Plan Overview

A Data Management Plan created using DMPonline

Title: Implementation of OCTN2 deficiency in the Newborn Screening Program: is it usefull and feasible?

Creator: Loek Crefcoeur

Principal Investigator: Loek Crefcoeur

Data Manager: Loek Crefcoeur

Affiliation: UMC Utrecht

Funder: ZonMw (Nederlands)

Template: UMC Utrecht DMP

ID: 32873

Last modified: 04-12-2022

Grant number / URL: 543002003

Implementation of OCTN2 deficiency in the Newborn Screening Program: is it usefull and feasible?

1. General features of the project and data collection

1.1. Please fill in the table below.

ABR number	NL67683.041.19
METC number	19-234
Acronym study title	ODIN
Name Research Folder	OCTN2
Name Division	Kinderen
Name Department	Metabole ziekten
Partner Organization	Amsterdam UMC
Start date study	1-1-2019
Planned end date study	31-12-2022

1.2 Select the specifics that are applicable for your research (more than one option possible). If necessary, add text in the additional comment area.

- Prospective study
- Retrospective study
- Use of Questionnaires
- WMO
- Multicenter study
- Observational study

2. Data Collection

2.1 Give a short description of the data for your research, including

- **the source of the data**
- **what tools you use for Data Capture**
- **the type of data, the size of the data**
- **the format of the data**

To answer the research question "is the addition of OCTN2-deficiency to the Dutch newborn screening program useful and feasible" we need the following data from approximately 1000 (exact numbers are yet unknown, and will be discovered by this study) subjects who were, at one point in time, suspected of having OCTN2-deficiency:

- Biochemical data; blood and urine tests
- Clinical data; medical history, current clinical status, physical examination.

In subjects with confirmed OCTN2-deficiency diagnosis additional data will be captured:

- Biochemical data; blood and urine tests, skin biopsy, genetic testing.
- Clinical data; follow-up of clinical status

Study subjects	Date Source	Data Capture Tool	File Type	Volume (records, MB, GB, TB..)	Format
Human	EPD	Research Data Platform or EPD (for medical centres other than UMCU)	Quantative data	1000 records	.CSV
Human	eCRF	Castor EDC	eForms	1000 records	.CSV or .xls

2.2 Do you reuse data from other researchers or from the EHR?

- No (please specify)

In this retrospective study we will use pseudonymized data of the RDP.
Missing data will, if possible, be prospectively collected.

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

1. Existing health data of patients are provided by the Research Data Platform. Selection is made for patients that did not object (NL. geen bezwaar) against the use of their health care data for research. The data manager is authorized to link different datasets of the selected patientgroup and thus has access to the identifiable data such as patientID. Researchers not mentioned in informed consent will only have access to the pseudonymized research data.

Tabel 1 Data Access Table

RDP Datamart - personal data	Datamanager/researcher
RDP Datamart - pseudonimized data	PI, research team
Data from external sources (e.g. other hospital)	Datamanager/researcher (who also deals with contracts and paperwork) and PI from external hospital
Key table linking study specific IDs to Patient IDs	PI (with care relationship to patient), data manager
Linked pseudonymized research data set	PI, research team

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

All data collected from the EPD are protected by an audit trail.
Castor EDC (for eCRF-capture) is also protected by an audit trail.
When analysis is performed extractions will be made from sources mentioned above. Extraction scripts are saved on a secured drive. Frozen copies of the extraction are saved as mentioned in sub 5.2.

2.5 Specify costs involved in managing and storing your data.

Type of costs	**Costs**
datamanager time	0
analysis of lab data	n.a.
licence fees for questionnaires	900
hosting of data	1350
design of eCRF	0
medical devices costs	n.a.
storage costs	0
archiving costs	0

2.6 State if intellectual property rights (IPR) are applicable on your data collection and state which agreements will be or are made.

IPR is not applicable.

3. Legislation/ Data Protection Impact Assessment

Check the proper option below to give insight in whether you need to fill out this section or not.

- My study is a complicated multicenter study. I therefore need to fill out the elaborate DPIA and store the report of measures taken in the same folder as the DMP. I therefore skip this section and proceed to 4.1.

3.1 Describe the purpose of the collection of personal data and of the medical data.

We need to know name and address of participants in our research to be able to contact them for questionnaires and invite them for taking part in the research. Medical data are collected for analysis of disease outcome and prognostication.

3.2 Describe systematically how you will process privacy sensitive data in your research.

Data are pseudo-anonymized using a participant key. This key is password protected and stored on a special partition within UMC-Utrecht, which is only accessible to the researcher and the monitor.
Genetic data are collected in a seperate, password protected, file. This file only contains the pseudo-anonymized key and no other privacy sensitive data.

3.3 Describe how you manage the rights of study participants.

Right	*Example answers:*
Right of insight	research data are coded, but can be linked back to personal data, so we can generate a personal record. 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 can not be granted because this will 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 (Beveiligde Onderzoeks Map) to make sure that only authorized personnel have access to personal data. Also we make use of GCP compliant Data Capture software (Open Clinica).

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

For sending personal data to colleagues we use Surffilesender.
We have a Data Transfer Agreement with Amsterdam UMC, Erasmus MC, Radboud UMC, Maastricht UMC+ and UMC Groningen. For all, two copies were signed, one copy is stored in UMC Utrecht and the other is stored in the repespective research site.

4. Data Storage and Backup

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

The data will be stored within the firewall of the UMC on the network drive (L:\Onderzoek-FAO\OCTN2) in a folder protected by permission rights. We will need +/- 50 gig 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 t a backup is regularly made automatically by the 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. Castor EDC generates a codebook upon extraction.

5.2 Describe where you registered your research project, which standards you use for filenames and how you keep track of versions.

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.

We will be using SAS, version 9.4, for statistical analysis of the data. The scripts will contain comments, such that every step in the analysis is documented and peers can read why I made certain decisions during the analysis phase.

7. Data Preservation and Archiving

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

In view of the regulation for Clinical Trials, I need to save all data for at least 15 years with the goal to be able to go back to patient level. After finishing the project all documents and data (data package to be described later) are stored at the UMC Utrecht L:\Onderzoek-FAO\OCTN2 and maintained by P. Anbeek, datamanager.

7.2 Describe which archive or repository (include the link!) you will use for long-term archiving of your data collection once the project has ended and whether the repository is certified.

This will be done conform standards used by UMCU

7.3 Give the persistent identifier (the ISBN for your data) that you will use as a permanent link to your data collection.

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. Data cannot be reused without consulting the research team of metabolic diseases in UMCU.

8.2 Describe the related information that will be available with the data.

The publication will be open assessable. Also the study protocol and the datamanagementplan will be available.

8.3 Describe when the data or metadata will be available, under which criteria and for how long.

After finishing the project the metadata will be available. Data will not be available because of privacy legislation but the metadata will be open. In the event that peers like to reuse our data this can only be granted if the research question is in line with the original research question of our research. Every application therefore will be screened upon this requirement.

8.4 Describe where you will make your data findable and available to others. A link to the data repository or a link to the data should be provided.

1. As the data is privacy-sensitive, we publish the descriptive metadata in the data repository (DataverseNL), with a description of how a data request can be made. If granted, a data usage agreement is signed by the receiving party.