PEPFAR
U.S. President's Emergency Plan for AIDS Relief

Data for Accountability, Transparency and Impact Monitoring (DATIM)

MER Data Import Reference Guide Version 2

December 2018

U.S. Department of State
U.S. Office of Global AIDS Coordinator (OGAC)

Table of Contents

1   Document Scope

This guide will provide guidance to Data for Accountability, Transparency and Impact Monitoring (DATIM) data importers and explain key concepts in more detail. This guide will refer to detailed technical reference materials, but it will not provide detailed step-by-step instructions on how to transform data from an external system into a format suitable for import into DATIM due to the wide variability and differences in use cases. The basic principles of preparing data for import into DATIM will be the same, however, regardless of the details of the source system.

Currently, data submitted need to be validated first by DATIM Support (https://datim.zendesk.com/hc/en-us) prior to import. Any errors in the data can lead to delays. Ideally, data submitted to DATIM Support will have been tested thoroughly for both formatting as well as content errors, prior to submission. Data should be submitted to DATIM Support in advance of any data entry close reporting deadline to provide enough time for validation, resolution of any problems, and import testing. Please refer to the PEPFAR Data Calendar (https://datim.zendesk.com/hc/en-us/articles/115001940503-PEPFAR-Data-Calendar) for the exact date listed under the Import Deadline by reporting quarter.

Important Note: Data must be reported into DATIM regardless of reporting method i.e. data import or manual data entry.

> **Manual Data Entry in DATIM**
> Most users of DATIM enter data via the data entry user interface. Details for how to perform manual data entry can be found on DATIM Support (https://datim.zendesk.com/hc/en-us/sections/203333523-Login-Navigation-). The DATIM user interface enforces meta-data and validation rules.

2   Characteristics of DATIM MER Data

Data must be valid with respect to the destination sites, periods, indicators, disaggregation, and mechanisms. Therefore, all data values must be associated with four metadata dimensions, which describe different aspects of the data, as follows:

- **Where—**Organization unit (OU): This dimension describes the location of the data (i.e., where the data are captured).
  - Example: clinical facility, community site, or OU level
- **What—**Data element (indicator) + disaggregation: This dimension describes the phenomena to which the data value is attached, such as the Number of HTC Tests for Females 1-4. In this case, the data element has a disaggregation of Female 1-4. Data elements in DATIM typically have many different disaggregations with respect to age and gender. In addition, different classes of data elements exist, such as Targets, Results, DSD, and TA.
- **When—**Period: This dimension describes the time period of the data being reported.
  - Example: January through March 2016 or April through June 2016
- **Who—**Funding mechanism: Also known as the *"attribute option combination,"* this is an extra custom dimension in DATIM that describes for which implementing mechanism the data are being reported and refers to the Foreign Assistance Coordination and Tracking System (FACTS Info) funding mechanism ID.

All data in DATIM is associated with these four dimensions.

3   Roles and Responsibilities

Stakeholders	Roles and Responsibilities
S/GAC PRIME Team Lead	Provides approval for import data files to be imported into DATIM. Approval to import data files is only provided after an import data file has met all DATIM requirements (i.e., file submission to DATIM Support, successful completion of all validation and testing, and receipt of HQ SI approval).
DATIM Support Team	The DATIM team has developed a validation script to perform system validation on MER data. On receiving a file for import, the DATIM Support Team will run the validations scripts on the import file prior to attempting any import. If any issues are detected, the file will not be imported, and the participating agency will be notified of validation findings for the purpose of correction, re-validation, re- testing and re-submission of an updated file before submitting again to the DATIM Support Team. The DATIM Support Team is responsible for providing instructions on how to access and execute the validation scripts. The DATIM support team is available to provide assistance throughout this process. Stakeholders may contact the DATIM Support Team via the DATIM Support app in DATIM.
U.S. President's Emergency Plan for AIDS Relief (PEPFAR) Country Team	The PEPFAR Country Team is responsible for submitting the request—accompanied by evidence of successful MER data import in the data exchange testing environment (DEV-DE)—to the DATIM Support Team for MER file import into DATIM. Prior to that submission, these representatives are responsible for coordinating and managing their teams through the process of creating a properly formatted CSV import file, running the import file through the validation script, testing the import using DEV-DE.datim.org, and capturing results of the validation and test import **Recommended Pre-Import Activities** We recommend that the PEPFAR Country Team undertake pre-import activities to develop processes and create documentation to support ongoing data import, including: • Ensuring concurrence from the Inter-Agency, Agency, or Implementing Partner (IP) Lead • Structured coordination of importing efforts • Resources for the process (time, expertise, financial, human) • Documentation for data mapping, instructions, or rationale for mapping • Process for test file creation • Data import test plan and results • Country data exchange process flow diagram • Long-term data import/exchange strategy for phased implementation. • Ongoing documentation of challenges and lessons learned We recommend that time be planned for corrective action, planning, and updating these processes because they are considered iterative.

