DMP title

Project Name Targeted mass spectrometry-based quantification of autoantibodies in systemic rheumatic diseases - DMP title

Grant Title FWO â€“ TBM - T004821N

Principal Investigator / Researcher Xavier Bossuyt

Description Dit project beoogt een gevoelige en specifieke methode te ontwikkelen voor de detective van autoantistoffen bij patiÃ«nten met inflammatoÃ¯re myopathiÃ«n en systeemsclerose. InflammatoÃ¯re myopathiÃ«n en systeemsclerose zijn systemische reumatische aandoeningen waarbij autoantistoffen worden aangetroffen. Deze antistoffen hebben een diagnostische en prognostische waarde. Voor een aantal van deze antistoffen zijn betrouwbare geautomatiseerde testen beschikbaar. Voor andere antistoffen, evenwel, zijn er testen beschikbaar (vb line en dot blot), maar deze testen hebben enkele nadelen zoals een sub-optimale specificiteit of sensitiviteit. Voor nog andere testen zijn geen commerciÃ«le testen beschikbaar. Het doel van dit project is om een methode op punt te stellen die ons in staat stelt verschillende antistoffen gelijktijdig op te sporen op een betrouwbare manier. Daarvoor zal gebruik gemaakt worden van immuunprecipitatie gekoppeld aan identificatie van de immuungeprecipiteerde eiwitten door middel van massa spectrometrie.

Institution KU Leuven

1. General Information

Name applicant

Xavier Bossuyt

FWO Project Number & Title

FWO – TBM - T004821N

Targeted mass spectrometry-based quantification of autoantibodies in systemic rheumatic diseases

Affiliation

- KU Leuven

2. Data description

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

- Generate new 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).

Serum samples from several clinically well-defined cohorts (myositis, systemic sclerosis and (disease) controls) from UZ Leuven, UZ Gent [including Gent early myositis cohort] and OLV Aalst. Clinical information (i.a. diagnosis and clinical phenotype, routine laboratory results) of the patients will be documented. Serum samples will be pseudonimized.

- The datatypes that will be generated are protocols, methods and SOPs, which will be provided and stored as word or pdf documents. The raw data (mass spectrometry data, flow cytometry, antibody sequencing, monoclonal antibody production & characterization) will be stored as processed test results (excel files), the figures will be generated by e.g. Prism (GraphPad) or excel and stored as jpg- and pfzx- or excel-files. For monoclonal antibody procution, all activities and protocols are logged in an eLAB journal.

Types of data	Format	Volume	How created
MS run raw data	.RAW	3 TB	Q Exactive Plus of ELITE
Proteome Discover data	PDSTUDY File, BAK File, Thermo's Mass Spec Format, MSFVIEW File, PDRESULT File, PDRESULTVIEW File	360 GB	Proteome Discover
Scaffold data	Scaffold File	30 GB	Scaffold
Statistical processed data	Excel	2-6 GB	Excel, Statistica 8.1, Qlucore
Figures and graphs procesed data	.jpg, .pfzx	12 GB	Excel, GraphPad Prism, Qlucore
sera	sera	>2500	existing cohorts - see above
Protocols	.doc(x)	10 MB	Microsoft Word
Sequencing raw data	.ab1, .fasta	200 Mb	Received from provider
FACS data	.fcs	3Gb	FCS express
Figures, graphs and processed data	.jpg, .pfzx	5GB	Excel, Graphpad
Protocol and SOPs	.doc(x)	5000 Mb	MS Word

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:
Short description of the kind of personal data that will be used:
Privacy Registry Reference:
Short description of the kind of personal data that will be used:
- Age
- Gender
- Date (systemic rheumatic disease) diagnosis
- clinical phenotype
- Lab results (antinuclear antibodies and antibodies to extractable nuclear antigens)
- Serum samples

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

The project has been submitted to the Central Ethical Committee of UZ Leuven (Ethische Commissie Onderzoek UZ/KU Leuven) for approval on February 26th (reference number S65989). It has also been submitted to the ethical committee of OLV Aalst and will be submitted to the ethical comittee of UZ Gent.

