THE TWIN GROWTH PROJECT - PATHOPHYSIOLOGY, DIAGNOSIS AND OUTCOMES OF ISOLATED SELECTIVE FETAL GROWTH RESTRICTION IN MONOCHORIONIC TWIN PREGNANCIES

A Data Management Plan created using DMPonline.be

Creator: Liesbeth Lewi

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Liesbeth Lewi

Grant number / **URL:** 1804723N

ID: 198239

Start date: 01-05-2023

End date: 01-05-2026

Project abstract:

Objective: The objective is to develop a model to beter predict the outcome of selective fetal growth restriction (sFGR) in monochorionic diamniotic twin pregnancies based on specific ultrasound parameters throughout the pregnancy. Secondary objectives are to increase our knowledge of the pathophysiology of sFGR, fetal and neonatal survival, neonatal morbidity, neurodevelopment infant outcomes as well as psychological impact of a sFGR pregnancy on parents.

Study design: Prospective multicentre international cohort study between UZ Leuven (Belgium), LUMC (the Netherlands), Karolinska University Hospital (Sweden), BCNatal Hospital Clínic i Hospital Sant Joan de Déu, (Spain) and Mt Sinai Hospital Toronto (Canada).

Study population: sFGR monochorionic twin pregnancies in Leuven, Leiden, Stockholm, Barcelona and Toronto that meet the inclusion criteria. The twins (or twin in case of single fetal demise) and their parents will be seen during follow-up.

Last modified: 31-03-2023

THE TWIN GROWTH PROJECT - PATHOPHYSIOLOGY, DIAGNOSIS AND OUTCOMES OF ISOLATED SELECTIVE FETAL GROWTH RESTRICTION IN MONOCHORIONIC TWIN PREGNANCIES

Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

We will collect clinical variables of pregnancies complicated by sFGR. These are new and personal data. The raw clinical data will be entered encrypted in CASTOR EDC as ".sav". The data will be exported for analysis in SPSS or STATA (as ".csv"). The total expected volume will be between 10-100 GB.
Variables that will be collected are:

- Maternal baseline demographics and baseline characteristics of co-parent
- Characteristics of the pregnancy (including labour)
- Outcomes of the parental questionnaires (PCL-5, MAAS/PAAS, MPAS/PPAS, PARCA-R)
- Neonatal outcome
- Variables regarding follow-up: neurodevelopmental outcomes of the children at 2 years of age
- Ultrasound parameters, images and clips
- Placental evaluation (macroscopic- and microscopic evaluation, additional immunohistochemical staining)

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

1. Designation of responsible person

Provisional: Liesbeth Lewi Liesbeth.Lewi@uzleuven.be
The contract negations are still ongoing between the sites - either all sites will be joint-controllers or LUMC-Leiden will be the data-processor and the other sites will be data-controllers.

1. Storage capacity/repository
    - during the research Data will be stored in the CASTOR EDC
    - after the research: At the end of the study, the relevant raw data will be extracted from Castor database for long term archiving. Coded data will be stored in a secured DataSafe at the LUMC, Leiden. The principal investigators of each institution will have access to CASTOR EDC and SataSafe and ensure that the database is maintained efficiently and all information is current and accurate. The collected data will be stored for at least 20 years after the ending of the study in accordance with the Nederlands Code of Conduct for Research Integrity. For this, the LUMC has long-term storage with back-up available.Only the PIs of each site and persons involved in the research (e.g. PhD student, assigned by the PIs) will have access to the data. As these are privacy-sensitive data, they will not be shared in Open Access. We will publish our data in manuscripts in Open Access journals.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

Not applicable.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

Data management of these privacy-sensitive and personal data will be implemented according to Good Clinical Practice (GCP) guidelines and will comply with the General Data Protection Regulation (GDPR) 2016/679. Data will be handled confidentially and anonymously by coding. Patient data will be entered by way of an electronic Case Report Form (eCRF) in the central GCP proof internet-based CASTOR database to facilitate on-site data-entry. CThe stored data will be encrypted with an identification code (CASTOR will generate a coded ID per case), thus information identifying the patient or child will be omitted. Data will be collected and entered by the principal investigators at each centre.

