ACTwatch Lite

Analysis User Guide

Version 1.0

Table of Contents

[Introduction 4](#_Toc205905126)

[Purpose of this guide 4](#_Toc205905127)

[Who should use this guide 4](#_Toc205905128)

[What is included 4](#_Toc205905129)

[Core indicators overview 4](#_Toc205905130)

[Study objectives 5](#_Toc205905131)

[Market Survey tool & data collection 5](#_Toc205905132)

[Data export 6](#_Toc205905133)

[Data files and structure 6](#_Toc205905134)

[Data structure 6](#_Toc205905135)

[Merging Logic and Unique Keys 6](#_Toc205905136)

[.do File structure and workflow overview 7](#_Toc205905137)

[Naming and syntax conventions 8](#_Toc205905138)

[Syntax conventions and notation 8](#_Toc205905139)

[Data file naming 9](#_Toc205905140)

[Tracking modifications 9](#_Toc205905141)

[Step-by-step-guide 10](#_Toc205905142)

[Masterfile 10](#_Toc205905143)

[Step 0: Set up 11](#_Toc205905144)

[Step 0.1 Set folder paths and macros/ for all Stata cleaning and analysis activities 11](#_Toc205905145)

[Step 0.2 Load programs 11](#_Toc205905146)

[Step 0.3 Prepare master product lists 12](#_Toc205905147)

[Step 1: Cleaning 12](#_Toc205905148)

[Step 1.1 Outlet data preparation and cleaning 13](#_Toc205905149)

[Step 1.1.1. Outlet data preparation 13](#_Toc205905150)

[Step 1.1.2. Outlet data cleaning 14](#_Toc205905151)

[Step 1.2 Antimalarial data preparation and cleaning 17](#_Toc205905152)

[Step 1.2.1. Antimalarial data preparation 17](#_Toc205905153)

[Step 1.2.2. Antimalarial data cleaning 19](#_Toc205905154)

[1.2.3: Merge QAACT and NATAPP 24](#_Toc205905155)

[1.2.4: Save clean dataset 24](#_Toc205905156)

[Step 1.3. Malaria blood testing data preparation and cleaning 24](#_Toc205905157)

[Step 1.3.1. Malaria blood testing data preparation 24](#_Toc205905158)

[Step 1.3.2. Malaria blood testing data cleaning 25](#_Toc205905159)

[1.3.3: Generate quality assured RDT (QARDT) variable 27](#_Toc205905160)

[1.3.4: Save clean dataset 27](#_Toc205905161)

[Step 1.4. Supplier data preparation and cleaning 28](#_Toc205905162)

[Step 1.4.1. Supplier data preparation 28](#_Toc205905163)

[Step 1.4.2. Supplier data cleaning 29](#_Toc205905164)

[1.4.3: Export supplier list for field team (optional) 29](#_Toc205905165)

[Step 2: Data management 30](#_Toc205905166)

[Step 2.1. Complete documentation and set local macros for CPI and exchange rates 31](#_Toc205905167)

[Step 2.2 Create a single product dataset for analysis 31](#_Toc205905168)

[Step 2.3 Apply sampling weights to the dataset 32](#_Toc205905169)

[Step 2.4. Generate denominator variables 33](#_Toc205905170)

[Step 2.5. Generate outlet type category variables 33](#_Toc205905171)

[Step 2.6 Generate blood testing category variables 33](#_Toc205905172)

[Step 2.7 Generate antimalarial category type variables 34](#_Toc205905173)

[2.7.1. Generate core antimalarial classification variables 34](#_Toc205905174)

[2.7.2. Generate WHO pre-qualification and national registration status indicators 34](#_Toc205905175)

[2.7.3. Create dosage-specific QA variables 34](#_Toc205905176)

[2.7.4. Generate binary variables for commonly encountered antimalarials 34](#_Toc205905177)

[2.7.5. Generate severe malaria treatment variable 35](#_Toc205905178)

[2.7.6. Customize for country-specific needs 35](#_Toc205905179)

[2.7.7. Conduct quality checks 35](#_Toc205905180)

[Step 2.8. Generate availability (stocking) variables 35](#_Toc205905181)

[Step 2.9. Define and assign AETDs 35](#_Toc205905182)

[Step 2.9.1. Define the full-course treatment dose 36](#_Toc205905183)

[Step 2.9.2. Adjust for presence of salt 36](#_Toc205905184)

[Step 2.9.3. Calculate the number of units required for a full course 36](#_Toc205905185)

[Step 2.9.4. Standardize product size 36](#_Toc205905186)

[Step 2.9.5. Calculate the number of AETDs per package 36](#_Toc205905187)

[Step 2.9.6 Final review and save 37](#_Toc205905188)

[Step 2.10. Generate price variables 37](#_Toc205905189)

[Step 2.11. Generate volume variables 37](#_Toc205905190)

[Step 2.12. Generate provider interview variables 38](#_Toc205905191)

[Step 2.13. Finalize and save datasets for results generation 39](#_Toc205905192)

[Step 2.14. Sensitivity analyses 39](#_Toc205905193)

[2.14.1: Antimalarial Sensitivity Analysis 39](#_Toc205905194)

[2.14.2: Diagnostic Sensitivity Analysis 40](#_Toc205905195)

[Step 3. Results 41](#_Toc205905196)

[Step 3.1: Set up local definitions 41](#_Toc205905197)

[Step 3.2: Review syntax and refine for specific study indicators and desegregations 42](#_Toc205905198)

[3.2.1: Determine which indicator tables to generate 42](#_Toc205905199)

[3.2.2: Determine desired disaggregations for each indicator 42](#_Toc205905200)

[3.2.3: Footnotes for tables 42](#_Toc205905201)

[Step 3.x: Run syntax for each relevant indicator 42](#_Toc205905202)

[Step 3.3: Run syntax for the data flow diagram 43](#_Toc205905203)

[Next steps 44](#_Toc205905204)

[Appendices 45](#_Toc205905205)

[1. Appendix 1. List of .do files 45](#_Toc205905206)

[2. Appendix 2. Sampling weights guidelines 47](#_Toc205905207)

[3. Appendix 3. AETD definition and calculation guidelines 49](#_Toc205905208)

# Introduction

## Purpose of this guide

This user guide provides step-by-step instructions for conducting an ACTwatch Lite analysis using the provided Stata syntax files. It is designed to support users in preparing, cleaning, managing, and analysing private sector outlet survey data to generate standardized indicators related to the availability, price, and market share of malaria commodities.

When used in tandem with the ACTwatch Lite toolkit and survey instruments, this guide enables a trained user to implement a full analysis without requiring additional resources. This document is designed for users already familiar with the study tools, study design, and Stata software.

## Who should use this guide

This user guide is intended for users who:

* Are responsible for implementing or supporting the analysis of quantitative data collected from the market survey component of an ACTwatch Lite study
* Are familiar with the ACTwatch Lite methodology and survey tools
* Has access to Stata (version 14 or higher) and a working knowledge of the software

This document assumes that users are operating within a shared folder structure and have access to all required data and model syntax files including in the toolkit.

## What is included

This user guide complements a package of tools included in the ACTwatch Lite toolkit required for the quantitative market study component:

* [Stata syntax](https://psiorg.sharepoint.com/:f:/r/sites/ACTWatchLite/Shared%20Documents/2.%20Technical/0.%20Toolkit/ACTwatch%20Lite%20Toolkit%20v3%20-%20FINAL%20FOR%20WHO%20REVIEW/12%20Analysis%20syntax?csf=1&web=1&e=zoZtQK) (“.do” files) for all stages of data preparation, cleaning, and analysis
* [Indicator Table](https://psiorg.sharepoint.com/:x:/r/sites/ACTWatchLite/Shared%20Documents/2.%20Technical/0.%20Toolkit/ACTwatch%20Lite%20Toolkit%20v3%20-%20FINAL%20FOR%20WHO%20REVIEW/02%20Indicator%20table%20%26%20qualitative%20themes/ACTwatch%20Lite%20Indicator%20Table.xlsx?d=w08d5e35598e245819834c352b4b4a572&csf=1&web=1&e=15iJNP) which lists quantitative indicators assessed through market survey
* [Quantitative questionnaire](https://psiorg.sharepoint.com/:f:/r/sites/ACTWatchLite/Shared%20Documents/2.%20Technical/0.%20Toolkit/ACTwatch%20Lite%20Toolkit%20v3%20-%20FINAL%20FOR%20WHO%20REVIEW/11%20Quantitative%20data%20collection%20tool?csf=1&web=1&e=GH755i) ODK form for digital data collection
* [Master product lists](https://psiorg.sharepoint.com/:f:/r/sites/ACTWatchLite/Shared%20Documents/2.%20Technical/0.%20Toolkit/ACTwatch%20Lite%20Toolkit%20v3%20-%20FINAL%20FOR%20WHO%20REVIEW/07%20Product%20Masterlists?csf=1&web=1&e=01orDk) to search antimalarials and RDTs found in the field; and
* [Results workbooks](https://psiorg.sharepoint.com/:f:/r/sites/ACTWatchLite/Shared%20Documents/2.%20Technical/0.%20Toolkit/ACTwatch%20Lite%20Toolkit%20v3%20-%20FINAL%20FOR%20WHO%20REVIEW/13%20Results%20workbooks?csf=1&web=1&e=Ptwcat) with shell tables for each indicator

Together, these tools allow for country contextualization or customization and full reproducibility of the analysis for each study implementation.

## Core indicators overview

The ACTwatch Lite analysis produces a standard set of core indicators:

1. Market composition by outlet type
2. Availability of malaria commodities
3. Median sales volumes
4. Market share
5. Sale price to customers
6. Purchase price from suppliers
7. Stockouts; and
8. Additional provider interview information

These indicators are disaggregated by outlet type, and where applicable, additional geographic strata and/or urban-rural study areas. Model syntax is available for assessing each indicator. The study tools can be modified based on the specific scope of your study and syntax can be modified accordingly.

**For more detail, refer to the** [**ACTwatch Lite Quantitative Indicator Table**](https://psiorg.sharepoint.com/:x:/r/sites/ACTWatchLite/_layouts/15/Doc.aspx?sourcedoc=%7B08D5E355-98E2-4581-9834-C352B4B4A572%7D&file=ACTwatch%20Lite%20Indicator%20Table.xlsx&wdLOR=cE75D2C03-D0CB-48EA-A1DF-6C2591943E7E&action=default&mobileredirect=true)**.**

## Study objectives

ACTwatch Lite is a cross-sectional market survey of retail-level private health sector providers and suppliers. The study is designed to capture representative data of malaria commodity markets at the retail and wholesale levels. This data, often absent or inadequate, is needed to better understand drivers of market performance and target areas and channels for intervention, inform national and subnational malaria control strategies, and funding requests. The study is designed to be efficient and adaptable, enabling country teams to collect key market data without the burden of extensive custom development

## Market Survey tool & data collection

The quantitative questionnaire tool in the ACTwatch Lite toolkit is modular and can be tailored to meet national needs, while maintaining a core set of indicators. It is designed to be implemented using digital platforms compatible with the Open Data Kit (ODK) ecosystem, including SurveyCTO, KoboToolbox, and ODK Collect[[1]](#footnote-2). Modules include eligibility screening, provider interviews, product audits for antimalarials and rapid diagnostic tests, and the collection of supplier information.

Forms are tailored for use by trained enumerators with a digital data collection device (e.g., smartphone, tablet) and may be locally adapted based on the study scope and context. Data are collected through a census of all outlets within each of the sampled study areas. The data output from the market survey feeds directly into the analysis syntax described throughout this document.

**Refer to the** [**ACTwatch Lite Toolkit Manual & Implementation Guide**](https://psiorg.sharepoint.com/:f:/r/sites/ACTWatchLite/Shared%20Documents/2.%20Technical/0.%20Toolkit/ACTwatch%20Lite%20Toolkit%20v3%20-%20FINAL%20FOR%20WHO%20REVIEW/01%20Manual%20and%20implementation%20guide?csf=1&web=1&e=f71Jjh) **as well as** [**the Protocol**](https://psiorg.sharepoint.com/:f:/r/sites/ACTWatchLite/Shared%20Documents/2.%20Technical/0.%20Toolkit/ACTwatch%20Lite%20Toolkit%20v3%20-%20FINAL%20FOR%20WHO%20REVIEW/04%20Protocol%20template%20and%20references?csf=1&web=1&e=Tj7OTo) **template for more detail.**

## Data export

Once data collection is complete, data are exported from the ODK platform in long format. This format is critical to ensure compatibility with the ACTwatch Lite Stata syntax, which has been developed specifically to work with repeat group structures in the survey. Wide-format data will not be compatible.

The export includes separate files for outlet-level responses, antimalarial audits, RDT audits, and supplier details. These are then imported into Stata and processed using the standard cleaning and management scripts provided.

# Data files and structure

## A screenshot of a computer AI-generated content may be incorrect.Data structure

Figure 1. Data structure

The ACTwatch Lite survey produces four core datasets (Figure 1. Data structure):

1. the outlet-level dataset, which contains screening and provider interview data;
2. the antimalarial product audit dataset;
3. the malaria rapid diagnostic test (RDT) audit dataset; and
4. the supplier dataset.

