Plan Overview

A Data Management Plan created using DMPonline

Title: BIO-002

Creator: Callistus Iwuagwu

Principal Investigator: Dr Ruth Payne

Project Administrator: Jee-Sun Cho PhD

Affiliation: The University of Sheffield

Funder: Medical Research Council (MRC)

Template: MRC Template

Project abstract:

Malaria in humans is caused by five species – Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale, Plasmodium malariae and Plasmodium knowlesi. P. falciparum causes the most
morbidity and mortality of the Plasmodium species, accounting for an estimated 241 million cases
of malaria and 627,000 deaths worldwide in 2020. Almost half of the world's population are at risk
of malaria, with sub-Saharan African populations at highest risk of acquiring malaria: approximately 95% of cases are estimated to occur in the World Health Organisation (WHO) African Region. Children under five years of age are the most severely affected, accounting for 77%
of deaths from this infection (1).
Since 2000 massive efforts have been made to increase distribution of commodities to prevent
malaria across Africa, with a significant reduction in malaria deaths. It is estimated that 1.7 billion
malaria cases and 10.6 million malaria deaths have been averted since 2000 with most of those
averted, 82% of the cases and 95% of the deaths, in Africa.(1) However, challenges to the success
of current strategies to combat malaria (such as insecticide-treated nets, indoor residual spraying,
and antimalarial drugs) include: the development of resistance of Anopheles mosquitoes to certain
insecticides; the development of resistance of malaria parasites to chemotherapeutic agents (2);
the absence of a gametocidal drug suitable for mass administration (3); and the risk of re-importation of malaria into geographic regions previously cleared of malaria using environmental
elimination measures.
The Roll Back Malaria (RBM) Partnership was launched in 1998 by the WHO, the United Nations

Children’s Fund (UNICEF), the United Nations Development Programme (UNDP) and the World Bank. A major goal of the RBM Partnership is to support the development of a vaccine against malaria as a key future strategy for reducing mortality from malaria. The development of an effective vaccine may indeed be necessary for the greater goal of global eradication of malaria (4).
The updated Malaria Vaccine Technology Roadmap calls for the development of a vaccine against
P. falciparum and P. vivax by 2030, that will have protective efficacy of at least 75 percent against
clinical malaria, suitable for administration to appropriate at-risk groups and development of vaccines to reduce malaria transmission suitable for administration in mass campaigns (5).

ID: 121888

Start date: 15-05-2023

End date: 19-01-2025

Last modified: 19-04-2023

Grant number / URL: admin.ox.ac.uk

Copyright information:

BIO-002

0. Proposal name

0. Enter the proposal name

Phase I clinical trial to assess the safety and immunogenicity of the malaria vaccine candidate RH5.1 soluble protein in Matrix-MTM using two dosing regimens

1. Description of Data.

1.1 Type of Study

This is an open-label, single-centre Phase I trial to assess the safety and immunogenicity and efficacy of the candidate malaria vaccine RH5.1 formulated in adjuvant Matrix-M.

1.2 Types of Data

1. Physical samples: These may include tissue samples, blood samples, urine samples, etc. from participants to measure antibody responses to the RH5.1 vaccine.
2. Environmental data: This may include physical obs such as the participant's temperature, heart rate, systolic and diastolic BP, and other physical and chemical parameters.
3. Paper records: These may include participants' medical records, case report forms (CRFs), laboratory reports, informed consent forms, and protocol-related documents.
4. Genetic data: This may include the participant's genomic DNA, mitochondrial DNA, RNA sequencing data, etc.

1.3 Format and scale of the data

The format and scale of the data, including physical samples and paper records, may include the following:

Physical samples:

Format: Biological samples (blood as whole blood, serum and plasma). Urine is collected and tested but not stored.
Scale: 336 blood samples (24 participants * 14 blood samples each).
Storage conditions: Samples collected for ‘safety blood’ will be labelled with patient identifiers and processed in the NHS laboratories in Sheffield Teaching Hospitals. Results from these samples will be available on the local results reporting system (ICE) and participants will consent to this. These samples will be destroyed in the standard timeframes and manner for haematology and biochemistry samples in the NHS laboratory (i.e. no long-term storage).
Source: Sheffield Teaching Hospitals NHS Foundation Trust, NIHR Sheffield Clinical Research Facility, O Floor, Royal Hallamshire Hospital, Glossop Road, Sheffield, South Yorkshire, S10 2JF.
Ethical considerations: Informed consent, Ethics approval. Participants may opt in or out specifically to allow for future research to be done on their samples (subject to approval from relevant ethical committee).

