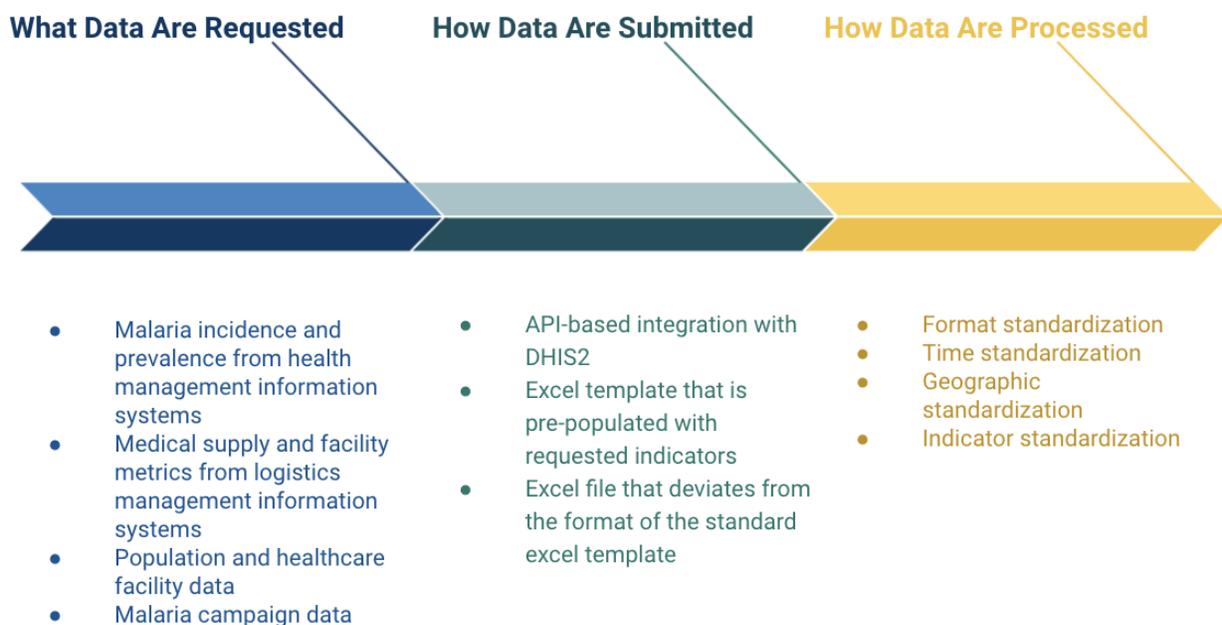# Malaria Quarterly Report Data Pipeline Overview: How Quarterly Report Data Are Processed in M-DIVE

Updated July 2022

This document provides a high-level introduction to the PMI Malaria Quarterly Report data set, and how it is processed on M-DIVE. Where applicable, we link to more detailed documentation.

The Malaria Quarterly Report (QR) is the primary data resource in M-DIVE for understanding malaria incidence, burden, and progress metrics in PMI-supported countries. It includes data provided by countries each quarter, based on what they have available from their existing data collection systems. This overview document includes:

1. What data are requested for submission and made available as part of the QR process
2. How data are submitted from its various source systems
3. What processing steps are run in M-DIVE to combine the data into a single resource

What Data Are Requested

- Malaria incidence and prevalence from health management information systems
- Medical supply and facility metrics from logistics management information systems
- Population and healthcare facility data
- Malaria campaign data

How Data Are Submitted

- API-based integration with DHIS2
- Excel template that is pre-populated with requested indicators
- Excel file that deviates from the format of the standard excel template

How Data Are Processed

- Format standardization
- Time standardization
- Geographic standardization
- Indicator standardization

1. What Data Are Requested for Submission

Each quarter, PMI requests that member countries submit key data elements that are used to calculate:

- Malaria incidence and prevalence using data from health management information systems (HMIS)
- Medical supply and facility metrics from logistics management information systems (LMIS)
- Population, community, and healthcare facility indicators
- Malaria campaign indicators (Indoor Residual Spraying, Seasonal Malaria Chemoprevention, etc.)

The submission of these requested indicators is referred to in M-DIVE as QR Data. The data reported in M-DIVE is aligned with the country's geographical or administrative units, which allows the combination of QR data with other types of data, such as survey, MOP planning, and climate data, as reflected in the Malaria Quarterly Report or other sections of M-DIVE. However, these other types of data will usually have their own submission, review, and/or import process, separate from the quarterly submission of data by partner countries.

