**IDDO: Data Management Plan**

**Version 1.0**

**Clinical, Molecular and Pharmacological Trials Data**

**Infectious Diseases Data Observatory (IDDO)**



Suggested citation: **IDDO. (2017). Data Management Plan. Retrieved from:** <https://www.iddo.org/tools-resources/data-management-plan>

Version History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version number** | **Author** | **Revision(s) & reason for amendment** | **Release date** |
| V1.0 | Michael Otieno and Hazel Ashurst | Initial draft | 01/06/2017 |

For more information, contact:

[clinical@iddo.org](mailto:clinical@iddo.org)

**Infectious Diseases Data Observatory**

[**www.iddo.org**](http://www.iddo.org)

# Table of Contents

[Table of Contents 3](#_Toc484423012)

[List of Abbreviations 4](#_Toc484423013)

[1 Scope 5](#_Toc484423014)

[2 Introduction 5](#_Toc484423015)

[3 Data collation process 6](#_Toc484423016)

[4 Data submission process 7](#_Toc484423017)

[4.1 Online submission system 7](#_Toc484423018)

[4.2 Required variables 8](#_Toc484423019)

[4.3 Additional variables 8](#_Toc484423020)

[4.4 Study metadata 9](#_Toc484423021)

[5 Data dictionary 9](#_Toc484423022)

[5.1 IDDO database structure 10](#_Toc484423023)

[6 Data mapping and standardisation 10](#_Toc484423024)

[6.1 Key assumptions during mapping process 11](#_Toc484423025)

[6.1.1 Date of inclusion/ Date of event 11](#_Toc484423026)

[6.1.2 Days and hours of event 11](#_Toc484423027)

[6.1.3 Age 11](#_Toc484423028)

[6.1.4 Temperature 12](#_Toc484423029)

[6.1.5 Fever and Fever history 12](#_Toc484423030)

[6.1.6 Drug administration 12](#_Toc484423031)

[6.1.7 Biochemistry 12](#_Toc484423032)

[6.1.8 Symptoms 13](#_Toc484423033)

[6.1.9 Last day of follow up 13](#_Toc484423034)

[6.2 Other disease specific considerations 13](#_Toc484423035)

[7 Data cleaning 13](#_Toc484423036)

[7.1 Unexpected results 13](#_Toc484423037)

[7.1.1 Single variable 13](#_Toc484423038)

[7.1.2 Combined variables 14](#_Toc484423039)

[7.2 Platform specific checks 14](#_Toc484423040)

[7.3 Consistency checks between patient level and study level data 14](#_Toc484423041)

[8 Data analysis 14](#_Toc484423042)

[9 Clinical data management report and patient book 15](#_Toc484423043)

[10 DMP Versioning 15](#_Toc484423044)

[11 Conclusion 15](#_Toc484423045)

[12 References 16](#_Toc484423046)

[Annex A: IDDO Repository Generic Variables Data Dictionary 17](#_Toc484423047)

# List of Abbreviations

DHAC Data Handling and Analysis Convention

DMP Data Management Plan

CDMR Clinical Data Management Report

EVD Ebola Virus Disease

IDDO Infectious Diseases Data Observatory

IPD Individual Patient Data

NGO Non-Governmental Organisations

NCP National Control Programmes

SAP Statistical Analysis Plan

SOP Standard Operating Procedure

VL Visceral Leishmaniasis

WHO World Health Organization

# Scope

The aim of Infectious Diseases Data Observatory (IDDO) data sharing platforms is to provide a data repository for researchers, investigators and other stakeholders involved in clinical research and policy formulation for infectious diseases such as malaria, Ebola and visceral leishmaniasis (VL) currently, and possibly others in the future. The platform offers an opportunity for researchers to store and retrieve individual patient data (IPD) in a standardised format. Access to comparable data at the individual patient level would in turn facilitate pooled data statistical analyses and facilitate optimisation of interventions for the treatment and control of these diseases.

The purpose of this Data Management Plan (DMP) is to present a clear and transparent methodology by which IDDO data platforms will transform and process the uploaded clinical data into standardised format in a manner that ensures reproducibility. It also provides a framework for discussing and developing such methodology.

The IDDO data management team currently manages three data sharing platforms namely:

1. Malaria data platform – a platform for pooled malaria disease data,
2. Ebola data platform – a platform for Ebola Virus Disease (EVD) data, and
3. Visceral Leishmaniasis (VL) data platform - a platform for pooling VL treatment data.

This DMP therefore outlines the baseline procedures applicable across all the existing and future platforms at IDDO given that the underlying IDDO infrastructure is shared across these platforms. Additional platform specific data handling and analysis conventions (DHAC) will be made available to address platform specific data management and analysis requirements.

# Introduction

Clinical trials in infectious diseases are conducted across the globe by different research groups in order to come up with safe and efficacious treatments that are affordable and easy to use by patients. The results of these trials provide the basic data needed by the World Health Organization (WHO) to publish and update international guidelines for use of these treatments. Most research groups, non-governmental organisations (NGOs) and national control programmes (NCPs) adhere to these guidelines; however study design modifications (ranging from minor to major) are often adopted by investigators to accommodate specific research questions, characteristics of the study site or simple logistical constraints.