Each dataset is exported in long format and structured to support one row per observation (e.g., one row per outlet, product, or supplier).

The outlet dataset serves as the primary reference file to which the product (antimalarial and RDT) and supplier datasets are linked using unique identifiers

## Merging Logic and Unique Keys

As stated above, all repeat-group data (antimalarial and RDT product audits and supplier listings) are stored separately and merged to the outlet-level data using a unique outlet ID generated by the ODK platform.

Table 1. Linking datasets together, below, shows each dataset’s unique IDs and IDs that can be used to link datasets together.

Table 1. Linking datasets together

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | Unique record ID within this dataset | ID that links to outlet dataset | Notes |
| Outlet | key | - | The outlet data is the main dataset to which the others should be linked |
| Antimalarial audit | amauditkey | parentkey | parentkey should be renamed to “key” before merge |
| RDT audit | rdtauditkey | parentkey |
| Suppliers | suppkey | parentkey |

Each dataset is prepared and cleaned separately. Outlet data cleaning should be run first so that any variables from this dataset that are merged to the others are clean. The master .do file is structured to run in this order.

In the data management syntax, product datasets (antimalarials and RDTs) are combined. Analysis .do files run on each of the final 3 datasets (outlet data, product data, and supplier data)

## .do File structure and workflow overview

The ACTwatch Lite analysis .do files are organized by analysis step into folders as demonstrated in Table 2. Organization of analysis .do files, below.

Table 2. Organization of analysis .do files

|  |  |
| --- | --- |
| **Step 0 Set up** | Initialization scripts and global macros. |
| **Step 1 Cleaning** | Scripts for cleaning outlet, product, and supplier datasets |
| **Step 2 Data management** | Scripts for creating analysis variables and merging datasets |
| **Step 3 Results tables** | Scripts for generating output data tables |

All scripts are ultimately managed through a single master file (**ACTwatch Lite\_MASTERFILE.do**) (Figure 2. ACTwatch Lite analysis workflow). This .do file runs the full workflow, from set up to output generation. Before executing the Masterfile, users must customize each referenced .do file by inputting study-specific information (e.g., file paths, local treatment guidelines, etc.) where prompted.

***A diagram of a workflow

AI-generated content may be incorrect.***

Figure 2. ACTwatch Lite analysis workflow

## Naming and syntax conventions

### Syntax conventions and notation

ACTwatch Lite syntax uses a set of conventions to guide users in navigating and customizing code. Specific markers are included to highlight areas requiring attention.

While much of the syntax is pre-written, many sections require direct input from the user, as would be expected for any large study.

The symbol $$$ is used throughout the syntax to indicate a line that must be modified by the user. These typically include file directories or paths, dataset names, macros, and other study-specific parameters (e.g. first line treatment definitions, national treatment guidelines, etc.).

For example, the following standard syntax is provided to facilitate the importation of outlet-level data from a primary CSV:

$$$ insheet using "${datadir}/raw/[OUTLET MAIN DATASET FILENAME]", names clear

The ‘$$$’ is flagging to the user that a modification is required. In this case, the user should remove the “$$$” and add the.csv file name that they wish to import. An example of how the syntax should be modified, where the .csv file is named “outlet\_main\_1”, is as follows:

insheet using "${datadir}/raw/outlet\_main\_1.csv", names clear

In addition, the characters \*!!! are used to flag important instructions or troubleshooting notes and blocked comments, demarcated using /\* ... \*/, are used to explain more complex sections of the code. Some example syntax from pilot studies has been retained for reference and is indicated by /\* EXAMPLE\*/ throughout.

### Data file naming

As users progress through the analysis, datasets are saved with suffixes to track progress and prevent overwriting. Raw datasets are saved as initially exported. Once the preparation scripts are run, the suffix "\_prepared" is added. After cleaning, datasets are saved with the suffix "\_clean". This naming convention is applied consistently across outlet, audit, and supplier datasets.

### Tracking modifications

It is essential to document any major changes in an analysis log file or within the syntax using comment blocks. Outputs from Stata should also be copy-pasted into the do file to track where modifications have been made during cleaning using /\* ... \*/

# Step-by-step-guide

## Masterfile

**Purpose:** This .do file is designed toset up the ACTwatch Lite analysis for a given country/implementation and then run all subsequent .do files that prepare, clean, and manage market survey data. Running this .do file culminates in the automatic production and output of data tables for every study indicator (Figure 2. ACTwatch Lite analysis workflow). Note that it will not be possible to run this master .do file in full until all cleaning and analysis steps are completed by the analyst, and in fact it may not be necessary to run this file in full. It does though act as a high-level tracker for the study’s cleaning and analysis .do files and should be kept updated.

The MASTERFLIE.do file has two sections, A and B, detailed in Table 3. Masterfile sections and analysis syntax step overview.

Table 3. Masterfile sections and analysis syntax step overview

|  |
| --- |
| **Section A** |
| * Initializes Stata with standard display, version and memory settings. * Sets specific global parameters (e.g., study country, study year) for the analysis. * Sets the folder location for each user so that all later .do files may be run by multiple users via the same syntax. To complete this step, ensure folder pathways are tailored for each user. **It is recommended that shared folders (OneDrive, or similar) are used for all data and analysis files.** |
| Section B |
| * Runs separate .do files that correspond to Step 0-3 of the data analysis process:  |  |  | | --- | --- | | Step 0 Set up | This step sets specific file pathways and parameters for the study and runs programs needed for the analysis. | | Step 1 Cleaning | This step uses a series of outlet data, antimalarial audit and RDT audit .do files to prepare raw data, convert .csv files to Stata data files (.dta), add variable labels, check unique identifiers, and append and merge data as necessary. | | Step 2 Data management | This step ensures consistency, transparency, and reproducibility in data management and analysis by merging datasets; creating sampling weights; creating antimalarial categories, denominators, and stocking category variables; and calculating adult equivalent treatment dose values. | | Step 3 Results tables | This step sets up the analysis for each key indicator, creating tables and figures used to report ACTwatch Lite data. | |

**It is important to note that the preparation, cleaning, management, and analysis .do files ALL REQUIRE SIGNIFICANT MODIFICATION and inputs before they can be run.**

The master .do file, as well as all subsequent or dependant .do files, are annotated to indicate where inputs are needed from the user (the notation ‘$$$’ indicates lines in the syntax that require input or review). While the MASTERFILE allows for the automatic production of outputs from the study, this file should only be executed after all other steps (setup, cleaning, data management, results) and corresponding .do files have been completed by the user. After all subordinate .do files have been reviewed and customized, the MASTERFILE can be run to conduct all steps in the full analysis, should this be needed. Steps and .do files for each are detailed in the sections below.

***Remember***

*Any changes made to the datasets during data preparation, cleaning or management must be documented in the .do files by copy-pasting relevant Stata output and including annotations from the analyst explaining what has changed and why, as needed*

## Step 0: Set up

**Purpose**: This .do file initializes the analysis environment by defining pathways, macros and parameters needed throughout the workflow. Users will specify study-specific inputs such as the country code and study year, and AM/RDT product master lists. These setup files ensure that all subsequent scripts can be executed consistently by multiple users (Table 4. Step 0. Set Up .do file names).

Table 4. Step 0. Set Up .do file names

|  |  |
| --- | --- |
| 0.1\_MACROS | Defines global macros used throughout the analysis. |
| 0.2\_PROGRAMS | Loads user-defined programs to be called by subsequent .do files |
| 0.3\_PRODUCTLISTS | Prepares the master product lists for antimalarials and RDTs |

**Steps:**

### Step 0.1 Set folder paths and macros/ for all Stata cleaning and analysis activities

Creates and assigns macros to sub-folders used by Stata for cleaning and analysis processes. **All but the final macro should not be modified**, as later .do files will rely on them in their unaltered state. However, the final macro should be updated to include the file pathway to the antimalarial and RDT master list files, which were used during data collection as the searchable datasets that auto-populate product audit information.

### Step 0.2 Load programs

In this step, a set of Stata programs are defined that can be used later by the users to support pharmaceutical data validation by checking consistency in reported strengths of active pharmaceutical ingredients across various dosage forms. Programs are tailored for products with one, two, or three APIs in solid (**ais\_tk**) and liquid (**ais\_nk**) forms, while **ais\_tu** and **ais\_nu** handle cases with unknown strengths. The **strtu** utility standardizes text entries. Each program flags discrepancies between recorded and expected strengths based on user-defined inputs, aiding in efficient quality control of drug audit data. **These programs should not be modified.**

### Step 0.3 Prepare master product lists

Product lists with known antimalarials and RDTs were used to auto-populate questionnaire instances to reduce fieldworker time and effort (and error) caused by manual product audit input. Instead, fieldworkers can search the database of known products for each audited product during data collection. This step converts these product lists from .csv to .dta format to be used during analysis. To ensure the .dta datasets are fit for purpose, this step also cleans variable names, standardizes text to uppercase, and renames ID variables (**amcode** for antimalarials and **rdtcode** for RDTs). It further checks for duplicate IDs and counts observations before saving cleaned datasets to specific directories. In completing this step, you are also creating a set of standardized and deduplicated master lists for further data validation or analysis in the future.

**Outcome**: The outcomes of the 0.Setup .do files are a set of macros for each sub-folder in the analysis, a set of programs loaded into Stata to assist with data cleaning, and standardized antimalarial and RDT product lists for subsequent analysis.

## Step 1: Cleaning

**Overview:** The overall approach to cleaning ACTwatch Lite data is a conservative one, in which most variables and values/cases should not require modification. Variables are checked for common, known errors (e.g., those known to occur frequently during fieldworker data entry) and inconsistencies between values. The most common areas that require cleaning are the antimalarial and RDT audit sections of the questionnaire, which include detailed, technical information on pharmaceutical products, collected by fieldworkers who have been trained to gather these data from the packaging. The data collection process allows for the collection of product photos that can also be viewed by the users to resolve inconsistencies in the dataset where they are found.

**Purpose:** This step involves executing multiple Stata .do files to process and clean raw survey data. These scripts handle data preparation tasks such as converting .csv files to .dta files, labelling variables, ensuring data integrity through checks on unique identifiers, and merging or appending datasets for analysis, as well as data cleaning steps that require the user to check variables for each question from the quantitative questionnaire and clean data as needed.

There are four datasets output by the standard ACTwatch Lite ODK form. These are the (1) outlet-level data from the census and provider questionnaire, (2) antimalarial data from antimalarial product audits, (3) malaria blood testing data from diagnostic audits, and (4) supplier information reported by retail outlets on their main suppliers of antimalarials and RDTs (Table 5. Step 1. Cleaning .do file names). Each of these datasets are cleaned separately. Steps 1.1-1.4 correspond to each of these datasets.

Table 5. Step 1. Cleaning .do file names

|  |  |
| --- | --- |
| 1.1\_CLEANING-OUTLET | Imports, prepares, and cleans data captured for each outlet visited during the study, as well as screening and provider data for outlets where this was conducted |
| 1.2\_CLEANING-ANTIAMALARIAL | Imports, prepares, and cleans audit data for antimalarials |
| 1.3\_CLEANING-RDT | Imports, prepares, and cleans audit data for RDTs |
| 1.4\_CLEANING-SUPPLIER | Imports, prepares, and cleans supplier repeat group data for antimalarials and RDTs. |

**Steps:**

### Step 1.1 Outlet data preparation and cleaning

#### Step 1.1.1. Outlet data preparation

##### 1.1.1.1: Import data sets

This step imports the primary .csv dataset into Stata and converts it to .dta format.

##### 1.1.1.2: Label variables

Variable labels are then applied using a dedicated .do file. To run the labeling .do file, ensure the appropriate directory is set in the syntax.

à Note that the outlet labelling .do file requires review and editing to ensure that the labels are correct for the study details. A new label .do file may be outputted directly from the ODK software package being used.

##### 1.1.1.3: Prepare field types

In this step, field types are prepared. Each string, date, and time/date variable from the ODK questionnaire should be manually inserted and modified by the user. Once this is complete, the existing syntax will automatically drop non-essential variables (e.g., notes variables), eliminate supplementary forms that contain no data, format date and date/time fields, ensure text fields are imported as strings, and trim and convert string values to uppercase.

##### 1.1.1.4: Check and correct duplicates

Outlet ID duplicate observations are flagged for resolution in this step. Duplicate form IDs may occur where tablets have only partially uploaded forms to the ODK servers, for example. Example code is provided to drop true duplicates, as required.

##### 1.1.1.5: Check links

Each outlet may have up to four linked datasets: antimalarial audits, RDT audits, antimalarial suppliers, and RDT suppliers. In this step, the user should check that each outlet has the unique ID variables (UIDs or keys) to link to the datasets relevant to that outlet (e.g., if the outlet has RDTs there should be an **rdtaudit** link).

