Plan Overview

A Data Management Plan created using DMPonline

Title: Uitkomsten van NEC bij extreem prematuren

Creator: Maud Lindeboom

Principal Investigator: Jan B.F. Hulscher, Maud Lindeboom

Contributor: Otis C. van Varsseveld

Affiliation: UMC Utrecht

Template: UMC Utrecht DMP

Project abstract:

As of 2010 premature infants with a gestational age of 24 and 25 weeks are actively treated in the Netherlands. Mortality and morbidity are significant, with only 25% of infants surviving without disabilities. One of the most dreadful early complications is necrotising enterocolitis (NEC). NEC is an acute inflammation of the bowel often leading to perforation and death when left untreated. It occurs in some 10-15% of all neonates with a gestational age of 24/25 weeks (Aarnoudse-Moens, 2017). Surgery is necessary in ~50% of NEC cases. In surgical NEC, survival rates rarely exceed 60%, with significant long-term morbidity both in terms of gastrointestinal problems as well as in neurodevelopment.
However, little is known about the short- and long-term results after surgical treatment for NEC in the most vulnerable children: those born at 24/25 weeks. Therefore, we aim to perform a nationwide cohort analysis (going back until 2010) of all Dutch neonates born with a gestational age of 24 and 25 weeks who developed surgical NEC. Main outcome measures are both short term mortality and morbidity and neurodevelopmental outcome at the age of two. These data will provide clinicians and parents at least some guidance into the decision whether to proceed to surgery in these extremely vulnerable infants.

ID: 114112

Start date: 01-01-2019

End date: 01-01-2023

Last modified: 18-07-2023

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

Uitkomsten van NEC bij extreem prematuren

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)*	20-182
DEC number *(only for animal-related research)*	
Acronym/short study title	NEC
Name Research Folder	20-182_NEC
Name Division	Pediatric Surgery
Name Department	Surgery
Partner Organization	University Medical Center Groningen
Start date study	01-01-2019
Planned end date study	01-01-2023
Name of datamanager consulted*	Dax Steins
Check date by datamanager	10-01-2023

1.2 Select the specifics that are applicable for your research.

- Retrospective study
- Non-WMO
- Multicenter study

All pediatric surgical centers in the Netherlands collaborate, therefore it will be a nationwide analysis.

2. Data Collection

2.1 Give a short description of the research data.

For this retrospective multicenter study, all patients with a gestational age of 24/25 weeks and born between 1-1-2010 and 1-10-2022 with radiologically or surgically proven NEC Bell's stage ≥2a will be included. The main objective is to analyse mortality and morbidity of all surgical NEC cases. Other retrospective analysis include: a retrospective analysis, up to 1-10-2022, of neurodevelopment at the age of two years of all surgical NEC cases in neonates born at 24/25 weeks of gestational age between 1-1-2010 and 1-10-2022, and a cost-effectiveness study

Patient data will be manually extracted from the hospital electronic patient files (EPD; HiX) and stored in an Excel spreadsheet in a pseudonymized fashion, keeping the key file strictly separate.

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format	Storage space
Human	40	EPD (HiX)	Microsoft Excel	Quantitative	.xlsx	0-10GB

2.2 Do you reuse existing data?

- No, please specify

Preliminary literature review showed that international data exists on the topic, but no Dutch, nationwide data is available. This kind of data has not been collected in a database before at UMCU. However, all required primary data can be obtained from the EPD.

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

Type of data	Who has access
Direct identifying personal data	Local investigator UMCU (Dr. M.Y.A. Lindeboom) with care relationship to patients, Datamanager UMCU
Key table linking study specific IDs to Patient IDs	Local investigator (Dr. M.Y.A. Lindeboom), Datamanager UMCU
Pseudonymized data	Local investigator (Dr. M.Y.A. Lindeboom), PhD Datamanager UMCU

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

Pseudonymized health data from patients will be collected in an electronic Case Report Form (eCRF) in the form of an Excel spreadsheet which will be stored in the secured Research Folder Structure of the UMC Utrecht. Importantly, personal data is stored separately from other research data and adequate access and control rights are in place. In other participating sites, data and documentation will be stored accordingly. Data collection will be frozen before analysis and versions will be recorded.

#	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 datamanager		x	
2.	Data storage		x	
3.				
4.				
5.				

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 will be made available, in a secure and pseudonymized manner, to a collaborating university medical center (UMCG) as a service to the research community. A data sharing agreement defining all the specifics will be signed by both parties before starting data collection for this project.