To confound matters further, the analysis of clinical efficacy data may vary depending upon whether the data come from a stand-alone drug study or a comparative drug trial. Such variations in the design, methodology, analysis and presentation of clinical drug trials can lead to substantial bias in the derived estimates of drug efficacy as illustrated by Verret *et al* (2009). The interpretation of aggregated data from published results of clinical drug studies is therefore fraught with confounding factors and these aggregated data cannot be used reliably for assessing geographical or temporal trends.

Over the last two decades there have been several revisions of guidelines and statistical approaches for the analysis of clinical trials. Through different disease-specific data sharing platforms provided by IDDO, collation of raw data from clinical trials will help to ensure that any future methodological changes can be accommodated and presented with minimal effort.

In order to achieve the above, IDDO is working towards these goals as follows:

1. Facilitating the inclusion in a **data repository** of results from clinical drug efficacy studies carried out by research groups, NGOs and NCPs;
2. Creating **standardised processes** to facilitate the collation of diverse datasets from studies taking place around the world;
3. Optimising **analytical tools** to accommodate this diversity and increase comparability of results between heterogeneous studies;
4. Developing standardised methods and systems to improve the **quality** of source data in future studies.

# Data collation process

IDDO aims to facilitate two processes: to give researchers the tools to collect, clean and analyse their own data, and to transform clinical data from a diverse range of studies into a standardised format that can be derived from almost any database structure, so that data from different studies can be pooled and analysed collectively in a standardised manner.

To achieve these goals, a series of steps are followed:

1. **Upload clinical data**

Upon accepting the terms of submission on IDDO websites, study investigators will proceed to upload individual anonymised patient data from clinical studies via IDDO’s web portal. However, the investigators can also opt to send their data directly to IDDO data managers who can proceed to upload the data on their behalf. Once the data has been uploaded a personal data review takes place, which ensures that the data submitted has been fully anonymised. Once this has been completed the curation begins with a curator entering study meta-data into the IDDO custom Chassis II program

1. **Data transformation and standardisation**

Once data is uploaded through the IDDO web portal, it is transformed into a standardised format as described in **section 6** under **Data mapping and standardisation**.

1. **Consistency checks**

Submitted data are checked forinconsistencies, unexpected values and missing values as described in **section 7** under **Data Cleaning**.

1. **Data revision**

Any inconsistencies discovered during data cleaning will be communicated back to the data submitter for correction. In case there is no revision from the data submitter, the data items will be marked as missing hence neutralising their impact on the analysis.

1. **Data analysis of individual study data**

A uniform disease specific analytical methodology and reporting is applied during data analysis of individual study data to provide consistent data summaries and estimates of efficacy. These can be provided as an automated report to the data contributor.

1. **Data presentation on IDDO Explorer**

If the data contributor agrees, the study data summaries may be presented on an IDDO Explorer, an interactive, online tool which will allow users to perform custom queries of more than a hundred studies and visualise the results using dynamic mapping.

IDDO notes that other researchers may take different approaches to data management, particularly with regard to defining and managing protocol deviations. It is important to stress that it is inevitable that the IDDO-derived efficacy estimates may vary to some degree from analyses performed by the data submitter.

These differences do not reflect a value judgment as to which analytical approach is correct. The decisions are made only to apply standardised methodologies and minimise bias on geospatial and temporal trends derived from the many studies for the respective platforms compiled in the Data Repository.

# Data submission process

The IDDO data repository receives data from different clinical studies that have been obtained in accordance with any laws and ethical approvals applicable in the country of origin. It is the responsibility of the data contributors to ensure that those requirements have been met prior to data sharing.

Data can be submitted by the investigators directly via the online submission module or through online transfer to the IDDO data management team for inclusion into the repository.

## Online submission system

IDDO provides disease specific online portals for submission of datasets into the shared data repository. Data contributors must accept the Terms of Submission during the submission process.

The data submission steps are:

1. **Log onto the disease specific online submission portal**: IDDO provides online submission portals for malaria, EVD and VL data platforms. Each portal is accessible via the IDDO website. Investigators are expected to create an account before gaining access to any of these portals.
2. **Register a study**: contributors enter their study title and each study is assigned a unique identifier. Contributors are required to tick one or more boxes to indicate the type of data submitted (e.g.: clinical, safety, ECG, PD/PK).
3. **Permissions**: each study may have any number of administrators, assigned by the original data contributor. An administrator can access the study, upload files and edit supplied information.
4. **Files**: contributors are asked to submit data files and supporting documentation including a data dictionary, protocols and publications. IDDO data sharing platforms accept data from a variety of file formats such as STATA, CSV, MS Access, MS Excel, R, SPSS, SAP and many others.
5. **Publications:** contributors can provide the PubMed ID, citation or DOI for publications relating to the submitted data. For unpublished studies, the trial protocol, study report, and trial registration number are desired.
6. **Acknowledgements**: the names of acknowledged individuals and institutions will appear in the study details displayed on the IDDO Explorer.
7. **Study info**: data contributors are asked to provide information on the study site and study design. They may enter this information themselves or provide protocols and publications which will be used by IDDO data managers to extract the relevant data. This includes information on: randomisation, treatment blinding, method of treatment concealment and treatment allocation, sequence generation, supervision status, laboratory techniques used among others.

During the submission process the investigators are expected to provide in the submitted datasets, required variables, including, but not limited to, the data items highlighted below.

