets with commercial potential include:

- Clinical and preclinical data: these datasets will support product validation and attract partnerships for further clinical trials and commercialization.
- Models and algorithms: licensing of AI-driven models and algorithms to companies developing smart ventilators or critical care systems.

We work closely with the KU Leuven tech transfer office (LRD) to protect and manage IP with plans to make data available through the KU Leuven RDR when appropriate. Our strategy is to pursue non-exclusive licensing to ensure broad market access and avoid barriers to entry.

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

No

2 Documentation and metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

A README file will be provided with each of the datasets indicating what the data represents and how they were generated. This fil will be kept in the same directory as the data itself.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse.

No, technical and analytical methods used to generate the data will be documented in sufficient detail to allow for independent reproduction. These will include software version numbers, segmentation algorithm requirements, computational requirements, etc. When depositing data in a repository, the final data set will be accompanied by tis information in the form of a README.txt document. This file will be located in the top-level directory of the dataset and will also allow the data to be understood by other members of the

research group and add context to the dataset for further use.

3 Data storage and back-up during the research project

Where will the data be stored?

- Personal network drive (I-drive)
- OneDrive (KU Leuven)
- Other (specify below)

During the measurements, all data will be stored on a local drive. Regular backup onto a KU Leuven owned OneDrive repository and an external drive will be made with set reminders to aid the researchers. These backups onto OneDrive ensure that if something were to happen to the local computer, only very recently (max 2 working days) measured data would be lost.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution
- Personal back-ups I make (specify below)

During the entire process of data measurements, back-ups are ensured via OneDrive as well as at least three back-ups onto an external drive.

Is there currently sufficient storage and backup capacity during the project? If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

Normally enough storage will be available. In the case a shortage occurs, extra storage can be acquired through KU Leuven.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

- The external drive has a password, only known by the researchers.
- The computer performing the measurements requires the specific login credentials from the researchers and is physically behind a locked door. Remote access to this computer is only possible after multi-factor authentication.
- The online storage is secured though OneDrive itself.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

No costs are expected; if they, however, do arise a part of the research budget will be allocated to cover them. The only potential cost would be the need to purchase additional storage via KU Leuven.

4 Data preservation after the end of the research project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans.

Where will these data be archived (stored and curated for the long-term)?

Large volume storage (long term for large volumes)

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

This will be around €173.58 per year per TB and will be covered by KU Leuven.

5 Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available.

Yes, as restricted data (upon approval or institutional access only)
Given the nature of the data generated in the VICAR project, including sensitive clinical data and the potential for intellectual property protection, the data will not be made publicly available without restrictions. Any sharing of data will be subject to approval by the PI and institutional guidelines, including the MTA/DTA. Relevant datasets may be available after evaluation and approval by KU Leuven's LRD to ensure compliance with intellectual property and ethical requirements.

If access is restricted, please specify who will be able to access the data and under what conditions.

Once data files associated with a publication are released, they can be used by anyone to generate new results, provided they reference the original publication and do not use the data for commercial purposes. Other data will only be released upon requests, after an embargo period following publication. Access must be requested from a KU Leuven representative responsible for the data and model repositories, with written approval from the PI, adherence to institutional guidelines, and evaluation by LRD.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain per dataset or data type where appropriate.

Yes, ethical aspects
Ethical considerations related to the use of sensitive data from clinical or animal studies may impose restrictions, requiring evaluations and approvals before data sharing.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

- KU Leuven RDR
- Other data repository

This will be decided upon later in the project, once a clearer view on the amount of data is possible.

When will the data be made available?

- Upon publication of research results
- Other

The data will be made available after publication via the required link in the publication or upon requests and after an embargo period after publication, upon decision of the PI and after evaluation by LRD.

Which data usage licenses are you going to provide? If none, please explain why.

Yet to be decided.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

No

What are the expected costs for data sharing? How will these costs be covered?

The expected cost is not yet defined and will depend on the amount of data that is shared. The cost, however, will be covered by KU Leuven.

6 Responsibilities

Who will manage data documentation and metadata during the research project?

Prof. dr. Marc Sabbe, PI of the research project, is responsible for the data documentation and metadata. He delegates this task to ir. Anaïs Langlois as she is working full-time on this project and is performing the measurements.

Who will manage data storage and backup during the research project?

Prof. dr. Marc Sabbe is responsible for the data storage and back-up during the project and at least 10 years after the end of the research project.

Who will manage data preservation and sharing?

Prof. dr. Marc Sabbe is responsible for ensuring data preservation for at least 10 years after the end of the research project and the sharing of the data as outlined in the DMP.

Who will update and implement this DMP?

The end responsibility to update and implement this DMP lies with prof. dr. Marc Sabbe.

7 Logbook DMP

Version	Time	Notes
1	29/12/2024	Initial DMP