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 consulted the division datamanager 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?**
Patient ID	To be able to trace back to source of the health data (pseudo ID) in case of missing or erroneous health data
Data on NEC surgery and short- and long-term outcomes (retrospective from EPD): 30-day mortality; 2-year neuropsychological outcomes, Gestational age, birth weight, gender, presence of a patent ductus arteriosus, other congenital anomalies, age at NEC development, date of NEC development, Bell stage (severity of NEC), age at surgery, date of surgery, type of surgery performed (resection, anastomosis or ostomy, open/close procedure etc), ascites, NICU stay, hospital stay, need for subsequent surgery, complications of NEC/surgery.	We need these data to answer our research questions
Gender, gestational age	To describe the study population and as potentially relevant risk factors for mortality or morbidity

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

- Study-specific informed consent

Only in the case of a deceased child we will apply the 'no objection' instead of informed consent, in accordance with the Dutch law Artikel 7:458 BW; to minimize the psychological burden for parents.

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

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 objection: informed consent applies.

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.

Health data from patients will be extracted from the EPD and pseudonymized by the local investigator (dr. M.Y.A. Lindeboom) in an electronic Case Report Form (eCRF) in the form of an Excel spreadsheet. This file will be stored in the secured Research Folder Structure of the UMC Utrecht. Importantly, personal data is stored separately from other research data and adequate access and control rights are in place. In other participating sites, data and documentation will be stored accordingly.

The key file linking pseudo-ID to the patient ID will be kept within the UMCU and only be accessible by the local investigator of the UMCU (Dr. M.Y.A. Lindeboom) and the datamanager of the UMCU.

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

We have a signed Data Transfer Agreement with UMC Groningen. The agreement is stored at location: *location will follow.*

4. Data Storage and Backup

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

Digital files will be stored in the secured Research Folder Structure of the UMC Utrecht. For analysis, pseudonimized data will be entered/stored in REDCap. Personal data is stored separately from other research data and adequate access and control rights are in place. The key-linking table will not be shared with other participating sites.

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, data is stored on the data share(s) of one/more DRE workspace(s). DRE data shares are stored on Azure according to LRS standard, ensuring data is replicated thrice in the same Azure data center. Additionally, for version history, DRE takes daily snapshots of the data share, and retains these snapshots for 30 days rolling.

5. Metadata and Documentation

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

The Excel spreadsheet in the secured UMCU research folder structure will include a data dictionary: each variable and its values will be explained.

5.2 Describe your version control and file naming standards.

In the secured Research Folder Structure of the UMC Utrecht, raw data files will be kept in a separate folder from analysis files. Also, newer versions of the same file will be named with "v[number]" and a date of editing "yyyymmdd".

For analysis files, 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 and older versions are moved 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.

The general analysis plan can be found that will be applied in the collaborating center (UMCG) can be found in the study protocol at the same location as the Data Sharing Agreement (see 3.5). IBM SPSS Statistics 28 will be used for analysis.

We will compare the outcome of neonates born at 24 weeks with the outcomes of neonates born at 25 weeks, using Chi2 statistics for categorical data and t-test/Mann Whitney U for continuous data. We will search for prognostic factors using logistic regression analysis with variables significant in univariate analysis.

After correcting for prematurity, Bayley scores 70-84 (<-1,0 SD) will be classified ‘mild impairment’, scores 55-69 (<-2,0 SD) ‘moderate impairment’, scores < 55 (<-3,0 SD) ‘severe impairment’. In children with cerebral palsy motor function will be classified following the ‘Gross motor function classification system’. Grade I denotes a mild impairment, II/II a moderate impairment and IV/V severe impairment

7. Data Preservation and Archiving

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

In the UMCU, currently no specific analyses are planned to be applied. No specific documents are needed to reproduce the collection of the pseudonymized health data. Of course, the raw data, the study protocol and the data sharing agreement will be stored on UMC Utrecht network drives.
The data package that will be kept at the UMCG (in their appropriate research folder structure) after analyses will contain: the raw data (pseudonymized), the study protocol describing the methods and materials, the analysis plan, 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 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 raw data, study protocol and DSA will be stored at the UMC Utrecht Research Folder Structure and is under the responsibility of the Principal Investigator of the research group.

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

I cannot publish the dataset in an external repository. Therefore, I do not have a PID.

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. In this light, the research data may be reused by peers.

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, a data request can be made by contacting the corresponding author or principal investigator 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 informed consent signed by the study participants. Every application therefore will be screened upon this requirement. If granted, a data usage agreement is signed by the receiving party.

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

All data and documents contained in the UMCG data package mentioned in 7.1 may be shared under restrictions, taking into account the conditions of the UMCU-UMCG data sharing agreement. IBM SPSS Statistics software will be needed to reuse the data.

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

(Meta)data will be available as soon as the article is published and will be available for the minimum period of 15 years mentioned at 7.2.

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

It will be stated with the published article that data is available upon reasonable request.