## Required variables

The dataset and/or accompanying documents (e.g. data dictionary, protocols, and publications) must contain the following information:

1. Unique patient identifier
2. Date of inclusion and days or dates of follow up visits
3. Treatment received
4. Dosing received or dosing protocol
5. Patient age or date of birth
6. Patient weight
7. Parasitaemia and species on day 0 and during follow up (for Malaria data platform)

Presence of parasites on day 0 and during follow-up (for VL)

Note: missing values must be distinguishable from zero parasites in blood smear results

1. Temperature on day 0 and during follow up
2. Study location

## Additional variables

Additional variables where available will be necessary for submission in order to improve analysis and study report generation. These include data variables on:

1. Patient baseline data
2. Patient medical history data
3. Clinical symptoms including assessments and examinations data
4. Laboratory data
5. Concomitant medication data
6. Adverse Events data
7. PCR results if available

## Study metadata

As part of study data submission, the investigators will be expected to provide additional study metadata which will provide more information about the study. This information includes:

1. Basic information on study site and study design:
   1. Location of study site
   2. Study design - whether randomised, longitudinal
   3. Whether pregnancy data is collected or not
   4. Treatment arms
   5. Methodology used
2. Other study reports and documentation such as:
   1. Inclusion and exclusion criteria
   2. Protocol
   3. Data dictionary
   4. Publications (including reports)

# Data dictionary

The IDDO data dictionary contains clinical and pharmacology variables required for generation of efficacy and safety outcomes for all disease platforms. The data dictionary contains both generic variables and disease specific variables for all of the IDDO platforms.

The dictionary is divided into different sections as shown below:

1. Meta-data
2. Subject
3. Treatment
4. Haematology
5. Biochemistry
6. Clinical
7. Adverse Events
8. Concomitant Medication
9. Outcome
10. Molecular
11. Concentration

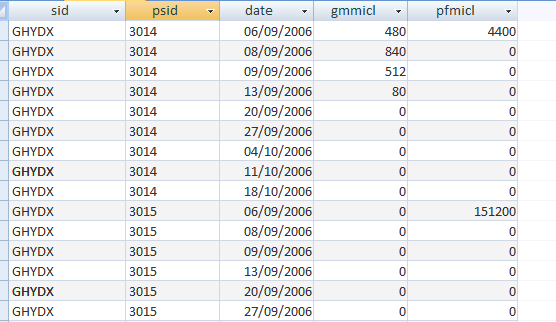
Each section contains variable names, descriptions, units, lower limits and higher limits and data types. A copy of the IDDO data dictionary for generic variables is attached in Annex A. Disease specific variables are annexed in DHAC for respective disease data platforms.

## IDDO database structure

The IDDO data repository contains data in a number of tables corresponding to data dictionary sections shown above. Each of these tables is linked to each other by a unique study identifier, patient identifier and date of inclusion.

Each line in a table is one event with study identifier, patient identifier, date of inclusion and date and time of event. The subject table shows one line per patient for non-longitudinal studies and may have multiple lines for longitudinal studies, each corresponding to one episode of infection. The treatment table shows one line per dose, per patient. If dose was administered on several occasions and this information is available there will be more than one line per patient, each with corresponding date and time of event. The example below, shows a fragment of a parasitology table which has a line for each blood smear result of a patient.

Study identifier patient identifier date of event gametocytaemia parasitaemia (falciparum)



# Data mapping and standardisation

All data mapping and standardisation procedures are described in a series of data management Standard Operating Procedures (SOPs).

Submitted data are extracted from source datasets and transformed into a standardised format, allowing single study analysis, generation of a study report and, with permission, visualisation of summary study information on IDDO Explorer. The transformed data is stored in a disease specific data repository for future pooling with other related datasets and meta-analyses.

Source data are submitted in various file formats such as flat files, with one line per patient, or as multiple relational databases. The extraction process transforms all source datasets into a standard format. Variables from the source dataset, equivalent to variables in the IDDO data repository, are extracted and mapped onto corresponding variables in the repository and finally saved into one of the underlying repository tables.

There exist generic tables in the repository in which variables that are common across all disease platforms are stored. Non-generic disease specific data goes into separate tables specific to those disease platforms within the repository.

All non-generic string variables are systematically recoded into the controlled terminology depending on disease specific DHAC. An audit trail records and saves the complete data extraction and transformation process. Once transformed, the contributor or their designees may download the derived dataset in the standard IDDO format that is amenable to other datasets submitted to IDDO, for offline analysis.

## Key assumptions during mapping process

### Date of inclusion/ Date of event

If the date of inclusion or visit dates per patient are not specified, i.e. there are no dates in the dataset, a theoretical date of inclusion - January 1 [year the study started] is used for each patient. This date is used to tag the patient to a specific year and to derive the subsequent dates of follow up based on the relative event times provided (day x, hour x).

### Days and hours of event

These are automatically calculated as number of days from a day of inclusion. If time is given, takes into account date and time. If time is not given hours are calculated as numbers of days multiplied by 24.

### Age

In general, study recorded age is used (converted to age in years with decimal points), but if date of birth is available, age is calculated as (date of enrolment - date of birth)/365.25 and recorded with decimal place.

### Temperature

Converted to Celsius, but not converted between different measurements methods (oral, axillary, tympanic, rectal). Method of measurement is recorded as meta-data.

### Fever and Fever history

During data curation, fever is mapped directly from the source data into a **Fever** variable which records its presence or absence as yes or no respectively. In cases where the fever variable is missing from the source data but temperature data is available, **Fever** will not be derived from temperature. Fever is mapped as a multi-day variable for all the available study visits.