Participant Demographic Data: This includes personal information of the participants (e.g., age, gender, ethnicity, address, email address, date of birth, phone numbers.). This data may be collected on an electronic case report form using the Oxford Vaccine group REDCap database and stored in a structured database format, such as SQL or REDCap's native format, or a paper case report forms (CRFs) transcribed from REDCap. Identifiable data will be held in the CRF manager database - separate from RedCap to allow for the smooth running of the trial. Only age, DOB, gender, ethnicity and email address will be recorded in RedCap itself. The email address is present to facilitate the eDiary surveys.

Clinical Data: This includes medical history and vital signs (temperature, blood pressure and heart rate) results of 'bedside tests' such as urinalysis and pregnancy tests. These will be recorded on the paper checklist as well as in RedCap. If a paper printout is provided by a machine it will be stuck to the checklist.

Laboratory Data: This includes data from various laboratory tests (e.g., Full blood count, serum chemistry, blood borne virus screening), as well as data from the analysis of physical samples (e.g., blood results, urine, etc.). Laboratory data may be collected in electronic formats on REDCap and stored in a structured database format, such as SQL or REDCap's native format, or in paper formats.

Adverse Events and Safety Data: This includes data on adverse events, serious adverse events, and other safety-related data. Adverse events data may be collected on paper CRFs or in REDCap and stored in a structured database format, such as SQL or

REDCap's native format.

Immunogenicity and Efficacy Data: This includes data on the immune response to the candidate malaria vaccine, as well as data on the vaccine's efficacy in preventing malaria. Immunogenicity and efficacy data may be collected from laboratory tests or from electronic/paper CRFs.

Data Scale:

CRFs: Each participant's Case Report Form, including demographic information, previous medical conditions, and test results, may require up to 50 MB of data storage.
Clinical trial documents: Consent forms, questionnaires, and case report forms may require 5-10 MB of data storage per participant.

Laboratory data: Assay results may require 1-5 MB of data storage per participant.
Imaging data: MRI or CT scans may require 50-100 MB of data storage per participant.
Based on these estimates, the total amount of data generated by the trial could be around 1-2 GB for EMRs, 120-240 MB for clinical trial documents, 24-48 MB for laboratory data, and 1.2-2.4 GB for imaging data, resulting in a total estimated data volume of 2.3-4.6 GB.

However, the actual data volume could be higher or lower depending on the specific details of the trial.

2. Data collection / generation

2.1 Methodologies for data collection / generation

Methodologies for Data Collection/Generation:

Participant Recruitment: Participants will be recruited from a single centre in Sheffield and will be screened for eligibility criteria which will include healthy adults aged 18-50 who are at low risk for malaria. Eligible participants will be provided with detailed information about the study, and written informed consent will be obtained before enrolment.

Study Design: This is an open-label, single blinded Phase I trial, which means that the investigators will know which dose of treatment each participant is receiving but the participant will be blinded.

Study Procedures: Participants will receive the candidate malaria vaccine RH5.1 formulated in adjuvant Matrix-M. The amount of RH5.1 given will be different dependent on treatment group. Safety and immunogenicity will be assessed through a range of laboratory tests and clinical assessments. Efficacy will be evaluated by monitoring the incidence of malaria infection.

Data Collection: Data will be collected through a variety of methods, including laboratory tests, clinical assessments, interviews and questionnaires. Laboratory tests will be conducted to evaluate safety and immunogenicity, and clinical assessments will include physical exams and monitoring of adverse events. Interviews and questionnaires will be used to collect data on participant demographics, medical history, and malaria exposure. Specifically:

Physical examination: Participants will undergo a physical examination at baseline and at various time points throughout the trial to assess their safety and any adverse events.

Blood tests: Blood samples will be collected at baseline and at various time points throughout the trial to assess immunogenicity and efficacy.

Questionnaires: Participants will be asked to complete questionnaires at various time points throughout the trial to assess their experience with the vaccine and any adverse events.

Data capture and storage: Data will be captured and stored using REDcap. This system allows for secure, web-based data capture and management, and ensures data quality and security.

Personal details: Personal details will be collected from participants, including name, age, and contact information. However, to protect participants' privacy, data will be pseudonymized using an alpha-numerical ID system (BIO002-01-001). Personal details will only be kept for as long as necessary for the research.