2. How Data Are Submitted

Data are submitted in one of three (table based) forms described below. See here for detailed information describing how QR data should be submitted and standard indicator definitions.

1. A **direct database push connection** between the country data system and M-DIVE. For example, M-DIVE supports an Application Programming Interface (API) based integration with DHIS2 that countries can use to submit their data directly in an automated manner. *This is the preferred method for data submission, as it preserves data in the most consistent and reliable format.*
2. Data provided in a **PMI-provided Excel template** or entered into **a customized excel worksheet,** which is structured and pre-populated with the reporting time period, expected administrative subdivisions, and requested indicators on malaria incidence and medical facility logistics for each country.
3. Data provided in **a spreadsheet or other file** as exported from the data source system, but that does not directly match either the API data format or the specific PMI submission template.

In some cases, countries submit data through a combination of the above formats. For example, DHIS2 systems may not contain LMIS data, so a country may use the direct database connection for their HMIS indicators and a supplemental Excel template submission for LMIS indicators.

3. How Data Are Processed

Because the data are coming from many different countries and data systems, which have been developed and customized for various needs, the data need to be processed and standardized in adherence to the metadata format before it can be presented in a consistent way within M-DIVE. To do this, the submitted data goes through our M-DIVE QR Data Pipeline, which consists of the following steps to standardize:

1. Format standardization
2. Time standardization
3. Geographic standardization
4. Indicator standardization

Format standardization: Data may be stored in many different formats in the different source data systems, or submitted in different Excel table formats. The first step is to standardize the structure of all submitted data so that it follows a consistent format for all countries, indicators, and time periods. Specifically, the data must first be converted from many different sheets or columns layout so that there is one unique row for each geographic area, time period, and indicator provided.

Time standardization: Almost all QR indicators are expected to be reported in monthly time periods. In some cases, the source data may track these values at different time intervals, such as weekly or bimonthly. The M-DIVE QR pipeline will transform some of these data into monthly format, when it is expected that the result would still be reasonably accurate and representative. Depending on the level of reporting, indicators are time standardized by aggregating or disaggregating data through one of three operations:

- Summed or divided (e.g. weekly data are combined into monthly data)
- Replicated (e.g. quarterly data are copied into each month instead of divided across months of the quarter. Population data, which is submitted at the annual level, is copied into each month instead of divided across months in the year.)
- Omitted (e.g. semi-annual and annual data are not divided evenly across 6 or 12 months).

Geographic standardization: Data systems may have different names or definitions for the geographical areas they represent, especially when working with smaller regions that may or may not match to the official administrative subdivisions. Through a process called "Admin Alignment", we match the different geographic locations/names submitted to a list of standard geographies recognized in M-DIVE. This ensures that data from different sources can be aligned and drawn on maps with the same boundaries, even when there are differences in spelling and naming conventions between sources.

Indicator standardization: PMI has identified a list of key indicators that are requested and collected through the QR process. Some of these indicators are already very common and

precisely defined across data systems, but others have important nuances and variations across sources and countries. In some cases, the same indicator may be submitted under a different name, interpretation, definition, and category. In other cases, indicators might be subdivided or disaggregated to a different level than the current standard version, such as when confirmed cases are reported separately based on demographic strata. To be able to compare indicators across different countries and time periods, source data indicators are renamed to use the M-DIVE standard label, or combined as necessary to match the most consistently available version.

If you have any questions, please submit them to Alaine Knipes (wpv8@cdc.gov) and Misun Choi (mchoi@usaid.gov) on the PMI Surveillance & Informatics team, or to support@civisanalytics.com. If you would like to learn more about additional topics on M-DIVE, please click the links below for further reading.

Important Resources

- Quarterly Report Data Submission Instructions & Indicator Definitions
- Malaria Quarterly Report Quality Control Processes Overview in M-DIVE
- Malaria Quarterly Report Time Standardization
- Malaria Quarterly Report Indicator Standardization
- Malaria Quarterly Report Format Standardization
- Malaria Quarterly Report Geographic Standardization
- Malaria Quarterly Report (M-DIVE access required)
- M-DIVE Help Center (M-DIVE access required)