	The documents created and used by countries for planning and implementing data exchange will serve as valuable resources for other countries. Countries with well-documented and set processes should consider sharing their resources.
	Note: The current policy requires each OU to submit a single data import file. This means that if multiple implementing partners of an OU plan to use DATIM data import services, they must coordinate with their country U.S. Government (USG) points of contact (POCs) to create a single consolidated data import file for submission to DATIM Support. The country USG POCs should submit a DATIM Support ticket to seek technical support from the DATIM data import team. The in-country POC should submit the final consolidated data import file to DATIM Support.
HQ SI Advisor	Prior to importing the MER data import file submitted by the PEPFAR Country Team, the DATIM Support Team will request approval to import the file from the HQ SI Advisor. Upon receiving this approval, the file will be imported into DATIM.
Implementing Partners	IPs seeking to import data must express their interest to their in-country SI contacts and see their consent. After consent is given, IPs must work with their own developers, or others familiar with data import. Please note that an IP can complete the testing and validation of data import files individually, but after completing the validation tests, the importer should coordinate with the USG POC to create a single consolidated data import file for the concerned OU.

4   Data Import into DATIM

Data import is the process of importing appropriately formatted files into DATIM using the Import-Export application; only the DATIM Support Team has access to this app in DATIM. The graphic below illustrates the steps necessary to perform a successful MER data import into DATIM.
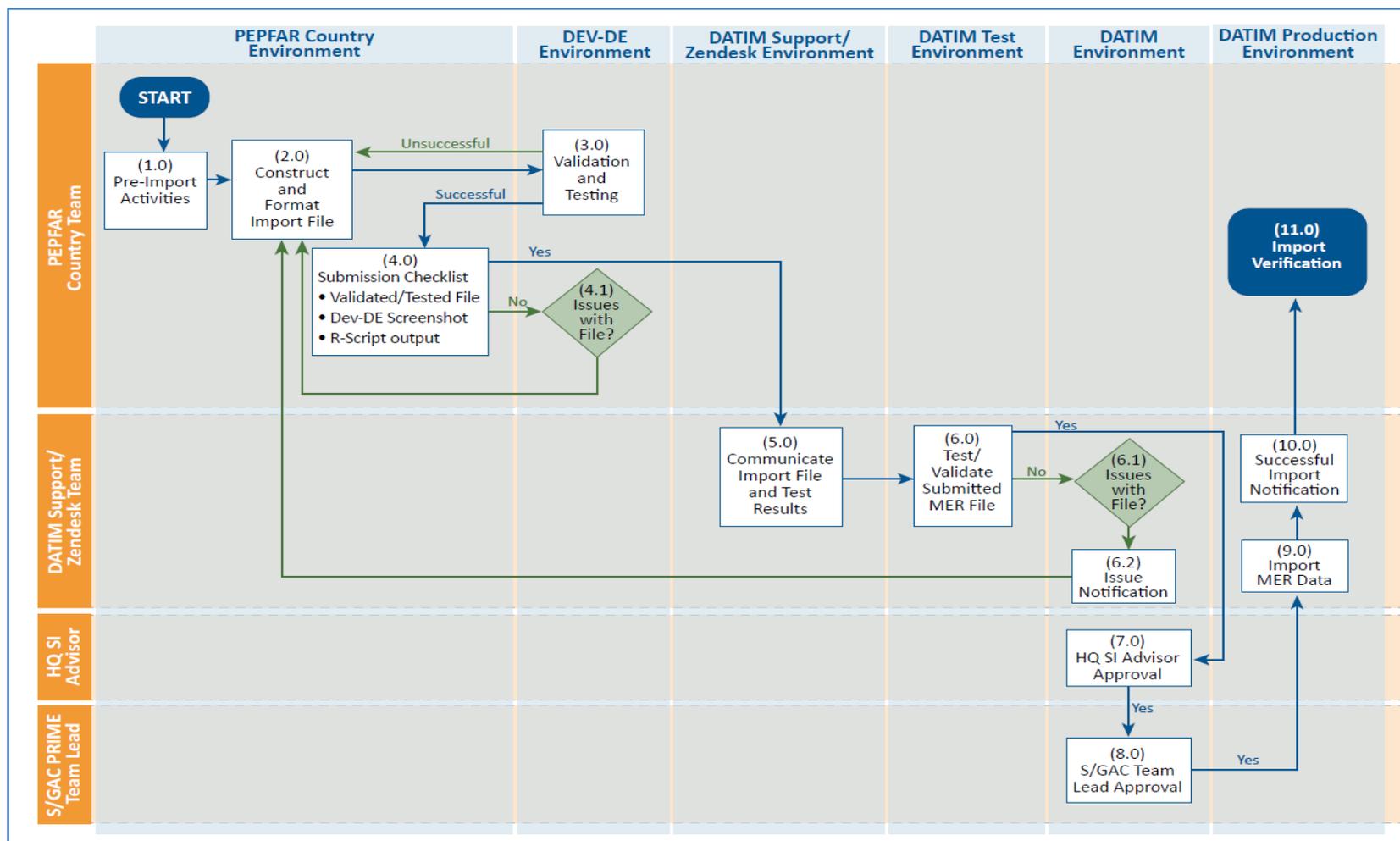
A detailed description of these steps provided in the next sections.