##### 1.1.1.6: Correct known errors highlighted during data collection

In this step, the user will have the opportunity to correct known errors highlighted during data collection, typically flagged in the ‘**comments**’ variables of the ODK form questionnaire or tracked by the study team through their QA procedures. Examples might be fieldworkers misidentifying the study area code for a particular day or outlet, for example.

##### 1.1.1.7: Drop variables not required for analysis

This step drops ODK questionnaire variables that are not required for cleaning and analysis. A list of variables is provided in the syntax. However, the user should modify the list as necessary, according to their study. Generally, users should drop ‘**calc’** or ‘**group’** fields (variables created by the ODK software during data collection) and other variables that are products of the data collection software. A conservative approach to dropping variables is recommended—no variables of potential use in cleaning or analysis should be dropped.

##### 1.1.1.8: Save prepared data

This step will save and prepare the data as a clean .dta file. The variables **key** and **outletid** uniquely identify forms. The variable **outletid** uniquely identifies outlets. Both are needed for conducting the one-to-many merge of the outlet data to multiple records in the antimalarial and RDT audits and suppliers’ datasets.

#### Step 1.1.2. Outlet data cleaning

The outlet data set is long with respect to outlets (i.e. has one case/ line in the dataset per outlet). This dataset contains all the screening and geographic variables, all the provider interview information and the UIDs linking to the audits and supplier datasets. In this step, each variable in the outlet dataset is reviewed, checked and cleaned.

##### 1.1.2.1: Interviewer comments

In this step, the user should review all interviewer comments that could require cleaning of outlet-level variables. Flagged comments should be cleaned in their respective cleaning step. There are two relevant comments sections to be checked for this step, end of questionnaire comments and business practice comments.

##### 1.1.2.2: Metadata

This step entails checks for data quality assurance related to date range and interview duration. Out of range dates should be investigated and pilot data should be removed (if they haven’t been already). Average interview time should be reviewed, and outlier values should be flagged and investigated.

It is recommended that each of the checks performed in this section are also done during data collection as well as after data collection completion, during the analysis phase. This way, potential issues may be spotted and corrected while data collection is still on-going.

##### 1.1.2.3: Census module

This step is designed to help the user review, check, and clean each variable in the census module section. This includes variables for team, data collectors, location, geocoordinates, outlet type, and outlet name. Syntax and additional guidance are provided for each variable to help the user sense check the data.

##### 1.1.2.4: Screening and eligibility module

This step facilitates the review, check, and cleaning of screening and eligibility variables. The examination of variables for screening, inclusion criteria, eligibility, consent, and wholesale/outlets engaging in re-sale is facilitated by syntax and additional guidance within the .do file.

After these variables have been cleaned, an outlet status variable will be generated, which can be sense checked and cleaned as required. Further variables are created to flag instances where interviews were started or interrupted and where antimalarial and RDT audits are incomplete. This step culminates in the creation of a final interview status variable, **finalIntStat**, which will be referenced and finalized during the cleaning process.

##### 1.1.2.5: Provider module

This set of steps is designed to review, check, and clean the provider module of the quantitative questionnaire. It is broken into seven sections, which are laid out below. Sections include outlet characteristics, staff characteristics, quality control and compliance, respondent malaria knowledge, outlet tech/digital access and use, outlet participation in monitoring, and business practices.

For the provider module, the default setting for each of the seven sub-sections is that these questions are only asked at 1) outlets that have consented to participate and 2) outlets that have not been coded in the above section as ‘informal.’

If the user’s survey employs different eligibility criteria, the syntax **prov\_int=1** & **inf!=1** should be updated. Similarly, if sub-sections are added or removed, the numbering of checkpoints in the user’s survey may be different and will need to be updated, as required.

###### 1.1.2.5.1: Section 1: Outlet characteristics

In this step, variables related to outlet characteristics will be reviewed and cleaned, as necessary. Variables in this sub-section relate to outlet opening time, year established, number of employees, existence or appropriateness of licensure/registration, and history of government inspection.

###### 1.1.2.5.2: Section 2: Staff characteristics

In this step, variables related to staff characteristics will be reviewed and cleaned, as necessary. Variables in this sub-section include staff qualifications and recent receipt of malaria training.

###### 1.1.2.5.3: Section 3: Quality control and compliance

In this step, variables related to quality control and compliance will be reviewed and cleaned, as necessary. Variables in this sub-section include antimalarial and RDT storage practices, as well as the availability of microscopy equipment.

###### 1.1.2.5.4: Section 4: Respondent malaria knowledge

In this step, variables related to respondent malaria knowledge will be reviewed and cleaned, as necessary. Variables in this sub-section correspond to several malaria knowledge questions asked of the respondent about first-line drug recommendations, RDT knowledge, RDT use, necessity of a positive confirmatory test prior to antimalarial vending, and antimalarial use with negative malaria RDT result.

###### 1.1.2.5.5: Section5: Outlet tech/digital access and use

In this step, variables related to the outlet’s access to, and use of technology, digital tools, and related necessary infrastructure will be reviewed and cleaned, as necessary. Variables in this sub-section include recent use of tech/digital tools including whether the outlet had water, electricity, any phone, mobile phone, wifi, and tablet/computer. This section also probes as to whether the provider is interested in digitizing certain functions in the future.

###### 1.1.2.5.6: Section 3: Outlet participation in monitoring

In this step, variables related to outlet participation in monitoring will be reviewed and cleaned, as necessary. Variables in this sub-section include those related to monthly malaria case reporting practices as well as the forms/tools used to report that malaria data.

###### 1.1.2.5.7: Section 3: Business practices

In this step, variables related to business practices will be reviewed and cleaned, as necessary. Variables in this sub-section include those about online and in-person retail of antimalarials and RDTs; types of customers (individual, wholesaler etc.); types of suppliers, procurement, payment and stability; wholesale business practices; and business networks.

##### 1.1.2.6: Antimalarial audit module – outlet-level variables

This step will assist in the review and cleaning of outlet-level variables within the antimalarial audit module. Most of the antimalarial data is captured in the "repeat group" of the form and stored in a separate dataset (antimalarial audit dataset). Those variables will be cleaned separately later in the process. Variables addressed in this step relate to currently stocked antimalarials in the outlet, audit completeness, and antimalarial stock-outs.

##### 1.1.2.7: Blood testing audit module -outlet-level variables

This step will assist in the review and cleaning of outlet-level variables within the RDT audit module. Most of the RDT data is captured in the "repeat group" of the form and stored in a separate dataset (RDT audit dataset). Those variables will be cleaned separately later in the process. Variables addressed in this step relate to currently stocked RDTs in the outlet, RDT completeness, RDT stock-outs, and microscopy details.

##### 1.1.2.8: Supplier details module – outlet-level variables

This step will assist in the review and cleaning of outlet-level variables within the supplier details module. Most of the supplier data is captured in the "repeat group" of the form and stored in a separate dataset (supplier dataset). Those variables will be cleaned separately later in the process. Variables addressed in this step relate to details of main commodity suppliers.

##### 1.1.2.9: Save clean dataset

This step ensures that a clean .dta dataset is saved that includes all updates made during step 1.1.2.

**Outcome**: The outcome of this .do file is a clean .dta dataset with information on each outlet visited during the study, as well screening and provider interview data for outlets where this was conducted. The clean outlet dataset will be used for outlet-level analyses and merged to product and supplier-level analyses later in the process.

### Step 1.2 Antimalarial data preparation and cleaning

#### Step 1.2.1. Antimalarial data preparation

##### 1.2.1.1: Import datasets

This step imports the primary .csv dataset into Stata and converts it to .dta format.

##### 1.2.1.2: Label variables

Variable labels are then applied using a dedicated .do file. To run the labeling .do file, ensure the appropriate directory is set in the syntax.

à The antimalarial audit labeling .do file requires review and may need to be updated to match the data collection tool to ensure that all variables in the audit are correctly labelled.

##### 1.2.1.3: Prepare field types

This step prepares fields to facilitate subsequent analysis. First, the user should manually list string variables, which can be copied from ODK software .do files. Once this is complete, the existing syntax will automatically drop non-essential variables, remove accents from letters, and trim and convert string values to uppercase to make analysis easier.

##### 1.2.1.4: Drop tablet and non-tablet values where these have already been combined into a single variable by ODK form

The ODK form captures information from tablet and non-tablet antimalarials separately. The form then auto-combines variables. Separate ‘**tab**’ and ‘**ntab**’ are not needed and can be dropped here. Review provided syntax to ensure that no necessary variables are included in the drop command.

Other variables that are not required for this analysis should be dropped. Due to the structure of the ODK questionnaire, there are separate sets of variables for tablet and non-tablet products, as well as products found and entered through the search function versus manually. Only a complete set of combined variables, generated elsewhere in the syntax, are required for analysis. The rest may be dropped using the provided syntax.

##### 1.2.1.5: Combine searched and manually entered product information

Like the above step, the ODK form captures variables separately for products that were searched and pre-populated versus those that were manually entered during data collection. Distinct ‘**\_search**’ and ‘**\_manual**’ variables are not needed and can be dropped at this stage. The user will then be prompted to create integer variables for product strengths, merge suspension formulation details and de-string active ingredient names. The user will then recalculate and recode the dose form categorical variable ‘**a3\_category’** so that it will run correctly with the active ingredient program.

##### 1.2.1.6: Check and correct duplicates

In this step, the user will check for duplicates using the unique record variable **key** and resolve as required.

##### 1.2.1.7: Correct known errors highlighted during data collection

During data collection, known issues (e.g., duplicate or incomplete/incorrect entries, erroneous IDs) should have been flagged by data collection teams. In this step, the user will either correct all flagged issues or note elsewhere in the syntax where issues have been resolved.

##### 1.2.1.8: Create cleaning variables to maintain original values

To avoid mass recodes or loss of original values, use the provided syntax to create ‘**\_orig**’ variables to store these data before cleaning. The **\_orig** variables can then be used to back-compare with cleaned data as needed.

##### 1.2.1.9: Merge outlet-level data

Now, outlet-level data will be merged with the antimalarial audit dataset. This is done by merging using the ‘**key**’ variable, which is unique to every outlet form.

##### 1.2.1.10: Merge and replace updated masterlist information

If the antimalarial MASTERLIST file was updated or if any changes were made during data collection, in this step, the user will merge the database to the product data and replace information. The merge is done using **amcode** as the key variable, which is unique to each antimalarial in the masterlist.

In completing this step, the user must be sure to check the merge against the ‘**fillmethod**’ variable (this variable indicates whether the antimalarial product information came from the database or was manually entered by a data collector). Any changes or additions should be made only after careful review of the associated product images.

##### 1.2.1.11: Save prepared data

This step ensures that a clean. dta dataset is saved that includes all updates made during step 1.2.1.

#### Step 1.2.2. Antimalarial data cleaning

##### 1.2.2.1: Import prepared antimalarial dataset

In this step, the user will open the prepared dataset if it is not already in use.

##### 1.2.2.2: Generate sum of strengths

The **sumstrength** variable is generated, which is the sum of the antimalarial strengths. This variable is used for subsequent cleaning as it allows for products to be more easily identified in many cases (i.e. where artemether 20mg lumefantrine 120mg has been entered during data collection as "lumefantrine 120mg artemether 20mg" the sum of strengths variable has the same value (i.e. 140) in both cases.

Before generating the **sumstrength** variable, the user should quickly check the manually entered product strengths for any obvious errors (e.g. AL 20/120 that has been entered as 21/120, with sumstrength 141).

##### 1.2.2.3: Confirm all products are antimalarials

In this step, the user should follow prepared syntax to identify non-antimalarials. The user should add results or make changes based on careful review of product images or of the same products elsewhere in the dataset to avoid dropping a true antimalarial observation. Common products that are confused with antimalarials include antibiotics and herbal medicines.

##### 1.2.2.4: Clean brand names

In this step, it is important for the user to carefully check manually entered brand names to ensure they are as clean and consistent as possible. This cleaning step may include manually correcting obvious errors and/or reviewing product photos. The analyst should use their preferred approach for cleaning string variables in Stata (eg., regexm). This step can be laborious, but care must be taken to avoid mass recodes/replacements, and a product-by-product approach is advised.

##### 1.2.2.5: Clean manufacturers

This step is similar to the previous step. The user must carefully check manually entered manufacturer names to ensure they are as clean and consistent as possible. This cleaning step may include manually correcting obvious errors and/or reviewing product photos. The analyst should use their preferred approach for cleaning string variables in Stata (eg. regexm). This step can be laborious, but care must be taken to avoid mass recodes/replacements, and a product-by-product approach is advised.

##### 1.2.2.6: Check brand/manufacturer consistency

In this step, the user confirms that all manufacturers have been correctly cleaned and relate directly to the antimalarial product. In general, a given brand should only be made by one manufacturer. However, this is not always the case. Confirm via product photo to ensure that the dataset accurately reflects brand/manufacturer combinations.

