

WHONET – InFARM Data Validation Report



WHO Collaborating Centre for
Surveillance of Antimicrobial
Resistance

Boston, February 2026

Table of Contents

1. Background	2
2. About this document	2
3. Differences between Option A and Option B reports	3
4. Report overview.....	3
5. Missing data checks	3
6. Validation checks	4
7. Consistency checks	4

1. Background

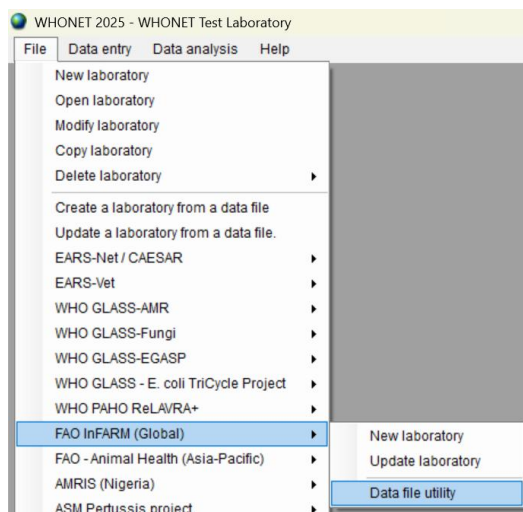
InFARM stands for the “International FAO Antimicrobial Resistance Monitoring System”. More information can be found at the following URL:

<https://www.fao.org/antimicrobial-resistance/resources/database/infarm/en/>

The InFARM protocol is subdivided into two categories: Model A (isolate-level data) and Model B (aggregate statistics). Countries can choose between submitting more granular data with Model A or aggregate statistics with Model B. Countries make these determinations using the InFARM web portal and then use the corresponding settings in WHONET to generate the data files for upload to the portal.

2. About this document

This document describes the various types of messages contained in the data validation reports, which are optionally generated by WHONET during the InFARM data export processes. You may also generate the data validation report from an existing InFARM data file, regardless of the source of the file, using a utility included with the WHONET software. This utility is available on the “File”, “FAO InFARM (Global)” menu as shown below.



3. Differences between Option A and Option B reports

The InFARM Option A and Option B data files have differing levels of detail. Option A is an isolate-level export, where Option B only includes summary statistics. This causes the data validation reports generated by the system to also differ.

For example, if you've chosen Option A, the data validation report will provide you with precise details for the offending isolates, such as the isolate identifier. If you choose Option B, the validation report will contain information pertaining to groups of isolates with the same issue (there may be many isolates that were summarized into a single InFARM Option B row).

If you need to export Option B data only to the InFARM platform, you may generate an export and validation report for Option A temporarily to obtain the higher resolution associated with the Option A report. For example, a given issue in the Option B report will usually also be present in the matching Option A report, just with greater detail because the isolate level information is available.

4. Report overview

The data validation report is divided into 3 types of issues: missing data summary, validity checks and consistency checks. Each of these issue types will have two Excel data sheets associated with them, one for a summary of the issues and another providing the details. If there are no issues found for a given category, then these Excel worksheets will be omitted. For example, if there are no problems with the mandatory data fields in your data, then your data validation report will not contain these worksheets.

If your data file passes all the data checks, then your data validation report will contain a single worksheet informing you that there were no issues discovered. The WHONET team has modeled our validation report around the same types of issues that are raised by the InFARM platform with the intent that you are able to review and improve your local data as needed before attempting to submit.

5. Missing data checks

Certain data fields are mandatory according to the InFARM data protocol. The purpose of this check is to identify isolates that have missing mandatory data values. If there are problems such as this in your data file upon export, you will receive summary and details section like those shown below.