4.1 Constructing Appropriately Formatted Files for Import File

A correctly formatted .CSV file containing MER data should be generated. Only the following columns are required in the CSV file (order is important):

Dataelement	Period	OrgUnit	CategoryOptionCombo	AttributeOptionCombo	Value

The header in the CSV file is optional. DATIM does not make assumption that there is a header and verifies the first line of the document. If it the first line is a header, an error will be included, saying that "dataelement" (name of the first column header) is not a valid data element and can be ignored. Alternatively, the header line can be omitted from the CSV file.

4.1.1 Data Elements and Category Option Combos/Disaggregations

MER data in DATIM must have a data element and disaggregation (also known as a category option combination). Data element and category option combos/disaggregations can be reported as either code or UID (Unique Identifier). Please refer to the codelist on the DATIM Support page (https://datim.zendesk.com/hc/en-us/articles/115002334246-DATIM-Data-Import-and-Exchange-Resources) for the appropriate dataelement and CategoryOptionCombos/Disaggregations codes and UIDS.

Many data elements have multiple levels of disaggregation (age, gender, service type, etc.). DATIM also contains a number of legacy data elements, which are present in the system but may not currently be active for data entry. Please make sure to reference the correct codelist for the corresponding reporting period.

The dataelement and CategoryOptionCombos/ Disaggregations codes and UIDS are organized on the DATIM Support page by reporting period (year/version), type (results, targets, narratives, SIMS, etc.), dataset type (e.g., community, facility), and the code list in various formats (HTML, JSON, CSV and XML).

Note: The ultimate reference for this are the data entry screens of DATIM itself. If there are any invalid data elements or disaggregations present in the import file, a conflict will be reported during testing. Code lists have been made available to help you to develop mapping files for data elements and disaggregations below. In addition, validation rules will validate the data for data elements that have an option set, and if the value provided does not match a valid option, the data will be considered to be invalid.

If you unsure about which version of the code lists to use for your data, contact the DATIM Support Team for further guidance.

4.1.2 Periods

Each data value in DATIM must have a valid period. MER results are currently reported on a quarterly basis, MER targets are reported once a financial year. All periods refer to calendar periods. The time period refers to when the data are captured. DATIM uses the DHIS2 internal date format for designating reporting periods. All periods in DATIM data exchange files must provide a valid DHIS2 period identifier. In particular, DATIM uses financial year October for country operating plan (COP) targets and PEPFAR quarters for MER results. Thus, Jan–March 2015 should be entered as "2015Q1," and October 2016–September 2017 should be entered as "2016Oct."

