# Plan Overview

*A Data Management Plan created using DMPonline*

**Title:** Three dimensional radiological analysis of Hip dysplasia on Computed tomography (CT) scans

**Creator:**Milou Hüsken

**Affiliation:** UMC Utrecht

**Template:** UMC Utrecht DMP

**Project abstract:**

Rational: 3.3%-5.4% of the general adult population has dysplastic hips based on the definition of a reduced lateral center edge angle on anterior-posterior (AP) radiographs. The pathophysiology and treatment of hip dysplasia are based on the three-dimensional (3D) aspect of the dysplastic hip. However, the disease is still classified and treated based on two-dimensional (2D) imaging.

Objective: The aim of this study is to describe 3D aspects and parameters of the dysplastic hip based on Computed tomography (CT) scans.

Design/population: Retrospective study in which CT scans from adult patients with known hip dysplasia will be analyzed.

Study endpoint: The most common radiological parameters for the hip joint will be described in a population with adult hip dysplasia, to define the relevant parameters and to describe the differences with a healthy population. A statistical shape model will be used to describe the average dysplastic hip and the parameters that will define the dysplastic hip the most.

**ID:** 98890

**Start date:** 15-02-2023

**End date:** 15-02-2025

**Last modified:** 28-03-2023

# Three dimensional radiological analysis of Hip dysplasia on Computed tomography (CT) scans

## 1. General features

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

| | |
|---|---|
| DMP template version | 30 (don't change) |
| ABR number *(only for human-related research)* | |
| METC number *(only for human-related research)* | |
| DEC number *(only for animal-related research)* | |
| Acronym/short study title | DysplaCT |
| Name Research Folder | 22U-0170_DysplaCT |
| Name Division | Heelkundige specialismen |
| Name Department | Orthopedie |
| Partner Organization | St Maartens Kliniek Nijmegen |
| Start date study | 15-02-2023 |
| Planned end date study | 15-02-2025 |
| Name of datamanager consulted* | Nivard Koning |
| Check date by datamanager | 28-3-2023 |

**1.2 Select the specifics that are applicable for your research.**

- Multicenter study
- Observational study
- Non-WMO
- Retrospective study

## 2. Data Collection

**2.1 Give a short description of the research data.**

| Subjects | Volume | Data Source | Data Capture Tool | File Type | Format | Storage space |
|---|---|---|---|---|---|---|
| Human | 200 | PACS/CT | CT-scanner | Imaging | .dcm | 150-1000 GB |
| Human | 200 | eCRF | SPSS | measurements | .sav | <1GB |
| Human | 30 | publiced literature | N.A. | PDF | .pdf | <1GB |

**2.2 Do you reuse existing data?**

- Yes, please specify

In this retrospective study, we will reuse CT scans that have already been made as part of a diagnostic process or for pre-operative planning and postoperative evaluation. CT scans from adult patients with known hip dysplasia from the Sint maartens kliniek Nijmegen will be analysed and compared to existing liturature.

The liturature is also a reuse of data, but only the publicly available results and conslusions will be used to compare our data to: Florkow MC, Willemsen K, Zijlstra F, Foppen W, van der Wal BCH, van der Voort van Zyp JRN, Viergever MA, Castelein RM, Weinans H,

van Stralen M, Sakkers RJB, Seevinck PR. MRI-based synthetic CT shows equivalence to conventional CT for the morphological assessment of the hip joint. J Orthop Res. 2022 Apr;40(4):954-964. doi: 10.1002/jor.25127. Epub 2021 Jul 12. PMID: 34191351; PMCID: PMC9291600.

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

| Type of data | Who has access |
|---|---|
| Raw CT-scans (without identifying data)* | N. Koning (data manager)<br>B.C.H. van der Wal (Principal investigator)<br>M.F.T. Hüsken (Coordinating researcher)<br>J. Magre (researcher)<br>V. Arbabi (researcher) |
| Measurements derived from CT-scans | N. Koning (data manager)<br>B.C.H. van der Wal (Principal investigator)<br>M.F.T. Hüsken (Coordinating researcher)<br>J. Magre (researcher)<br>V. Arbabi (researcher) |

**Explanation:**

* The raw CT-scans will be ripped of all direct and indirect identifying data by the radiology department of the St Maartens clinic Nijmegen. The CT scans will be coded according to the order of entry. There will not be a key-linking table to enable patient re-identification.

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

Data/measurements derived from the CT-scans will be kept and analyzed in SPSS. A codebook is automatically generated from the column names. When analyzing in SPSS, every step will be saved in syntaxes, with the addition of the date.