	A	B	C	D
1	Message	Data field	Number	
2	A required field is missing.	ORIGIN	194	
3	A required field is missing.	SPECIES	194	
4				

< > 1 - Missing data summary 2 - Missing data details

In this example, generated with fabricated data, you can see that there are 194 issues with the ORIGIN data field and the same number with the SPECIES data field. Selecting the "Missing data details"

worksheet instead, you will be provided with a complete list of the problematic isolates. Note that the “Sample number” and “Isolate number” fields correspond to the encrypted values present in the InFARM data export. The “Sample number – Local code” and “Isolate number – Local code” fields should reflect your local values for these data fields in the WHONET data files. This is very useful information for locating the original isolates, since your local isolate numbers should be unique as a condition of InFARM reporting.

	A	B	C	D	E	F	G
1	Message	Data field	Laboratory	Sample number	Isolate number	Sample number - Local code	Isolate number - Local code
2	A required field is missing.	ORIGIN	TST	39592984783725817096	90175401966063911204	365229762036	900734284916613
3	A required field is missing.	ORIGIN	TST	79289613242022899785	82092704415326566459	884590881789	1153068232146312
4	A required field is missing.	ORIGIN	TST	99853024412880168150	36609615641550439863	583767441621	1830382547142848

If there are only a few isolates which are missing data, then this can usually be corrected by editing the existing WHONET data file using the “Open data file” feature found on the “Data entry” menu. However, if there seems to be a high proportion of isolates which are missing key data, then there may be a more fundamental improvement that can be made. For example, if a facility uses the BacLink software to generate the WHONET data files, then perhaps there is a missing code definition corresponding to the ORIGIN or SPECIES fields which eventually lead to missing data in the InFARM export. In cases such as this, you can typically regenerate your WHONET data file with BacLink again after you have defined or corrected the codes, reducing the issues identified in a subsequent InFARM export using the corrected data.

6. Validation checks

The second category of issues detectable by the reporting system is for determining whether the data values provided are valid. For example, certain InFARM fields, such as MICROORG (isolated organism) and SPECIMEN (specimen type) must use values from their respective lists indicated in the InFARM protocol documentation. There are also validity checks to ensure that the XCOORD and YCOORD are within the expected ranges, a check to ensure that free text fields do not exceed the maximum length specified by InFARM, and a check to ensure that at least one antimicrobial was tested for a given isolate.

	A	B	C
1	Message	Data field	Number
2	Invalid values	MICROORG	3
3	No antimicrobial susceptibility testing	All antibiotics	1
4	The data value exceeds the maximum allowed length.	ORIGIN_NOTES	1
5	The provided data value was not in the expected range.	XCOORD	1
6	The provided data value was not in the expected range.	YCOORD	1

Please refer to the InFARM documentation for specific details regarding valid codes, field lengths and acceptable numeric ranges.

7. Consistency checks

The final category of issues presented by the report checks to ensure consistency between certain data fields according to the InFARM protocol. The SPECIMEN, ORIGIN, SPECIES and MICROORG field have

restrictions imposed by the protocol which dictate that they are only relevant for certain surveillance programs.

There is an additional check which ensures that the antimicrobial interpretations provided in the data file are consistent with the INT_CRITERIA field. For example, if the data row indicates that it is reporting clinical breakpoints, then the antimicrobial interpretation columns should have corresponding values (R, I, S, NS). If the data row indicates that it is for ECOFFs, then the interpretation columns should have values of WT and NWT.

Finally, the system will ensure that specimen dates fall within the data year that was provided while creating the Option A export. Because the specimen date is an optional data field, it may be left blank. However, if it has been provided, this check will ensure that it falls within the expected year. This final check is not applicable for Option B data files because the specimen date is not one of the data elements for that option.

	A	B	C
1	Message	Data field	Number
2	The data value is not valid for the chosen surveillance program.	SPECIMEN - ANIMPH	18
3	The data value is not valid for the chosen surveillance program.	ORIGIN - ANIMPH	4
4	The antibiotic interpretation is inconsistent with the interpretation criteria.	CLIN	1