Data Pseudonymisation: Participants will be assigned an alpha-numerical ID, and personal data will be kept separate from pseudonymised data in a different database. Personal data will be kept only for as long as necessary for the research. This will be held in the 'CRF manager' software.

Data Organisation: Data will be organised in an RDBMS structure that reflects the study procedures and data collection methods. For example, data related to laboratory tests will be stored in an instrument labelled "Blood Results D(test day)," and files will be named in a logical, concise, and informative way.

Data Quality Assurance: To maintain consistency and ensure data quality, good practices will be employed, such as version control, naming conventions, peer reviewing data, and using controlled vocabularies where possible. All data will be reviewed for quality and consistency while being entered into REDcap. Data generated by different people or equipment will be reviewed and inputted by two people to ensure consistency and accuracy.

Overall, data collection in this Phase I trial will employ a variety of methods to evaluate the safety, immunogenicity, and efficacy of the candidate malaria vaccine RH5.1 formulated in adjuvant Matrix-M. Data will be collected, organised, and pseudonymised securely, and data quality will be assured through good practices such as maker checker reviewing, and using controlled vocabularies where possible.

2.2 Data quality and standards

In this trial, data will be collected through various methods, including mechanical experiments, interviews, and questionnaires.

Data Quality and Standards to be adhered to:

Accuracy: The data collected must be accurate, reliable and valid. Any errors should be detected and corrected promptly.

Completeness: All necessary data should be collected and recorded, ensuring that no important information is missing.

Consistency: The data should be consistent across all participants and over time.

Timeliness: Data should be collected in a timely manner to ensure that it is up-to-date.

Confidentiality: Personal data should be kept confidential and secure, and only accessible by authorized personnel.

Data Collection:

Data will be collected and stored logically using REDcap, a secure web-based platform for data capture and management. Questionnaires will be designed and administered using the REDcap tool as well as other sample data, and participants will be invited to complete the questionnaire online. For interviews, a consent form will be obtained from the participants before the interview. Interviews will be conducted in person. Personal details will be collected only if necessary and will be pseudonymized using an alpha-numerical ID.

Data Organisation:

Data will be organized in an RDBMS table structure with separate tables for questionnaires, interviews, and trial data. On the front end, the data would be collected and entered into the different folder-like instruments for different activities such as;

Eligibility
Screening
Blood results
followup
randomization/vaccine administration

Data Consistency and Quality Control:

To ensure consistency and quality of data collection, the following processes will be implemented:

Repeat Samples or Measurements: Multiple samples or measurements will be taken over the course of the trial to ensure accuracy and reliability.

Standardized Data Capture or Recording: A standard protocol will be developed for data capture and recording to ensure consistency across all participants and over time.

Data Entry Validation: Data entry will be validated to ensure accuracy and completeness.

Peer Review of Data: All data will be peer-reviewed for accuracy and completeness while being entered into the database.

Controlled Vocabularies: A controlled vocabulary will be used for all questions in the questionnaire to ensure consistency in data collection.

All processes of repeat samples or measurements, standardized data capture or recording, data entry validation, peer review of data, and representation with controlled vocabularies will be documented to ensure transparency and reproducibility of the data. The documentation will include the date, time, and personnel responsible for each step of the process. Any discrepancies or errors will be documented and addressed promptly.

3. Data management, documentation and curation

3.1 Managing, storing and curating data

Data Capture and Storage:

The study will use the REDCap (Research Electronic Data Capture) system database hosted at The University of Oxford by the Oxford Vaccine Group for data capture and storage. REDCap is a secure web-based application for building and managing online surveys and databases. REDcap stores data in a secure SQL database and all data will be backed up on a regular basis to prevent loss or corruption of data. All data will be entered directly into the REDCap database by authorized research personnel using secure login credentials.

Data Backup and Restoration:

The REDCap database will be backed up on a daily basis to ensure data is protected. Backups will be stored offsite in secure, encrypted storage. Data restoration will be tested regularly to ensure a seamless process in case of any data loss.

Data Management:

Data management will involve the oversight, organization, and coordination of data collection, documentation, storage, and quality control. Data entry will be checked for completeness, accuracy, and consistency by two designated personnel. The data will be managed by a dedicated data management team from The CIRG group at The University of Sheffield who would be responsible for ensuring the accuracy, completeness, and integrity of the data. Data will be entered into the REDcap database in a timely and consistent manner, and all data will be reviewed for quality control. Any errors will be corrected promptly, and data will be locked for analysis once all data cleaning and quality control procedures have been completed.

