DMP title

Project Name TB-AI: advanced analytics for the management of traumatic brain injury - DMP title

Grant Title C321071

Principal Investigator / Researcher Prof. Geert Meyfroidt

Project Data Contact geert.meyfroidt@kuleuven.be

Description Traumatic Brain Injury (TBI) is a global health problem, representing the first cause of death in young adults and a major cause of disability across all ages. Treatment of TBI mostly aims at preventing and treating secondary brain injury, a series of inflammatory, ischemic, and mechanic events that can follow the initial traumatic event, leading to additional brain damage or even death. For this, invasive neuromonitoring is applied. We developed a prototype decision-support software, called TB-AI, that uses advanced analytics to display and interpret neuromonitoring signals in order to provide valuable information for the treatment of severe TBI to clinicians. Among others, the TB-AI gives early forewarning for impeding catastrophic events, i.e. future life-threatening elevations in intracranial pressure. The software is the first of its kind and has the potential to revolutionize clinical management of TBI. The goal of this project is to test clinical safety and accuracy of this software in a multi-center clinical study. The TB-AI software will be used in 4 Europeans intensive care units (ICU). Clinicians will be asked to consult the results of the software at pre-defined moments, based on this consultation they can decide whether or not to adjust therapy. During the trial, several data will be collected, such as, among others, recordings of the output of the software, information on the clinical management of the patients and patients' outcome. The data collected will allow to quantify the accuracy of the software and its clinical safety. This Data Management Plan describes how data will be generated, organized, stored, and backed up throughout the duration of the project and after its completion.

Institution KU Leuven

1. General Information

Name of the project lead (PI)

Prof. Geert Meyfroidt

Internal Funds Project number & title

C321071. TB-AI: advanced analytics for the management of traumatic brain injury.

2. Data description

2.1. Will you generate/collect new data and/or make use of existing data?

- Generate new data
- Reuse existing data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

Type of Data	Format	Volume	How created
Demographic data	.pdf, .readme files, .xlsx, .csv, .sql	10-20KB	Demographic information that are routinely available in the personal files of the patient, such as age, sex, relevant medical history. The collected data will be strictly limited to the absolute minimum data that are necessary to characterize the patient. Data will be pseudonymised before the start of the data analysis.

Clinical characteristic at admission	.pdf, .readme files, .xlsx, .csv, .sql	10-20KB	Metrics that quantify the clinical status of the patient at admission are routinely collected in an intensive care unit and are available in the patient electronic health record system. These metrics normally quantify the severity of illness and the probability of death or complications. The collected data will be strictly limited to the absolute minimum data that are necessary to characterize the patient. Data will be pseudonymised before the start of the data analysis.
Monitoring data	.readme files, .xlsx, .csv, .sql, .npy	30-50GB	Recordings of continuous intracranial pressure (ICP), mean arterial blood pressure (MAP) and partial tissue oxygenation (PbtO2) monitoring.
Daily treatment data	.pdf, .readme files, .xlsx, .csv, .sql	2-5 MB	Recordings of the patient treatment, such as treatment intensity level, surgery, use of specific interventions to lower intracranial pressure or doses of medications. Part of these data are routinely collected in the intensive care unit and stored/can be extracted from the patient's electronic health record, additional information on the patient's treatment will be collected specifically for this study through questionnaires to the clinicians. The collected data will be strictly limited to the absolute minimum data that are necessary to characterize the therapy strategy for the patient.
Output of the TB-AI software	.readme files, .xlsx, .csv, .sql	10-30 GB	Output of the TB-AI that is displayed to the clinician. This is automatically collected by the software.

Outcome data	.readme files, .xlsx, .csv, .sql	10-20KB	Recordings of mortality in the intensive care unit (ICU) or in the period that follows the ICU stay. Results of the patient's standardize questionnaire for the assessment of the long-term neurological status of the patient.
User satisfaction questionnaire	.readme files, .xlsx, .csv, .sql	10-20KB	Output of the user satisfaction questionnaires that will be provided to the clinicians at the end of each patient stay.
Programming code for data analysis (digital)	py, .ipynb or similar	1GB	Code that will be used to analyse the data of the study and extract the final results.
Results	npy, xlsx, .csv, .docx, .pdf, .jpeg or similar	2-5MB	Outcome of the project. It includes all the documents, figures and tables that report and summarize the scientific findings on the primary and secondary outcomes of the study.

3. Ethical and legal issues

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

Yes, we will collect health data of the included patients (in a pseudonymised form), and the name and contact of the patient or patient's family for the assessment of the patients long-term outcomes (3 to 6 months after discharge from the intensive care unit).

Patient indentifiers will be stored in a separate patient log at the local site.

3.2. 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, add the reference to the formal approval by the relevant ethical review committee(s).

Yes

Ethical issues may concern the collection of personal data, specifically health data. The protocol of this study and the related data management plan will be submitted to the medics ethical committee (EC) of UZ Leuven for approval. Additional ethical approval will be requested from all the participating centers.

Given our experience with previous similar studies we do not foresee problems from the EC of UZ Leuven in the approval of the collection of the planned data. To minimize potential ethical issues we will restrict the data collected to the minum necessary to answer the pre-defined research questions.