In handling of fever history, IDDO has defined a variable “*Fever in the last 24 hours?”* and source variables of fever history must be within this time frame to be deemed equivalent.

### Drug administration

Each drug dose is recorded in the treatment table with date and time of administration, number of tablets given, tablet size, and mg per kg dose given. Additionally, total mg per kg dose given is recorded which can either be calculated from individual doses or could be supplied by the source dataset. Variables for drug vomiting within one hour of administration (yes/no) and re-dosing (yes/no) for each event of drug administration are included in the IDDO clinical dictionary.

If the dose was re-administered because of vomiting, the extra dose is not included in the total dose calculation. If none of these variables are given in the source data, the protocol treatment schedule is used to calculate the target dose for each patient and this is stored in the variable totalmgperkg in the repository.

In cases where other dose related variables are missing in the source data, such variables are calculated based on other available study information, if possible. For example if dose is given per each administration then the total dose is calculated automatically; if number of tablets is given for each dose, the dose is calculated based on the size of tablets given.

### Biochemistry

The following units are used: Alanine transaminase (units/L) ; Albumin (g/dL) ; Alkaline phosphatise (units/L); Aspartate aminotransferase (units/L); Total bilirubin (mg/dL); Calcium (mg/dL); Creatinine (mg/dL); Glucose (mg/dL); Potassium (mmol/L); Sodium (mmol/L); Total protein (g/dL); Hemoglobin (g/dL); Hematocrit (%); Neutrophils (%); Platelets (billion /L ); Lymphocytes (%); White blood cells (per microlitre); Monocytes (%)

### Symptoms

Presence of symptoms is recorded as binary variables so that if grade is provided in the study data set it will be converted into binary variable yes /no.

In cases where splenomegaly and/or hepatomegaly data is collected, if the same are noted in centimetres in the source dataset, they will be converted to a binary (yes/no) variable. For example, if the source dataset notes splenomegaly >0 cm, this is converted to the variable splenomegaly=yes. The rationale for this approach is that the size of organomegaly is often poorly quantified by clinical examination. It is the presence or absence of organomegaly, rather than its degree, that appears to have greater clinical relevance.

### Last day of follow up

Last day of follow-up is generated automatically as the last day when the parasitaemia was measured.

## Other disease specific considerations

Additional data handling conventions are defined for each IDDO disease platform and as such reference should be made to the specific disease (DHAC) document for more details.

# Data cleaning

Submitted data are checked for **inconsistencies**, **unexpected values** and **missing values**.

The full list of data checks will be given in the Data Handling and Analysis Convention (DHAC) documents. If identified, these values are communicated with data contributor, along with the patient identifier, for possible correction from source documents. Resubmitted, corrected, and/or missing values will be used to update the Data Repository. If corrections cannot be made, the unexpected results are transformed to missing values.

In order to transform all datasets into a consistent format for potential combined analysis, variables are constrained within limits or ranges set by the analysis program as described below.

## Unexpected results

### Single variable

During data cleaning, a number of single variable range data checks are conducted on all studies. Data are checked against extreme values which are deemed incompatible with the range of values observed in clinical trials.

### Combined variables

In addition to these single variable data checks, the IDDO edit checklist also lists cross form or combined variable checks on the submitted studies data. For example, in the case of a patient age and weight discordance, such inconsistencies will be raised with the data contributors for resolution, and when no correction is available from the data contributors, both weight and age will be converted to a missing value.

## 

## Platform specific checks

In addition, further validation checks are defined for each disease platform according to each disease’s variables specifications.

## Consistency checks between patient level and study level data

A number of checks are performed to ensure that the study level and patient level data are consistent. These include:

1. Ensuring that the patient level treatment regimen is described in the study level data
2. Ensuring that the site at which a patient is treated is described in the study level data with associated latitude and longitude
3. Ensuring that the dates of inclusion for patients fall with the start and stop dates of the trial as described in the study level data
4. PCR done and PCR results attached
5. Duration of follow-up and data available from follow-up visits

# Data analysis

Analysis datasets are derived from the deposited data (this includes, for example, calculation of mg/kg, patient dose, generation of outcome, calculation of parasite clearance half-life) using standard methodology for all studies. The methodology is documented in the DHAC.

Individual patient meta-analyses based on the pooled studies from the IDDO repository are organised within the Study Group concept ( <http://www.wwarn.org/working-together/study-groups> ) where experts are drawn from amongst the data contributors in order to carry out specific analyses on questions of interest. The study groups operate according to the defined process outlined in the IDDO Study Group protocol and all analyses are carried out according to the *a priori* developed study group specific Statistical Analysis Plan, for example: <http://www.wwarn.org/tools-resources/gametocyte-carriage-study-group-statistical-analysis-plan>

# Clinical data management report and patient book

As IDDO receives clinical data collected from diverse clinical data management systems, data transformation occurs while importing data from source formats into standardised format in the IDDO data repository. When the data curation and import is complete, IDDO data managers will, if requested, generate a clinical data management report (CDMR) for the submitted data from the data repository and send back the CDMR to the data contributor for their verification and records.

