DMP title

Project Name FWO G097322N (ADNEX) - DMP title

Grant Title G097322N

Principal Investigator / Researcher Ben Van Calster

Description In essence, this project entails research topics to further understand and improve the ADNEX risk model. ADNEX estimates the risk that an ovarian tumor is benign, borderline malignant, stage I invasive, stage II-IV invasive, or secondary metastatic. Validation studies report very good model performance, and the model is being recommended by professional societies and being introduced in ultrasound scanning machines. The model can be further improved, however, and its performance can be studied in more detail. We aim to address the following objectives: 1. Can we understand and describe the heterogeneity in performance between centers for ADNEX and other models like RMI? (WP1) 2. How can we update the original ADNEX model to make it better applicable to the true target population and to different centers with the best possible risk estimates? (WP2) 3. How does ADNEX work in the hands of less experienced examiners in comparison to other models like RMI? (WP3) 4. ADNEX is the result of a fruitful combination of applied and methodological research on risk prediction. It is important to continue this. We therefore want to address methodological questions, such as elucidating the benefits of selecting predictors based on clinical utility (i.e. taking the burden of measuring the predictors into account), studying methods to better deal with and assess heterogeneity when using machine learning algorithms for binary and multiclass prediction. (WP4) The study uses data of completed or ongoing international multicenter observational studies. This implies that no new patient data will be collected for the sole purpose of this project.

Institution KU Leuven

1. General Information

Name applicant

Ben Van Calster (supervisor), Dirk Timmerman (co-supervisor)

FWO Project Number & Title

The ADNEX risk model for ovarian cancer diagnosis: improving performance across different hospitals and examiners (G097322N)

Affiliation

- KU Leuven

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: observational numeric data on clinical, demographic, ultrasound, and histological parameters from patients with an ovarian tumor.

Format: .xls and .txt.

Size: max 1GB.

How created: through ultrasound examination, lab analysis of blood samples, and pathological examination of excised tissue from surgery.

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: G-2020-2523

Short description of the kind of personal data that will be used: We process data related to tumor characteristics as measured by ultrasound, age, CA125 level, and tumor histology.

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

IOTA5 study: amendment to ML5344 (EC UZ / KU Leuven)
IOTA7 study: S59207 (EC UZ / KU Leuven)
Retrospective use: S64709, PRET G-2020-2523

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

4. Documentation and metadata

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

Data dictionaries will be generated to explain the meaning of each variable in the datasets. For the IOTA5 data this has already been done.

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.

- No

We will generate hte metadata that is required by the Metadata model of the KU Leuven RDR (Research Data Repository): title of dataset, names/e-mail/ORCID of responsible authors, a summary of the purpose/nature/scope of the dataset, appropriate keywords, technical format, access rights,

5. Data storage and backup during the FWO project

Where will the data be stored?

1. The time-stamped master copy of the data will be kept on our research unit central storage facility. Copies can be made and kept on personal devices.
2. We will use OneDrive for active use of the data during the project.

How is backup of the data provided?

The data will be stored on the 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

Data volume is limited (<1GB).

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

We do not expect any costs.

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

Data are stored in the university's secure environments to prevent access by unauthorized persons.

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 at least 10 years.

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

The data will be stored on the university'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?

We do not expect any costs.

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)?

- Yes. Specify:

Data sharing is not currently foreseen because of ethical (this was not part of the consent given by participants) and legal (personal data) reasons. Post hoc data sharing for research purposes (i.e. no commercial use) may be discussed, leading to controlled access described in targeted data sharing agreements.

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

Data sharing is not currently foreseen because of ethical (this was not part of the consent given by participants) and legal (personal data) reasons. Post hoc data sharing for research purposes (i.e. no commercial use) may be discussed, leading to controlled access described in targeted data sharing agreements.

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

- Other (specify):

Data is stored using secure options form the university. Reuse by KU Leuven supervisors or collaborators will be done after ethics approval (retrospective study).

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?

Currently, only study collaborators can access the data.

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

No costs are expected.

8. Responsibilities

Who will be responsible for data documentation & metadata?

The (co-)supervisors.

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

The (co-)supervisors.

Who will be responsible for ensuring data preservation and reuse ?

The (co-)supervisors.

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

The PI bears the end responsibility of updating & implementing this DMP.