WWARN Data Upload Report

[NAME OF STUDY] ([STUDY ID])

WWARN

[Automated Date]

**Introduction**

WWARN would like to thank you, your collaborators and colleagues, for contributing data on [date of submission] from the study entitled [study title] to the WWARN project. Combining data across countries and time will be the only way to track effectively the emergence and spread of antimalarial drug resistance, leading to improved health policy and effective malaria control and containment measures.

WWARN has curated and transformed your submitted data set(s) to a standard, defined format. Key variables, required for analysis of the data, have been derived according to a series of procedures details of which are available in the WWARN Molecular data management and statistical analysis plan (DM-SAP). This important document available at [link] is referred to throughout this report.

The purpose of this report is to describe the study and sample data that have been derived from your data set and used for analyses and to summarize the results. The results presented here may differ from your results due to different inclusion and exclusion criteria and methods of analysis. We would be happy to discuss any variances with you.

Data requiring contributor review before being imported into the WWARN data repository are included in a Data Query spreadsheet (attached). This table includes data inconsistencies, missing data and data transformations requiring your validation. Please see Appendix A for the data query procedure.

This study report was produced following [number of revisions] data revisions, the last of which was received on [date of accepted revision].

Table of Contents

[1. Study design, description of data sets and key variables 4](#_Toc289867301)

[1.1 Study design 4](#_Toc289867302)

[1.2 Imported variables 4](#_Toc289867303)

[1.3 Excluded variables 4](#_Toc289867304)

[2. Quality control 4](#_Toc289867305)

[2.1 Data consistency 4](#_Toc289867306)

[2.2 Data transformations 5](#_Toc289867307)

[2.3 Data description 5](#_Toc289867308)

[3. Outputs 5](#_Toc289867309)

[3.1 Sample Profile 5](#_Toc289867310)

[3.1.1 With Automatic Data correction 6](#_Toc289867311)

[3.1.2 Without Automatic Data correction 6](#_Toc289867312)

[3.2 Baseline Characteristics 6](#_Toc289867313)

[3.2.1 With Automatic Data Correction 6](#_Toc289867314)

[3.2.2 Without Automatic Data Correction 6](#_Toc289867315)

[3.3 Prevalence of resistance genotypes 6](#_Toc289867316)

[3.3.1 With Automatic Data Correction 6](#_Toc289867317)

[3.3.2 Without Automatic Data Correction 7](#_Toc289867318)

[3.4 Day 0 resistance genotypes and clinical outcomes 7](#_Toc289867319)

[3.5 Resistance genotypes on day of treatment failure 7](#_Toc289867320)

[3.6 Resistance genotypes and in vitro outcomes 7](#_Toc289867321)

[4. Conclusions 7](#_Toc289867322)

[Appendix 8](#_Toc289867323)

[A. Data query procedures 8](#_Toc289867324)

[B. Data consistency criteria 8](#_Toc289867325)

[C. Deviation criteria 8](#_Toc289867326)

[D. Imported variables 9](#_Toc289867327)

[E. Variables not imported 9](#_Toc289867328)

[F. Unexpected results 9](#_Toc289867329)

[G. Study deviations 9](#_Toc289867330)

[H. Data transformations 9](#_Toc289867331)

[I. Prevalence Tables (with automated correction) 9](#_Toc289867332)

[J. Prevalence Tables (without automated correction) 9](#_Toc289867333)

# Study design, description of data sets and key variables

## 1.1 Study design

This study is a [study type] with data on molecular antimalarial resistance markers. Data were collected from [number of subjects] subjects in [number of sites] sites. Molecular data were included for the following markers:

[TABLE MARKER LIST]

## 1.2 Imported variables

WWARN follows a standard procedure for importing variables from submitted data sets.

From data set[s] [data set file names], a set of variables listed in Appendix D were imported into the WWARN data repository. Listed in the Appendix are the WWARN header name, and in cases requiring review from the contributor, the original (source) header name. These headers are included in the Data Query spreadsheet for your review at the end of this report.

## 1.3 Excluded variables

We did not import [number of unused variables] variables from your data set. See Appendix E for a complete list. These variables are not currently required for WWARN statistical analysis.

# Quality control

WWARN conducts systematic audits on submitted data sets using standard procedures (see WWARN DM\_SAP). These audits provide a detailed profile of the data used in the WWARN analyses.

## 2.1 Data consistency

Data are checked for inconsistencies or unexpected results that may otherwise influence baseline characteristics, prevalence estimates or other types of analysis. A list of variables audited for inconsistencies can be found in Appendix B.

Appendix F lists all identified unexpected results, with the date of the event, patient identification number and additional explanations of identified anomalies. Inconsistencies are listed for your review in the Data Query spreadsheet.

Table 1 summarizes the occurrence and frequency of inconsistencies identified in the submitted data set.