The CDMR contains the following information:

1. **Basic description of the study** including; the total number of participants and by therapeutic group.
2. **Information on which source variables** were extracted.
3. **Systematic audits**
4. **Data consistency**: the numbers of unexpected results and deviations (Section 7.1) are tabulated with an annexed, detailed list of cases (patient number, date, day, and patient specific data).
5. **Data description**: a table presenting the percentage of participants with all unexpected results.
6. **Trial profile**: displaying the total number of included patients and study deviations, by study arm. The remaining patients are numerated by study site and arm: When a patient has two or more deviations, only the first chronological deviation is taken into account.
7. **Baseline characteristics**: from day 0 values.
8. **Outcome**
9. In addition to the CDMR, IDDO can provide data contributors with a detailed Patient Book which contains history of each participant in a one page summary with the resulting deviations or efficacy endpoints. The Patient Book has as many pages as there are patients in the study.

# **DMP versioning**

The final version of this DMP is 1.0. If amendments are required to the DMP, the first version of the revised DMP will be assigned a version number of “1.1” together with the date of issue.

# Conclusion

This DMP is an evolving document and aims to be in line with current WHO guidelines on managing data on treatment and drug efficacy studies for the different platforms currently implemented by IDDO. Managing data from diverse sources for diverse disease platforms can be a very complex endeavour and as a result, IDDO has further developed platform specific data handling and analysis conventions to address platform specific data management needs.

Despite complexities in managing different data sharing platforms, IDDO commits to availing high quality data for pooled analysis. However, it is foreseen that some cases will arise where no international consensus is available in handling some aspects of data in the uploaded datasets. IDDO will seek advice from a broad group of experts and choose an approach based on their inputs.

There will always remain parts of the methodology which are contentious. IDDO will continue to encourage feedback on these issues and will endeavour to incorporate suggestions into future versions or bring major issues into a wider forum for open discussion. Comments should be directed to [clinical@iddo.org](mailto:clinical@iddo.org).

# References

1. Verret WJ, Dorsey G, Nosten F, Price RNThe effect of varying analytical methods on estimates of anti-malarial clinical efficacy. *Malaria Journal* 2009; 8:77.

