Plan Overview

A Data Management Plan created using DMPonline

Title: Malaria, diabetes and metabolic syndrome: investigating the double burden of disease

Creator: Katja Wyss

Principal Investigator: Anna Färnert

Data Manager: Katja Wyss

Affiliation: Karolinska Institutet

Funder: Swedish Research Council

Template: Swedish Research Council Template

Project abstract:

The aim of this collaborative project is to investigate whether diabetes and obesity affects the risk of severe malaria, and whether these underlying conditions affect susceptibility to chronic asymptomatic parasitaemia and antibody mediated immunity to malaria in adults. Many sub-Saharan African countries are facing an increase in non-communicable diseases (NCDs) such as diabetes, obesity and metabolic syndrome due to lifestyle changes. At the same time infectious disease such as malaria remain highly endemic in these areas. A recent study in Sweden detected a markedly increased risk of severe malaria in travellers with diabetes and obesity. Whether this is also found in endemic areas is yet to be confirmed. The purpose of this project is to establish a collaboration between institutions in Cameroon and Sweden; and to develop research projects addressing the interplay between NCDs and malaria. We will perform a hospital-based study of comorbidities impact on malaria severity, and a cohort study assessing malaria parasite prevalence and immunity in diabetics. The proposed collaboration builds on complementary expertise that will enable novel interdisciplinary findings of high public health importance. The collaboration includes workshops, research visits and planned PhD, and will be mutually beneficial for the research and development of both partners.

ID: 92254

Start date: 31-01-2022

End date: 31-12-2024

Last modified: 24-01-2022

Grant number / URL: VR 2020-09454

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

Malaria, diabetes and metabolic syndrome: investigating the double burden of disease

General Information

Project Title

Malaria, diabetes and metabolic syndrome: investigating the double burden of disease

Project Leader

Anna Färnert

Registration number/corresponding, date and version of the data management plan

92254

Version

1

Date

2022-01-22

Description of data – reuse of existing data and/or production of new data

How will data be collected, created or reused?

The two studies will be performed at the District Hospital Dschang and the Bafoussam Regional Hospital in Western Cameroon where patients with malaria and diabetes will be included for collection of clinical data and blood samples at multiple time points.

For the first study, adult patients diagnosed at the two hospitals with both severe and non-severe malaria will be included after informed consent, and clinically evaluated for diabetes and metabolic syndrome (according to IDF criteria). Weight and height will be measured for body mass index (BMI). Routine blood samples will be drawn as part of the regular management of patients with malaria, but in addition to routine lab parameters below, clinical chemistry will include fasting plasma glucose, glycated haemoglobin (HBA1c), lipids, ESR, liver enzymes, bilirubin, creatinine and haemoglobinopathies such as sickle cell trait. HIV test will also be offered to all patients. For patients admitted to hospital, fasting plasma glucose, parasitaemia and insulin levels will be measured on a daily basis. Malaria parasite density will be assessed by microscopy of blood films stained with Giemsa, according to normal routines.

In the second study, a cohort of diabetics and healthy comparators will be followed prospectively during a year for assessment of asymptomatic and symptomatic malaria as well as immune responses to malaria infection. Adult patients with type 2 diabetes, with and without the metabolic syndrome, will be identified within the outpatient setting and invited to participate in the study. Community comparators matched on age and sex will be identified in the patients' family or in the home community to control for malaria transmission. Data on demographics and malaria exposure, including bed net use and housing, will be acquired through a structured questionnaire created with the form designer of the EPI info software. Testing for asymptomatic malaria parasitaemia will be performed by microscopy and PCR four times per year and include lab tests and evaluation of body mass index. Serum biochemistry will include same parameters as described for study 1 as well as insulin concentration. In addition, plasma samples will be used to analyse antibody responses to *P. falciparum* whole parasite extract and a panel of antigens using ELISA and a Luminex assay including a panel of merozoite antigens. Venous EDTA blood will be stored frozen as plasma and packed cells.

What types of data will be created and/or collected, in terms of data format and amount/volume of data?

Clinical data, such as vital signs and routine blood chemistry results from patients hospitalized for malaria, or weight and height and blood pressure for the diabetic cohort, will be directly entered into the Sensivo platform accessible from both Cameroon and Sweden. Demographic data (such as sex, age, educational level, information on bed net usage) from the questionnaires will also be extracted to the Sensivo platform. Some demographic and clinical data as well as results from certain laboratory and the antibody analyses will be stored in Excel files at the Center for Molecular Medicine (CMM) on the Karolinska University Hospital site in Stockholm. We expect there to be approximately 300 and 500 individuals respectively in the two studies. For each individual approximately 50-80 variables will be collected, which would result in an estimated 1000kB per file. At a later stage, part of the data from the Sensivo platform and Excel files will be imported to STATA for statistical analyses.

Documentation and data quality

How will the material be documented and described, with associated metadata relating to structure, standards and format for descriptions of the content, collection method, etc.?

Personal data connected to study participants, such as names, birthdates, addresses and telephone numbers will be coded with a key stored in a locked safe and under password protection on the CMM server. All data associated with each individual will be traceable with this code. The data collection process as well as the methods for antibody analysis are documented in the lab SOP , Klara (KI risk assessment), and will be documented in the KI Electronic lab notebook (ELN). The data analysis process and statistical codes will be described in a STATA log-book format and shared in the ELN.

How will data quality be safeguarded and documented (for example repeated measurements, validation of data input, etc.)?

Handling of patient material and relevant sample information will be noted in the KI ELN together with results. Experiment validation of antibody analyses by ELISA and a Luminex assay include data quality standards with negative and positive control samples in repeated experiments. In the prospective study with longitudinal sampling of antibodies at four time points, the individual will be used as its own control.