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

None

THE TWIN GROWTH PROJECT - PATHOPHYSIOLOGY, DIAGNOSIS AND OUTCOMES OF ISOLATED SELECTIVE FETAL GROWTH RESTRICTION IN MONOCHORIONIC TWIN PREGNANCIES

DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

THE TWIN GROWTH PROJECT - PATHOPHYSIOLOGY, DIAGNOSIS AND OUTCOMES OF ISOLATED SELECTIVE FETAL GROWTH RESTRICTION IN MONOCHORIONIC TWIN PREGNANCIES

GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

THE TWIN GROWTH PROJECT - PATHOPHYSIOLOGY, DIAGNOSIS AND OUTCOMES OF ISOLATED SELECTIVE FETAL GROWTH RESTRICTION IN MONOCHORIONIC TWIN PREGNANCIES

FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Clinical variables	• Maternal demographics and characteristics of co-parent • Pregnancy characteristics of the pregnancy • Outcomes of the parental questionnaires (PCL-5, MAAS/PAAS, MPAS/PPAS, PARCA-R) • Neonatal outcome s • Neurodevelopmental outcomes of the children at 2 years of age	New	Digital entered in Castor EDC database	Observational	.sav;.cvs	<1 GB
Ultrasound variables	• Gestational age • Fetal growth • Fetal Dopplers • Cord insertion site • UV measurements • AA measurement • Cardiac measurements • Brain measurements	New	Digital entered in Castor EDC database	Observational	.sav; .cvs	<1 GB
Ultrasound images	• Cardiac • Brain	New	Digital uploaded in Castor EDC database	Observational	.JPEG; .TIFF;.avi; .mp4; .Mov; not exhaustive	10-100 GB
Placental images	• Macroscopic • Histology	New	Digital uploaded in Castor EDC database	Observational	JPEG; .TIFF;not exhaustieve	10-100 GB

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

Not applicable

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, human subject data

Data management will be implemented according to Good Clinical Practice (GCP) guidelines and will comply with the General Data Protection Regulation (GDPR). Data will be handled confidentially and anonymously by coding. Patient data will be entered by way of an electronic Case Report Form (eCRF) in a central GCP proof internet-based database to facilitate on-site data-entry. CASTOR EDC was chosen as our data management system. The stored data will be encrypted with an identification code (CASTOR will generate a coded ID per case), thus information identifying the patient or child will be omitted. Data will be collected and entered by the principal investigators at each centre. The principal investigators will ensure that the database is maintained efficiently, and all information is current and accurate. The study will be submitted to the Ethical Committee of UZ Leuven and is registered at the Clinical Trial Center of UZ Leuven (S67514) once the clinical trial and collaboration agreements are approved by all participating centers (in progress contact person Marie Gillot - legal adviser CTC).

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes

- Maternal baseline demographics and baseline characteristics of co-parent
- Characteristics of the pregnancy (including labour)
- Outcomes of the parental questionnaires (PCL-5, MAAS/PAAS, MPAS/PPAS, PARCA-R)
- Neonatal outcome
- Variables regarding follow-up: neurodevelopmental outcomes of the children at 2 years of age
- Ultrasound parameters, images and clips
- Placental evaluation (macroscopic- and microscopic evaluation, additional immunohistochemical staining)

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- Yes

The contracts are currently still under negotiation between the legal departments of Leuven-Leiden-Barcelona and Toronto (contact: Marie Gilliot - legal advisor-CTC) .