4.1.3 Organization Units

OU is typically the site where the data are captured. Examples are site, facility, community, district, and county. This should not to be confused with PEPFAR OU. Facility, site, community, and OU codes and UIDs are available in the current code list available on the DATIM Support page (https://datim.zendesk.com/hc/en-us/articles/115002334246-DATIM-Data-Import-and-Exchange-Resources).

N**OTE:** To access these codes, users must be logged into DATIM Support because this is a password-protected resource.

4.1.4 Attribute Option Combo

This refers to the funding mechanism for which the data apply. An example is "25147," which refers to the FACTS Info mechanism ID.

FACTS Info is the source for the DATIM implementing mechanism. Data from FACTS Info are regularly synchronized with DATIM. All mechanisms that are present in a given data import file must be active for the period for which data are being imported. After a mechanism is "promoted" or made active in FACTS Info, it will take up to 24 hours for the data to be updated in DATIM. Please ensure that all mechanisms for which you will be reporting data are active in FACTS Info prior to attempting to test any data in DEV-DE. If any mechanisms are not active or data need to be reported against mechanisms to which the user does not have access, validation errors will occur.

The "Agreement End Date" attribute in FACTS Info may inhibit a mechanism being visible or available in DATIM. The date must be valid for data to be entered appropriately. **NOTE:** IPs DO NOT have access to this system. Funding agencies should provide the mechanism ID to partners interested in data exchange.

Funding mechanisms can be reported using code or UID should be reported. Funding mechanism UIDs and code are available at the DATIM Support page (https://datim.zendesk.com/hc/en-us/articles/115002334246-DATIM-Data-Import-and-Exchange-Resources). If codes are used for the funding mechanism, in the More Options section of the data import interface, select "Code" for "ID scheme (all objects)" and "UID" for "Org unit ID scheme."

Mechanism *

* These code lists are **open to the public** i.e. you **do not** need a DATIM log in to access them

Concepts	HTML	JSON	CSV	XML
Mechanism Attribute Combo Option UIDs	HTML	JSON	CSV	XML

Back to the top.

Mechanisms partners agencies OUS Start End

View of mechanisms, partners, agencies, OUs and start and end dates for each mechanism.

mechanism	code	uid	partner	primeid	agency	ou	startdate	enddate
00000 De-duplication adjustment	00000	X8hrDf6bLDC	Dedupe adjustments Partner		Dedupe adjustments Agency			
00001 De-duplication adjustment (DSD-TA)	00001	YGT1o7UxfFu	Dedupe adjustments Partner		Dedupe adjustments Agency			
00100 - PEPFAR-MOH align: MOH Data	00100	QCJpv5aDCJU				Global		
00200 - PEPFAR-MOH align: PEPFAR Data	00200	TRX0yuTsJA9				Global		
10000 - HAIVN	10000	ffafhQZ4umI	Harvard Medical School of AIDS Initiative in Vietnam	8927	HHS/CDC	Vietnam	2013-09-01	2019-09-30
10001 - VINACAP Follow-on	10001	yu7T7qTxIV4	TBD	2997	HHS/CDC	Vietnam	2013-09-01	2014-09-30
10002 - ASPH Fellowship Program	10002	sta4zq1waoe	Association of Schools of Public Health	3733	HHS/CDC	Vietnam	2014-09-30	2014-09-30
10004 - APHL's Partnership with HHS/CDC to assist PEPFAR Build Quality Laboratory Capacity_799	10004	c777je5vxxV	Association of Public Health Laboratories	1923	HHS/CDC	Nigeria	2013-09-01	2019-09-30
10015 - National Blood Transfusion Service (NBTS)_997	10015	MM7T199ppfl	National Blood Transfusion Service of Nigeria	8851	HHS/CDC	Nigeria	2013-09-01	2019-09-30
10019 - Safe Blood for Africa Foundation	10019	wd6llybHwVn	Safe Blood for Africa Foundation	238	HHS/CDC	Nigeria	2013-09-01	2019-09-30
10044 - MUHAS-SPH	10044	qe9Ke06Xp1m	Muhimbili University College of Health Sciences	1716	HHS/CDC	Tanzania	2014-09-30	2014-09-30
10074 - Partnership with HHS/CDC to Assist PEPFAR-built Quality Laboratory	10074	tnUPdgffjDN	Association of Public Health Laboratories	1923	HHS/CDC	Caribbean Region	2013-09-01	2014-09-30

4.1.5 Value

This is the data value being reported. It can be numeric or text. Narratives are typically reported as text within the value field.

