

DMP title

Project Name : Optimal surgical planning or gynaecological patients.

Project Identifier u0047968

Grant Title C24M/21/039

Principal Investigator / Researcher Thierry Van den Bosch

Project Data Contact thierry.vandenbosch@uzleuven.be

Description Surgery for myometrial lesions and for deep endometriosis are amongst the most common gynecological procedures. A reliable preoperative evaluation is pivotal to tailor the most appropriate procedure, estimate the surgical complexity, select the most appropriate surgical team, and provide correct patient counselling. This will optimize patients' outcome, avoiding unnecessary or inappropriate procedures and minimize operation time and costs for both patients and health service. In myometrial lesions, the preoperative differentiation between benign (myoma) and malignant (sarcoma) lesions is of utmost importance: the former is often managed conservatively or treated by laparoscopic resection and morcellation, while the latter requires radical surgery by an oncological surgeon, morcellation being absolutely contraindicated. In endometriosis, correct detection and mapping of deep endometriosis lesions will determine the indication and the radicality of the surgery. The aim of our project is to develop innovative management algorithms for both myometrial lesions and deep endometriosis. As first line evaluation, alongside the patient's history and clinical examination, we will define relevant sonographic features, based on our recent consensus papers on the systematic ultrasonographic evaluation of the myometrium and of endometriosis. As second line evaluation, alongside MRI, we will assess the added value of novel liquid biopsy biomarkers.

Institution KU Leuven

1. Data Description

What data will you collect or create? Fill out the table below and/or describe.

Type of data	Format	Volume	How created?
basic patients' characteristics MUSA	digital data: e.g. age, parity, body mass index and symptoms, selected from a drop down list	1500 cases 15 characteristics	-collected by the clinician at the start of the ultrasound examination -pseudonymized -collected in an online encrypted data collection system: the Clinical Data Miner (see below)
basic patients' characteristics IDEA	digital data: e.g. age, parity, body mass index and symptoms, selected from a drop down list	1500 cases 22 characteristics	-collected by the clinician at the start of the ultrasound examination -pseudonymized -collected in an online encrypted data collection system: the Clinical Data Miner (see below)
Myometrial lesions: sonographic features reported according the MUSA-terms and definitions	digital data: e.g. measurement of the lesion; descriptive data, e.g. echogenicity, selected from a drop down list	1500 cases 50 characteristics	-collected by the clinician after the ultrasound examination -pseudonymized -data collected in Clinical Data Miner

Deep endometriosis: sonographic features reported according the IDEA-terms	digital data: e.g. measurement of the lesion; descriptive data, e.g. echogenicity, selected from a dropdown list	1500 cases 50 characteristics	-collected by the clinician after the ultrasound examination -pseudonymized -data collected in -Clinical Data Miner
ultrasound images	2D images (.png) 3D volumes (.vol); 4D View uncompressed	400 cases average of three 2D images (~ 120kB per image) and one 3D volume (~ 4050MB per volume) per case	-pseudonymized -on an external hard disk, dedicated to the study and kept locked at the UZ Leuven departement
liquid biopsy results MUSA	digital data about the presence and concentration of different proteins, immune cells and ctDNA	<ul style="list-style-type: none"> • 100 cases • 3 different group of markers: proteins, immune cells and ctDNA (defining the number of relevant markers is part of the exploratory part of the study) 	-pseudonymized -blood collection, processing at the laboratory for tumor immunology (KU Leuven; head prof dr An Coosemans)
Liquid biopsies IDEA Type: plasma, aliquots van 500µl, +/- 20 aliquots/patient	biobank LEERM-labo (G-PURE research group) S59006 prof Joris Vriens (KU Leuven, Department of Development and Regeneration)	target 300ptn	-pseudonymized -blood collection, processing at the laboratory for tumor immunology (KU Leuven; head prof dr An Coosemans)

Histology: All patients with endometriotic lesion	clinical workflow (KWS) as well as biobank LEERM-labo (G-PURE research group) S59006	1500 cases for MUSA and 1500 cases for IDEA	-collected by the clinician in KWS -pseudonymized -data collected in Clinical Data Miner
---	---	--	--

Do you intend to reuse existing data?