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

A data dictionary (notebook) will be produced in CASTOR EDC describing all variables. We will export our data dictionary from CASTOR. Imaging data from microscope analyses will be saved, if possible, in Recommended Metadata for Biological Images (REMBI).
All syntaxes used in data cleaning and analysis (including annotation describing the goal of processing steps) will be stored to facilitate replication.
A readme.txt with a list of all available files and a description of their contents will be created at the end of the project, before archiving the data.
Supporting information is provided in CASTOR EDC to the clinicians that will add data the database

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

No metadata standard will be used. Where possible SNOMED coding will be used for clinical variables. Data from patient files include diagnosis information using ICD10 coding. A data dictionary (notebook) will be produced in CASTOR EDC describing all variables. We will export our data dictionary from CASTOR. Imaging data from microscope analyses will be saved, if possible, in Recommended Metadata for Biological Images (REMBI).

3. Data storage & back-up during the research project

Where will the data be stored?

Patient data will be entered by way of an eCRF in CASTOR EDC (a central GCP proof internet-based database) to facilitate on-site data-entry. The stored data will be directly encrypted with an identification code (CASTOR will generate a coded ID per case) on a secured network disk, thus information identifying the patient or child will be omitted. After completion of the study the coded data together with the trials management file will be stored in the secured DataSafe at the LUMC, Leiden.

How will the data be backed up?

CASTOR EDC backups are made four times a day and stored at another geographical location by Castor. The backup files are kept for fifteen days. Reserve copies can be restored but this might add additional costs.
Every day a copy is made of any changed files in DataSafe. A full back-up of data is made once per week. Also, monthly and quarterly backups are made so that earlier data can also be restored if necessary.

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

Both Castor EDC and DataSafe have sufficient storage and backup capacity for the estimated 10-100 GB of digital data that our project will generate.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Patient data will be entered directly by way of an eCRF in a central GCP proof internet-based database to facilitate on-site data-entry. The collected data will be directly coded when stored. CASTOR will generate a coded ID per case. In that way, the data cannot directly be traced back to the patients and personally identifiable data will not be handed over to investigators that do not have a therapeutic alliance with the patient. Personal data are only processed by the researchers or by those who fall directly under their authority. Security is guaranteed with login names, login codes and encrypted data transfer.

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

Use of CASTOR with large file uploading is 1800 Euro/year (excl VAT) – for duration of study. The costs will be divided by the participating sites. Datasafe is provided at no extra cost by the LUMC, Leiden.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

At the end of the study, we will extract the raw data from CASTOR EDC and save these and the processed data for long-term archiving in DataSafe. In line with the Netherlands Code of Conduct for Research Integrity, raw and processed data will be stored for a period of at least 20 years. For this, the LUMC has long-term storage with back-up available.

Where will these data be archived (stored and curated for the long-term)?

The data will saved in DataSafe provided by the LUMC, Leiden.

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

Storage in DataSafe is provided by the LUMC for the purpose of the study.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, …)

If access is restricted, please specify who will be able to access the data and under what conditions.

We have privacy-sensitive data, and therefore, we will not share this data in Open Access. Data will be shared in agreement with GDPR, legislation and agreements in place.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- Yes, Privacy aspects

The collected data are personal and privacy-sensitive.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

After completion of the study, the data will be archived at LUMC in the DataSafe database.

When will the data be made available?

The data can be made available after publication of the research results.

Which data usage licenses are you going to provide? If none, please explain why.

The data can be used for future research in agreement with GDPR, legislation and institutional agreements in place.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

We will, if possible, publish the metadata online with a persistent identifier.
If the image data are released in tissue-IDR, then a data DOI will be minted following public release.

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

We will not use on Open Access repository. The data will be stored in DataSafe, which is provided by the LUMC.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

The principal investigators and their assigned collaborators of the participating sites

Who will manage data storage and backup during the research project?

The principal investigators and their assigned collaborators of the participating sites through CASTOR EDC

Who will manage data preservation and sharing?

currently part of the contract negotiations if all sites will be joint-data controllers or LUMC will act as data processor and the other sites will be controllers.

Who will update and implement this DMP?

Liesbeth Lewi principal investigator at UZ Leuven