Data Documentation:

Documentation of the data will be conducted using standardized procedures and templates. All data will be annotated with appropriate metadata including the date of collection, the data collector, and any other relevant information. A data dictionary will be developed to document the meaning and structure of the data elements in the database.

Data Curation:

Data curation involves the active and ongoing management and maintenance of data throughout its lifecycle. It includes documentation, quality control, metadata management, and data sharing.

In the short to medium term, data will be regularly reviewed and analyzed for quality control purposes, and relevant findings will be shared for publication as per the study's publication policy. Access to data will be restricted to authorized personnel only, and any requests for data sharing will be considered on a case-by-case basis and will be subject to ethical clearance and confidentiality requirements.

3.2 Metadata standards and data documentation

Metadata Standards and Data Documentation:

Metadata is a key aspect of data management, as it enables research data to be used and understood by others outside of the research team. The following metadata standards and data documentation will be used for the BIO-002 trial

Study Protocol: A detailed description of the study protocol will be documented, including information on the study design, methodology, inclusion and exclusion criteria, and data collection procedures.

Data Dictionary: A data dictionary will be created on REDCap to provide a detailed description of the variables in the study, including data types, coding schemes, and any specific instructions for coding.

Variable Definitions: Detailed definitions for each variable will be documented, including any assumptions or decisions made during data collection or analysis.

Instrument Metadata: Metadata about the instruments used to collect data will be documented, including the name of the instrument, version number, and any modifications made to the instrument.

Provenance Information: Information about the provenance of the data, including data sources, data transformations, and data cleaning processes, will be documented.

Codebook: A codebook will be created in REDCap to document the coding schemes used in the study, including any coding rules or conventions.

Analytical and Procedural Information: A detailed description of the analytical and procedural information used in the study will be documented, including any software or tools used for data analysis.

Data Sharing and Archiving Plan: A plan for data sharing and archiving will be created to ensure that the data is accessible and preserved for future use.

Security and Confidentiality: Information about the security and confidentiality of the data will be documented, including any measures taken to protect participant privacy and confidentiality.

Metadata produced about the data generated from the research will include information such as the study protocol, data dictionary, variable definitions, instrument metadata, provenance information, codebook, analytical and procedural information, and data sharing and archiving plan. This metadata will enable others to understand the methods used to generate the data, analyze the data, and use the data for future research purposes.

3.3 Data preservation strategy and standards

Introduction

The purpose of this document is to provide a data preservation strategy and standards documentation for this phase 1 study trial. The group will be using REDcap for data capturing and storing. The REDCap database includes a complete suite of features which are compliant with GCP, EU and UK regulations and Sponsor security policies, including a full audit trail, user-based privileges, and integration with the institutional LDAP server. The MySQL and PostgreSQL database and the webserver will both be housed on secure servers maintained by the University of Oxford IT personnel. The servers are in a physically secure location in Europe. Backups will be stored in accordance with the IT department schedule of daily, weekly, and monthly retained for one month, three months, and six months, respectively. The IT servers provide a stable, secure, well-maintained, and high-capacity data storage environment. RedCap™ is a widely-used, powerful, reliable, well-supported system. Access to the study's database will be restricted to members of the study team by username and password.

The data preservation strategy and standards will ensure the integrity, accessibility, and sustainability of the research data.

Data Preservation Strategy

The data preservation strategy for this study includes the following main steps:

1. Backups: Backups of the REDcap database will be made on a periodic basis (e.g. daily or weekly) to ensure data loss prevention in case of any unforeseen circumstances.
2. Data Retention Period: The research team will retain the data for a minimum period of 10 years after the completion of the study.
3. Data Sharing: Data-sharing options will be assessed annually to ensure that the research group has provided information-sharing opportunities.
4. Archiving: Paper copies will be kept of the informed consent and lab analyses. Other specialized data will be carefully electronically archived and kept accessible.
5. Storage: The University of Oxford will initially hold information provided during the pre-screening process but this will be downloaded and stored in Sheffield and will be deleted from the sponsor’s server at the end of the recruitment with the participants’ consent. Moreover, the University of Oxford will hold participant e-diaries and the identifiable information associated with this until the study results are published with participants’ consent. The access will be limited to the site research staff, sponsor staff, external monitor (delegated by sponsor) and the University of Oxford data managers.