##### 1.2.2.7: Country of manufacturer

As above, this step entails cleaning to ensure consistency across manufacturers and country of manufacture. While most manufacturers have only one country of manufacture, this is not always the case. Check consistency of country of manufacturer by manufacturer and brand name. Use photos when in doubt. “Other, specify” is often selected by fieldworkers when they do not recognize the country, and include a city instead, for example, Shanghai should be recoded to **amcountry**= China.

##### 1.2.2.8: Generic name – preparation

The generic name (**gname**) is the combination of active ingredients. This variable is generated here as it is needed for subsequent cleaning steps. In this step, the user will ensure there are no active ingredient errors, generate a generic name and recode/label with for consistency with provided syntax. Using standard **gname** syntax will facilitate further analysis later in the do files, so avoid altering the **gname** syntax and values unless these changes can be carried forward throughout the .do files.

##### 1.2.2.9: Dose form (three checks)

This step entails three separate checks to ensure that the dose form is accurate in the dataset.

**Suspensions:** Data collectors often confuse powder and liquid suspensions as well as liquid suspensions and syrups. The user should check all records for obvious errors and confirm with photos or other audit information while cleaning. Here are some cleaning hints to help the user execute this step:

* Most suspensions are powder.
* All artemisinin-combination therapy (ACT) suspensions should be powder. Document any chloroquine/quinine/other suspensions that may be liquid or powder.
* Note that if a product is liquid, but its packaging says “suspension” and/or says to shake before use, it should be marked as a liquid suspension, not a syrup.

**Injectables:** Another common data collection error is confusing liquid and powder injectables. The user should check all observations for obvious errors, using photos or other audit information to clean. Here are some cleaning hints to help the user execute this step:

* ACTs should not be injectables or syrups (document any cases if confirmed with product photos).
* Injections in ampoules are liquids. Injections in vials are powder.
* Powder injections are frequently co-packaged with an ampoule of solvent (to dissolve the dry powder injectable) – these should still be entered into the dataset as powder injectables as the active ingredient is a powder injection.
* Injectables are usually monotherapies and will likely be artesunate, artemether, artemether, chloroquine, or quinine.

**Other dose forms:** Finally, the user should check other dose forms. The user should check all observations for obvious errors, using photos or other audit information to clean. Here are some cleaning hints to help the user execute this step:

* Most ACTs should be tablets, granules or suspensions.
* Granules should only be ACTs.
* Artesunate tablets are rare.
* Suppositories audited in ACTwatch Lite studies to date have been mostly artesunate.
* Syrups are generally only chloroquine or quinine.
* Drops are rare, and mostly or always quinine.

##### 1.2.2.10: Generate number of active ingredients

In this step, the user will create a variable that indicates the number of active ingredients for each product called **ainum**.

##### 1.2.2.11: Strengths – ML/MG or MG

Product strength is an important variable for estimating several key indicators. Liquids (i.e., injectables, syrups, drops, and liquid suspensions) have their strengths notated as mg/ml. All other products have strengths notated as mg. In this step, the user will ensure that the strength correctly corresponds to each product.

##### 1.2.2.12: Strengths – tablets, suppositories, granules

In this step, syntax is provided to check consistency between active ingredients and strength (using ais\*\_ programs). Different strengths of a product with the same generic name (**gname**) are expected, but they should be consistent.

Syntax is provided to check strengths of tablets, suppositories, and granules. The syntax will highlight outliers by each generic product. The user should carefully move through each tablet, suppository, and granule generic name and review and clean the strengths as necessary. These checks are only required for manually filled antimalarials (i.e**., fillmethod**==2)

##### 1.2.2.13: Strengths – Syrups, suspensions, liquid injections

As above, the user should investigate and clean syrup, suspension, and liquid injection strengths in this step.

##### 1.2.2.14: Active ingredient/strength – consistency for syrups, suspensions, liquid injections

In this step, syntax is provided to check consistency between active ingredients and strength for syrups, suspensions, and liquid injections using the ais\_\* programs. These checks are only required for manually filled antimalarials (i.e**., fillmethod**==2)

In the provided syntax, the third argument takes the known active ingredient strengths in milligrams (mg), separated with spaces (no commas), and the fourth arguments takes the known active ingredient strengths in milliliters (ml), also separated with spaces. Across the two arguments, the first listed milligram strength corresponds with the first milliliter strength, the second milligram strength corresponds to the second milliliter strength, and so forth.

##### 1.2.2.15: Salt

Some active ingredients have strengths (in mg or mg/ml) for the active ingredient plus salt. In this step, the user should ensure that a salt is specified in cases where non-base strength was recorded for an active ingredient. A set of syntax is provided within this .do file based on salt information gathered from prior ACTwatch studies. However, where there is doubt, manually clean based on careful review of product images and/or information on the same product elsewhere in the dataset.

##### 1.2.2.16: Package type

This step ensures that each observation has the correct package type information. The type variables were defined in a specific way for data collection and the flow of the ODK form. In this step, these variables will be redefined using provided syntax to facilitate cleaning.

Next, the user should check package type by dose form to ensure those variables align. There are several common data collection errors highlighted in the .do file, along with syntax to help the user identify them.

##### 1.2.2.17: Package size

This step serves to clean and align the number of tablets, suppositories, vials, and so forth per **packagetype**. The user should check and edit package size information where there is enough evidence to do so. Here are some cleaning hints to help the user execute this step:

Check that package size and package type are consistent:

* Tablet, suppository and granules package type should be pot when package size is > 56 (unless exceptions found).
* Tablet, suppository and granule package type should be pack when size is <=56.
* Other products package type should be bottle when package size is > 10ml.
* Other products package type should be vial/ampoule when package size is <=5ml.

If the user identifies inconsistent cases in any of the above checks, then:

* Check product images.
* Review audit sheet comments (**nt\_15**).
* Search the dataset for other medicines with the same formulation, brand name and package size. Edit package type for consistency with these products. Consistency means that the package type is commonly associated with the given brand name and package size and is consistent with formulation (e.g., ampoules/vials correspond with liquid or powder injections).

Next, the user should check the size versus strength for ACTs by checking for consistency within brand, and/or looking at product photos. This step should be conducted for each antimalarial type found in the dataset. Note that it is a common error for data collectors to confuse size and strength for liquid injections. Confirm each product is correct and recode as necessary.

Suspensions can be very difficult to accurately collect data on for size/strength variables. Often, there is confusion around bottle size, dosage and strength. Therefore, a variety of values are likely to be input for strength, pack size etc. The user should pay close attention during this step to ensure that suspension information is accurate by scrutinizing each suspension generic in a systematic fashion.

##### 1.2.2.18: Products with regulatory code

In this step, users should edit variable names and add additional checks if more than one code is recorded.

##### 1.2.2.19: Expired products

This checks the expired product variable.

##### 1.2.2.20: Stocked out in the last three months

This step checks the products that were stocked out in the past three months.

##### 1.2.2.21: Amount sold

In this step, the number of each product sold in the previous seven days is calculated. Careful review of missing values should be conducted here.

##### 1.2.2.22: Price

This step calculates the retail price to customers for each antimalarial product sold, price paid to suppliers, price charged for resale, wholesale selling price, and price per amount (e.g. pack) for all antimalarial products.

##### 1.2.2.23: Photo availability

This step reviews the variable for availability of product photos

##### 1.2.2.24: Fixed-dose combination (FDC)

The FDC variable is needed to estimate the number of adult equivalent treatment doses (AETDs) per pack later during data management. AETDs are a standardized unit used throughout the ACTwatch Lite analysis using WHO-recommended dosing guidelines for a 60 kg adult. AETD is required for the price and market share indicators. See **Appendix 3. AETD definition and calculation guidelines** for more information on the definition of and calculations for AETDs.

Syntax is provided to check for consistency across FDC and a number of variables.

##### 1.2.2.25: Identify manually entered products matched to masterlist and add **amcode** from masterlist

This section ensures that manually entered products that were either always in the database and missed by data collectors or added during data collection are assigned the correct database product ID.

#### 1.2.3: Merge QAACT and NATAPP

If the antimalarial masterlist was updated or changes were made during data collection, this step merges the database to the product data and replaces the information.

##### 1.2.3.1: Generate a variable to identify which audited products are WHO pre-qualified (QAACT)

Generating the **QAACT** (quality-assured ACT) variable requires the WHO pre-qualified product list. The ACTwatch Lite masterlist of antimalarials includes approved antimalarials before 1 July 2024. However, this will need to be updated to reflect the most up to date list. The user should check WHO[[2]](#footnote-3) and Global Fund[[3]](#footnote-4) websites for the most recent list.

##### 1.2.3.2: Generate a variable to identify which audited products are nationally approved

This step will require a list of nationally approved antimalarials from the appropriate regulatory body. For example, in Nigeria, this was obtained through NAFDAC.

Ideally, the nationally approved antimalarials for the study country should be added to the ACTwatch masterlist of antimalarials. There should be a column added in that dataset that indicates which are nationally approved for this country.

##### 1.2.3.3: Antimalarial true – generate

This step generates a variable that indicates if the antimalarial audited was deemed valid. This step is required for search and manual fill antimalarials

#### 1.2.4: Save clean dataset

This step ensures that a clean .dta dataset is saved that includes all updates made during step 1.2.

**Outcome:** The outcome of step 1.2 is a clean .dta dataset of antimalarial products collected during the study (one row per product). This data is critical for multiple key indicators.

### Step 1.3. Malaria blood testing data preparation and cleaning

#### Step 1.3.1. Malaria blood testing data preparation

##### 1.3.1.1: Import datasets

This step imports the primary .csv dataset into Stata and converts it to .dta format. In completing this step, be sure to manually insert the RDT audit data filename.

##### 1.3.1.2: Label Variable

Variable labels are then applied using a dedicated .do file. To run the labeling .do file, ensure the appropriate directory is set in the syntax.

The RDT audit labeling .do file requires review and may need to be updated to match the data collection tool to ensure that all variables in the audit are correctly labelled.

##### 1.3.1.3: Prepare field types

In this step, the user should manually insert string variables into the appropriate location in the .do file. Provided syntax will ensure that the text fields were imported as strings, drop missing observations, and trim and convert string text to uppercase.

##### 1.3.1.4: Merge searched and manually entered product information

In this step, searched and manually entered product information are merged into the dataset. Specific information merged include that related to RDT parasite, antigen, and manufacturer variables.

##### 1.3.1.5: Check and correct duplicates

In this step, RDT audit duplicate observations will be flagged for resolution. The user should investigate and drop true duplicates, as necessary.

##### 1.3.1.6: Correct known errors highlighted during data collection

In this step, the user will have the chance to correct known errors highlighted during data collection, typically flagged in the ‘**comments’** variables of the ODK form questionnaire.

##### 1.3.1.7: Merge outlet-level data

Outlet-level data will be merged into the RDT audit dataset in this step. The user should cross-check products in the RDT dataset with corresponding outlet records and drop outlets that do not stock RDTs.

##### 1.3.1.8: Save prepared data

This step ensures that a clean .dta dataset is saved, that includes all updates made thus far during step 1.3.1.

#### Step 1.3.2. Malaria blood testing data cleaning

##### 1.3.2.1: Check reported completeness

In this step, the user will check the dataset for reported completeness. Provided syntax will flag incomplete audits, which can be addressed and edited as required.

##### 1.3.2.2: Check comments

Any RDT audit dataset comments that require cleaning of RDT-level variables should be flagged and cleaned in the respective cleaning step. In addition to reviewing comments during analysis, supervisors should routinely check comments during data collection to address issues as they arise.

##### 1.3.2.3: Fill method

In this step, RDT audit fill method (**fillmethod**) should be reviewed, sense-checked and cleaned, as required.

##### 1.3.2.4: RDT code

This step ensures all RDT entries are coded correctly. Users should resolve any missing or inconsistent RDT codes during this step.

##### 1.3.2.5: Manufacturer and country

This step begins with recoding ‘other-specify’ country and manufacturer responses to valid ones. The validation process is required for both search and manual fill RDTs.

The user will harmonize the manufacturer responses retaining the original manufacturer variable in case it needs to be referenced in subsequent cleaning steps. This is only required for manual fill RDTs. If upon review some of the manufacturer names are actually in the search database, then the user should harmonize the manufacturer name to the exact string in the search database. The analyst should use their preferred approach for cleaning string variables in Stata (eg., **regexm**). This step can be laborious, but care must be taken to avoid mass recodes/replacements, and a product-by-product approach is advised.

Next, users must check consistency between brand name and manufacturer for manual fill RDT observations. Some brands may have more than one manufacturer. The user should sense check and edit as needed so that there is consistency within brand names

Finally, the user will harmonize country of manufacture with the manufacturer. As above, some manufacturers have more than one country of operation. The user should sense check and edit as needed so that there is consistency between manufacturer and country.

##### 1.3.2.6: Brand name

In this step, the RDT brand name variable is harmonized. The user should retain the original brand name variable in case it needs to be referenced during subsequent cleaning steps. This is only required for manual fill RDTs. The analyst should use their preferred approach for cleaning string variables in Stata (eg., **regexm**). This step can be laborious, but care must be taken to avoid mass recodes/replacements, and a product-by-product approach is advised.