***Table 1: Summary of data inconsistencies***

[TABLE\_DATA\_INCONSISTENCIES]

## 2.2 Missing data

A review of missing data is required to identify potential biases that may affect the study results. Table 2 documents the missing data from the submitted study. If some of these data are retrievable (e.g. from other source files or study documents), then these can be submitted as revised data using the spreadsheet provided at the end of this report.

***Table 2: Summary of data description***

[TABLE\_MISSING\_DATA]

## 2.3 Data transformations

Contributed data are transformed to fit the standard WWARN format of coding variables. Appendix H lists data transformations requiring review from the contributor before these values are imported into the WWARN database. These are also included in the Data Query spreadsheet for your review.

# Outputs

The following outputs have been produced from the submitted data using WWARN standard analysis procedures (see Molecular DM\_SAP). The results are displayed in two formats: exactly as received without any data correction, and after automatic data correction. In the latter case, all data inconsistencies (identified in Appendix F) have been transformed to missing values.

After reviewing these alternative outputs, you may leave your data set as received or resubmit revised data according to the procedure in Appendix A.

Once we have received your approval we will present the corrected results on WWARN Explorer

## 3.1 Sample Profile

The sample profile below summarizes sample sizes at each site in the study.

### 3.1.1 With Automatic Data correction

***Table 3: Sample profile with automatic data correction***

[TABLE\_SAMPLE\_PROFILE\_CORRECT]

### 3.1.2 Without Automatic Data correction

***Table 4: Sample profile without automatic data correction***

[TABLE\_SAMPLE\_PROFILE\_INCORRECT]

## 3.2 Baseline Characteristics

The baseline characteristic table below summarizes key features of the sample

population in the submitted data set.

### 3.2.1 With Automatic Data Correction

***Table 5: Baseline characteristics with automatic data correction***

[TABLE\_BASELINE\_CHARS\_CORRECT]

### 3.2.2 Without Automatic Data Correction

***Table 6: Baseline characteristics without automatic data correction***

[TABLE\_BASELINE\_CHARS\_INCORRECT]

### 3.3 Prevalence graphs of resistance genotypes

Prevalence and sample size of each marker is given below, grouped by location. See Appendix B of the Molecular DMSAP for a list of markers available for display in this report. Note that invariable positions (such as *pfcrt* 73V) are not shown in the prevalence graphs but are shown in the detailed prevalence tables (Appendix I and J). Combinations of markers, or haplotypes, such as the dhfr triple mutant, are not shown in the report but are available for display in WWARN Explorer. Prevalence is calculated from all samples in cross-sectional surveys or day 0 samples in efficacy studies.

### 3.3.1 With Automatic Data Correction

[INSERT CORRECTED PREVALENCE CHARTS]

### 3.3.2 Without Automatic Data Correction

[INSERT INCORRECT PREVALENCE CHARTS]

# Appendix

## A. Data query procedures

A Data Query spreadsheet of missing data, data inconsistencies and transformations requiring review has been provided to you as a an Excel file with more details in order to enable changes if you wish. Revisions and corrections can be made in the last column of the spreadsheet. Please return the completed spreadsheet to [molecular@wwarn.org](mailto:molecular@wwarn.org).

The spreadsheet is composed of the following columns:

1. STUDY\_ID: Study ID that WWARN uses to identify your study.

2. PATIENT\_ID: Patient’s study ID from your data set.

3. HEADER: Header name in WWARN data format (attached).

4. ROW: Row number from WWARN data format (attached).

5. SAMPLE\_COLLECTION\_DATE: Date on which the event took place.

6. CATEGORY: Data query type (inconsistency, unexpected or missing, optional or required…)

7. DESCRIPTION: Description of the specific data query and the headers/values involved

8. REASON: Reason that the query is being made

9. ACTION: Explanation of action required (if any) and how your data will be managed if no corrections or revisions are made. For your reference, this report includes results of baseline characteristics and genotype tables using:

a) The unmodified data set (without Automatic Data Correction)

b) The data set with these listed transformations performed (with Automatic Data Correction)

10. CORRECTION: This blank column is for your usage. Please add corrected, revised or additional results corresponding to the observation listed in each row.

## B. Data consistency criteria

WWARN data checks

Date of inclusion < January 01 1980

Sample collection date < January 01 1980

Date of inclusion > current date

Sample collection date > current date

Sample collection date < Date of inclusion

Age < 0.08 years

Age > 120 years

Unusual genotype: not found in WWARN list of expected genotypes

## C. Deviation criteria

N/A

## D. Imported variables

[APPENDIX\_IMPORTED\_VARIABLES]

## E. Variables not imported

[APPENDIX\_NOT\_IMPORTED\_VARIABLES]

## F. Unexpected results

[APPENDIX\_UNEXPECTED\_RESULTS]

## G. Study deviations

N/A

## H. Data transformations

[APPENDIX\_DATA\_TRANSFORMATIONS]

## I. Prevalence Tables (with automated correction)

[INSERT AUTOCORRECTED PREVALENCE TABLES HERE]

## J. Prevalence Tables (without automated correction)

[INSERT RAW PREVALENCE TABLES HERE]