Preservation Standards

The following are the formal preservation standards that will be used:

1. Compliance with ethical requirements: The data preservation will comply with regulatory authorities such as the Health Insurance Portability and Accountability Act (HIPAA) and other ethical and privacy standards.
2. Data Quality: The quality of data will be maintained by performing ongoing data validations and cleaning, including authenticating data entry by a second person.
3. Consistency measures: Measures will be taken to ensure uniformity and consistency in data collection and procedures such as double-checking methodologies.
4. Storage Security: Data will be stored in secure environments to ensure that there is limited to no unauthorized access.
5. **Data protection: The trial will comply with the UK General Data Protection Regulation (UK GDPR) and Data Protection Act 2018. The study protocol, documentation, data and all other information generated will be held in the strictest confidence. No information concerning the study or the data will be released to any unauthorised third party, without the prior written approval of the Sponsor.**

Data not retained
There is no data that will not be retained, given the importance of all data, including negative data to the integrity of the study. However, personal information such as names and other identifying information will be removed to ensure that the research is secure and completely de-identified.

4. Data security and confidentiality of potentially disclosive personal information

4.1 Formal information/data security standards

Introduction:
This plan outlines the formal information/data security standards that will be implemented to ensure that the data is secure.
Information Security Policies:
All staff and researchers involved in the project must complete the appropriate Information Security Training, and adhere to the University's Information Security Policy and Data Protection Policy. More secure system policies may be defined where necessary, especially when patient data is involved. Departments may also have their information security policies.

The following measures will be implemented to ensure the security of the data:
Access Control: Access to the data will be restricted to authorized personnel only. Passwords will be used to control access to the database, and each user will have a unique username and password.
Data Encryption: All data will be encrypted during transmission and storage to ensure confidentiality and prevent unauthorized access.
Data Backup: Regular backups will be taken to ensure that data is not lost in the event of a hardware failure or other unforeseen circumstances.
Data Retention: Data will be retained for only as long as necessary, and any personal data will be deleted once it is no longer needed.
Physical Security: The server where the data is stored will be physically secured to prevent unauthorized access.

Audit Trails: All activities relating to data access and modification will be logged, and an audit trail will be maintained to track any changes made to the data.

Security Text for Low-Risk Data:
For low-risk data that does not include personal data relating to human participants, the following security text should be used:
"The data will not include personal data relating to human participants. The University's Information Security Policies will be abided by at all times."
Security Text for High-Risk Medical Data:
For high-risk medical data that is highly confidential and critical to the clinical treatment of patients, the following security text should be used:
"We recognize that this data is highly confidential and critical to the clinical treatment of patients. Therefore, a project-specific security policy has been developed in conjunction with the University's Information Security Team [link to project policy]."

4.2 Main risks to data security

The main risks to the data security plan for the Phase I trial can be summarized as follows:
Data breaches: Unauthorized access to the personal data of human participants or accidental loss of data could occur during data capture, storage, or processing.
Cyber-attacks: Hackers or malicious actors could attempt to compromise the security of the REDcap platform, where the data is captured and stored.
Insider threats: Authorized users with access to personal data could intentionally or unintentionally misuse or disclose the information.
Physical theft or loss: Physical theft or loss of devices or hard copies containing personal data could occur.

The level of risk for these threats could range from low to high, depending on the sensitivity of the data, the number of people with access to the data, and the effectiveness of the controls in place.

To manage these risks, the Medical Research Council guidance recommends implementing appropriate technical and organizational measures, such as:

Encryption of personal data in transit and at rest, to prevent unauthorized access in case of data breaches or cyber-attacks.

Access controls, such as passwords, two-factor authentication, and role-based access, to limit access to personal data to authorized personnel only.

Regular backups and disaster recovery plans to ensure that personal data can be restored in case of data loss or corruption.

Physical security measures, such as secure storage facilities and tracking of devices containing personal data, to prevent theft or loss.

User training and awareness programs to ensure that all personnel handling personal data are aware of their responsibilities and the risks involved.

Regular audits of data access and usage, to ensure compliance with consent and security conditions.

In addition, the REDcap platform has built-in features for data access and control, such as user access levels, audit trails, and automatic logging of user activity. These features can help to mitigate the risks of insider threats and unauthorized access to personal data.

Overall, the risks to the confidentiality and security of information related to human participants can be managed effectively through a combination of technical and organizational measures, user training, and regular audits.

5. Data sharing and access

5.1 Suitability for sharing

Suitability Sharing Plan for Data Sharing and Access for an Open-label, Single-centre Phase I Trial to Assess the Safety and Immunogenicity and Efficacy of the Candidate Malaria Vaccine RH5.1 Formulated in Adjuvant Matrix-M

Data Sharing Plan Overview:

The aim of this data-sharing plan is to ensure that data from the open-label, single-centre Phase I trial to assess the safety and immunogenicity and efficacy of the candidate malaria vaccine RH5.1 formulated in adjuvant Matrix-M is shared in a manner that is consistent with ethical, legal, and regulatory requirements while maximizing the potential for reuse and impact.

