FOSCOLLAB DATA QUALITY WORKING GROUP

GUIDANCE DOCUMENT FOR THE INTEGRATION OF DATA IN FOSCOLLAB

World Health Organization

WHO/HSE/FOS/2015.2

1. Background

Food safety is a global issue requiring global solutions. A resolution[1] was passed at the 63rd World Health Assembly in May 2010 reaffirming this and outlining key actions needed to advance food safety in the coming years. One of the key elements to ensure global food safety is the ability to share data from food monitoring and other sources to better inform the development of preventive control strategies and risk-based approaches to food safety.

2. Objectives of FOSCOLLAB

Current and future global food safety assessments require access to global data (both quantitative and/or qualitative). Quantitative information includes levels of occurrence of chemicals in various food matrices, growth curves of microbes, statistics on food production and food consumption. Qualitative information includes for example contact information of competent authorities, list of laboratory capacity and capabilities of analytical, method testing information and texts of national legislation.

The generalization of the web links of scientific laboratories and food authorities combined with the performance of the standard web crawlers (google,...) results in a large amount of available information. These web links give also access to specialized databases containing experimental results. The difficulty is therefore to extract the pertinent information under a format rapidly and easily manageable. There is a critical benefit for being able to extract and analyse the various pieces of information in real-time in a comprehensive manner .

The World Health Organization (WHO), through its Department of Food Safety and Zoonoses (FOS), initiated a project named FOSCOLLAB to improve ways of sharing food safety data and information to support risk assessment and decision-making in food safety.

> *FOSCOLLAB is a platform accessible from internet and displaying together within dashboards various data (quantitative and qualitative) and information (e.g. expert advice) useful for food safety professionals. FOSCOLLAB allows linkages between databases based on 4 criteria: food name, hazard name, country of origin and year for data generation.*

The overarching objectives for FOSCOLLAB are to assist food safety professionals by:

♦ Overcoming the fragmented availability of current food safety data and information

♦ Providing easy access to global food safety data and information

♦ Saving time by not needing to search and sift through numerous resources

♦ Bringing together data and information coming from animal/agriculture, food and human health areas to improve assessments linked to global public health

♦ Strengthening the underlying data sources by promoting awareness and increased utilization

---

1 WHA 63.3 Advancing Food Safety Initiatives (http://apps.who.int/gb/ebwha/pdf_files/WHA63-REC1/WHA63_REC1-P2-en.pdf)

The project aims to develop an interactive and user-friendly tool that will extract and present key globally available food safety data and information into a single, interactive access point. The tool is designed for food safety professionals to be the first place to go when having to assess a food safety risk, obtain information to enhance decision making, bring together data from multiple sectors (e.g. animal/agriculture, food, public health, economic indicators).

Apart from the technical challenge of FOSCOLLAB, the success of the project also depends on the quality of the data integrated in the platform. It is therefore essential for WHO to establish minimum quality criteria to evaluate the quality of new databases/data before integrating them into FOSCOLLAB.

The initial consultative process for this initiative was conducted from January to July 2011 and involved 131 participants from 64 Member States allowing for a broad representativeness of participants' specializations. Throughout the consultations, participants mentioned different sources of information they used (and what they would like to have access to) in their day-to-day work.

> *Overall, the review strengthened the notion that FOSCOLLAB should leverage existing data sets: If a strategic decision is made to integrate a new type of data and information, WHO should first determine if this is available in-house and decide whether it is feasible to integrate. If not, WHO would work with partners and Members States to establish what the most valuable contribution from external sources would be.*

Example:

For any food hazard, one of the pillars of the risk analysis process is the quality and the robustness of data supporting risk assessments and establishment of food standards. This is particularly relevant in an international context (e.g. Codex Alimentarius) when chemical and microbiological analytical results are generated by a number of national laboratories and reported to WHO and other international organizations.

Once compiled, these results can be aggregated to generate a regional or an international distribution for the hazard under consideration. The distribution of the occurrence of chemical in food should provide estimates of human exposure which in turn may allow for the calculation of the possible risk for public health. This distribution should represent the best data available and also reflect the current situation insofar as possible of the food production chain (food supply chain) and provide relevant information to help assess the feasibility and possible impact of new regulatory measures. The data can also provide insight into current practices and help identify opportunities for potential harmonization.