# Annex A: IDDO Repository Generic Variables Data Dictionary

| ID | Category Name | Column Description | Column Header Name | Column Name | Data Type |
| --- | --- | --- | --- | --- | --- |
|  | Subject | Patient identifier | pid | SubjectID | String |
|  | Subject | Age of the subject | ageyears | Age | PositiveRealNumber |
|  | Subject | Gender of the patient | gender | Gender | Enumerated |
|  | Subject | Height of the patient | height | Height | PositiveRealNumber |
|  | Subject | Weight of the patient | weight | Weight | PositiveRealNumber |
|  | Subject | Study site | site | StudySite | String |
|  | Subject | Did the patient consent to participate in the clinical trial? | cons | Consent | Boolean |
|  | Subject | Date of Consent | dcons | Consent Date | Date |
|  | Subject | In case of relapse: treatment regimen of last episode known? | last\_treat\_reg | last\_treat\_reg | Enumerated |
|  | Subject | Mid upper arm circumference of the patient | muac | MUAC | PositiveRealNumber |
|  | Subject | Last day of follow up | lastdayfup | LastDayFUP | PositiveRealNumber |
|  | Subject | Date of inclusion of the patient in the trial | dateinc | DateInclusion | Date |
|  | Subject | Time of inclusion of the patient in the study | timeinc | TimeInclusion | Time |
|  | Subject | Body mass index of the patient | bmi | BMI | PositiveRealNumber |
|  | Subject | Information on the pregnancy status of the patient | pregnancy | Pregnancy | Boolean |
|  | Subject | Gestation in weeks | ega | EGA | PositiveRealNumber |
|  | Subject | Any reported symptoms of malaria | sympmal | Sympmal | Boolean |
|  | Subject | Healthy volunteer status | healthy | Healthy | Boolean |
|  | Subject | Boolean | smoke | Smoke | Enumerated |
|  | Subject | Boolean | relativetime | RelativeTimes | Enumerated |
|  | Clinical | Heart rate | heart\_rate | hr | PositiveRealNumber |
|  | Clinical | History of fever in the past 24 hours | feverhist | FeverHistory | Boolean |
|  | Clinical | Location of thermometer to measure body temperature | thermo | Thermometer | Enumerated |
|  | Clinical | Body temperature | temp | Temperature | PositiveRealNumber |
|  | Clinical | Fever status | fever | Fever | Enumerated |
|  | Clinical | Information about the patient's diarrhoea? | diarrhea | Diarrhoea | Boolean |
|  | Clinical | Presence of splenomegaly | spleen | Splenomegaly | Boolean |
|  | Clinical | Presence of hepatomegaly | liver | Hepatomegaly | Boolean |
|  | Clinical | Occurrence of vomiting | vomit | Vomiting | Enumerated |
|  | Clinical | Presence of malaria symptoms | sympmal | SymptomaticMalaria | Enumerated |
|  | Clinical | The pulse rate of the patient | pulse\_rate | PulseRate | PositiveRealNumber |
|  | Clinical | The systolic blood pressure of the patient | systolic\_bp | SystolicBP | PositiveRealNumber |
|  | Clinical | The systolic blood pressure of the patient | diastolic\_bp | DiastolicBP | PositiveRealNumber |
|  | Clinical | The respiration rate of the patient | respiration\_rate | RespirationRate | PositiveRealNumber |
|  | Haematology | Basophils | bas | basophils | PositiveRealNumber |
|  | Haematology | Eosinophils | eos | eosinophils | PositiveRealNumber |
|  | Haematology | PMN | PMN | polymorphs | PositiveRealNumber |
|  | Haematology | Red Cell Count | rbc | rbc | PositiveRealNumber |
|  | Haematology | Haemoglobin measure | hb | Haemoglobin | PositiveRealNumber |
|  | Haematology | Haematocrit (Percentage of red blood cells in blood) | ht | Haematocrit | Percentage |
|  | Haematology | Platelet count | pt | Platelets | PositiveRealNumber |
|  | Haematology | WBC Count (All the white cell types are given as an absolute number per litre) | wbc | WBC | PositiveRealNumber |
|  | Haematology | Lymphocytes % | lymph | Lymphocytes | Percentage |
|  | Haematology | Neutrophils % | neu | Neutrophils | PositiveRealNumber |
|  | Haematology | Monocytes % | mono | Monocytes | Percentage |
|  | Biochemistry | BUN | bun | BUN | PositiveRealNumber |
|  | Biochemistry | Magnesium | magnesium | Magnesium | PositiveRealNumber |
|  | Biochemistry | Patient's glucose levels | glucose | Glucose | PositiveRealNumber |
|  | Biochemistry | Patient's sodium levels | sodium | Sodium | PositiveRealNumber |
|  | Biochemistry | Patient's potassium levels | potassium | Potassium | PositiveRealNumber |
|  | Biochemistry | Patient's calcium levels | calcium | Calcium | PositiveRealNumber |
|  | Biochemistry | Patient's creatinine levels | creatinine | Creatinine | PositiveRealNumber |
|  | Biochemistry | Patient's bilirubin levels | bilirubin | Total bilirubin | PositiveRealNumber |
|  | Biochemistry | Patient's alanine transaminase levels | alatrans | Alanine transaminase | PositiveRealNumber |
|  | Biochemistry | Patient's aspartate aminotransferase levels | aspamtrans | Aspartate aminotransferase | PositiveRealNumber |
|  | Biochemistry | Patient's alkaline phosphatase levels | alkphos | Alkaline phosphatase | PositiveRealNumber |
|  | Biochemistry | Patient's albumin levels | albumin | Albumin | PositiveRealNumber |
|  | Biochemistry | Patient's total protein levels | totprot | Total protein | PositiveRealNumber |
|  | Treatment | Treatment being investigated | treat | TreatmentArm | String |
|  | Treatment | Name of drug X | trt6 | Drug name | Enumerated |
|  | Treatment | Name of drug X | trt4 | Drug name | Enumerated |
|  | Treatment | Name of drug X | trt5 | Drug name | Enumerated |
|  | Treatment | Name of drug X | trt2 | Drug name | Enumerated |
|  | Treatment | Name of drug X | trt3 | Drug name | Enumerated |
|  | Treatment | Name of drug X | trt1 | Drug name | Enumerated |
|  | Treatment | ATC code X | atc5 | ATC Code | Enumerated |
|  | Treatment | ATC code X | atc4 | ATC Code | Enumerated |
|  | Treatment | ATC code X | atc6 | ATC Code | Enumerated |
|  | Treatment | ATC code X | atc1 | ATC Code | Enumerated |
|  | Treatment | ATC code X | atc3 | ATC Code | Enumerated |
|  | Treatment | ATC code X | atc2 | ATC Code | Enumerated |
|  | Treatment | Dose of drug per administration | dos6 | Dose value | PositiveRealNumber |
|  | Treatment | Dose of drug per administration | dos5 | Dose value | PositiveRealNumber |
|  | Treatment | Dose of drug per administration | dos4 | Dose value | PositiveRealNumber |
|  | Treatment | Dose of drug per administration | dos3 | Dose value | PositiveRealNumber |