After brand names are harmonized against one another, they must next be harmonized against the manufacturer.

##### 1.3.2.7: Parasite and antigen

This step involves cleaning fields related to parasite species and antigen targets of study RDTs. The user should check these values for internal consistency and reference photos and other similar observations, as necessary.

##### 1.3.2.8: Self-administered proportion

In this step the user will assess what proportion of RDTs are self-administered, sense check the data and clean as needed.

##### 1.3.2.9: Amount sold/distributed (range)

This step checks the range of RDTs sold/distributed. It is required for both search and manual fill RDTs values. Outliers will be confirmed later, in the data management syntax.

##### 1.3.2.10: Stock out (proportion)

The step assesses the proportion of stock outs, which the user should sense check and clean, as required.

##### 1.3.2.11: Retail price

This step checks the range for diagnostic retail price for adults and children. It is required for both search and manual fill RDTs. This step also assesses the take-away test retail price range for adults and children. This too is required for both search and manual fill RDTs.

##### 1.3.2.12: Wholesale price

This step checks the range for wholesale prices. It is required to determine the price per test, as respondents were asked about the number of tests purchased and the total price for all tests purchased. RDTs with a wholesale number or price of don't know (code 998) should be removed from this check. This step is required for both search and manual fill RDTs.

##### 1.3.2.13: Retail and wholesale price – outliers

In this step, retail and wholesale price outliers are assessed. Users should sense check and clean values (including checking for missing values), as required. This step is required for both search and manual fill RDTs.

#### 1.3.3: Generate quality assured RDT (QARDT) variable

This step helps the user generate and populate a **QARDT** (quality-assured RDT) variable. Syntax for preparing the RDT master list dataset (section 0.5) must be completed before this section can be completed.

##### 1.3.3.1: Generate the QARDT variable

This step will generate the **QARDT** variable, which is coded based on brand name, manufacturer name, country of manufacturer, antigen and parasite species. To successfully complete this step, be sure to have completed all of section 1.3.2.

##### 1.3.3.2: Recode searched products using QQA column in the RDT database

In this step, the quality assured RDT database is merged with the RDT datasets. For study countries with many RDTs, especially many manually entered RDTs, refer to model syntax for additional ways of assuring quality. However, provided syntax should be comprehensive of all WHO and nationally approved RDTs at a minimum.

##### 1.3.3.3: RDT true – generate

This step generates a variable that indicates if the RDT audited was deemed valid (**rdt\_true**). It is required for both search and manual fill RDTs.

#### 1.3.4: Save clean dataset

This step ensures that a clean .dta dataset is saved that includes all updates made during step 1.3.

**Outcome:** The outcome of step 1.3 is a clean .dta dataset with cleaned data for each RDT product identified during the study (one row per product). These data are essential for multiple study indicators.

### Step 1.4. Supplier data preparation and cleaning

#### Step 1.4.1. Supplier data preparation

This step involves importing, preparing, and cleaning the supplier records related to antimalarial medicines and RDTs. Provided syntax produces a dataset in long format, with one supplier per row (with repeated outlet data for outlets reporting more than one supplier). In the final list exported for field teams, true duplicates are automatically removed, but manual cleaning may be needed to identify suppliers reported by multiple outlets and therefore duplicated in the list.

##### 1.4.1.1: Import datasets

This step combines (appends) data on antimalarial and RDT suppliers (which are saved as separate csv files by the ODK form).

##### 1.4.1.2: Append supplier data

In this step, antimalarial and RDT supplier datasets are appended to one another.

##### 1.4.1.3: Label variables

In this step, a set of automatic syntax is provided to label variables.

##### 1.4.1.4: Prepare field types

In this step, the user should manually add string variable names to the local macro to be modified. Provided syntax will then automatically ensure that the text fields are imported as strings, drop missing observations, and trim and convert string text to uppercase.

##### 1.4.1.5: Check and correct duplicates

Duplicate unique IDs may occur when uploading data from tablets to the server. This step allows the user to investigate and drop true duplicates.

##### 1.4.1.6: Correct known errors highlighted during data collection

In this step, the user will address any errors noted during data collection or in the comments related to the supplier data.

##### 1.4.1.7: Merge outlet-level data

In this step, outlet-level data will be merged with the supplier dataset. Outlets without suppliers will be dropped from the dataset.

##### 1.4.1.8: Save prepared data

This step ensures that a clean .dta dataset is saved that includes all updates made during step 1.4.1.

#### Step 1.4.2. Supplier data cleaning

##### 1.4.2.1: Check and clean each variable

In this step the user will check each variable in the supplier dataset and clean as required. Variables include source (**source**), supplier name (**suppl**), supplier type (**supp2** & **supp3**), supplier location (**supp4 – supp6**), and supplier contact information (**supp7**).

##### 1.4.2.2: Save clean dataset

This step ensures that a clean .dta dataset is saved, which includes all updates made during step 1.4.2.

#### 1.4.3: Export supplier list for field team (optional)

For some ACTwatch Lite studies, supplier information reported from retail level outlets will be used to inform the field team where to capture additional data from higher levels of the supply chain during data collection. Syntax is provided in this step that can be used to export a list of reported suppliers if needed.

To complete the above, additional manual cleaning is likely required. This may include dropping suppliers reported multiple times, renaming variables for clarity and dropping non-essential variables. This step culminates in exporting a supplier list to Microsoft Excel

**Outcome**: The outcome of this .do file is a single, clean list of reported RDT and antimalarial suppliers that can be shared throughout data collection with field teams to guide additional interviews at these supplier/wholesale level outlets.

## Step 2: Data management

**Overview:** Once the datasets have been cleaned, the next step involves using cleaned data to create key variables for analysis that can be used to generate results tables for each indicator (in Step 3). This includes applying weights, generating derived variables, merging the cleaned datasets, and so forth. These steps transform the raw survey responses into a structured format (typically binary variables) suitable for producing availability, price, and market share indicators.

The overall approach to data management is to provide a reproducible, systematic approach to setting key macros, merging product data, weighting the dataset, generating variables for analysis, and conducting sensitivity analysis. This step should not involve any changes to the values of existing, clean variables – those changes, if necessary, should have been completed and documented during the cleaning process (Step 1).

**Purpose:** This step involves executing multiple Stata .do files to prepare an analytic dataset by merging datasets and constructing derived variables. The .do files in this step generate weights, indicator variables, denominators, and so forth from the cleaned data. These .do file should be modified and run once full clean datasets are available.

Note that there are multiple areas for analyst input during this step, and careful attention should be paid to ensuring the correct information is inputted and changes made to .do files, where necessary. The data management stage will output datasets for analysis. More information on each data management step .do files can be found below, in Table 6. Step. 2 data management .do file names.

Table 6. Step. 2 data management .do file names

|  |  |
| --- | --- |
| 2.1\_LONG-DATASET | Appends cleaned product datasets and merges them with outlet data to create a dataset in long format with respect to products (i.e. outlet data is repeated if >1 audited product found). |
| 2.2\_WEIGHTS | Applies sampling weights, defines strata and rural/urban variables, and creates weight variables for market share and all other indicators. More information on weighing data can be found in **Appendix 2. Sampling weights guidelines** |
| 2.3\_DENOMINATORS | Generates key binary variables that define analysis denominators based on interview status |
| 2.4\_OUTLET-CATEGORIES | Creates standardized outlet category variables for disaggregating results by outlet type. |
| 2.5\_ BLOODTEST-CATEGORIES | Identifies and creates category variables for the top 3 RDT manufacturers for disaggregating results by RDT type |
| 2.6 ANTIMALARIAL-CATEGORIES | Generates binary antimalarial category variables for common standard types |
| 2.7\_STOCKING | Generates binary stocking variables to flag antimalarial and RDT availability in each outlet |
| 2.8\_AETD | Calculates Adult Equivalent Treatment Doses for antimalarials. More information onAETDs can be found in **Appendix 3. AETD definition and calculation guidelines** |
| 2.9\_MACROS\_AND\_PRICE.do | Generates standardized retail and wholesale price variables |
| 2.10\_VOLUME.do | Computes reported sales volume variables |
| 2.11\_PROVIDER.do | Summarizes provider-level characteristics for analysis |
| 2.12\_FINALDATASETS.do | Creates final analysis-ready datasets for indicator tabulation |
| 2.13\_SENSITIVITYANALYSIS.do | Runs sensitivity checks for outlier detection |

### Step 2.1 Create a single product dataset for analysis

In this step, a single long-format analytic dataset is created for users by combining product-level and outlet-level data.

First, the user will generate a new variable, **productype**, to classify each row as a tablet-sachet-or-granule (TSG) product, non-TSG product, diagnostic, or stockout, based on product codes and other identifying fields.

Next, the user will append the cleaned RDT dataset to the cleaned antimalarial dataset. Then, merge the full set of outlet-level variables to the appended product dataset using the **outletid** variable. Note that some outlet variables may have already been merged to product datasets during the cleaning step; however, this step ensures that all outlet-level variables are included.

Review the merge results using the \_**merge** variable. All product records (antimalarials and diagnostics) must match to an outlet, while outlets without products (e.g., screened out or stocked out) are not required to match to products but should be checked and confirmed to be ineligible. Use the browse and list commands provided to review unmatched records and confirm they are expected. Address any issues and verify that the cleaning process was correctly implemented.

Next, complete some basic data management. Recode all missing values from numeric variables to system missing. Create a flag for one entry per outlet. Check that all rows (products) have been assigned a product type. Ensure that stock-outs are correctly recorded. Add any other data management steps and syntax specific to this survey, then save the long format dataset.

### Step 2.2 Apply sampling weights to the dataset

In this step, sampling weights are applied to the cleaned analytic dataset to ensure that analyses accurately reflect the survey design. First, the sample weights must be calculated in the Excel weighting tool:

|  |
| --- |
| * **The ACTwatch Lite weighting tool can be found in the toolkit folder “12 Analysis syntax\2\_MANAGEMENT”. The file name for the template is “**[**AwL Weighting Tool.xlsx**](https://psiorg.sharepoint.com/:x:/r/sites/ACTWatchLite/Shared%20Documents/2.%20Technical/0.%20Toolkit/ACTwatch%20Lite%20Toolkit%20v4/12%20Analysis%20syntax/2_MANAGEMENT/AwL%20Weighting%20Tool.xlsx?d=wc11ef357c9df4da085c7c30efec7b9ad&csf=1&web=1&e=Jwobeg)**"** * The weighting tool uses information gathered during the sample design process combined with information about how the study was actually conducted (including booster outlets/ areas, any replacement PSUs, etc), and population data at the PSU level. |

The user can then import the sampling weights from the Excel file and use Stata to convert and save them as a Stata dataset.

The user will then merge this file with the cleaned antimalarial/RDT dataset using the cluster variable (typically **c4**). Next, generate variables for sampling strata and booster sample status (if applicable), followed by indicator variables for each stratum and combined rural/urban-strata groups for disaggregated analysis.

Two weight variables are created in this step: one for market share indicators (**wt\_marketShare**) and another for all other indicators (**wt\_allOutlet**). The market share weights (for use in market share and market composition indicator calculations) exclude any outlets that were in a booster sample. The main sampling weight (for all other indicators) is in the **wt\_allOutlet** variable and should include all outlets in the study.

Finally, a finite population correction (**FPC**) factor is generated to account for the often-large proportion of PSUs sampled in an ACTwatch Lite study. This is calculated within each stratum by dividing the number of PSUs sampled in the stratum by the total number of PSUs in that stratum.

The weights, FPC and geographical unit variables are then used to declare the survey design for the dataset using the **svyset** command.

The **nOut** variable, which flags one observation per outlet – used in this long dataset for any outlet-level estimations – is regenerated to ensure one observation per outlet is correctly flagged.

### Step 2.4. Generate denominator variables

This step generates key binary variables that define the analysis denominators used throughout the results tables. These variables categorize outlets based on eligibility, screening status, and interview completion, allowing users to consistently calculate proportions and disaggregate results by relevant groups.

This syntax should not require modification, however the **finalIntStat** variable, which captures the final interview outcome for each outlet, should be reviewed and corrected, if needed.

The user will begin by loading the cleaned, weighted analytic dataset.

Next, the user will generate binary variables to indicate whether the outlet was successfully screened (**screened**), started a full interview (**interviewed**), or was fully enumerated in the dataset (**enum**). Based on these classifications, define binary categories (yes/no) for each group to simplify later analysis, such as calculating proportions or disaggregating results.

The user will conclude by saving the dataset once all variables have been created and reviewed.

### Step 2.5. Generate outlet type category variables

This step creates standardized binary variables that classify each outlet into categories used for disaggregating results in the analysis tables. The standard outlet types set for the toolkit are private facilities (for-profit and not-for-profit), pharmacies, drug shops (e.g., PPMVs in Nigeria), laboratories, wholesalers, informal vendors, and general retailers. All categories should be reviewed and edited to reflect the outlet types relevant to this study implementation.

The user will begin by loading the cleaned, weighted analytic dataset.

