Plan Overview

A Data Management Plan created using DMPonline

Title: Genome wide association study bij acute retina necrose

Creator: Joke de Boer

Principal Investigator: Joke de Boer

Data Manager: Joke de Boer, Dax Steins

Affiliation: UMC Utrecht

Funder: Erasmus MC

Template: UMC Utrecht DMP

Project abstract:

Acute retina necrose (ARN) is een ernstige visusbedreigende infectie van het netvlies met herpes simplex virus (HSV) en/of varicella-zoster virus (VZV). Indien deze oculaire infectie niet vroegtijdig intensief behandeld wordt, kan de infectie uitbreiden naar het andere oog wat tot bilaterale blindheid kan leiden. Een deel van de ARN patiënten heeft ooit een herpes encefalitis (HE) doorgemaakt wat een potentieel dodelijke ziekte is. De relatie tussen beide HSV/VZV-geïnduceerde ziektebeelden doet zowel een mogelijk verhoogde infectiviteit van het veroorzakende herpesvirus virus voor hersenen en/of het oog als een mogelijke stoornis van de aangeboren afweer vermoeden. Dit is bij HE in detail bestudeerd, maar bij ARN nog niet onderzocht. In deze pilot study willen wij met behulp van een nieuwe moleculair biologisch platform, genome-wide association study (GWAS), onderzoeken of er vergelijkbare dan wel nieuwe mutaties in het DNA van het virus en/of ARN patiënt aanwezig zijn. Patiënten met ARN die eerder oogvocht en/of bloed hebben afgestaan aan de biobank uveitis van het UMCU als een broad consent hebben getekend worden geselecteerd voor deze studie. De GWAS en analyse van de uitkomsten zal verricht worden in het Erasmus MC. Indien er mutaties gevonden worden zal er een uitgebreider onderzoek met een grotere ARN patiëntengroep verricht worden. De uitkomsten zullen bijdragen aan het identificeren van (post-HE) ARN patiënten met een ernstig beloop van een HSV/VZV infectie welke therapeutische en preventieve consequenties kunnen hebben.

ID: 71874

Last modified: 17-06-2021

Copyright information:

Genome wide association study bij acute retina necrose

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)*	21-264
DEC number *(only for animal-related research)*	
Acronym/short study title	GWASARN
Name Research Folder	21-264_GWASARN
Name Division	DHS
Name Department	Ophthalmology
Partner Organization	Erasmus MC
Start date study	
Planned end date study	01-04-2023
Name of datamanager consulted*	D. Steins
Check date by datamanager	24-02-2021

1.2 Select the specifics that are applicable for your research.

- Fundamental / translational study
- Multicenter study
- Biobank approval needed: "uitgifteprotocol"

Centrale Biobank Uveitis: 12-514/C, TcBio number: 21-264

2. Data Collection

2.1 Give a short description of the research data.

UMCU is not directly involved in this research project. We shall only provide biomaterial and associated medical data (gender, age, diagnosis, virus type, and medical history).
The study aims to investigate whether there are genetic mutations in the herpes virus as well as the patient, which may increase the risk of acute retinal necrosis. For this study, we shall use biomaterial of 10 ARN patients whose biomaterial are stored in the UMCU's Centrale Biobank Uveitis: 12-514/C

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format	Storage space
Human	10	Biobank (12-514/C)	N/A	N/A	N/A	N/A
Human	10	EPD (HIX)	Exel	Quantitative	.xlxs	0-1 GB

2.2 Do you reuse existing data?

- Yes, please specify

In this retrospective study, we reuse data from HiX (made available for research by the research data platform) and the Central Biobank Uveitis.

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

1. 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.

Experimental data from patients will be collected in Excel.

#	**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 datamanager	x		
2.	Storage	x		
3.	Archiving			Erasmus MC

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.

Erasmus MC is the owner of all collected data for this study. There is a material transfer agreement (MTA) with Erasmus MC.

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

2. I will process personal data. I have consulted the division datamanager and I do will fill out a mini DPIA, including 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 demographics (Gender, Age)	to identify a relation with risk for the disease
Medical background (Diagnosis, ?	encephalitis yes or no HSV or VZV virus
Biobank material (Blood samples, ..?	DNA

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

- Other, please explain

Broad consent registered under 12-514/C (Centrale Biobank Uveitis)

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

1. The data are pseudonymized and the linking table to personal data is saved at the site in a secure research folder. An authorized person manages the linking table, can re-identify study participants when necessary and deliver, correct or delete the data.

Right	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 have a Material Transfer Agreement with Erasmus MC. The agreement is stored at location: L:\Onderzoek\Oogheelkunde\21-264_GWASARN\B_Documentation\6_Contracts
We will transport pseudonymized data with colleagues at Erasmus MC. For that we use SURFDrive with encryption.

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.

N/A

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.

Analysis is performed in Erasmus MC.
WGS is applied to detect variants that are linked to disease progression in individual consented ARN patients. We compare variants of each individual WGS sample against 500 well and elderly WGS runs. Variants are annotated using ANNOVAR including pathogenicity scores e.g CADD scores as well as frequencies in control cohorts (GnomAD, 1K). Furthermore, VarElect will be used to prioritize candidate genes based on diseases or phenotypes of interest.

This data is kindly provided to prof. Peter J. van der Spek (Erasmus MC) by Dr. Eric Topol (Scripps). Hereto, we are able to enrich for rare variants present in the genome of each ARN patient. Once multiple variants in the same candidate gene appear, those candidate genes are followed up with functional studies in both the departments of Immunology and Viroscience of the Erasmus MC (Rotterdam).

Pathway analysis of candidate gene mutations are explored in the commercially available pathway database IPA from Qiagen. This allows us to explore whether mutations that accumulate in genes are part of a known pathway based on previously described Protein-Protein interactions.

7. Data Preservation and Archiving

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

This wil be performed in Erasmus MC

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 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.

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)

The PI's of the Erasmus MC will decide about sharing data.

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

NA

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

- (Meta)data will be available after completion of project (with embargo)

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

I will publish my data in the ArrayExpress repository (www.ebi.ac.uk/arrayexpress/). The repository does not have a data seal of approval, nor issues persistent identifiers. However, it is very established, international, and a standard go to as a source of information in the field of functional genomics. ArrayExpress is indexed by Thomson Reuters Data Citation Index and by SCOPUS. Datasets submitted directly to ArrayExpress are known as good sources for reuse as they are curated by a team of specialist biological curators. Data is collected according to MIAME standards.