|  | Treatment | Dose of drug per administration | dos2 | Dose value | PositiveRealNumber |
|  | Treatment | Dose of drug per administration | dos1 | Dose value | PositiveRealNumber |
|  | Treatment | Total daily dose of drug | dostot5 | Daily dose | PositiveRealNumber |
|  | Treatment | Total daily dose of drug | dostot6 | Daily dose | PositiveRealNumber |
|  | Treatment | Total daily dose of drug | dostot1 | Daily dose | PositiveRealNumber |
|  | Treatment | Total daily dose of drug | dostot2 | Daily dose | PositiveRealNumber |
|  | Treatment | Total daily dose of drug | dostot3 | Daily dose | PositiveRealNumber |
|  | Treatment | Total daily dose of drug | dostot4 | Daily dose | PositiveRealNumber |
|  | Treatment | Total daily dose of drug in mgs per Kg | mgperKg3 | Daily mgs per Kg | PositiveRealNumber |
|  | Treatment | Total daily dose of drug in mgs per Kg | mgperKg4 | Daily mgs per Kg | PositiveRealNumber |
|  | Treatment | Total daily dose of drug in mgs per Kg | mgperKg5 | Daily mgs per Kg | PositiveRealNumber |
|  | Treatment | Total daily dose of drug in mgs per Kg | mgperKg6 | Daily mgs per Kg | PositiveRealNumber |
|  | Treatment | Total daily dose of drug in mgs per Kg | mgperKg1 | Daily mgs per Kg | PositiveRealNumber |
|  | Treatment | Total daily dose of drug in mgs per Kg | mgperKg2 | Daily mgs per Kg | PositiveRealNumber |
|  | Treatment | Amount of dose of drug taken per Kg | dosing6 | Dosing | PositiveRealNumber |
|  | Treatment | Amount of dose of drug taken per Kg | dosing5 | Dosing | PositiveRealNumber |
|  | Treatment | Amount of dose of drug taken per Kg | dosing4 | Dosing | PositiveRealNumber |
|  | Treatment | Amount of dose of drug taken per Kg | dosing3 | Dosing | PositiveRealNumber |
|  | Treatment | Amount of dose of drug taken per Kg | dosing2 | Dosing | PositiveRealNumber |
|  | Treatment | Amount of dose of drug taken per Kg | dosing1 | Dosing | PositiveRealNumber |
|  | Treatment | Unit of dose of drug taken | dosunit6 | Dose unit | Enumerated |
|  | Treatment | Unit of dose of drug taken | dosunit4 | Dose unit | Enumerated |
|  | Treatment | Unit of dose of drug taken | dosunit5 | Dose unit | Enumerated |
|  | Treatment | Unit of dose of drug taken | dosunit2 | Dose unit | Enumerated |
|  | Treatment | Unit of dose of drug taken | dosunit3 | Dose unit | Enumerated |
|  | Treatment | Unit of dose of drug taken | dosunit1 | Dose unit | Enumerated |
|  | Treatment | Formulation of dose of drug | dosfrm6 | Dose formulation | Enumerated |
|  | Treatment | Formulation of dose of drug | dosfrm5 | Dose formulation | Enumerated |
|  | Treatment | Formulation of dose of drug | dosfrm2 | Dose formulation | Enumerated |
|  | Treatment | Formulation of dose of drug | dosfrm1 | Dose formulation | Enumerated |
|  | Treatment | Formulation of dose of drug | dosfrm4 | Dose formulation | Enumerated |
|  | Treatment | Formulation of dose of drug | dosfrm3 | Dose formulation | Enumerated |
|  | Treatment | Was the administration of the treatment supervised? | superv | Supervision | Enumerated |
|  | Treatment | Was fat given with the treatment? | fat | Fat | Enumerated |
|  | Treatment | Frequency of dose of drug | dosfrq6 | Dose frequency | Enumerated |
|  | Treatment | Frequency of dose of drug | dosfrq5 | Dose frequency | Enumerated |
|  | Treatment | Frequency of dose of drug | dosfrq2 | Dose frequency | Enumerated |
|  | Treatment | Frequency of dose of drug | dosfrq1 | Dose frequency | Enumerated |
|  | Treatment | Frequency of dose of drug | dosfrq4 | Dose frequency | Enumerated |
|  | Treatment | Frequency of dose of drug | dosfrq3 | Dose frequency | Enumerated |
|  | Treatment | Route of administration of drug | route4 | Dose route | Enumerated |
|  | Treatment | Route of administration of drug | route3 | Dose route | Enumerated |
|  | Treatment | Route of administration of drug | route2 | Dose route | Enumerated |
|  | Treatment | Route of administration of drug | route1 | Dose route | Enumerated |
|  | Treatment | Route of administration of drug | route6 | Dose route | Enumerated |
|  | Treatment | Route of administration of drug | route5 | Dose route | Enumerated |
|  | Treatment | How much fat (in grams) was given with the treatment? | fatamt | Fat amount | PositiveRealNumber |
|  | Treatment | Description of the fat | fatdes | Fat description | String |
|  | Treatment | Patient vomited within an hour of treatment | vomdrug | Vomit | Boolean |
|  | Treatment | Dose was repeated after patient vomited | dosrepeat1 | DosRepeat | Boolean |
|  | Treatment | Dose was repeated after patient vomited | dosrepeat2 | DosRepeat | Boolean |
|  | Treatment | Dose was repeated after patient vomited | dosrepeat3 | DosRepeat | Boolean |
|  | Treatment | Dose was repeated after patient vomited | dosrepeat4 | DosRepeat | Boolean |
|  | Treatment | Dose was repeated after patient vomited | dosrepeat5 | DosRepeat | Boolean |
|  | Treatment | Dose was repeated after patient vomited | dosrepeat6 | DosRepeat | Boolean |
|  | Concentration | Name of concentration measured | concname2 | Concname | Enumerated |
|  | Concentration | Name of concentration measured | concname3 | Concname | Enumerated |
|  | Concentration | Name of concentration measured | concname1 | Concname | Enumerated |
|  | Concentration | Name of concentration measured | concname6 | Concname | Enumerated |
|  | Concentration | Name of concentration measured | concname4 | Concname | Enumerated |
|  | Concentration | Name of concentration measured | concname5 | Concname | Enumerated |
|  | Concentration | Name of metabolite concentration measured | metconcname6 | Metconcname | Enumerated |
|  | Concentration | Name of metabolite concentration measured | metconcname5 | Metconcname | Enumerated |
|  | Concentration | Name of metabolite concentration measured | metconcname4 | Metconcname | Enumerated |
|  | Concentration | Name of metabolite concentration measured | metconcname3 | Metconcname | Enumerated |
|  | Concentration | Name of metabolite concentration measured | metconcname2 | Metconcname | Enumerated |
|  | Concentration | Name of metabolite concentration measured | metconcname1 | Metconcname | Enumerated |
|  | Concentration | Time defined in the protocol | cprotocoltime6 | Cprotocoltime | Time |