FOSCOLLAB, in addition to displaying the most recent available data, aims to support risk assessments and descriptions of food safety problems by way of networks of national/regional institutions that collect and use data. WHO and other international organizations are compiling data collected through various initiatives and working groups e.g. the GEMS/Food Network[2] . FOSCOLLAB integrates data and information from a number of existing sources databases (namely the JECFA and JMPR databases on chemical risk assessments, the WHO database on Collaborating Centres and the GEMS Food databases on food consumption and chemical occurrence in food) within WHO and will continue to incorporate additional data sources e.g. Codex database on MRLs for pesticide residues in food.

2 http://www.who.int/entity/foodsafety/areas_work/chemical-risks/GEMS_institutions.pdf

In previous WHO initiatives, risk assessors expressed their requests in terms of the type and amount of data needed for risk assessment (for example HOF[3] and COFOCO[4]).

However, there are no standards or guidelines specific to FOSCOLLAB on the minimum quality of the national individual data included in the various databases to ensure the data captured is comparable, comprehensive, and fit-for-purpose. Developing and implementing expert recommendations along these lines will result in more robust data and information, thus improving the usability and reliability of the FOSCOLLAB data.

3 Declaration of interest

The Secretariat informed the Meeting that all experts participating in the Data Quality Expert Working Group (DQEWG) had completed declaration of interest forms and that no conflicts had been identified. Dr Pisan Pongsapitch declared his employment as an official of the National Bureau of Agricultural Commodity and Food Standards in the Ministry of Agriculture and Cooperatives in Thailand. This institution is responsible for submitting data to WHO and other International Institutions. The Working Group confirmed that this declaration should not be considered as a conflict of interest.

4 Terms of reference of the working group

This working group of data quality experts was established in February 2013 to research existing comparable data quality standards and systems used by other organizations, both public and private to select the most appropriate practices aligned with the goals of FOSCOLLAB.

Using FOSCOLLAB as a baseline, the expert working group will begin to address the larger discussions of quality and harmonization of food safety data and information to identify and create guidelines and recommendations to serve in particular the purpose of food risk assessment. The first results of the working group can lead to further investigation of a particular sector (i.e. microbiological hazards, pesticides, etc.).

Specific outputs of the working group will include:

- ◆ Review of existing procedures in place on other data platforms and selection of the most appropriate practices to be utilized (ex. eLEXNET, Foodrisk.org, WHONET, etc.).

- ◆ For each of the priority data and information sources previously identified to be integrated in FOSCOLLAB, propose expert recommendations on criteria for the platform, including:

  - ◆ Generic criteria to be applied to all sources for exclusion
    - ◆ Rubric; checklist, etc.
    - ◆ For both WHO and Non-WHO sources

  - ◆ Type of data and information from each selected source to be displayed by FOSCOLLAB
    - ◆ Depending on type of data (lab results, methods, sampling/sample size, survey data…)
    - ◆ For both WHO and Non-WHO sources

3 http://www.who.int/entity/foodsafety/publications/gems_hof/en/index.html
4 http://www.who.int/entity/foodsafety/publications/gems_cofoco/en/index.html

5 Laboratory data

5.1 Introduction

FOSCOLLAB has the potential to integrate analytical data from multiple data sources. By integrating multiple sources of data, FOSCOLLAB helps to overcome the challenges of accessing disparate sources of data. As a result, it has the possibility to allow for more effective use of data by food safety professionals and authorities. FOSCOLLAB should enable users to access food safety data quickly, promote more comprehensive analysis of data, promote the improvement and harmonization of data standards and best practices and the sharing and acceptance of data.

The working group focused most of its attention on identification and selection of parameters related to laboratory data. Such data may relate, among other things, to the identity of the analyte itself, namely what is being analysed, the amount detected, the matrix in which it was detected, and where and when the sample was collected.

5.2   Examples

Examples of analytical data include laboratory data on food contaminants and microbiological hazards.

