Plan Overview

A Data Management Plan created using DMPonline

Title: Retrospective evaluation of the outcomes of preputioplasty, performed in children with phimosis

Creator: Jikke Bosveld

Affiliation: Other

Template: UMC Utrecht DMP

Project abstract:

Phimosis is a condition of the foreskin in which the foreskin cannot be retracted by gentle manipulation. Traditionally, circumcision was the treatment of choice for boys with pathological phimosis. In recent years, there has been increasing resistance to circumcision. Consequently, there is an increasing trend towards performing a preputioplasty. According to the guideline database of the Netherlands, the literature currently does not provide a definite answer as to which approach is superior. More research is needed to determine the results of a preputioplasty.

ID: 76966

Start date: 03-05-2021

End date: 30-08-2021

Last modified: 11-06-2021

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

Retrospective evaluation of the outcomes of preputioplasty, performed in children with phimosis

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)*	N/A
METC number *(only for human-related research)*	21-376/C
DEC number *(only for animal-related research)*	N/A
Acronym/short study title	-21-376- Preputiumplasty
Name Research Folder	
Name Division	DHS
Name Department	Pediatric Urology
Partner Organization	N/A
Start date study	03-06-2021
Planned end date study	N/A
Name of datamanager consulted*	D. Steins
Check date by datamanager	11-05-2021

1.2 Select the specifics that are applicable for your research.

- Monocenter study
- Retrospective study
- Non-WMO

2. Data Collection

2.1 Give a short description of the research data.

The primary objective of this study is: evaluating the results of a preputioplasty, expressed as retractability of the foreskin, and examing whether certain determinants such as age at time of surgery or presence of symptoms affect (long-term) outcomes. As secondary objectives, we want to investigate what complications may occur during a preputioplasty. Moreover, does the usage of topical corticosteroids, affect the outcome. We shall retrospectively collect health care data from patients with phimosis, who have undergone a preputioplasty at the Wilhelmina Children's Hospital (WKZ) between the 1st of January 2011 and the 1st of January 2020. The data will be collected by the datamanager through the Research Data Platform (RDP). The datamanger will continue to extract the data until a set is created with the appropriate characteristics. Researchers who have a care relationship with the patient will further determine eligibility based on the patient's records (unstructured data). If a patient does not meet the inclusion criteria or is characterized by an exclusion criteria, the patient will be removed from the data set. Also, the researcher who has a care relationship with the patient will extract information from HIX to describe the desired determinants and to complete the data.

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format	Storage space
Human	> 100	EPD (HIX)	Research Data Platform (RDP)	Quantitative	.xlsx	0-10 GB
Human	> 100	EPD (HIX)	Excel	Quantitative	.xlsx	0-10GB

2.2 Do you reuse existing data?

- Yes, please specify

In this retrospective study, we use pseudonymized data, which is made available for research by Research Data Platform (RDP).

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

My division datamanager receives a datamart from the Research Data Platform (RDP) that contains direct identifying personal data and pseudonymized data. The datamanager is authorized to link different datasets of the selected patient group and has access to personal data.The key table linking study specific IDs to patient IDs is available to the datamanager and members of the research team with a care relationship to the patient. Other members of the research team receive a pseudonymized dataset and have no access to direct personal data or the key table.

Type of data	**Who has access**
Direct identifying personal data	Research team with care relationship to patient, Datamanager
Key table linking study specific IDs to Patient IDs	PI (with care relationship to patient), Datamanager
Pseudonymized data	Research team, Datamanager

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

The data is first collected using a script created by the data manager and from here manually supplemented from HIX. In an excel file, these data will come together to form one clear dataset. This excel file will be stored in the research folder. When the data is complete, it will be frozen before analysis. No personal data will be transported outside the network

#	**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** (please specify)
1.	Time of datamanager	x		
2.	Storage	x		
3.	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. 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?
Date of operation, age at time of surgery, type of operation, the use of corticosteroids preoperative, degree of phimosis preoperatively, follow-up period	To describe our study population.
Diagnose of balanitis xerotica obliterans (BXO)/lichen sclerosis/scarring	As secondary objective we want investigate whether lichen sclerosis/BXO/scarring affects the outcome.
Short-term and long-term complications	To investigate which complications arise from a preputioplasty.
Parent/patient satisfaction and cosmetic result from a surgical point of view	To evaluate the results of a preputioplasty.
Outcome of preputioplasty expressed as retractability of the foreskin.	To evaluate the result of a preputioplasty.
Symptoms present, surgeon who perfomed the preputioplasty, history of surgery	Investigate whether certain determinants influence the results.

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

- No objection, please explain

1. **Why:** We have the legal right to process personal data because we abide by the exception rules and researchers without a treatment relationship work with a pseudonymised dataset.
2. **Who:** The no-objection check will be performed by the devision datamanger.
3. **When:** The researchers receive the dataset on the same date as when the objection check is performed by the datamanager.

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

The researchers may use pseudonymized data in this research without consent based on the no-objection rule. This is permitted because the research meets the following four characteristics:

1. The processing is necessary for the purpose of scientific research,
2. The research serves a public interest,
3. Obtaining explicit consent would require a disproportionate amount of effort given the expected number of patients to be included, which is approximately 300.
4. The implantation shall be carried out with such guarantees that the privacy of the person concerned is not disproportionately affected.

Confidentiality will be maintained at all times; participant information will not be disclosed to third parties. For this study, the dHS datamanager will first identify potential eligible patients using the established inclusion criteria. Additionally the principal investigator who has a care relationship with the patient will further determine patient eligibility based on the patient's records (unstructured data). The principal investigator will extract desired determinants from HIX to complete the dataset. The extracted research data will be coded by the dHS datamanager with a key-linking table for patient re-identification and stored in a secure Research Folder Structure (RFS) on the UMCU network drive of my division (L:\Onderzoek\Kinderchirurgie). Direct identifiable data, including the key-linking table, will be stored separately from the research data using the RFS for access control. Unlike the principal investigator, the research team will only have access to the pseudonymized data.

Ultimately, there are two datasets that will be used. There will be a dataset with linkage table for the principal investigator who has a care relationship with the patient and there will be an pseudonymized dataset for researchers without a treatment relationship.

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

We use the secure Research Folder Structure (RFS), which ensures that only authorised people have 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.

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.

For the data collected in Excel, we created a codebook of our research database. My division's datamanager adds descriptive and technical metadata to the SAS script (.sav format) for the Research Data Platform data collection. We do not yet use metadata standards. The final synthesis of the analyses is kept in the research folder.

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.

6. Data Analysis

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

Most likely, this study will be a discriptive analysis. Once the data has been collected, an analysis plan will be written outlining what data will be used and what statistical analysis will be carried out in what software. The analysis plan will be kept in the project folder for future reference.

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 script to process the data, the data management plan (DMP), the research folder request form and the application form for the provision of data.

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.

To Be Determined (TBD)

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?

To Be Determined (TBD)

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

To Be Determined (TBD)

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

To Be Determined (TBD)

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

To Be Determined (TBD)