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

No IP restrictions will be claimed. If appropriate and relevant, we will share and transfer the obtained results on a non-profit basis. There are no restrictions on the freedom to operate for utilization of the project results. Patient data will not be transfered to other laboratories.

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?

MTA with Onze Lieve Vrouw Ziekenhuis in Aalst and with UZ Gent university will be set up. There will be no restrictions for dissimination of the data.

4. Documentation and metadata

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

Experimental data will be ordered in maps and (read-me) files with explanations will be provided. Protocols per experiment will be foreseen (word). The maps will be accessable to the PI and the study team. The maps will be organized such that it allows interoperability.

Clinical data: Patient data will be pseudonimized for the main researchers and anonimized for other researchers and clinicians. Patient number will be linked to serum sample.

SOURCE: Mass spectrometry data, flow cytometry data, sequencing, production & characterization of antibodies sourced from autoimmune parients

TYPE: mass spectrometry data (autoantigen identification), antibody sequences, antibody expression plasmids, transfected HEK293 cells, purified antibodies. Experimental methods and measurements.

FORMAT: sera stored at -20°C, expression plasmids stored at -80°C. Purified antibodies stored at -20°C. Experimental measurements in database in MS Word, Excel, Prism, FCS express for flow cytometry, jpg, pfzx, RAW, PDSTUDY File, BAK File, Thermo's Mass Spec Format, MSFVIEW File, PDRESULT File, PDRESULTVIEW File, Scaffold File,

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

Excel or word will be used for documentation of data and how the data are organized.

REDCap.

5. Data storage and backup during the FWO project

Where will the data be stored?

The generated data of this project will be stored on the server of UZ Leuven and/or on LVS (Large Volume Storage) (L-drive) of the KU Leuven.

How is backup of the data provided?

The data will be stored on the UZ Leuven university hospital servers and/or KU Leuven 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

The storage capacity of the server of UZ Leuven and of the KU Leuven (LVS - Large Volume Storage) L-drive) provides sufficient storage volume for our generated data.

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

We expect no costs for data storage and back up during the project at UZ Leuven. At KU Leuven, large volume stroage is used to cost-efficiently store large amounts of data for long periods. The price for large volume storage at KU Leuven amounts to € 113.84 per year per Terrabyte. This includes primary stoage €62.54 and mirror storage €51.30.

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

The data will be stored at the hospital's and the university's secure environment for private data. The patient data is pneumonized for the main researchers, but anonymized for the colleagues of the lab. The key will be stored in a separate file (separate from the experimental data).

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 (raw data, (statistical) processed data, methodologies, patient data,...) will be retained for the expected 5 year period after the end of the project. If too much space is needed to store all the data, the raw data generated by the mass spectrometry devices will not be maintained after 5 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 5 years after the end of the project, 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?

Data will be stored on a UZ Leuven server (free of cost at this moment) and/or at a KUL server (estimated costs: 113.84 euro per year and per Terrabyte).

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:

During the research, clinical data of patients and (disease) controls will be collected and stored. This data will not be shared.

MTAs are set up with OLV Aalst and with UZ Gent. There will be no restrictions regarding sharing of the data.

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

If successfull and relevant, the methodologies and the data will be published in scientific journals. All personal data regarding the patients will not be available.

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

- Upon request by mail

When a data sharing agreement is needed (advisable), we will seek assistance from the CTC (Clinical Trial Center).

Data will be shared for requests that are justified.

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?

Access will be granted upon written request to the creators (PI / study team) of the dataset. Commercial reuse is not allowed.

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

If relevant, we will share and transfer the obtained results to scientists/clinicians and other laboratories on a non-profit basis.

As the data will be electronic, we do not foresee major costs.

8. Responsibilities

Who will be responsible for data documentation & metadata?

The PI and his team will be responsible.

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

The PI and his team (Doreen Dillaerts) will be responsible.

For data related to monoclonal antibody generation, the responsible person for the data storage is Prof. Paul Declerck

Who will be responsible for ensuring data preservation and reuse ?

The PI and his team will be responsible.

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