Data may be recorded and reported in various ways, such as through different analytical methods measuring the same or similar analytes and reporting different units. Additionally data may be collected for different purposes, such as compliance with legal requirements, risk assessment supporting risk management and/or policy development, or for research purposes.

All these factors can lead to apparently similar data not being comparable. This working group emphasized the minimum criteria that is mandatory for a full characterization of an analytical result for data included in FOSCOLLAB. The working group also recognized other criteria which can be useful to know but their absence would not justify the rejection of data.

5.3 Criteria to include laboratory data sources into a FOSCOLLAB dashboard

5.3.1 Criteria

♦   Regional dimension and temporal coverage: Data will ideally be part of an already-existing international data repository and provide regional and temporal coverage.

♦   Avoid duplication and overlaps: The data source should be equal to the original source and be extracted from as close as possible from the original data producer i.e. analytical laboratory. In case the data source is compiling data from various databases, the identification of the original data producer is necessary.

♦   Filling identified data gaps and enhance global picture: the data sources complementary of data already integrated should be considered in priority in order to provide a more global picture.

In addition, any analytical data sources need to meet the requirements outlined in section 5.6 for data elements.

5.4   Selection of sources

Recognising that the framework should also establish common quality parameters for both microbiological and chemical hazards, the working group first identified existing data standards that were in widespread used.   Of particular note were the report of the WHO GEMS 'Working Group on data reporting for hazards occurring in food' (GEMS-HOF), the EFSA Standard Sample Description for Food and Feed, the US FDA's Electronic Laboratory Exchange Network (eLEXNET), and international academic journals.

The group identified similarities and overlaps between the existing data standards and distilled these into a core set of internationally-applied data elements taking into account the availability of data on a global level. This core set of elements was used as a starting point to further guide the discussion of the group.

5.5   Selection of data

Unsurprisingly, these existing data standards were established by developed economies with sophisticated laboratory networks and resources to collect large and useful data sets. The group considered whether data would still be useful for food safety officials if not all data elements were present. The group suggested a three-tiered hierarchy of importance, ranging from:

♦   <u>Tier 1 (base)</u> elements, the 'bare minimum' for inclusion of analytical data

♦   <u>Tier 2</u> elements, which add significantly to the value of the data and whose absence may preclude the use of the data by food safety professionals

♦   <u>Tier 3</u> elements, which add context and richness to the data but are not necessary for FOSCOLLAB's users.

It was recognized that in order to encourage the least advanced data owners, Tier 1 should promote the integration of analytical data generated under good laboratory practices but without necessarily reaching the optimum quality to be used for risk assessment purposes (e.g. national representativeness of the sampling plan, official accreditation) However, the working group emphasized that improving the quality of analytical data worldwide should remain an important goal and that WHO should encourage all data owners to provide FOSCOLLAB with Tier 2 and Tier 3 data elements.

5.6   Classification of data elements

Elements were classified in 4 categories namely "Sampling", "Methods of analysis", "Quality assurance", and "Analytical results". For each category the elements were divided into 3 Tiers, the first (Tier 1) being mandatory for accepting data into FOSCOLLAB the other ones (Tier 2 and Tier 3) being desirable.

The base elements (Tier 1) are used to link various databases together at a functional level. If these elements are not present, FOSCOLLAB will be unable to link up to the data source.

Tier 2 and Tier 3 data elements were those that add value, and may be critical to some users but whose absence should not necessarily prevent the inclusion of data for use in an international context. For example, product identifiers are critical in cases of international food recalls but are of less value when assessing dietary exposure on a country or regional basis. In the latter case, where available data are often very limited, poor or incomplete data is a better alternative to no data at all.

The list of Tier 2 and Tier 3 data elements is not exhaustive, as the possibilities are broad and the usefulness of data elements can be dependent on end user needs.

The group identified that all relevant data elements could be broadly grouped under the following categories to represent the progression of approaches from obtaining the sample to evaluating the sample.

5.6.1 Tier 1 (base)

Table 1. Display for each category the essential elements to be present in the data source to accept import in FOSCOLLAB.