Note: If there are any double quotes in the value field, these double quotes must be escaped to avoid issues with the CSV file format. The entire reported value should be enclosed in double quotes, and there should be a double quote added to escape the double quote in the value.
For example, in the case of a Narrative being reported in the value field, if your narrative reads: This is a "narrative" with a double quote, the double quotes must be escaped using the following method:
"This is a ""narrative"" with a double quote."
Additional guidance on double quotes in CSV files can be found in CSV Memo.

4.2 Data Import Validation and Testing

All MER data import files must be validated using (1) R scripts and (2) tested using DATIM's data exchange testing (DEV-DE) environment.

Important Note: The PEPFAR Country Team will need to perform validation and testing prior to submitting the MER data import file to DATIM Support and provide evidence of completion of these two steps in the submission to DATIM Support.

Validation and testing will perform checks for the following:

- Valid data set assignments for OUs (e.g., so that MER facility data are not imported into the community level)
- Active mechanisms
- Invalid data element/disaggregate combinations based on the current data entry form by dataset
- Duplicated data (OU/DE/DISAGG/PERIOD)

The above list not exhaustive. For a complete list, follow the steps outlined in 4.2.1 and 4.2.2 to perform hands-on testing and validation on a test MER data file.

4.2.1 Testing In Dev-DE

DEV-DE is a dedicated server and is a copy of the actual DATIM production system that can be used for testing purposes. It is simply a testing environment that can be used to test data import files prior to submission for import into the actual production system, without affecting the data stored in production.

Please note the following while testing in the DEV-DE environment:

- DO NOT conduct any testing in DATIM PROD.
- User credentials will be the same in DEV-DE as they are in DATIM.
- Data should only be submitted to the DATIM Support Team after you have successfully imported MER data into DEV-DE.
- Submitting data that have not been able to be imported into DEV-DE will lead to delays of your import.
- Although the exact number of values will vary based on the actual import, when importing data with the Import/Export module of DATIM, no conflicts should be reported (similar to Figure 2). If you receive any conflicts from the server, they should be resolved prior to submitting your data to the DATIM Support for further validation.
- If code is used when importing the file, expand "More options" and select "Code" as the "Data element ID scheme."

Use the Data Import/Export app in Dev-DE to conduct this testing.

4.2.2. Validation Rules and R Scripts

The DATIM development team has created a set of scripts in the R programming language that can be used to validate data outside the DEV-DE environment. The validation scripts and instructions on how to set up and use them are at: https://github.com/jason-p-pickering/datim-validation. The PEPFAR Country Team should perform validations testing using the validation scripts provided by DATIM.

Below is an example of the validation error output from the running this script on a MER file; the validation script user will need to add code (e.g., write.csv (vr_violations,file="my_violations.csv")) to the end of the script to save the output to a CSV file.

Data validation rules define how data values of different data elements should relate to each other. As a simple example, the number of positive HTC tests should not exceed the number of HTC tests. Validation rules are evaluated for a given OU – period – attribute option combination. Thus, data are not validated for two different periods or for two different mechanisms. Data validation rules provide additional restrictions and quality checks on the data that can be imported into DATIM. Users should be careful in choosing the correct period for validation and restrict themselves to a relatively small number of OUs to analyze data for at the same time. Attempting to validate too many sites over multiple time periods may result in a very long processing time.