The user will then use the variable **outcat2** to generate binary (yes/no) variables for each outlet type. These are essential for later analysis and the choice of which outlet categories to create here should be informed by the planned disaggregations for tables and figures.

Next, the user will generate grouped categories (e.g. **for**: Formal private sector total; **inf**: Informal sector total; **prt**: All private outlets; **tot**: All retail outlets)

Note that the outlet binary variable should all be 3 letters in length, lowercase, and be assigned short but clear labels.

Finally, the user will save the dataset once all variables have been created and reviewed.

### Step 2.6 Generate blood testing category variables

This step creates manufacturer-specific binary variables for RDTs, allowing disaggregation of RDT related indicators by top manufacturer.

The user will begin by loading the cleaned, weighted analytic dataset.

Then, the user will use the **rdtmanu** variable to tabulate the most frequently reported RDT manufacturers among diagnostic products (**productype** == 3). Next the user will identify the top three manufacturers based on frequency.

Next, the user will generate a separate binary variable for each of the top three manufacturers using regexm() to match manufacturer names and label each accordingly. Afterward, the user will create additional variables for **rdtmanu\_other** (other named manufacturers not in the top three) and **rdtmanu\_dk** (products where the manufacturer is unknown)

The user must be sure to review the manufacturer spellings carefully when writing the regexm match conditions to ensure consistency and avoid misclassification. Save the dataset once all variables are generated and labeled. These variables will be used in later analysis tables related to diagnostic product availability and market share by manufacturer.

### Step 2.7 Generate antimalarial category type variables

This step generates standardized variables to classify antimalarial products for disaggregation by outlet type for core indicators.

Before running this .do file, the user should:

* Obtain and review the country’s first-line treatment guidelines.
* Review and update the WHO prequalification lists.
* Review and update the country’s list of approved antimalarials (if available).
* Define additional relevant or country-specific categories (as needed).

To complete this step, the user will load the cleaned, weighted analytic dataset.

#### 2.7.1. Generate core antimalarial classification variables

In the step, the user should start by generating core treatment categories, including any ACT (**anyACT**), artemisinin monotherapy, non-artemisinin therapies, and prophylaxis products. A broad antimalarial flag (**anyAM**) should also be created. These categories are based on the antimalarial active ingredient(s) (**gname**) and formulation.

#### 2.7.2. Generate WHO pre-qualification and national registration status indicators

Next, the user will classify ACTs by regulatory approval status – including WHO prequalification (**QAACT**) and national registration (**natapp**). Further disaggregation includes whether they are first-line treatments or non-first-line treatments according to national policy (**flact** and **naact**).

#### 2.7.3. Create dosage-specific QA variables

In this step, dosage-specific categories are created to support pediatric and adult pricing and availability analyses. Pediatric formulations are flagged using variables such as **pqaact** and **pact\_fl**, while adult formulations use **aqaact** and **aact\_fl**. These are determined based on pack size, strength and formulation but should be confirmed.

#### 2.7.4. Generate binary variables for commonly encountered antimalarials

Additional variables are generated for specific compounds that are frequently analysed across countries. These include, for example, artemether lumefantrine (**AL**)**,** artesunate amodiaquine (**ASAQ**), dihydroartemisinin piperaquine (**DHAPPQ**), artesunate (**AS**), artemether (**AR**), chloroquine (**CQ**), sulfadoxine-pyrimethamine (**SP**), SP-amodiaquine combinations (**SPAQ**), and quinine (**QN**), with further disaggregation into oral (**oralQN**) and injectable (**injQN**) forms. Oral artemisinin monotherapies (**oartmono**) are also flagged due to their relevance for monitoring regulatory compliance.

#### 2.7.5. Generate severe malaria treatment variable

Next, a variable should be created to identify the first-line treatment for severe malaria used in the country (e.g., first-line injectable or rectal artesunate or artemether), following national guidelines. The user should consult the research manager to confirm which additional or country-specific product types should be included and update the syntax accordingly.

#### 2.7.6. Customize for country-specific needs

In this step, the user will review the drug list for each country: Some classifications (e.g., mefloquine, SP, or non-artemisinin therapies) may only be relevant in specific countries. The user should consult the research manager to confirm which non-standard or country-specific classifications are needed.

#### 2.7.7. Conduct quality checks

All classification variables should be reviewed for accuracy and consistency using tabulations and spot checks. In this step, the user will cross-check these variables with active ingredients, formulations, and pack descriptions. The user should then run tab commands for all new variables, spot-check outliers or unexpected classifications using list or browse, and cross-check with **gname**, **a10**, **pack\_desc**, and other drug-identifying variables.

### Step 2.8. Generate availability (stocking) variables

This step creates outlet-level variables that indicate whether specific types of antimalarials and diagnostics were available at the time of the survey. These variables are derived from the product-level classification flags generated in earlier steps (e.g., **AL**, **AS**, **DHAPPQ**, **CQ**, **SP**) and used for calculating product availability among screened and eligible outlets.

Start by generating a flag for malaria diagnostic availability (**st\_test**), which captures whether an outlet had any form of malaria testing—either RDTs or microscopy—available.

Next, generate flags for antimalarial availability by categories defined in the previous .do file. The user should review and adapt the list of antimalarial categories used in this step based on the specific country context and product availability observed in the field (for example **st\_anyAM**: any antimalarial product; **st\_act**: any ACT; **st\_qaact**: quality-assured ACT, etc.)**.** Additional variables may be added or removed as required.

Each stocking variable is constructed at the outlet level by collapsing product audit data and checking whether at least one relevant product was available in the outlet using egen max().

After generating each stocking variable, the syntax uses tab commands for quick checks and saves the updated dataset.

### Step 2.9. Define and assign AETDs

This step calculates Adult Equivalent Treatment Doses (AETDs) for all antimalarial products in the dataset. As a reminder, AETDs are a standardized unit used throughout the ACTwatch Lite analysis using WHO-recommended dosing guidelines for a 60 kg adult. The use of AETDs allows for comparisons to be made across different types of antimalarial products that have different strengths, pack sizes and treatment regimens by establishing the number of milligrams of active ingredient required to treat a 60kg adult with malaria. AETD is required for the price and market share indicators. See **Appendix 3. AETD definition and calculation guidelines** for more information on AETDs. Users should review syntax to ensure local products and newly introduced treatments are appropriately captured.

The following five steps are used to calculate AETD for each product found in the dataset.

#### Step 2.9.1. Define the full-course treatment dose

The user should start by generating the variable **fullcourse1**, which captures the total milligram dosage required to treat an adult with each antimalarial. This is based on WHO-recommended treatment guidelines and depends on the active ingredient (**gname**). For combination therapies, the artemisinin component is typically used as the reference.

If new treatments have been introduced in the country, the user should consult relevant literature and update the syntax accordingly. Preference should be given to peer reviewed scientific literature establishing 60kg adult dosage, where possible.

#### Step 2.9.2. Adjust for presence of salt

Next, the user must adjust product strength values to account for differences in active ingredient strengths where the mg strength is provided on the package as either a base strength (i.e. excluding any salt), or overall strength with salt (e.g. 10mg of quinine sulphate is equivalent to 8mg of quinine base), or between the same active ingredient but with different salts (e.g. 10 mg of quinine sulphate = 14 mg of quinine bisulphate = 8 mg quinine base). This ensures comparability between products with the same active compound but different chemical formulations. The syntax includes logic to identify and correct for common salt-based differences. Users should scan for invalid combinations or unusually low/high strength values and flag them for review and cleaning.

#### Step 2.9.3. Calculate the number of units required for a full course

Using the total milligram dose defined in **fullcourse1**, the user should now calculate how many units of the product (e.g., tablets, mL, ampoules) are needed to deliver the full course for that active ingredient (defined by **fullcourse1**). This value is stored in the variable **fullcourse2**.

#### Step 2.9.4. Standardize product size

In this step, a new variable is created to represent the standardized product size. For tablet-based products, **newsize** represents the number of tablets needed for one AETD. For non-tablet products (e.g., liquids, injections), equivalent measures such as mL or number of sachets are used. This step ensures all products are expressed in a comparable unit for dose-based analysis.

#### Step 2.9.5. Calculate the number of AETDs per package

Finally, the user will calculate the number of full AETDs contained in each product package using the variables defined above. This output is stored in the variable **package\_AETD**. The syntax includes automated checks for missing values and implausible entries (e.g., packages with fewer than 0.25 AETDs (possibly an error) or unusually high values (e.g., more than 10 AETDs). The user should flag and review any questionable values by checking product strength, formulation, and packaging details.

#### Step 2.9.6 Final review and save

After completing the above steps, the user should review the distribution of **package\_AETD** and cross-check outliers. Ensure that all relevant products are correctly classified and that missing, or incorrectly derived values have been addressed. Once validated, this dataset—with cleaned and standardized AETD values—is saved and used in subsequent steps to calculate pricing and market share indicators.

### Step 2.10. Set local macros for CPI and exchange rates and generate price variables

First, this step prepares for analysis any price indicators that require conversion to United States dollars (USD) or comparison across time. If the current implementation does not include USD conversion or trend analysis (i.e. if ACTwatch has not been conducted in the area or previous study data are not being compared), this step may not need to be adapted or executed. Note if USD or trend analysis is not being conducted, later data management steps that calculate USD prices and use inflation rates will need to be commented out of the syntax using /\* ... \*/.

The study uses CPI (consumer price index) and historic exchange rate data to standardize and compare price indicators over time in the Step 3. Results section, below. To ensure this analysis can be replicated, the user should first document data collection start and end dates, years included in the trend analysis (i.e. years of previous ACTwatch data), CPI values for both the first survey year and each year up to the current survey year, average exchange rates, etc. All source(s) for all values should also be documented.

Once documentation is complete, set the corresponding local macros in the syntax. These macros are used in downstream price analysis.

Next, this step calculates standardized price variables for antimalarial and diagnostic products. These prices are derived in local currency and USD, and are expressed per unit, per package, and/or per AETD. These outputs are used to generate price-related indicators in the ACTwatch Lite results tables.

**Before running this step, the user should confirm that the CPI and exchange rate macros have been correctly set in Step 2.1 (cpiFIRST, cpiCURRENT, exFIRST, exCURRENT, exDC) and that the AETD values (package\_AETD) have been generated in Step 2.9 or comment out sections that are not required (e.g. if USD amounts are not relevant).**

For each product, the syntax calculates selling prices depending on the study needs. The syntax begins by validating raw price values to exclude refused, “don’t know,” or system-missing responses. Users should also confirm that the range of acceptable price values is realistic based on the study’s local currency format and data collection tool.

Once valid prices are confirmed, price variables are generated using product-level sales price data and the previously calculated AETD values.

For studies that do not include USD conversion or trend analysis, only the local currency price indicators are required.

Lastly, the script uses the CPI and exchange rate macros to convert prices into inflation-adjusted local currency and USD. These standardized price variables are then used in the results output tables for price comparisons across time, country, or product types.

Be sure to review the distributions of the price data and spot-check extreme or missing values to ensure the logic and conversions are working as intended before saving the final dataset.

### Step 2.11. Generate volume variables

This step generates standardized volume variables required for market share and sales analysis. Using data on product quantities sold or distributed, along with AETD values, the syntax calculates outlet-level and product-level indicators that reflect how much of each antimalarial or diagnostic product was sold during the study period.

The user should start by reviewing microscopy and RDT volumes (**d3** and **r13**) for outliers. The user may need to return to this step following the sensitivity analysis (Step 2.13) in which the effects of volume outliers are assessed on the market share results. It may be necessary to investigate and correct volume outliers accordingly. Users are advised to revisit this section after completing sensitivity checks to exclude or recode outlier volume values.

Next, the syntax creates two types of volume-related variables:

* Volume flag variables (**vf\_\***): Indicate whether a product or product type was reported as sold or distributed (binary 0/1).
* Volume distributed variables (**vd\_\***): Represent the actual number of AETDs sold or distributed for each relevant product.

The AETD-based volume calculations are central to market share indicators, as they allow products of different strengths, formulations, and pack sizes to be compared on an equal footing. **Users should confirm that all products used in these calculations have a valid package\_AETD value generated in Step 2.9.**

Checks should be conducted by the user to verify the completeness of volume data and to ensure no implausible or inflated values remain.

### Step 2.12. Generate provider interview variables

This step creates a set of variables from the provider interview module that reflect provider knowledge, training, and malaria case management practices. These variables are used in descriptive tables and charts. Users should adapt this syntax based on how questions are phrased in the country’s provider module and ensure that response codes used in recoding match the questionnaire and label files.

This syntax file generates variables for key indicators from the provider interview. Each variable is typically binary (0/1), with 1 representing the presence of the attribute or correct response. Variables include:

* Provider reports that ACT is the most effective treatment
* Provider recommendation of antimalarial after a negative test
* Indicator variables for recommendation behavior post-negative test
* Denominator variable for this recommendation series
* Document stated reasons for providing antimalarials after a negative test
* Determine which outlets have at least one staff member with a formal health qualification
* Determine which outlets have at least one staff member with malaria training
* Supervision, support, and surveillance variables
* Mobile phone, internet access, and digital readiness
* Licensing and registration status

The syntax includes basic tabulations for verification after each variable is generated. After creating these variables, the user should use tabulations to verify distributions and check for unexpected or missing values before saving the dataset.