|  | Concentration | Time defined in the protocol | cprotocoltime5 | Cprotocoltime | Time |
|  | Concentration | Time defined in the protocol | cprotocoltime4 | Cprotocoltime | Time |
|  | Concentration | Time defined in the protocol | cprotocoltime3 | Cprotocoltime | Time |
|  | Concentration | Time defined in the protocol | cprotocoltime2 | Cprotocoltime | Time |
|  | Concentration | Time defined in the protocol | cprotocoltime1 | Cprotocoltime | Time |
|  | Concentration | Sample matrix, venous plasma, capillary plasma | matrix | Matrix | Enumerated |
|  | Concentration | Concentration of drug | concentration2 | Concentration | PositiveRealNumber |
|  | Concentration | Concentration of drug | concentration1 | Concentration | PositiveRealNumber |
|  | Concentration | Concentration of drug | concentration4 | Concentration | PositiveRealNumber |
|  | Concentration | Concentration of drug | concentration3 | Concentration | PositiveRealNumber |
|  | Concentration | Concentration of drug | concentration6 | Concentration | PositiveRealNumber |
|  | Concentration | Concentration of drug | concentration5 | Concentration | PositiveRealNumber |
|  | Concentration | Concentration of metabolite to three decimal places | metconcentration2 | Metconcentration | PositiveRealNumber |
|  | Concentration | Concentration of metabolite to three decimal places | metconcentration1 | Metconcentration | PositiveRealNumber |
|  | Concentration | Concentration of metabolite to three decimal places | metconcentration4 | Metconcentration | PositiveRealNumber |
|  | Concentration | Concentration of metabolite to three decimal places | metconcentration3 | Metconcentration | PositiveRealNumber |
|  | Concentration | Concentration of metabolite to three decimal places | metconcentration6 | Metconcentration | PositiveRealNumber |
|  | Concentration | Concentration of metabolite to three decimal places | metconcentration5 | Metconcentration | PositiveRealNumber |
|  | Concentration | Limit of quantification of concentration of drug | loq1 | LimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of concentration of drug | loq4 | LimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of concentration of drug | loq5 | LimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of concentration of drug | loq2 | LimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of concentration of drug | loq3 | LimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of concentration of drug | loq6 | LimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of metabolite concentration of drug | mloq3 | MetLimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of metabolite concentration of drug | mloq4 | MetLimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of metabolite concentration of drug | mloq1 | MetLimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of metabolite concentration of drug | mloq2 | MetLimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of metabolite concentration of drug | mloq5 | MetLimitOfQuantification | PositiveRealNumber |
|  | Concentration | Limit of quantification of metabolite concentration of drug | mloq6 | MetLimitOfQuantification | PositiveRealNumber |
|  | Concentration | Sample below limit of quantification | bloq3 | BelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | bloq2 | BelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | bloq1 | BelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | bloq6 | BelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | bloq5 | BelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | bloq4 | BelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | mbloq1 | MetBelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | mbloq2 | MetBelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | mbloq3 | MetBelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | mbloq4 | MetBelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | mbloq5 | MetBelowLimitOfQuantification | Boolean |
|  | Concentration | Sample below limit of quantification | mbloq6 | MetBelowLimitOfQuantification | Boolean |
|  | AdverseEvents | Action on Study Drug | AE\_action | action on study drug | Enumerated |
|  | AdverseEvents | Was concomitant medication given for this AE | AE\_concmed | concmed given | Boolean |
|  | AdverseEvents | AE Outcome | AE\_outcome | outcome | Enumerated |
|  | AdverseEvents | AE Relation to Study Drug | AE\_studydrug\_related | study drug related | Enumerated |
|  | AdverseEvents | AE Start Date | AE\_stat\_date | AE start date | Date |
|  | AdverseEvents | The term for the adverse event as described by the investigator | raw\_term | raw term | String |
|  | AdverseEvents | The meddra preferred term for the adverse event | preferred\_term | preferred term | String |
|  | AdverseEvents | The code for the meddra preferred term for the adverse event | preferred\_term\_code | preferred term code | String |
|  | AdverseEvents | The date on which the AE finished | AE\_stop\_date | stop date | Date |
|  | AdverseEvents | AE Start Time | AE\_start\_time | AE start time | Time |
|  | AdverseEvents | The time at which the AE finished | AE\_stop\_time | stop time | Time |
|  | AdverseEvents | Is the AE continuing? | AE\_continuing | continuing | Enumerated |
|  | AdverseEvents | The severity grade of the AE | AE\_severity\_grade | severity grade | Enumerated |
|  | AdverseEvents | Is the AE related to the use of primaquine | AE\_primaquine\_related | primaquine related | Enumerated |
|  | AdverseEvents | Is the AE related to the use of ACT | AE\_act\_related | act related | Enumerated |
|  | AdverseEvents | Is the AE actually an SAE | AE\_is\_SAE | is SAE | Enumerated |
|  | AdverseEvents | The type of SAE | SAE\_type | SAE Type | Enumerated |
|  | AdverseEvents | Other type of SAE | SAE\_other | SAE Other | String |
|  | AdverseEvents | Other type of SAE | SAE\_prescribed\_drug | SAE prescribed drug | String |
|  | AdverseEvents | Other type of SAE | SAE\_action\_taken | SAE action taken | Enumerated |
|  | AdverseEvents | Other type of SAE action taken | SAE\_other\_action\_taken | SAE other action taken | String |
|  | AdverseEvents | Other type of SAE | SAE\_outcome | SAE outcome | Enumerated |
|  | AdverseEvents | The date on which the patient died | date\_of\_death | date\_of\_death | Date |