Plan Overview

A Data Management Plan created using DMPonline

Title: Identification of patients at risk for scoliosis before disease onset: an elastographic study during growth in healthy and high scoliosis risk populations

Creator: Aaron Moens

Project Administrator: Steven de Reuver, Aaron Moens

Affiliation: Other

Template: UMC Utrecht DMP

Project abstract:

Rationale: most deformity in idiopathic scoliosis occurs in the intervertebral disc. As such, it plays an essential role in the initiation and progression of scoliosis. Apparent stiffness of the developing disc is believed to be important, but prospective data are limited since scoliosis patients are only identified after curve onset with secondary changes in the disc as a result. Hypothesis: already before disease onset, the intervertebral disc's apparent stiffness differs between scoliotics and healthy controls. Objective: first, a cross-sectional analysis with non-invasive, radiation-free, ultrasound shear-wave elastography to define apparent stiffness of the intervertebral disc in the normal population at the start of their growth spurt. Second, a prospective analysis of a population subset, known to develop scoliosis in 50% of the cases, that will undergo elastography assessment of the disc at inclusion, before onset of scoliosis. The final goal is to prospectively follow them through their standard outpatient clinic visits for scoliosis screening, and when after follow-up it is known who does and does not develop a scoliosis, retrospectively analyze if there was a difference in apparent disc stiffness at inclusion. Study design: cross-sectional/ prospective Study population: 'healthy' children aged 8-10 years and children with 22q11.2 DS aged 8-10 years Main study parameter: shear wave speed measured in a intervertebral disc Nature and extent of the burden and risks associated with participation, benefit and group relatedness: None, besides a 15-minute visit including a 1-minute ultrasound investigation.

ID: 81613

Start date: 01-09-2021

End date: 01-09-2025

Last modified: 07-09-2021

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Identification of patients at risk for scoliosis before disease onset: an elastographic study during growth in healthy and high scoliosis risk populations

1. General features

1.1. Please fill in the table below. When not applicable (yet), please fill in N/A.

DMP template version	29 (don't change)
ABR number *(only for human-related research)*	
METC number *(only for human-related research)*	TBD
DEC number *(only for animal-related research)*	
Acronym/short study title	ELASTO
Name Research Folder	xx-xxx_ELASTO
Name Division	Surgical specialties
Name Department	Orthopedic surgery
Partner Organization	Arts et Métiers ParisTech, Technische Universiteit Eindhoven
Start date study	01-09-2021
Planned end date study	01-09-2025
Name of datamanager consulted*	Dax Steins
Check date by datamanager	17-08-2021

1.2 Select the specifics that are applicable for your research.

- Monocenter study
- Non-WMO
- Clinical study
- Observational study
- Prospective study

2. Data Collection

2.1 Give a short description of the research data.

The goal is to measure shear wave speed in the intervertebral disc in ‘healthy’ patients and patients with 22q11.2 deletion syndrome, to determine after follow-up if shear wave speed differed at inclusion between children that do and do not develop a scoliosis.

Control group (n = 40): children aged 8-10 (20 males, 20 females), with no known spinal pathology, that will undergo an abdominal ultrasound for a random indication in the Wilhelmina’s Children Hospital (WKZ).

22q11.2 deletion syndrome group (n = 40): children aged 8-10 (20 males, 20 females), with no known spinal pathology (yet), that will get X-rays of the spine in the WKZ as is standard in the care of children with 22q11.2 deletion syndrome.

With a standard clinical ultrasound machine from Phillips, the elastography measurements will be done, multiple per patient (at least 3) are performed and the average shear wave speed is noted. The raw DICOM images will be stored and pseudonymized for scientific research in the UMCU's Research Imaging Architecture (RIA).

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format	Storage space
Human	80	Philips Ultrasound machine	Research Imaging Architecture (RIA)	Images	.dicom	0-10 GB
Human	80	Ultrasound machine outcome measures	Excel	Quantitative	.xlsx	0-10 GB
Human	80	EPD (HiX)	Excel	Quantitative	.xlsx	0-10 GB

2.2 Do you reuse existing data?

- No, please specify

For this prospective study, we will generate new data on this topic.

There are no other researchgroups with datasets on this topic that we can or will re-use in this research.

2.3 Describe who will have access to which data during your study.

Type of data	Who has access
Direct identifying personal data	Research team (Steven de Reuver, Aaron Moens), DHS datamanager
Key table linking study specific IDs to Patient IDs	PI (Moyo Kruyt), Datamanager
Pseudonymized data	Research team (Steven de Reuver, Aaron Moens), DHS datamanager

2.4 Describe how you will take care of good data quality.

#	Question	Yes	No	N/A
1.	Do you use a certified Data Capture Tool or Electronic Lab Notebook?		x	
2.	Have you built in skips and validation checks?			x
3.	Do you perform repeated measurements?	x		
4.	Are your devices calibrated?	x		
5.	Are your data (partially) checked by others (4 eyes principle)?	x		
6.	Are your data fully up to date?	x		
7.	Do you lock your raw data (frozen dataset)	x		
8.	Do you keep a logging (audit trail) of all changes?			x
9.	Do you have a policy for handling missing data?			x
10.	Do you have a policy for handling outliers?			x

2.5 Specify data management costs and how you plan to cover these costs.

#	Type of costs	Division ("overhead")	Funder	Other (specify)
1.	Time of datamanger	x		
2.	Time of radiologist/ ultrasound technician	x		
3.	Storage	x		
4.	Archiving	x		

2.6 State how ownership of the data and intellectual property rights (IPR) to the data will be managed, and which agreements will be or are made.