Data Sharing Policy:

Our data-sharing policy is consistent with the open science policy of the University of Sheffield and the requirements of our funders. We will strive to make all data collected in this study available to the research community as soon as possible. All data shared will be anonymized, and any confidential or sensitive data will be managed in accordance with ethical, legal, and regulatory requirements.

Data repository: The data generated from the trial will be stored and curated in REDcap which is a secure web application designed to support data capture for research studies. The REDcap database will serve as the primary source for storing data and will be managed by the research team.

Data Storage and Management:

All data will be captured and stored using REDCap which is a secure web application designed to support data capture for research studies. The REDcap database will serve as the primary source for storing data and will be managed by the Oxford Vaccine Group at The University of Oxford. The University of Oxford will initially hold information provided during the pre-screening process but this will be downloaded and stored in Sheffield and will be deleted from the sponsor’s server at the end of the recruitment with the participants’ consent.

Data Access:

We will provide open access to all non-sensitive data as soon as possible after the publication of our study findings. Data that is deemed confidential or sensitive will be made available through controlled access. Access to sensitive data will be provided only to researchers who have obtained ethical approval and will be governed by data-sharing agreements that include confidentiality and data protection clauses. We will also comply with any data-sharing requirements of our funders and regulatory bodies. The access will be limited to the site research staff, sponsor staff, external monitor (delegated by sponsor) and the University of Oxford data managers.

Data Licensing:

The need to store this information for longer, in relation to licensing of the vaccine will be reviewed every 5 years. Files will be confidentially destroyed when storage is no longer required. For effective vaccines that may be licensed, secure storage of research data may be required for at least 15 years after the end of the study, subject to adjustments in clinical trial regulations.

Confidentiality: The trial data will include confidential information that should be protected. Therefore, the research team will ensure that all data made available is properly anonymised to protect individual participants' identities. Confidentiality will be maintained as described in section 12.5 of the BIO-002 protocol v1.0 document.

Ethics: Participation of individuals in the trial will be voluntary, and they will be provided with detailed information about the study and the data-sharing arrangements. Participants signed consent forms will also clearly state the intended use of the data, and participants will be informed that their data may be shared.

Responsibility for Controlling Access:

The CI will be the data custodian with the responsibility for delegating the receiving, entering, cleaning, querying, analysing and storing of all data that accrues from the study. All non-identifiable (pseudo-anonymised) laboratory research data will either be generated, analysed and stored in Oxford OR generated and analysed in Sheffield/named collaborating sites and then transferred to

Oxford (where Oxford will maintain ownership/responsibility for these data). All pseudo-anonymised clinical data will be generated and analysed in Sheffield, with support and oversight from Oxford.

Conclusion:

This data-sharing plan outlines the policies and procedures that will be followed to ensure that data from the open-label, single-centre Phase I trial to assess the safety and immunogenicity and efficacy of the candidate malaria vaccine RH5.1 formulated in adjuvant Matrix-M is shared in a manner that is consistent with ethical, legal, and regulatory requirements while maximizing the potential for reuse and impact. By making data from this study openly accessible, we hope to contribute to the advancement of science and the development of effective malaria vaccines.

5.2 Discovery by potential users of the research data

Discovery:

Potential users of the research data for this study may find the data valuable for further analysis and research. The trial aimed to evaluate the safety and immunogenicity of the vaccine, as well as assess its potential efficacy in preventing malaria infection. The study recruited healthy adult volunteers who were at risk of malaria infection in the trial site area. The data collected from the trial would provide valuable insights into the safety and efficacy of the vaccine, which could be used to inform future malaria vaccine development and research.

Data Accessibility:

The study data will be captured and stored using REDCap, a secure web-based platform designed for online data capture, storage, and management. Potential users of the data outside of the study organization can access the data through the study website, which will have summary information (metadata) readily available. The data may also be available through the MRC gateway for population and patient research data or other databases and catalogues. The depository for the data will be widely accessible, subject to any legal, ethical, or privacy restrictions.

