ESTABLISHING AN MFL DATASET
MFL Resource Package Training

September 2019
# Agenda

<table>
<thead>
<tr>
<th>Topic</th>
<th>Estimated Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction to Establishing a Dataset</td>
<td>10 Minutes</td>
</tr>
<tr>
<td>2. Overview of Key Steps in Establishing a Dataset</td>
<td>40 Minutes</td>
</tr>
<tr>
<td>3. Review Governance Issues and Potential Challenges</td>
<td>10 Minutes</td>
</tr>
<tr>
<td>4. Group Exercises</td>
<td>30 Minutes</td>
</tr>
<tr>
<td>5. Action planning</td>
<td>30 Minutes</td>
</tr>
</tbody>
</table>
Session Learning Objectives

• At the end of today’s session participants will be able to:
  – Describe steps to establish an MFL dataset
  – Determine which sections of this module apply to your country
  – Identify gaps in your current MFL, if applicable
Review of Terms

• Minimum Data Content
• Data Specifications
• Signature Domain
• Validate
What is an MFL Dataset?
MFL Dataset

• Establishing an MFL Dataset involves:
  – Compiling the facility data you need from different sources
  – Validating the information obtained

• Depending on your current MFL status, you may need to:
  – Build a new dataset
  – Harmonize existing lists
  – Collect additional data to complete a list
  – Validate an existing MFL
Key Steps in Establishing an MFL Dataset

1. Determine What Data the MFL will Contain
2. Identify Available Facility Lists and Assess Their Content
3. Identify a Good List as a Starting Point
4. Address Gaps in Data
5. Collect New Data
6. Harmonize Lists into a Single List
7. Validate Data for Each Facility
Overview of Key Steps
I. Determine what data the MFL will contain

• Before you begin, determine:
  – What facility data you want included
  – What format the data should be in
  – Refer to Data Content module for more information
2. Identify Available Facility Lists and Assess Their Content

• Identify and assess any existing facility lists:
  – Refer to the MFL Assessment Module for guidance
  – Determine how complete and up-to-date the data are
  – Determine if the data in these lists supplement or build the MFL

• Consider:
  – Are available data complete?
  – Are any data missing?
  – Will you need to collect new data to fill gaps?
  – Are there any discrepancies across lists? Are they important?
  – Are data defined according to data specifications and standards?
3. Identify a good list as a starting point

- If an MFL already exists, even if there are gaps, you can skip this section
- If there is no list, consider the list information you have gathered to determine if there is a sufficient facility list to be used as a basis
- This list does not have to be complete, but should be:
  - Credible, i.e. information is from a trustworthy source
  - Include enough information about facilities to populate a substantial portion of the MFL
  - Contain data verified with in the last 5 years
3. Identify a good list as a starting point, cont’d

• If there are missing elements they can be collected separately

• You need to have confidence in the validity of data selected
  – If there are too many errors, incomplete information, or data quality concerns it may be best to start from scratch

• If lists are too outdated or incomplete, then you may need to start over and collect new data
4. Address gaps in data

1. Clearly document any gaps in the list

   – Gaps can be:
     • Missing data
     • Incomplete data
     • Data that appear erroneous

   – Examine gaps carefully so you know what additional data you need to create a comprehensive MFL
4. Address gaps in data, cont’d

2. Determine where to obtain data
   – You can pull data from another facility list
   – You can collect data from the facilities
   – You may need to do both

• When deciding to collect new data or pull from other lists, consider the level of effort needed and associated costs

• If missing data are not available in any current lists or if harmonization is not practical, you will need to collect new data
5. Collect new data

• You may need to collect new data if your current data do not meet the minimum data requirements

• Be clear on:
  – Types of facilities you are including
  – Geographic areas where data collection will occur
  – Data elements to collect with clear definitions for each
  – Data collection tools or measurement approaches to use
  – Data sources to use
5. Collect new data, cont’d

• Approaches to collecting new data include:
  
  – **Health facility assessment surveys:** examples include WHO Service Availability and Readiness Assessment (SARA), DHS Program’s Service Provision Assessment (SPA)
  
  • Coordinate with the facility assessment surveys to make sure they are collecting the data you need
  
  – **Targeted facility census:** send teams into the field to collect the data directly from facilities only on data elements needed
  
  – **District Health Information Officers:** enlist to collect information about facilities in their districts
  
  – **Crowd sourcing:** collecting information from large groups of people; may have data quality issues
6. Harmonize lists into a single list

- Establish a technical working group to oversee harmonization
- 2 main purposes:
  - Add additional facilities and their records to the MFL
  - Add supplementary data about facilities already in the MFL
- To harmonize lists:
  - Identify data to keep from each list
  - Compare data element definitions/data specifications across each list that will be used in the MFL
  - Estimate the proportion of facilities that can be matched electronically across lists
6. Harmonize lists into a single list, cont’d

– Next carry out the following steps:

1. Match facility records to create one record per facility
   – Organize facility data
   – Electronically match if there are a large number of facilities
   – For matched facilities, check data fields to verify signature domain to ensure the facility is the same across both lists
   – Manually match facilities that were not matched electronically
   – Combine information in matching records for one record per facility

2. Import data for new facilities not included in the MFL, verify the facility is not listed under another name
6. Harmonize lists into a single list, cont’d

– A common challenge during harmonization is discrepancies in facility-level data, including:
  • Spelling differences
  • Address or location differences
  • Differences in naming of facilities

– Next step is to create the unique facility identifier codes for facilities without them

– Then assess the new list for gaps
7. Validate data for each facility

- After data is collected, you need to validate the data for each facility
  
  • Determine if you need to resolve any data quality issues
  • This is an ongoing process throughout the life of the MFL
  • All data must be validated but accuracy and signature domain are most important

- Once data has been validated for each facility, assign unique identifiers

- Then a validated MFL is ready to share
Activity 1: Data Elements

- What are the main data elements to include?

- What should be used as the signature domain?

- How will you validate data?
Governance Issues to Consider
Governance Considerations

• Set up a technical working group tasked with follow up for the MFL dataset
  – Develop a work plan
  – Create a budget
  – Establish a timeline to aid managing the process
• Stakeholder engagement is critical at every step
• Stakeholder buy-in is important if you are harmonizing lists owned by different groups and to reach consensus on an authoritative list
Potential Challenges for the MFL Dataset
Potential Challenges

• Too many current facility lists exist
• Too many data elements included in the MFL
• Non-standard facilities or mobile facilities
• Addressing discrepancies between MFL and externally managed facility lists
Activity 2: Data Validation
Group Exercise 1

• In small groups, outline a data validation process, include resources to check the list data against, potential contacts for validation, and supervisors or officers who could perform additional data validation or collection.

• Point out any gaps or challenges in the data validation process, provide solutions if possible.

• At the end of the process, determine if it is too costly and if collecting new data would be a better option.

• Present process to the rest of the audience.
Activity 3: Steps for establishing an MFL Dataset
Group Exercise 2

• In small groups, determine which steps for establishing an MFL dataset need to be carried out for [COUNTRY].

• Present steps to include or omit and reasons why.

• As whole group, if possible, discuss any harmonization or differences between groups.

• Determine a final lists of steps to include for [COUNTRY].
Summary and conclusions