### Step 2.13. Finalize and save datasets for results generation

This step finalizes the cleaned and weighted data files for analysis and indicator generation. First, the full cleaned dataset containing both antimalarial and RDT product records is opened, compressed, and saved in its final format. If supplier data was collected, it is then merged into this full dataset using the key variable, and the merged file is saved again.

Next, the dataset is used to generate an outlet-level file. This is done by keeping one observation per outlet (using the **nOut** flag), ensuring that each outlet appears only once. This outlet-level dataset is then prepared for analysis by applying svyset using the outlet-level sampling weight (wt\_allOutlet), along with the stratification and finite population correction variables.

Finally, a separate RDT-only dataset is created by filtering the full product-level dataset to include only records marked as RDTs (**rdt** == 1). Like the outlet dataset, this subset is also survey-prepared and saved for use in RDT-specific analyses.

These datasets—full, outlet-level, RDT-only, and antimalarial-only—serve as the final cleaned inputs for results generation.

|  |
| --- |
| Note:  Results are generated using final datasets in **data/2\_MANAGEMENT**  Final datasets for reference and future analysis are saved in **data/final** |

### Step 2.14. Sensitivity analyses

This step identifies and adjusts for outliers in reported volumes of antimalarial and diagnostic products. Because volume values can substantially influence market share indicators, sensitivity analyses are conducted to test the robustness of results under different assumptions. Users will apply two outlier definitions—one from the Independent Evaluation (IE) as well as a more conservative definition—and compare the resulting market share estimates. Final datasets are saved based on the preferred definition.

#### 2.14.1: Antimalarial Sensitivity Analysis

The user will start by loading the full dataset and applying the correct survey settings. As this portion of the sensitivity analysis pertains only to antimalarials, the user should then restrict the dataset to include only antimalarial products (i.e. drop **drugcat**==4).

Next, the user will apply the IE outlier definition. This entails setting outlier volume values to missing using thresholds specific to the outlet types (e.g., >1000 units for private for-profit or private not-for-profit health facilities, >300 or >200 for informal providers). Once this is complete, the dataset should be saved, indicating that it is the IE-based sensitivity data.

With the above-mentioned dataset, the user should next generate market share summary tables. In doing so, the user will compute the weighted total volume for each antimalarial subtype, by outlet type. Results will then be exported into a Microsoft Excel spreadsheet (IE\_outlier\_table) using the putexcel command. This will provide a reference for volume distribution across subtypes and help the user assess the impact of outlier removal.

It is now time to apply a second, more conservative, outlier definition. To do so, the user should reopen the IE dataset and examine all cases where volume is greater than 100 AETDs, regardless of outlet type. Each case should then be flagged as "plausible" or "improbable" based on context (e.g., formulation, number of packs, product type). Volume observations flagged as “improbable” should be set to missing and the dataset saved as the conservative sensitivity dataset.

The user will next repeat the table generation process for the conservative dataset, including the exportation of market share results into the Microsoft Excel spreadsheet for comparison. Based on their review (including the percentage differences between scenarios), the user must choose which outlier strategy to employ for the final antimalarial market share dataset.

This step concludes with the user saving a final cleaned antimalarial dataset, reopening the selected sensitivity dataset (IE or conservative), compressing it, and saving it as the antimalarial\_data file.

#### 2.14.2: Diagnostic Sensitivity Analysis

To conduct the diagnostic sensitivity analysis, the user will essentially repeat the process outlined in 2.14.1.

To start, the user will keep only one record per outlet stocking microscopy and append the RDT data, in order to create a dataset that is long with respect to diagnostic products (i.e both microscopy and RDTs). Test type (microscopy vs. RDT) should be flagged and assigned the appropriate volume variable. The user will then recode volume outliers based on IE thresholds specific to outlet types (e.g., >1000 for health facilities, >500 for pharmacies, >200 for informal vendors). Finally, the user will save a cleaned dataset for diagnostics (rdt\_micro\_data). Should the user wish to conduct a more conservative approach with the RDT volumes data, we suggest replicating the approach taken with antimalarials. In the 3 ACTwatch Lite pilots to date, RDT volumes have been very low in the private sector, and this further sensitivity analysis has not been necessary.

Successful execution of step 2.14 results in two finalized datasets – one for antimalaria products and one for diagnostics – with volume outliers treated according to a clearly documented and justified approach. These datasets will be used for market share analysis.

## Step 3. Results

**Overview:** We have, where possible, automated the analysis of key indicators for ACTwatch Lite to ensure that indicators are generated in a standard, and reproducible way.

The Step 3. Results .do files comprise two groups:

1. Two local definition .do files that provide the input for the analysis (more information in step 3.1).
2. 26 ACTwatch Lite indicator analysis .do files – one per indicator spanning commodity availability, distribution, purchasing price, wholesale price, sales volume, market share, and stockouts, as well as product, outlet, provider, and supplier characteristics. These indicators do not, generally, require any modification except for choice of disaggregation and footnotes for tables (more information in Step 3.2).

The ACTwatch indicator analysis .do files produce weighted estimates that are automatically outputted to Excel files using the **putexcel** command. If variable names have been maintained from the toolkit Stata files, and the file structure has been set up in line with the toolkit, then once the user has made the necessary choices in the two local definition .do files and selected which disaggregations to produce, the indicator analysis .do files can theoretically be run with no further input from the user.

In reality, variable names may change, and errors may occur. We have annotated all of the indicator analysis .do files for the user to not only be able to understand what is happening, but also potentially modify them as needed. The annotations are full for the first disaggregation in each indicator analysis .do file (i.e. national level), but as the indicator analysis is largely the same for each disaggregation, we only include further annotations for novel elements.

**Steps:**

### Step 3.1: Set up local definitions

There are two local definition .do files provided for this analysis – one to set up analysis for outlets, suppliers, and antimalarial products (.do file 31a\_LOCAL DEFINITIONS antimalarials) and another to set up analysis for diagnostics (.do file 3.1b\_LOCAL DEFINITIONS diagnostics). These .do files are the main area for user input in the analysis.

The analysis has been designed to use a small number of different sets of variables (for example, a standard set of outlet types for all “product availability” indicators; and a standard group of antimalaria types for prices, or sales volumes).

The local definition .do files are where the user should specify which variables are to be used in the analysis. This should be the first step following completion of data cleaning and management (Steps 0-2).

### Step 3.2: Review syntax and refine for specific study indicators and desegregations

#### 3.2.1: Determine which indicator tables to generate

Accompanying this analysis guide are .do files written for all ACTwatch Lite indicators. However, while planning the user’s implementation and designing the corresponding quantitative questionnaire, researchers may have removed indicators. Similarly, while sufficient data may have been collected, researchers may decide not to present all indicators.

Therefore, as a first step, users should determine which indicators to run tables for. These decisions should be made in collaboration with researchers and other stakeholders involved in the user’s implementation.

#### 3.2.2: Determine desired disaggregations for each indicator

In each indicator analysis .do file, the following tables are set up to be produced:

1. National (overall across all strata)
2. By urban/rural
3. By strata
4. By strata and urban/rural

However, not all study implementations will require all levels of stratification, and not all indicators need to have all disaggregations, even if they are available. Some studies may have urban/rural as the only stratification, for example. In which case only table types i and ii should be produced.

We recommend discussing which indicators should be disaggregated by urban/rural or strata with key stakeholders to support this decision making. When that is complete, the user should comment out (using /\* …\*/) or delete syntax pertaining to unnecessary disaggregations before they are run. To facilitate this process, all indicator analysis .do files have each of the four possible disaggregation levels clearly indicated.

#### 3.2.3: Footnotes for tables

All tables produced for ACTwatch Lite have footnotes providing more detail on, for example, N’s, missingness, and/or non-valid responses. Each indicator analysis .do file is set up to automatically output these footnotes to Microsoft Excel sheets.

Most footnotes are disaggregated by outlet type. If the user’s ACTwatch Lite study has outlet types that differ from those outlined in the standard toolkit syntax, the footnotes will need to be modified to reflect that change. This will require that the user make manual edits to the syntax in the indicator analysis .do files.

### Step 3.x: Run syntax for each relevant indicator

*We have labeled this step 3.x to indicate that it does not correspond to a specific .do file. Instead, it relates to any relevant indicator analysis .do file that the user, in collaboration with the research team and implementation stakeholders, has chosen to run. See below for the comprehensive list of 26 possible ACTwatch Lite indicators.*

When Steps 3.1 and 3.2 are complete, the user may run syntax for all desired indicators to produce table outputs in Microsoft Excel.

The indicator analysis loops and putexcel command in Stata can be slow to run over a large set of variables and disaggregations. We recommend testing the indicator analysis outputs with either national-only estimates, or a small number of variables (products and outlets) to begin with to avoid wasting time before finding errors.

Below is a list of each of the 26 indicator analysis .do files. To learn more about each associated indicator, numerator, denominator, value reported, potential disaggregations and source questionnaire section, access the [ACTwatch Lite Indicator Table](https://psiorg.sharepoint.com/:x:/r/sites/ACTWatchLite/_layouts/15/Doc.aspx?sourcedoc=%7B08D5E355-98E2-4581-9834-C352B4B4A572%7D&file=ACTwatch%20Lite%20Indicator%20Table.xlsx&wdLOR=cE75D2C03-D0CB-48EA-A1DF-6C2591943E7E&action=default&mobileredirect=true).

|  |
| --- |
| 1.1\_Market Composition among antimalarial-stocking outlets |
| 1.2\_Market Composition among outlets with malaria blood-testing |
| 2.1\_Availability of antimalarial types in all screened outlets |
| 2.2\_Availability of antimalarial types in all antimalarial-stocking outlets |
| 2.3\_Availability of tests among all screened outlets |
| 2.4\_Availability of tests in all antimalarial-stocking outlets |
| 3.1\_Median sales volume of antimalarial AETDs |
| 3.2\_Median sales volume of AETDs of outlets with any sales |
| 3.3\_Median sales volume of malaria blood tests |
| 3.4\_Median sales volume of tests of outlets with any sales |
| 4.1\_Market share of antimalarials |
| 4.2\_Market share of malaria blood testing overall |
| 4.3\_Market share of antimalarials by brand and manufacturer |
| 5.1a\_Sales price of AETDs to customer - USD |
| 5.1b\_Sales price of AETDs to customer - Local currency |
| 5.2a\_Sales price of pre-packaged ACTs - USD |
| 5.2b\_Sales price of pre-packaged ACTs - Local currency |
| 5.3a\_Sales price of tests to customer - USD |
| 5.3b\_Sales price of tests to customer -local currency |
| 6.1b\_Purchase price of AETDs - Local currency |
| 6.2b\_Purchase price of RDTs - Local currency |
| 7.1&2\_Stock outs of antimalarials and RDTs |
| AI.1 product quality characteristics |
| AI.2 outlet characteristics |
| AI.3 provider characteristics |
| AI.4 supplier characteristics |

### Step 3.3: Run syntax for the data flow diagram

Finally, the user can run the flow diagram .do file (3.3\_FLOW-DIAGRAM). This syntax produces tables and outputs in Microsoft Excel that show final interview status code and number of participants for each stage (number of those who were screened, those who were eligible, those who completed a full interview etc.). This syntax produces one table with results for the national level and strata.

1. National-level results: across the whole geographic reach for which the study was designed to be representative.
2. Results by other strata: if the user has other geographical stratifications (e.g., region/ state).

Note that this table is different from other table .do files as it stacks the national and stratum-level results together in a single sheet (if stratum-level results are required).

## Next steps

All automated analysis for the core indicators is outputted to preformatted excel workbooks to create tables and figures that can ultimately be copy-pasted into the final report.

Careful review of the results at this stage is essential to identify any potential errors in the cleaning or analysis and fix them and re-run the syntax as needed.

