Error Resolution Guide: Metadata Template

Overview

This guide helps you identify and resolve common errors encountered when filling out the Metadata Collection Template. Each error type includes clear explanations, step-by-step solutions, and practical examples to ensure your data meets validation requirements for successful submission to the AGARI platform.

Purpose of this guide

When uploading data to the AGARI system, the validation system checks for data integrity and format compliance. This guide provides solutions for:

- Formatting errors in date fields
- Invalid dropdown selections
- Missing mandatory fields
- Data relationship violations
- File matching errors

Error Types and Resolutions

Error 1: Date must be in YYYY-MM-DD format

What this error means

The date entered does not follow the required ISO 8601 date format. All dates must include year, month, and day in a specific format.

How to fix it

- Locate the cell with the date error (it will be highlighted in red or shown in the error message)
- Delete the current value
- Enter the date in the format: YYYY-MM-DD
- Ensure all parts are present (year with 4 digits, month with 2 digits, day with 2 digits)

Incorrect Format	Correct Format	Issue
15/07/2024	2024-07-15	Wrong date separator
July 15, 2024	2024-07-15	Text format not allowed
2024	2024-07-15	Year only is incomplete
24-07-15	2024-07-15	Year needs 4 digits

[Insert Screenshot: Date field showing correct format]

Error 2: Each dropdown value must be a valid option from the list

What this error means

The value entered is not in the allowed list of options for this field. This happens when typing manually instead of selecting from the dropdown.

How to fix it

- Click on the cell showing the error
- Click the dropdown arrow that appears
- Select a value from the list provided
- If your desired value is not in the list, check the 'Picklist' sheet for the complete list of allowed values
- Never type values manually - always use the dropdown selector

Common mistakes

- Typing 'south africa' instead of selecting 'South Africa' from the dropdown
- Adding extra spaces before or after the value
- Using abbreviations (e.g., 'SA' instead of 'South Africa')
- Misspelling the option

[Insert Screenshot: Dropdown list showing selection process]

Error 3: Values must be a valid option from the list (Multi-select)

What this error means

One or more values entered in a multi-select field do not match the allowed options exactly. Multi-select fields allow multiple values but each must be exact.

How to fix it

- Check the cell note (hover over the cell) for the complete list of allowed values
- Type each value **EXACTLY** as shown in the list
- Separate multiple values with commas and NO SPACES
- Maintain exact capitalization, punctuation, and slashes

Examples

Incorrect	Correct
IncA, pSDH-1, pVC	IncA,pSDH-1,pVC
Spaces after commas	No spaces, exact match

[Insert Screenshot: Cell note showing allowed values]

Error 4: Value must be a number

What this error means

The field requires a numeric value but text or special characters were entered.

How to fix it

- Remove any letters or special characters
- Remove units of measurement (e.g., 'years', 'days', '%')
- Enter only whole numbers (integers)
- Do not include spaces or commas as thousand separators

Common examples

- Host Age: Enter '35' not '35 years'
- Number of Vaccine doses received: Enter '15' not 'Fifteen'

Error 5: Value is mandatory

What this error means

A required field has been left empty. Mandatory fields (highlighted in yellow) must have values for successful submission.

How to fix it

- Look for cells highlighted in yellow - these are mandatory
- Check the error message to identify which specific field is missing
- Enter an appropriate value in the empty field

Common mandatory fields for Cholera

- study_id
- isolate_id
- geo_loc_name_country
- sample_collection_date
- specimen_source_material_category
- biospecimen

[Insert Screenshot: Yellow highlighted mandatory fields]

Error 6: Date relationship errors

Types of date relationship errors

- Date must be before sample_receive_date
- Date must be after sample_collection_date

What these errors mean

Dates must follow a logical sequence. For example, a sample cannot be received before it was collected, and processing cannot occur before receipt.

How to fix it

- Review all date fields in the row
- Ensure dates follow logical order:
  - Collection Date → Received Date → Processing Date → Sequencing Date
- Check for data entry errors (e.g., typos in year)
- Verify dates with original records if necessary

Example of correct date sequence
- Sample collection: 2024-07-10
- Sample received: 2024-07-11
- Processing date: 2024-07-12
- Sequencing date: 2024-07-15

Error 7: Value must be unique

What this error means

The same value appears more than once in a field that requires unique identifiers. Common in isolate_id.

How to fix it
- Find and review all duplicate values
- Update duplicates with unique identifiers
- Consider adding suffixes (e.g., -A, -B) to differentiate samples

Tips for maintaining uniqueness
- Use a consistent naming convention
- Include location or date codes in IDs
- Maintain a master list of used IDs
- Use Excel's Data Validation to prevent duplicates during entry

[Insert Screenshot: Excel highlighting duplicate values]

Error 9: Fasta file name fasta header name must match record in the accompanying metadata

What this error means

The validation system requires that the fasta file name and fasta header name provided in the metadata must exactly match the accompanied fasta file.

How to fix it
- Open your FASTA file in a text editor.
- Check the sequence headers (lines starting with '>')
- Ensure the sequence headers in metadata exactly matches the header (after '>' symbol)
- Update either the metadata or FASTA file to ensure consistency.
- Check if the fasta file name matches the fasta file in the metadata.

Example

Sequence header: >NICD-01
Metadata Fasta header name: >NICD-01
✓ These match correctly

FASTA header: >NICD_01
Metadata isolate_id: NICD-01
✗ These don't match (underscore vs hyphen)

Best Practices for Error Prevention

Before data entry

- Download and use the latest template version
- Review the Instructions tab thoroughly
- Familiarize yourself with mandatory fields (yellow highlighting)
- Check the Picklist tab for allowed values

During data entry

- Fill mandatory fields first
- Use dropdown menus whenever available
- Double-check date formats (YYYY-MM-DD)
- Avoid copy-pasting from other sources without verification
- Save your work frequently

After data entry

- Use Excel's Data Validation to check for errors
- Sort by each column to identify outliers or blanks
- Check for duplicate IDs using conditional formatting
- Verify FASTA file headers match metadata IDs
- Run a test upload with a small subset of data first

Best Practices for Error Prevention

Quick Reference Guide

Error Type	Quick Fix	Prevention
Date format	Use YYYY-MM-DD format	Set cell format to Text before entry
Invalid dropdown	Select from dropdown list	Never type manually
Multi-select	Use commas, no spaces	Check cell notes for exact spelling
Not a number	Remove text/units	Enter numbers only
Mandatory field	Fill yellow cells	Complete all yellow fields first
Date sequence	Check chronological order	Verify dates against records
Duplicate ID	Make each ID unique	Use conditional formatting
FASTA mismatch	Match IDs exactly	Check headers before upload

Need Additional Help?

If you continue to experience errors after following this guide, please:
- Take a screenshot of the error message
- Note the row and column where the error occurs
- Contact your data management support team
- Provide the error details and this guide reference

Version 1.0 | November 2024