Our study is prospective. We do not intend to reuse existing data.

Do you use personal data (i.e. all data possibly identifying an individual)?

- Yes

Although we use patients' personal data, the data will immediately be pseudonymized. The data saved in the Clinical Data Miner (CDM) are pseudonymized, and for all further processing (statistical analysis) this metadataset will be used. It is impossible to identify an individual based on this metadataset.

2. Documentation and Metadata

Describe the documentation that will be created for the data. This section deals with the way in which you will document how the dataset was created and subsequently processed.

1. The basic patient characteristics are obtained by the clinician-researcher performing the ultrasound examination. The questions are shown on the web-based Clinical Data Miner and the answers are directedly recorded in the system by selecting from a drop down list.
2. The standardized report of the sonographic features according the MUSA and IDEA-terms and definitions are obtained by the clinician-researcher performing the ultrasound examination for myometrial lesions and deep endometriosis lesions respectively. The required features are shown on the web-based Clinical Data Miner and the answers are directedly recorded in the system by selecting from a drop down list.
3. The clinician-researcher performing the ultrasound examination downloads the ultrasound images and 3D volumes from the ultrasound machine after pseudonymization (the name and the hospital's KWS identifier is replaced by the automatically generated (by CDM) unique case number).
4. The liquid biopsies are processed by the dedicated laboratory technician according the guidelines of the laboratory for tumor immunology (KU Leuven; head prof dr An Coosemans) as outlined in the protocol. The pseudonymized results are recorded in the Clinical Data Miner by one laboratory technician under supervision of prof dr An Coosemans. For IDEA pseudonymized minimal clinical data collected in biobank database (FilemakerPro database)
Biopsies are being processed according to (published) SOPs(WERF-EPHECT papers:
DOI: 10.1016/j.fertnstert.2014.07.1208; DOI: 10.1016/j.fertnstert.2014.07.1208)
5. The histology results of the hysterectomy specimen are retrieved from the KWS by the clinician who performed the ultrasound and recorded as pseudonymized data in CDM by selection from a drop-down list.

Describe the metadata for the data. This section deals with metadata: information contained in your dataset about the research data.

The data saved in the Clinical Data Miner are pseudonymized, and for all further processing (statistical analysis) this metadataset will be used.

3. Ethical, Legal and Privacy Issues

Are there any ethical issues concerning the creation and/or use of the data?

Although the study is non-interventional, it deals with patients' data and hence is an ethical issue.

Did you consider all issues about copyrights and IPR?

Yes.

Are the collected data considered to be "data containing personal information" and are all the requirements about the collection of these data met?

The protocols (S62497/B322201941519 and S63045/B322201941402 for MUSA; and S63056 B322201941498 for IOTA) have been submitted and approved by the UZ Leuven ethics committee. This is specified in the protocol's DATA PROCESSING ANNEX ("DPA") and the Instructions for DPA, confirming the requirements about the collection of data containing personal information are met.

The study sponsor is the University Hospital UZ Leuven.

EU Standard Contractual Clauses for transfers outside EU:

https://ec.europa.eu/info/law/lawtopic/data-protection/data-transfers-outside-eu/model-contracts-transfer-personal-data-third-countries_en

4. Data storage and Backup during Research

How and where will the data be stored during research?

- In a cloud service offered by the university

All data will be collected in the **Clinical Data Miner** (CDM): an online encrypted data collection system with coded system for pseudonymisation of patient data in accordance with the European General Data Protection Regulation (GDPR), created by our research group with the Department of Electrical Engineering (ESAT), KU Leuven (JMIR Med Inform 2014 Oct 20;2(2):e28). When entering data, the user needs to enter a self-created patient identifier. Subsequently, a case identifier is automatically created, adding an additional layer of pseudonymization. When exported, the data only contain the case ID created by the program itself.