# Appendices

## Appendix 1. List of .do files

Table 7. List of .do files

|  |  |
| --- | --- |
| **Analysis step** | **File name** |
|  | ACTwatch Lite\_MASTERFILE |
| 0\_SETUP | 0.1\_MACROS |
| 0.2\_PROGRAMS |
| 0.3\_PROGRAMLISTS |
| 1\_CLEANING | 1.1\_CLEANING-OUTLET |
| 1.2\_CLEANING-ANTIMALARIALS |
| 1.3\_CLEANING-RDT |
| 1.4\_CLEANING-SUPPLIER |
| varlbl\_syntax\_amaudit |
| varlbl\_syntax\_outlet |
| varlbl\_syntax\_rdtaudit |
| 2\_MANAGEMENT | 2.1\_LONG-DATASET |
| 2.2\_WEIGHTS |
| 2.3\_DENOMINATORS |
| 2.4\_OUTLET-CATEGORIES |
| 2.5\_BLOODTEST-CATEGORIES |
| 2.6\_ANTIMALARIAL-CATEGORIES |
| 2.7\_STOCKING |
| 2.8\_AETD |
| 2.9\_MACROS\_AND\_PRICES |
| 2.10\_VOLUME |
| 2.11\_PROVIDER |
| 2.12\_FINALDATASETS |
| 2.13\_SENSITIVITYANALYSIS |
| 3\_RESULTS | 3.1a\_LOCAL\_DEFINITIONS antimalarials |
| 3.1b\_LOCAL\_DEFINITIONS diagnostics |
| 3.2\_MODIFY\_AND\_RUN\_TABLE\_SYNTAX |
| 1.1\_Market Composition among antimalarial-stocking outlets |
| 1.2\_Market Composition among outlets with malaria blood-testing |
| 2.1\_Availability of antimalarial types in all screened outlets |
| 2.2\_Availability of antimalarial types in all antimalarial-stocking outlets |
| 2.3\_Availability of tests among all screened outlets |
| 2.4\_Availability of tests in all antimalarial-stocking outlets |
| 3.1\_Median sales volume of antimalarial AETDs |
| 3.2\_Median sales volume of AETDs of outlets with any sales |
| 3.3\_Median sales volume of malaria blood tests |
| 3.4\_Median sales volume of tests of outlets with any sales |
| 4.1\_Market share of antimalarials |
| 4.2\_Market share of malaria blood testing overall |
| 4.3\_Market share of antimalarials by brand and manufacturer |
| 5.1a\_Sales price of AETDs to customer - USD |
| 5.1b\_Sales price of AETDs to customer - Local currency |
| 5.2a\_Sales price of pre-packaged ACTs - USD |
| 5.2b\_Sales price of pre-packaged ACTs - Local currency |
| 5.3a\_Sales price of tests to customer - USD |
| 5.3b\_Sales price of tests to customer -local currency |
| 6.1b\_Purchase price of AETDs - Local currency |
| 6.2b\_Purchase price of RDTs - Local currency |
| 7.1&2\_Stock outs of antimalarials and RDTs |
| AI.1 product quality characteristics |
| AI.2 outlet characteristics |
| AI.3 provider characteristics |
| AI.4 supplier characteristics |
| 3.3\_FLOW-DIAGRAM |

## Appendix 2. Sampling weights guidelines

Sampling weights should be applied when analysing ACTwatch Lite survey data to account for variations in selection probability due to the sampling design. We recommend involving a biostatistician in the sample design and weighting for the study. The sample weights should take account of:

1. Main Sample: Health areas (clusters) will be selected from sampling frames in each region using probability proportional to size. In each selected geographical unit, a census should be conducted of all outlets that are likely to sell or distribute antimalarials and/or provide malaria blood tests.
2. Booster Sample (if relevant): The geographical area for the outlet census can be expanded to include neighbouring areas, which will be randomly selected, for specific outlet types only (typically pharmacies and/or health facilities, where it is expected that these outlet types will be few in number, but for which estimates of key indicators are considered important by stakeholders)

ACTwatch Lite uses two types of sampling weights. One weight is for market share-based estimations and excludes any outlets from the booster sample (if used). The other weight (main weight) is for all other (non-market share based) estimations and includes outlets from the booster sample (if used). Where no booster sample has been done in a study, the market share and main weights will be equal for any given outlet. Where a booster sample has been included in the study, the specific outlet types that were included in the booster sample (typically health facilities and/or pharmacies) will have different values calculated for the two weights.

The sampling weights applied in the analysis are the inverse of the selection probability:

A group of numbers on a black background

AI-generated content may be incorrect.

Where:

Mα = Estimated population size of the cluster

ΣMα = Sum of estimated cluster population size across stratum

a = Number of selected clusters in the stratum

Sampling weights will be calculated at the study area (cluster) level and applied to all outlets within a given cluster, regardless of outlet type.

**Market share weights:**

Market share will be calculated using complete census data from the **primary (main) sample only** (i.e., the booster sample will not be included in the market share calculation). Main sampling weights should be determined using the sampling weight formula (Wi), where:

Mα = Estimated population size of the non-booster area (e.g. Ward)

ΣMα = Sum of estimated population size across stratum

a = Number of selected non-booster clusters in the stratum

**Main weights:**

Main weights are applied to all report indicators other than market share. If no booster sample was conducted, the main and market share weights will be identical for a given primary sampling unit (geographic area – ward, or equivalent). If a booster sample was conducted, the main and market share weights will be the same for any given non-booster outlet type in a given PSU (ward). For booster outlet types (typically pharmacies and/or private health facilities), the main sampling weight should be calculated using population data for the higher-level administrative unit from which the booster sample was taken. i.e. using the sampling weight formula (Wi), where:

Mα = Estimated population size of the higher-level administrative unit from which the booster

sample was taken (e.g. District)

ΣMα = Sum of estimated population size of the higher-level administrative unit from which the

booster sample was taken across stratum

a = Number of selected higher level administrative units from which the booster sample was

taken in the stratum

**Note**

A sampling frame based on population size is used to select the sample as generally there are no precise estimates of the total number of outlets per geographic or administrative unit that may be eligible for a market survey (i.e. including both formal and informal outlets). The key assumption in using population figures for sampling and weighting is that the distribution of outlets and the flow of malaria commodities through them are correlated with population size. The sampling approach and subsequent weighting may be adapted in cases where complete data is available on number of outlets within geographic areas.

**Finite population correction**

A finite population correction (FPC) is applied to the study estimates to account for the high proportion of health areas selected without replacement (where this exceeds 5% of the total). The FPC affects the standard errors of the estimates but does not alter the point estimates.

**Using weights and FPC**

Once calculated (during data cleaning/management) the sample weights and FPCs are saved to the dataset and used for all standard indicator estimation during analysis using Stata survey settings (svyset command). Details are provided in the accompanying Stata .do files.

## Appendix 3. AETD definition and calculation guidelines

This section outlines the definition and calculation information for AETDs used in ACTwatch Lite. Note that for the antimalarials listed in this document, standard calculations are performed automatically in the accompanying .do files for the study. These calculations will, however, require review and updates should new treatments, or new treatment guidelines be identified.

**Definition**

Antimalarial drugs are manufactured in various forms, with different active pharmaceutical ingredients, dosage forms, strengths, and pack sizes. ACTwatch Lite uses the adult equivalent treatment dose (AETD) as the standard unit for pricing and sales/distribution analyses. **An AETD is defined as the number of milligrams (mg) of an antimalarial drug required to treat an adult weighing 60 kilograms (kg).** For each generic antimalarial drug, the AETD is based on the number of mg recommended in the WHO guidelines for treating uncomplicated malaria in areas with low drug resistance. When the WHO treatment guidelines do not address a specific generic, the AETD is defined as based on any available peer-reviewed research or the treatment recommended by the product manufacturer for a 60 kg adult.

Although using AETDs may oversimplify and overlook some of the complexities of drug consumption and utilization, this approach has been chosen because it standardizes drug dosages across different drug types and countries (where dosages can vary). This standardization allows for consistent comparisons of prices and volumes based on the AETD.

Additional Considerations:

* For combination therapies containing two or more active antimalarial ingredients (co-formulated or co-blistered), the concentration of the primary active ingredient is used. For any artemisinin-containing medications, we consider the artemisinin compound to be the main ingredient for ACT AETD calculations.
* Fixed-dose combinations (FDC) are typically assumed to have a 1:1 ratio of tablets, unless otherwise specified during fieldwork or on the manufacturer’s website.
* Sulfamethoxpyrazine-pyrimethamine is assumed to provide the same full adult dose as sulfadoxine-pyrimethamine.

**Calculation**

Information on the drug concentration and unit size indicated on the product packaging is used to calculate the total quantity of each active ingredient in the package. The number of AETDs per unit is then determined. For monotherapies, the number of AETDs is calculated by dividing the total amount of active ingredients in the unit by the AETD (i.e., the total number of milligrams required to treat a 60 kg adult). For combination therapies, the number of AETDs is calculated by dividing the total quantity of the active ingredient used to determine the AETD (the primary active ingredient) by the AETD.

Table 8. Adult equivalent treatment dosage (AETD)

|  |  |  |
| --- | --- | --- |
|  | | |
| Antimalarial Generic  [Ingredient used for the dose in mg of AETD **bolded**]. | Dose used to calculate 1 AETD\* | Source |
| Amodiaquine | 1800mg | WHO model form, 2008 |
| Artemether | 960mg | WHO use of antimalarial drugs, 2001: Note: this includes a recommended loading dose of 4 mg / kg on the first day, followed by a six-day course of 2 mg / kg once daily. |
| **Artemether**-Lumefantrine | 480mg | WHO guidelines for malaria, 30 November 2024. Geneva: World Health Organization; 2024. <https://doi.org/> 10.2471/B09146. License: CC BY-NC-SA 3.0 IGO |
| **Artemisinin**-Naphthoquine | 2400mg | WHO use of antimalarial drugs, 2001 |
| **Artemisinin** - Piperaquine | 504mg | Thanh NX, Trung TN, Phong NC et al. 2012. Efficacy and tolerability of artemisinin-piperaquine (Artequick®) versus artesunate-amodiaquine (Coarsucam™) for the treatment of uncomplicated Plasmodium falciparum malaria in south-central Vietnam. Malaria Journal, 11 :217. |
| **Arterolane**- Piperaquine | 450mg | Patil C, Katare S, Baig M, Doifode S. Fixed-dose combination of arterolane and piperaquine: a new perspective in antimalarial treatment. Ann Med Health Sci Res. 2014 Jul; 4(4):466-71. doi: 10.4103/2141-9248.139270 |
| Artesunate | 960mg | WHO use of antimalarial drugs, 2001; Note: this includes a recommended loading dose of 4 mg / kg on the first day, followed by a six-day course of 2 mg / kg once daily. |
| **Artesunate**-Amodiaquine | 600mg | WHO guidelines for malaria, 30 November 2024. Geneva: World Health Organization; 2024. <https://doi.org/> 10.2471/B09146. License: CC BY-NC-SA 3.0 IGO |
| **Artesunate-**Mefloquine | 600mg | WHO guidelines for malaria, 30 November 2024. Geneva: World Health Organization; 2024. <https://doi.org/> 10.2471/B09146. License: CC BY-NC-SA 3.0 IGO |
| **Artesunate**- Sulfadoxine-Pyrimethamine | 600mg | WHO guidelines for malaria, 30 November 2024. Geneva: World Health Organization; 2024. <https://doi.org/> 10.2471/B09146. License: CC BY-NC-SA 3.0 IGO |
| **Atovaquone**-Proguanil | 3000mg | WHO guidelines for malaria, 30 November 2024. Geneva: World Health Organization; 2024. <https://doi.org/> 10.2471/B09146. License: CC BY-NC-SA 3.0 IGO |
| Chloroquine | 1500mg | WHO guidelines for malaria, 30 November 2024. Geneva: World Health Organization; 2024. <https://doi.org/> 10.2471/B09146. License: CC BY-NC-SA 3.0 IGO |
| **Dihydroartemisinin-**piperaquine | 360mg | WHO guidelines for malaria, 30 November 2024. Geneva: World Health Organization; 2024. <https://doi.org/> 10.2471/B09146. License: CC BY-NC-SA 3.0 IGO; Note: AETD under the new 2015 guidelines is now 480 mg, whereas 360 mg was indicated in the previous guidelines. Product availability for prepackaged adult DHA PPQ in 2015 was still most often 360 mg administered over a 3-day cycle on a total of 9 tablets (40/320). |
| **Dihydroartemisinin-**Piperazine-Trimethoprim | 256mg | Manufacturer's guidelines  (Artecxin - Medicare Pharma; Artecom - Ctonghe) |
| **Dihydroartemisinin**-sulfadoxine-pyrimethamine | 360mg | Manufacturer's guidelines  (Dalasin - Adams Pharma) |
| Hydroxychloroquine | 2000mg | Manufacturer's guidelines  (Plaquenil - Sanofi Aventis) |
| Mefloquine | 1000mg | WHO model form, 2008 |
| Quinine | 10408mg | WHO model form, 2008 |
| Sulfadoxine-pyrimethamine | 1500mg | WHO model form, 2008 |
| **Artesunate**- pyronaridine | 600mg | WHO guidelines for malaria, 30 November 2024. Geneva: World Health Organization; 2024. <https://doi.org/> 10.2471/B09146. License: CC BY-NC-SA 3.0 IGO |

*\*mg required to treat a 60 kg adult*

1. While the ACTwatch Lite toolkit is server-agnostic, researchers used SurveyCTO during ACTwatch Lite piloting to host questionnaires administered by the data collection teams on Android tablets. Data from submitted forms were then stored on the SurveyCTO server. These data were then downloaded in .csv format. Data were then imported into Stata [↑](#footnote-ref-2)
2. <https://extranet.who.int/prequal/medicines> [↑](#footnote-ref-3)
3. <https://www.theglobalfund.org/en/sourcing-management/quality-assurance/medicines/> [↑](#footnote-ref-4)