Important measurements will always be done twice and checked by others. Data will be kept in different versions with the addition of the date after the filename, as mentioned above, therefore the last version will be fully up-to date and the last additions or lost data can easily be retrieved.

| # | 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 data manager | x | | |
| 2. | Storage and archiving | | | x* |
| 3. | Data Capture Tool license fee (Castor) | x | | |

**Explanation:**

* The cost of the storage in the Research Imaging Architecture (RIA) will be funded through a larger project on hip dyplasia

**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.**

Sint Maartens Kliniek Nijmegen is the owner of the raw CT scans that originate from their hospital. The UMC Utrecht is allowed to use the CT scans to collect data for this study and to analyse the dysplastic pelvic for future treatment developments. 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. This is contractually agreed upon in a Data Transfer Agreement.

## 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

The original scans are tranfered and stored without personnel data nor re-identifying key. Name, date of birth or any other high risk specifics from these patients will not be available to the researchers.

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

| Which personal data? | Why? |
|---|---|
| CT scans | To answer our research question |

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

- No objection, please explain

Written consent will not be obtained on ground of general interest and general no-objection, because this retrospective study will use CT-scans without identyfing data. The scans will be from up to 20 years back and the effort of reaching them will not weigh out the burden for these patients, knowing the CT scans will already have been made. A “Objection check” will be performed before transferring the scans to the research team. The scans will be ripped from identifying data before transfered and the researcher from de UMC Utrecht will therefore have no access to the personal data.

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

The data will be transferred and stored using the Research Imaging Architecture (RIA) of the UMC Utrecht. The raw data in terms of the CT scans will be handled confidentially and will be coded according to the order of entry. CT-scan from the Sint Maartens Clinic will be transferred without identifying data. Only the clean Pelvic CT scans will be available for the research team. Confidentiality will be maintained at all times, participant information will not be available for the researchers and can therefore not be disclosed to third parties. Derived data and measurements will be reused and stored in secure research folder on network disc of the division. All generated (meta)data will be stored in a designated secure research folder with RFS for access control. There wil be no key-linking table to enable patient re-identification, so patient data will never again be accessible.

| | |
|---|---|
| Right of Access | We have to refuse participant's right of access. It wil not be possible, because there will not be a key to re-identify patients. |
| Right of Rectification | We have to refuse participant's right of rectification. It wil not be possible, because there will not be a key to re-identify patients. |
| Right of Objection | A “Objection check” will be performed during selection of the patients and before transferring to the UMC Utrecht. |
| Right to be Forgotten | 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.**

There will be no linking-key, to re-identify the CT/scans. The CT scans will be tranfered without personnel data and stored using RIA,

a centralized and secured storage for radiological research data in the UMC Utrecht, designed for this specific purose. Derived data and metadata will be stored in a designated secure research folder with RFS for access control.

**3.5 Describe how you ensure secure transport of personal data and what contracts are in place for doing that.**

The CT scans from the St maartens Clinic in Nijmegen will be derived of directly identifying data by the radiology department in Nijmegen. No key will be kept. Transportation of the CT scans will be facilitated by the Research Imaging Architecture (RIA) of the UMC Utrecht. This is a central, safe storage for research immaging. RIA-eXchange (Nextcloud) can be used for savely transferring the data.

## 4. Data Storage and Backup

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

Raw data in terms of the original scans will be transferred without identyfing data and stored in the hospitals Research Imaging Architecture (RIA). The derived data will be stored on the standard IT-system of the UMC Utrecht, in the secured Research Folder Structure with restricted access for the research team only. The scans and data will not be accessible outside of the Hospital network. All versions of the databases will be saved before editing with the addition of the current date to directly backup previous versions.

**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). For analysis, all versions of the databases will be saved before editing with the addition of the current date to directly backup previous versions.

## 5. Metadata and Documentation

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

For the data collected in SPSS we use metadata standards for the codebook and analysis.
Metadata that will be generated are the following:
- Data of the measurements and researcher taking the measurements as described in section 7.1 of the study protocol.
- Syntax for the data analysis
- Data analysis results
- Publication on the study results.

**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 date with initials after each edit, for example 02-02-2022, MH. 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 date in the filename. The most recent version is kept and older versions are moved to a folder OLD.

## 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 is stored in the project folder, 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 (scans), the study protocol describing the methods and materials, The derived data (measurments), the script to process the data, the 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 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. The data package with measurements (not the raw CT-scans) will be published in the UMC Utrecht repository (DataverseNL).

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

The publication will be open accessible. The study protocol and anonymized data, needed to reproduce findings will also be available. A DOI will be assigned to all publications.

## 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. Our processed genetic data can be of interest for other Europeans researchers in the field.

**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?**

- Yes (please specify)

As the data is privacy-sensitive, we publish the descriptive metadata, the derived measurements and models. Patient data and the raw CT scans will not be shared.

**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 accessable. 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 as soon as article is published

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

The data package with measurements (not the raw CT-scans) will be published in the UMC Utrecht repository (DataverseNL). Dataverse is open source software developed by Harvard University. Results and meta data will be published in an article.