The **Clinical Data Miner** (CDM) has already been used for other multicenter studies at UZ Leuven: e.g. the International Ovarian Tumor Analysis (IOTA)-study, including over 6000 cases and lead by prof dr Dirk Timmerman and the International Endometrial Tumor Analysis (IETA)-study, including 3000 cases and coordinated by prof dr Dirk Timmerman and prof dr Thierry Van den Bosch.

Which back-up procedures are in place?

The **Clinical Data Miner** server follows best practices with regard to security. It runs on a hardened Linux server with a restrictive firewall. Security updates are installed automatically. The data on the server are backed up to a remote server on a daily basis.

Describe the data security procedures and who has access to the data.

Our web server uses encrypted channels using SSL / HTTPS for all client-server communication. Clients that initiate an unencrypted connection are automatically and immediately redirected to an encrypted channel.

All client-server communication uses SSL / HTTPS. This is also the case when data are transmitted outside the Partners firewall.

5. Data selection and Preservation after Research

What is the long-term preservation plan for these dataset(s)?

All data are preserved in the **Clinical Data Miner**

(<https://cdm.esat.kuleuven.be/CDM/#home>) for at least 5 years after the end of the research. The data are stored in CDM on the cloud (Electronic Data Capture) system (host placed in the EU). A backup of the data is made on daily base, on a different site. The backup itself is encrypted. The study manager controls the access to the study data during and after the research.

The engineer responsible for the CDM is Willem Mestdagh, ESAT – STADIUS (Stadius Center for Dynamical Systems, Signal Processing, and Data Analytics), KU Leuven, Belgium (willem.mestdagh@esat.kuleuven.be)

Data Selection: Which data will have long time value for the research and will be preserved?

The metadata in CDM will be kept for at least 5 years after the end of the research.

6. Data Sharing

Are there any restrictions for sharing the data?

Patients' data

The pseudonymized data will only be shared as metadata with the KU Leuven biostatisticians (head Prof Ben Van Calster).

To have access to the Clinical Data Miner as study participant (user), permission had to be granted by the study coordinators (Thierry Van den Bosch and Dirk Timmerman). The coordinator can also revoke a user's grant.

Only the study coordinator and the user can see the list of patients featuring the patient ID (KWS-identifier) and the unique CDM patient number. However, a user cannot see the data of other users. The metadata, used by the statistician only provides the unique CDM-generated patient number. <https://cdm.esat.kuleuven.be/CDM/#home>

Researchers' (users') personal data

The users' personal data collected and stored by the CDM include

- **Username:** *The username given upon registration by the user him/her-self.*
- **Firstname:** *Used to identify the user.*
- **Lastname:** *Used to identify the user.*
- **Phonenumber:** *Could be used to take contact with the user when having problems.*
- **Email:** *Used for account-verification and to give in the case that you are department coordinator updates about new registrations in your realm.*
- **Realm:** *The realm (e.g. department) you belong to, is used to find your colleagues and to regroup your entries (per realm) together.*

Personal data is never passed on to third parties for commercial purposes by CDM.

If there are no restrictions, which mechanisms will be in place to assure that the data are discoverable, accessible and intelligible? not applicable

How will you share the data?

- Other, specify

The pseudonymized metadata from the Clinical Data Miner can be downloaded in xls format (e.g., for statistical analysis).

With whom will the data be shared?

- Within the research unit only

The data will be shared within the research unit only, to allow statistical analysis of the data.

7. Responsibilities and Resources

Who is responsible for Data Management during the project? This will be the person who might receive questions on the data management aspects of the research project.

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

Which additional resources are needed for the execution of the Data Management Plan?

The maintenance of the Clinical Data Miner is under the responsibility of Willem Mestdagh, ESAT – STADIUS (Stadius Center for Dynamical Systems, Signal Processing, and Data Analytics), KU Leuven, Belgium (willem.mestdagh@esat.kuleuven.be). The maintenance costs are integrated in the study funding proposal.

Did you read the KU Leuven Data Management Policy? (find the link to the policy in the guidance).

- Yes

The DMP has been included in the study protocol in addendum (**DATA PROCESSING ANNEX (“DPA”) – Instructions**), submitted to and approved by the UZ Leuven Ethics Committee.