3.3. Does your research possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

Yes

Potential intellectual property arising from this work will be managed according to the suggestions from the LRD office of KU Leuven. Given that UZ Leuven and KULeuven are the sponsors of the study the IP will remain within KU Leuven.

3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the

data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?

No.

4. Documentation and metadata

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

The protocol of the clinical trial and detailed documentation on the methodology for data collection will be published on a peer-reviewed scientific journal (preferrably open-access).

To facilitate the potential reuse of data, monitoring and clinical data will be stored and labeled in a standardized way. Every collected data will be associated with a patient study number (pseudonymisation) and timestamp to facilitate the further analysis of data.

Instructions related to the interpretation of the output of the TB-AI software will be recorded and made available.

The rationale behind the scientific and practical choices that will be taken during the design and setting-up of the study will be documented.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

No

In the study we will collect several data of different format from each enrolled patient. Every collected data will be associated with a patient study number (pseudonymisation) and timestamp to facilitate the further analysis of data.

We will also create a SQL database with the collected data to provide a structured storage of data. Data with different format will be collected in different SQL tables and linked together through the patient study number. A brief description of such SQL database will be created. Apart from this SQL dataset, data will be always available in the original format which includes : .xlsx, .csv, .pdf, .readme files, .npy, .sql.

The protocol of the clinical trial and detailed documentation on the data collection will be published on a peer-reviewed scientific journal (preferrably open-access).

5. Data storage and backup during the project

5.1. Where will the data be stored?

During data collection, temporary copies of the collected monitoring data will primarily be stored in the computer where the software will run in a password protected folder. In this temporary phase, data backup will be requested on a monthly base on an external drive which will be password protected.

The remaining data will be collected on a centrally manteined and backedup server, also password protected.

After merging of the datasets from the different centers, a copy of the dataset will be stored on the KU Leuven Archive Storage space (K-Drive) for long-term preservation and backup. The K-drive does not allow data modification nor cancellation.

During data processing and analysis, a copy of the data will be stored temporarely on the personal computer of the researcher in charge of the data analysis.

5.2. How will the data be backed up?

Data stored on the KU Leuven Archive K-Drive is managed, maintained, and backed up by KU Leuven IT services. Specifically, mirror copies of the stored data are made immediately upon upload, for safety backup purposes. The K-Drive is a write-once/read many drive, which means that data that are stored on the K-drive cannot be modified or cancelled.

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

Yes, the KU Leuven K-Drive has sufficient storage capacity for the outlined project.

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

The annual cost of K-Drive storage is 11.384 € for illimited storage volume per year. This cost and capacity include the performance of mirror copies of the stored data, for safety backup purposes. We expect that 150GB will be sufficient to store all data generated as part of the project. These costs will be covered by the budget of the laboratory of intensive care medicine.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Data will be always stored on a password protected envirorment , possibly with access log. Data stored on KU Leuven-managed personal computers are protected via password access to the computers, as set up by the KU Leuven IT Department. Off-site access to K-Drive data is available from KU Leuven personal computers and data access points, and is password protected. Access to files in the laboratory of intensive care medicine (LICM) lab K-Drive folders is limited to members of the LICM lab.

6. Data preservation after the end of the project

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

All raw and processed data will be retained for at least 10 years after the end of the project on the K-drive storage space.

6.2. Where will these data be archived (= stored for the long term)?

All data will be stored on the university's central servers for at least 10 years.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

The cost per year of the KU Leuven K-drive is of 11.384 euros. Therefore we expect a total cost of at least 113.84 euros for data storage during the period of 10 years.

This cost will be covered by the Laboratory of Intensive Care Unit at KU Leuven.

7. Data sharing and re-use

7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

We do not foresee any factor that would prevent the sharing of the data, but every scenario will be preventively discussed with the Research and Development (LRD) office of KU Leuven.

7.2. Which data will be made available after the end of the project?

After the end of the project, data can be made available to external reasearch groups upon agreement on the planned reuse.

7.3. Where/how will the data be made available for reuse?

- Other (specify):

Data will be send through encrypted email through the KU Leuven partner Belnet, which allows to send high-size data. The files are stored only temporarily and can be accessed only by authorised users.

7.4. When will the data be made available?

- Upon publication of the research results

7.5. Who will be able to access the data and under what conditions?

Access to the data from external reserach groups will be considered after a request is submitted explaining the planned reuse. Commercial reuse will be allowed under contractual agreement, only if in line with the TB-AI valorization path and only after consultation with the KU Leuven Research and Development (LRD) office.

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

No direct costs are foreseen for the sharing of the data. Indirect costs may be related to the

working hours that will be dedicated to the revision of the scientific proposals for the planned reuse and the working hours dedicated to the meetings with the external reserach groups and the transfer of data.

8. Responsibilities

8.1. Who will be responsible for the data documentation & metadata?

Prof. Geert Meyfroidt

8.2. Who will be responsible for data storage & back up during the project?

Prof. Geert Meyfroidt

Dr. Giorgia Carra

8.3. Who will be responsible for ensuring data preservation and sharing?

Prof. Geert Meyfroidt

8.4. Who bears the end responsibility for updating & implementing this DMP?

The end responsibility for updating and implementing the DMP is with the supervisor Prof. Meyfroidt.