The collection of clinical data will be carried out by one of the investigators and then reviewed by 1-2 responsible researchers in order to detect any inconsistencies or outliers. In addition, the Sensivo platform is specially designed to detect inconsistent entries or outliers.

The questionnaires will be filled out by specially assigned medical officers after an introductory training, and completed questionnaires will be reviewed by responsible researcher on site.

Before the statistical analyses are run, the data will also be checked for errors and outliers within the statistical program used (STATA).

Storage and backup

How is storage and backup of data and metadata safeguarded during the research process?

Patient questionnaires will be stored in locked cabinets in the project office in the University of Dschang, accessed only by the principal investigator on site. Clinical information and data from questionnaires will be entered into the Sensivo Platform, where data initially is stored. A copy of the raw data in Excel format will be exported from Sensivo and together with analysed data files and statistical codes stored in a designated project folder on the CMM server, which is protected by a firewall and frequently backed up. Analysis scripts will also be stored on the CMM server and displayed in the KI ELN (Electronic Logbook). ELN is backed up by KI servers.

How is data security and controlled access to data safeguarded, in relation to the handling of sensitive data and personal data, for example?

Only researchers responsible for data collection and data analyses will have access to the data on the Sensivo Platform. Data files stored at the CMM server with clinical data or personal identifiable information will be password protected and only accessible by the project leader and main investigators. Researchers involved in data analyses will only have access to pseudonymised data.

Legal and ethical aspects

How is data handling according to legal requirements safeguarded, e.g. in terms of handling of personal data, confidentiality and intellectual property rights?

All data handling is done in accordance with the General Data Protection Regulation (GDPR), the legal framework for the collection and processing of personal information in the European Union (EU). A collaboration agreement between Karolinska Institutet and University of Dschang in Cameroon has been prepared in accordance with the Order of Decision and Delegation at KI. A data sharing and material transfer agreement has been signed by the project leaders at both study sites (Prof Anna Färnert and Prof Simeon Choukem).

The amount of data collected from patients is only what is needed for analyses. All study subjects will have received both written and oral study information before agreeing to participate in the study by signing a written consent. The electronic version of the data collected will be created in the Sensivo platform which ensures that data can be handled securely and without unnecessary exposure of personal information.

Names, birth dates and any equivalent to personal identification numbers used in Cameroon, as well as other sensitive information will be removed after clinical data is added to the electronic database and replaced with study ID numbers. All analyses will be done on group level and the study results will be presented so that individual patients cannot be identified.

Patient information will be fully traceable through a key connecting study ID numbers with patient identifiable information. This key will be stored on the CMM server and only accessible by the project leader. Since all patient data is coded, patient-specific information can be traced via the study key. If a study participant wants to leave the study after inclusion, the correct information and samples will be identified via the key and removed /destroyed accordingly.

How is correct data handling according to ethical aspects safeguarded?

The study will be done in accordance with the Ethical Committee in Cameroon as well as the Swedish Ethical Review Authority.

Accessibility and long-term storage

How, when and where will research data or information about data (metadata) be made accessible? Are there any conditions, embargoes and limitations on the access to and reuse of data to be considered?

Analysed data is made available following publication in scientific journals. Underlying raw data, since subjected to GDPR, will be made available upon request after ensuring compliance with relevant legislation and KI guidelines. When data is shared, patient ID's will not be included.

In what way is long-term storage safeguarded, and by whom? How will the selection of data for long-term storage be made?

Data is stored long-term at CMM servers that are backed-up daily and maintained by CMM IT. Data associated with publications will be stored for a minimum of 10 years. Patient questionnaires will be stored for 10 years in locked, fire-proof cabinets in the University of Dschang.

Will specific systems, software, source code or other types of services be necessary in order to understand, partake of or use/analyse data in the long term?

Scripts for statistical analyses will be necessary to use and reanalyse the data in the long term, and will be stored at the KI ELN with KI backup systems.

How will the use of unique and persistent identifiers, such as a Digital Object Identifier (DOI), be safeguarded?

DOI's will be provided when data is published and made available electronically by the publisher. publishing houses. KI has a central database with DOIs of all the published articles, which is backed up regularly.

Responsibility and resources

Who is responsible for data management and (possibly) supports the work with this while the research project is in progress? Who is responsible for data management, ongoing management and long-term storage after the research project has ended?

The main responsibility for data management and storage both during and after project finalization is held by the project leader. The project leader, based at Karolinska Institutet and main supervisor to the planned PhD student at the project site, will together with the assistant supervisors (one from KI and one from Dschang University), be responsible for educating the PhD student and other researchers involved in the project about data management and storage. The investigators in Cameroon will be responsible for data management and long-term storage at the study site. The project leaders at KI will have a supportive role in data management while the project is in progress in Cameroon. A Data Transfer Agreement between Dschang University and Karolinska Institutet will regulate the transfer and use of data.

What resources (costs, labour input or other) will be required for data management (including storage, back-up, provision of access and processing for long-term storage)? What resources will be needed to ensure that data fulfil the FAIR principles?

The project has been funded by the Swedish Research Council/Swedish Research Links Network Grant, the Swedish Foundation for International Cooperation in Research and Higher Education (STINT) grant as well as an internal Sustainable Development Goal (SDG) project grant from Karolinska Institutet. These funds will cover the cost for data management and storage at CMM, supporting the IT departments responsible for data storage and back-up. The research group has a lab manager who has the main overview over sample handling. Data sharing will be handled by the corresponding author for the published results. Storage on ELN and publication ensures fulfilment of FAIR principles as well as the usage of commercial programmes for generating and analysing data.