DMP title

Project Name G0B1922N - KOTK - DMP title
Grant Title G0B1922N
Principal Investigator / Researcher Hilde Bosmans
Institution KU Leuven

1. General Information

Name applicant

Hilde Bosmans

FWO Project Number & Title

G0B1922N - KOTK

Affiliation

- KU Leuven

Multi-centric project together with UGent, Universiteit Antwerpen

2. Data description

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

- Reuse existing data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Type of data	Format	Volume	How created
Low dose chest CT scans	.dicom	400 MB per patient	Routinely acquired on CT scanner in radiology department for patient care
Standard chest CT scans	.dicom	400 MB per patient	Routinely acquired on CT scanner in radiology department for patient care

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- Yes

Privacy Registry Reference: Application to clinical trial center UZLeuven will be done in the frame of this retrospective study.
Short description of the kind of personal data that will be used:
Pseudonymized CT images of the chest
Patient age, gender, weight, BMI
Radiology and pathology report with information on cancer diagnosis, size, location and type of lesions.

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

Approval was not yet obtained. Application to clinical trial center UZLeuven will be done in the frame of this retrospective study.

Does your work 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?

- No

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 are in place?

- No

This is a multicentric project with UGent and Universiteit Antwerpen. We foresee the set-up of data transfer agreements for the sharing of CT data.

4. Documentation and metadata

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

Dicom chest CT images will be collected and the personal data will be removed from the dicom header. Patient age, weight, gender and BMI will be retained and completed with information of radiology and pathology report. This will be collected in an electronic case report form.

A database server (DATABASE), which will handle data accumulation and storage,
A PACS/DICOM server (PACS), where DICOM images will be collected,
A processing server (ANALYSIS)

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- Yes

dicom standard for chest CT images is used with pseudonymization of personal data following GDPR rules.

5. Data storage and backup during the FWO project

Where will the data be stored?

Data from KULeuven will be stored on the secured hospital network of UZLeuven data.

When data is shared with other universities, data transfer agreement will be made and secured data transport and storage will be described and established. (Example: Liquidfiles for data transfer)

How is backup of the data provided?

The data will be stored on the hospital's and university's central servers with automatic daily back-up procedures

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

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

200 EUR per TB per year

Data security: how will you ensure that the data are securely stored and not accessed

or modified by unauthorized persons?

Since we will be working with pseudonymized data, the data will be stored in the hospital's secure environment for private data. The link file to the personal data will also be stored in the hospital's secure environment but a different location than the pseudonymized images.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All data will be retained for the expected 5 year period after the end of the prioject including the pseudonymized images.

The core of the project also includes the making of a testing database for future testing and validation of AI algorithms for lung cancer detection; this database will be stored for minimum 10 years.

Where will the data be archived (= stored for the longer term)?

The data will be stored on the hospital's central servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

200 EUR per TB per year

7. Data sharing and reuse

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)?

- No

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

The aggregated study data e.g. the metric for image quality control of low dose chest CT images, conversion factors for size-specific chest CT dosimetry. The annotated database of chest CT images with lung cancer for testing and validation of computer aided detection algorithms will not be made available to the outside world in order to guarantee independent testing of future computer aided detection algorithms.

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

Data will be available on request after signing a data sharing agreement or will be made available via scientific publications.

When will the data be made available?

- Upon publication of the research results

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

Most data will be available via scientific publications.

Image data will be available on request after signing a data sharing agreement.

The annotated database of chest CT images with lung cancer for testing and validation of computer aided detection algorithms will not be made available to the outside world in order to guarantee independent testing of future computer aided detection algorithms.

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

No expected costs for data sharing.

8. Responsibilities

Who will be responsible for data documentation & metadata?

PhD student Kwinten Torfs

Post-doc Dimitar Petrov

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

PhD student Kwinten Torfs

Post-doc Dimitar Petrov

Who will be responsible for ensuring data preservation and reuse ?

Hilde Bosmans

Walter De Wever

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

The PI (Hilde Bosmans and Walter De Wever) bears the end responsibility of updating & implementing this DMP.