Data Sharing Policy:

The study's policy on data sharing will be published on the study website and other relevant platforms. The policy will detail how data can be accessed, who can access it, and any restrictions or limitations that may apply. The policy will also outline the study's approach to ensuring the confidentiality and privacy of study participants and their data. The study will adhere to established guidelines and regulations on data sharing, including obtaining appropriate permissions and ethical clearance from relevant authorities.

5.3 Governance of access

Governance of access data plan:

Data capturing and storage:

Data will be captured and stored using REDcap. This will ensure the data is securely stored and accessible to authorized personnel only. The data will be monitored for quality and completeness on a regular basis to ensure that data is accurate and up-to-date.

Access to data:

Access to data will be granted only to authorized personnel who are involved in the conduct of the trial, and those who have a legitimate interest in the data. Access will be granted on a need-to-know basis, and all users will be required to sign a data access agreement before being granted access.

Decision-making:

The Principal Investigator (PI) and the Chief Investigator (CI) will make the decision on whether to supply research data to a potential new user. The decision will be based on the user's legitimate interest in the data and their ability to maintain confidentiality and comply with data protection regulations.

Independent oversight:

The trial will be conducted in compliance with MRC policy. Independent oversight will be provided by Appledown Clinical Research Ltd which will review data access requests and monitor the use of data to ensure compliance with ethical standards.

Community database:

Research data will be deposited in an identified community database or repository, such as the MalariaGEN data repository. This will ensure that the data is widely accessible to the research community and can be used to advance research in the field of malaria vaccine development. Access to the data will be granted in accordance with the data access plan described above. Processed blood samples will be temporarily held in the University of Sheffield laboratories/Biorepository prior to transfer to Oxford. Immunology analyses for this study will take place at the University of Sheffield, the University of Oxford and at collaborating laboratories. Aliquots for analyses to be undertaken in Sheffield will remain there, all other aliquots (serum, plasma, whole blood and PBMC) will be shipped to Oxford and stored at the Department of Biochemistry laboratories, University of Oxford, UK. Samples for analysis of in vitro GIA (secondary immunological endpoint)may be shipped to the following collaborative laboratory;

• NIH/NIAID Laboratory of Malaria and Vector Research, Malaria Immunology Section, GIA Reference Centre, 12735 Twinbrook Parkway, Twinbrook III, Room 3W-13, Rockville, MD 20852, USA.

5.4 The study team's exclusive use of the data

The University of Oxford will initially hold information provided during the pre-screening process but this will be downloaded and stored in Sheffield and will be deleted from the sponsor's server at the end of the recruitment with the participants' consent. Moreover, the University of Oxford will hold participant e-diaries and the identifiable information associated with this until the study results are published with participants' consent. The access will be limited to the site research staff, sponsor staff, external monitor (delegated by sponsor) and the University of Oxford data managers.

The Sheffield Teaching Hospitals NIHR Clinical Research Facility will keep all other identifiable/source data from participants collected during the study initially for 5 years after the study has finished. They will be responsible for these data and with consent, they may use these data to invite participants to take part in future local research studies. These data will be limited to the minimum personal data necessary. Once the study has been completed, all documents would be archived in a secure facility. All essential documents will be retained for a minimum of 5 years after the study has finished. The Sponsor will securely store the pseudo-anonymised research data, and the Sheffield CRF will securely store any source data including any research documents with personal information, such as consent forms. The need to store this information for longer, in relation to licensing of the vaccine will be reviewed every 5 years. Files will be confidentially destroyed when storage is no longer required. For effective vaccines that may be licensed, secure storage of research data may be required for at least 15 years after the end of the study, subject to adjustments in clinical trial regulations. Financial information will be stored for the duration of the study before being archived with the other research documents. Financial information does not remain on the finance systems beyond the financial year following the end of the study as per Sheffield Teaching Hospitals' financial policy.

Where specific written consent is provided, samples will be retained as described in sections 8.2.2 and 8.3 for future use in ethically-approved studies. This will include the retention of copies of the participant consent until the destruction or depletion of the sample(s) in question to meet HTA traceability requirements during the storage period and confirm future use is in line with the consent provided. Consent forms held for this purpose will be stored securely and separately from the samples themselves and from other participant data by the Sheffield Clinical Research Facility. All samples stored will be labelled with the participant's study identification (ID) number, which cannot identify the study participant directly but is linkable to other research databases (e.g., questionnaires, clinical assessments, logbooks) generated by the main study. The subject identification log linking the study participant ID number to the name of the participant and the consent form itself will be maintained with access limited to authorised research team members only in Sheffield.