5   Submit the MER Data File for Import into DATIM

Once validation and testing are successfully completed, the submission process begins. The following is a submission checklist for items required for submission of the MER data file for import into DATIM:

- ✓ A validated and tested MER data file
    - ✓ Include the format of the file (CSV, JSON, or DXF); CSV is the preferred format.
    - ✓ Include the coding scheme that has been used (UIDs or codes)
- ✓ Screenshots and outputs from successfully running validation scripts and testing in Dev-DE of the MER import data file
    - ✓ The screenshot should show output from the DEV-DE environment, indicating that the data were imported **without any conflicts** reported from the server.
- ✓ (Optional) Include any reports of validation rule violations. If validation rule violations should be ignored, please include a statement of justification why.

After you have the items listed above, proceed with communicating and submitting the MER file for import into DATIM.

5.1   Communicate Data Import File and Test Results to the DATIM Support Team

Submission of MER data files for import into DATIM Prod (datim.org) is done via DATIM Support. Please take the following steps to request import of your MER data into the DATIM production environment (datim.org).
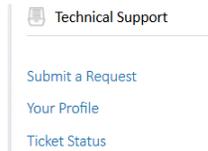
As a reminder, the DATIM Support Team may be contacted at any time during this process if you have any questions.

1. Navigate to DATIM.org and select the "DATIM Support" icon found under "Apps" in DATIM.
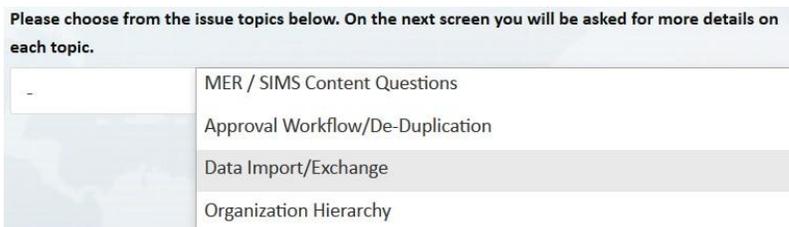
Datim Support

2. Select "Submit a Request," found under the Technical Support section.

Technical Support

Submit a Request
Your Profile
Ticket Status

3. This will take users to a new page where they can select a topic. From the dropdown menu, select "Data Import/Exchange."

Please choose from the issue topics below. On the next screen you will be asked for more details on each topic.

_	MER / SIMS Content Questions
	Approval Workflow/De-Duplication
	Data Import/Exchange
	Organization Hierarchy

4. After selecting "Data Import/Exchange," additional questions will appear. Answer them accordingly, and be sure to complete the mandatory fields or the request cannot be submitted. Under "Description," indicate the plan for importing data into DATIM and include a screenshot of the output from the DEV-DE environment, which shows that the data were able to be imported with no conflicts reported from the server and the outputs from running the validation scripts on the import file.

Import summary

Import count

Type	Count
Imported	202477
Updated	0
Ignored	1

Conflicts

Element	Description
DE_DSD	Data element not found or not acccessible

Figure 1: Example of Data Import Summary

5. Click on "Submit" when ready to send your request to DATIM Support.

5.2   MER Data Import into DATIM (performed exclusively by the DATIM Support Team)

Upon receiving your submission through the submitted DATIM support ticket, the DATIM Support Team will follow these steps:

1. Test and validate the submitted MER data file using the same R Scripts and server (DEV-DE). If any issues are detected, the file **will not** be imported.

    a. The PEPFAR Country Team will be notified of testing validation issues and will need to take corrective action to address validation issues flagged by the DATIM Support Team.

    b. After the PEPFAR Country Team has correct these issues, it will need to repeat the above process with the newly corrected import file (i.e., re-validate, re-test, and re-submit).

2. After all validation and testing issues have been verified by the DATIM Support Team as fixed, the file will be ready for import to DATIM production. The import into DATIM production will be performed by the DATIM Support Team.

3. The DATIM Support Team will seek approval from both HQ SI advisor and S/GAC Team Lead to proceed with importing the MER data file.

4. Upon receipt of approval from the HQ SI Advisor and the S/GAC Team Lead, the DATIM Support Team will import the MER data file into DATIM.

5.2.1   Verifying Import into DATIM

The DATIM Support Team will inform the PEPFAR Country Team after the import is complete. The PEPFAR Country Team can view the data using Analytics app in DATIM (may need to wait one hour to see the data in Analytics).

The DATIM Support ticket can be closed after verification that the data have been imported.