UMC Utrecht is and remains the owner of all collected data for this study. The data is collected in a relatively large patient group and is very valuable for further,
broader studies in Europe. It may for example be used to find study subjects for future treatment studies. Our data cannot be protected with IPR, but its value will be
taken into account when making our data available to others, when setting up Research Collaborations and when drawing up Data Transfer Agreement(s).

3. Personal data (Data Protection Impact Assessment (DPIA) light)

Will you be using personal data (direct or indirect identifying) from the Electronic Patient Dossier (EPD), DNA, body material, images or any other form of personal data?

- Yes, go to next question

I will process personal data. I have checked the full DPIA checklist and I do not have to complete a full DPIA. I therefore fill out this DPIA light and proceed to 3.1.

3.1 Describe which personal data you are collecting and why you need them.

Which personal data?	Why?
Name and email address of participants	To be able to invite participants for taking part in the research and to send them the patient information
Sex, age, length and weight	To describe our study population
Ultrasound images	Main parameter of study question.

3.2 What legal right do you have to process personal data?

- Study-specific informed consent

3.3 Describe how you manage your data to comply to the rights of study participants.

The data are pseudonymized and the linking table to personal data is saved. An authorized person manages the linking table, can re-identify study participants when necessary and deliver, correct or delete the data.

Right	Example answers
Right of Access	Research data are coded, but can be linked back to personal data, so we can generate a personal record at the moment the person requires that. This needs to be done by an authorized person.
Right of Rectification	The authorized person will give the code for which data have to be rectified.
Right of Objection	We use informed consents.
Right to be Forgotten	In the informed consent we state that the study participant can stop taking part in the research. Removal of collected data from the research database cannot be granted because this would result in a research bias.

3.4 Describe the tools and procedures that you use to ensure that only authorized persons have access to personal data.

We use the secured Research Folder Structure that ensures that only authorized personnel has access to personal data, including the key table that links personal data to the pseudoID.

3.5 Describe how you ensure secure transport of personal data and what contracts are in

place for doing that.

We will not transport any personal data outside the UMCU network drives.

4. Data Storage and Backup

4.1 Describe where you will store your data and documentation during the research.

The digital files will be stored in the secured Research Folder Structure of the UMC Utrecht. We will need +/- 50 GB storage space, so the capacity of the network
drive will be sufficient. Paper dossiers will be stored safely in a locked cabinet in a locked room in the UMC Utrecht. A project specific procedure is in place for
access to the paper dossiers. Documentation of this procedure is stored in the Research Folder Structure.

4.2 Describe your backup strategy or the automated backup strategy of your storage locations.

All (research) data is stored on UMC Utrecht networked drives from which backups are made automatically twice a day by the division IT (dIT).

5. Metadata and Documentation

5.1 Describe the metadata that you will collect and which standards you use.

We do not use metadata standards yet. The data from UPOD will be delivered including a data dictionary. For every variable this data dictionary contains an explanation of the values.

5.2 Describe your version control and file naming standards.

We will distinguish versions by indicating the version in the filename of the master copy by adding a code after each edit, for example V1.1 (first number for major
versions, last for minor versions). The most recent copy at the master location is always used as the source, and before any editing, this file is saved with the new
version code in the filename. The file with the highest code number is the most recent version. Every month, we will move minor versions to a folder OLD. The
major versions will be listed in a version document (projxVersDoc.txt), stating the distinguishing elements per listed version.

6. Data Analysis

6 Describe how you will make the data analysis procedure insightful for peers.

I have written an analysis plan in which I state why I will use which data and which statistical analysis we plan to do in which software. The analysis plan will be stored in the project folder after creation, so it is findable for my peers.

7. Data Preservation and Archiving

7.1 Describe which data and documents are needed to reproduce your findings.

The data package will contain: the raw data, the study protocol describing the methods and materials, the SPSS scripts to process the data, the SPSS scripts leading to tables and figures in the publication, a codebook with explanations on the variable names, and a 'read_me.txt' file with an overview of files included and their content and use.

7.2 Describe for how long the data and documents needed for reproducibility will be available.

Data and documentation needed to reproduce findings from this non-WMO study will be stored for at least 15 years.

7.3 Describe which archive or repository (include the link!) you will use for long-term archiving of your data and whether the repository is certified.

After finishing the project, the data package will be stored at the UMC Utrecht Research Folder Structure and is under the responsibility of the Principal Investigator of the research group. When the UMC Utrecht repository is available, the data package will be published here.

7.4 Give the Persistent Identifier (PID) that you will use as a permanent link to your published dataset.

I will be using a DOI-code and will update this plan as soon as I have the code.

8. Data Sharing Statement

8.1 Describe what reuse of your research data you intend or foresee, and what audience will be interested in your data.

The raw data can be of interest for other researchers or for spin off projects. Currently, we do not know IF anyone is interested in our data. But if so, most likely other researchers are interested in the raw ultrasound elastography measurements (shear wave speed), which could be shared if laws/regulations make that possible.

8.2 Are there any reasons to make part of the data NOT publicly available or to restrict access to the data once made publicly available?

- No, all data generated in this project will be made publicly available without any restrictions

8.3 Describe which metadata will be available with the data and what methods or software tools are needed to reuse the data.

The publication will be open assessable. The study protocol and this Data Management Plan will also be available.

8.4 Describe when and for how long the (meta)data will be available for reuse

- (Meta)data will be available upon completion of the project

8.5 Describe where you will make your data findable and available to others.

n/a