All non-identifiable (pseudo-anonymised) laboratory research data will either be generated, analysed and stored in Oxford OR generated and analysed in Sheffield/named collaborating sites and then transferred to Oxford (where Oxford will maintain ownership/responsibility for these data). All pseudo-anonymised clinical data will be generated and analysed in Sheffield, with support and oversight from Oxford.

5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions

To ensure the confidentiality and security of personal, sensitive, and confidential data used in the Phase I trial, the following plan will be implemented:

Data Handling: All personal, sensitive, and confidential data collected during the trial will be handled with utmost care and attention to ensure that the participants' privacy is protected. Data will be collected using the REDcap system, which is secure and compliant with data protection regulations.

Data Pseudonymisation: All personal and identifying data will be pseudonymised using customised record ID's during the trial to protect the participants' identities. The data will only be accessible by authorised personnel with a legitimate need to access it.

Data Storage: The personal and identifying data will be securely stored on password-protected computers and servers with restricted access. The data will be encrypted to prevent unauthorised access.

Data Sharing: Participants will be informed of the possibility of data sharing during the consent process, and their consent will be sought for data sharing. The data will be anonymised or aggregated before sharing to minimise the risk of identification. Copyright permissions will be obtained before sharing any data.

Data Retention: Identifiable data will be stored for a maximum of 10 years after the end of the trial, as per the ethics approval. After that period, the data will be securely deleted.

Ethical and Legal Considerations: All data handling procedures will be in compliance with ethical and legal requirements, including GDPR. The University's ownership and licensing policies for research data will be adhered to, and any collaborators or funders will be consulted regarding data-sharing rights.

Data Licence: Any data shared will be licensed to ensure that users are informed about how they can use the data. The applicable license will be determined based on the type of data shared and the intended use.

5.6 Regulation of responsibilities of users

The main responsibilities of external users may include:

1. Protecting the confidentiality and privacy of research participants by ensuring that the data is stored securely and access is restricted only to authorized personnel.
2. Using the data only for the purposes specified in the data sharing agreement and not for any other purposes without the express permission of the data provider.
3. Acknowledging the source of the data in any publications or presentations and complying with any restrictions on publication or dissemination of the data.
4. Ensuring that the data is used in accordance with all applicable laws and regulations, including data protection and privacy laws.

5. Taking responsibility for any misuse of the data or breach of confidentiality, and reporting any such incidents to the data provider as soon as possible.

6. Responsibilities

6. Responsibilities

Chief Investigator	Dr Angela M Minassian Oxford Vaccine Group Centre for Clinical Vaccinology and Tropical Medicine, University of Oxford, Churchill Hospital, Old Road, Headington, Oxford, OX3 7LE Email: angela.minassian@bioch.ox.ac.uk
Principal Investigator	Dr Ruth Payne Department of Infection, Immunity and Cardiovascular Disease Faculty of Medicine and Dentistry, University of Sheffield Beech Hill Road, Sheffield, S10 2RX
CTU	Oxford Vaccine Group, University of Oxford, Centre for Clinical Vaccinology and Tropical Medicine (CCVTM), Churchill Hospital, Oxford, OX3 7LE, United Kingdom
Sponsoring Institution	University of Oxford Research Governance, Ethics and Assurance Team, Research Services University of Oxford Joint Research Office Boundary Brook House Churchill Drive Headington OX3 7GB Tel: 01865 616480 Email: ctrg@admin.ox.ac.uk
Independent Monitor	Appledown Clinical Research Ltd. Orchard End, Greenlands Lane, Prestwood, Bucks, HP16 9QX Email: alexandrabrush@gmail.com
Local Safety Monitor	Dr Paul Collini Senior Lecturer and Honorary Consultant in Infectious Diseases Department of Infection, Immunity and Cardiovascular Disease Faculty of Medicine and Dentistry, University of Sheffield Beech Hill Road, S10 2RX

7. Relevant policies

7. Relevant institutional, departmental or study policies on data sharing and data security

Policy	URL or reference
Data Management Policy and Procedures	TuoS Research data management (RDM)
Data Security Policy	TUoS Data Security Policy
Data Sharing Policy	*e.g. a study policy of sharing research data*
Institutional Information Policy	
Institutional Ethics Policy:	Ethics Policy
Other	

8. Author and contact details

8. Author of this Data Management Plan (Name) and, if different to that of the Principal Investigator, their telephone & email contact details

Author of this Data Management Plan: Callistus Iwuagwu
Telephone: +44 1234567890
Email: c.iwuagwu@sheffield.ac.uk