6   Post-Import Activities

6.1   Move Imported Data through DATIM Approval Workflow

Data imported into DATIM must still follow the required approval workflow. When data are imported into the production system, they will be in the "Pending" state. IPs should submit their data to the agency, which can then approve and submit the data to inter-agency space. DATIM Support can also help you to move the data that have been imported in bulk up the approval chain. If requested, this operation should be supported with approval from the USG PCO, authorizing DATIM Support to approve the data in bulk. Clear instructions should be provided to which level the data should be submitted.

6.2   Data Deduplication

Once data has been imported and approved to inter-agency level, if there are any duplicates, they will also need to be resolved by the PEPFAR country team. In terms of planning, it is important to consider that the de-duplication process can be very lengthy depending on the number of duplicates.

6.3   Data Import Cleaning Guidance

Please read the below guidance for each of the data import cleaning options. Note that you are required to submit separate sets of files for each quarter, preferably in two separate DATIM Support tickets.

6.3.1   Preferred Approach for Data Updates: Submission of New/Updates + Delete File

If you are submitting updated data or new data, you should submit them as a single file. Clearly indicate in the name of the file that this is intended for updates. Use the following naming convention for this file:

New data: Country_YYYYQnimport

Updated data: PEPFAR_Q1_DataImport_Update_06232016.

If you intended to delete data, and know what data you want to delete, you should prepare another file with the data values that should be deleted, clearly indicating in the file name that these data should be deleted. Use the following naming convention for this file:

PEPFAR_Q1_DataImport_DeleteData_06232016

If you are able to submit data using this option and the data has been validated at your end, prior to submission, with the DATIM validation scripts, DATIM Engineering anticipates that this operation should be able to be completed within 48 hours of submission.

6.3.2 Bulk Removal of Data

If you are unable to provide a file that contains values that should be completely deleted from the database, or if you intended to submit a complete replacement of your data, you need to request at least three weeks in advance of the submission of your completed data file, that DATIM Engineering remove your data from the database. Use the following naming convention for this file:

PEPFAR_Q1_DataImport_BulkDataDelete_06232016

This type of operation requires a custom SQL script, which will operate directly against the database, and because of this, requires both development time, a security review, and testing. Even if your data are not ready, you should indicate your intent to remove the data as early as possible, but not later than three weeks prior to the re-submission of your replacement data.

6.3.3 Selected Removal of Data

In general, this operation is not supported except in very exceptional cases (please contact the DATIM Support Team through DATIM Support).

Selected removal of data refers to a request to remove certain set of data elements or values for certain mechanisms or OUs only (i.e., not bulk removal).

6.3.4 Data Import Cleaning Option Summary Table

#	Task Option	Description	Required Lead Time for Engineering Team
1	Submission of new/updates + delete file	• If you are submitting updated data and/or new data, you should submit them as a single file. Clearly indicate in the name of the file that this is intended for updates. • If you intend to delete data, and know which data you want to delete, should prepare another file with the data values that should be deleted, clearly indicating in the file	• If data have been validated at your end, prior to submission, with the DATIM validation scripts, the DATIM Engineering team anticipate that this operation should be completed within 48 hours of submission.

		name that these data should be deleted.	
2	Bulk removal of data	• If you intended to submit a complete replacement of your data, you need to request at least three weeks in advance of the submission of your completed data file, that DATIM Engineering remove your data from the database.	• You are required to submit the request at least three weeks in advance of the submission of your completed data file.
3	Selected removal of data by DATIM Engineering	• In general, this operation is not supported by the DATIM Engineering team.	Not applicable

7   Resources

Several reference and guidance materials further explaining the DATIM data import are available on the DATIM Support page under the DATIM Data Import and Exchange Resources page.

PEPAR data exchange stakeholders are also encouraged to participate in the PEPFAR Data Exchange Implementer Community call, which occurs on the first three of Wednesdays of every month at 9 a.m. Eastern Standard Time. The call schedule every month is as follows:

- First Wednesday of the month: SIMS data import
- Second Wednesday of the month: MER data import
- Third Wednesday of the month: SIMS data import

The PEPFAR Data Exchange Implementer Community page contains details on upcoming call dates and agenda items, previous call recordings and notes, call-in details, as well as information on how to sign up for the PEPAR Data Exchange listserv.