	Element Category			
	Sampling	**Methods**	**Quality Assurance**	**Analytical Results**
Tier 1 (BASE)	**Geographic location where the sample was collected:** A country where the sample was collected; sometimes not the same as a country where the sample was analyzed or the country that submitted the data.		**Laboratory Accreditation:** Is the information available? Yes or no	**Analyte(s) Detected:** The substance detected in the analysis must be reported. It is very important to clearly indicate what substances are being reported especially when dealing with metabolites and/or derivatives. The substance (analyte) detected is normally the same as the substance (analyte) tested. However, when they are different, clear explanation should be included e.g. when the test included aflatoxin B1, B2, G1 & G2 but only aflatoxin B1 was detected and reported. The data submitted should always indicate the analyte name related to the result of analysis submitted.
	When sample was collected: This element can also be identified as sampling date			**Result Description:** Qualitative and/or quantitative
	Product Description: The specific food being sampled (Food code for FOSCOLLAB)			**Result Units**

5.6.2 Tier 2

Table 2. Display for each category the desirable elements to be present in the data source to improve its use through FOSCOLLAB.

	Element Category			
	Sampling	**Methods**	**Quality Assurance**	**Analytical Results**
Tier 2	**Sampling point :** Point in the food chain (farm, port, restaurant, etc.) where the sample was collected	**Test Method Name**	**Description of Quality Management System/ISO 17025:** Important information may be available about the integrity of the laboratory results reported, even if the laboratory is not accredited to ISO[5] 17025, such as other official accreditation, internal quality assurance (e.g. SOP[6] , method validation, participation in proficiency testing)'. This alternative element is useful in providing the status of both accredited and non-accredited laboratory.	**Tertiary analysis (ex. PFGE[7])**
	Sample Condition: indicates whether sample condition meets acceptable ISOstandards or equivalent for analysis Yes or No	**Method Description:** The specific test method the laboratory employed for the test must be described. This can be done by referring to the name or code of standard method (e.g. ISO, AOAC[8]) or modified version of the standard method or a short description of the method of analysis employed by the laboratory.	**Scope of accreditation:** Additional detail about the scope of accreditation can be very useful if it is available.	**Laboratory Identification**

5 ISO: International Standard Organization

6 SOP: Standard Operational Procedures

7 PFGE: Pulsed-field gel electrophoresis

8 AOAC : Association of Official Agricultural Chemists

Tier 2 (contd)

	Element Category			
	Sampling	**Methods**	**Quality Assurance**	**Analytical Results**
Tier 2 **(contd)**	**Sample collection, prep and analysis:** Important context can be added to sampling information by also reporting the sample size, including units, and the sample's representativeness. Where the user is interested in knowing the prevalence of an analyte, or knowing that an analyte is not present with an estimated level of confidence, information about the representativeness of the sample will be very important. In some cases, this is not necessary, for example, where the user of FOSCOLLAB is only seeking an indication of the presence of an analyte	**Product Portion**		**Current PFGE Flag** Indicates whether laboratory that performed the PFGE is the same as the laboratory that performed main sample analysis
	Country of origin of the sample : Country of origin is necessary in identifying the country where contamination occurred	**Analyte(s) Tested**		
	Why sample was collected: Outbreak investigation, recall verification, compliance, random sampling/surveillance, monitoring, baseline studies…etc…	**Sample size (micro) or Limit Of Quantification/LOQ (chemical):** Are conditionally mandatory in absence of result		**Laboratory Serotype Description:** laboratory assigned description that identifies a group of closely related microorganisms distinguished by a characteristic set of antigens

5.6.3 Tier 3

Table 3. Display for each category the additional elements to be present in the data source to enhance the precision of data characterization.

	Element Category			
	Sampling	**Methods**	**Quality Assurance**	**Analytical Results**
Tier 3	**Lab Product Code:** alphanumeric string that laboratory uses to define product		**Accreditation bodies:** Accreditation body	**Confirmation Test:** indicates whether test per-formed was confirmative
	Action Description: action taken based on labora-tory result; e.g.. International Health Regulation (IHR) risk assessment/notification			**Confirmation Test Lab:** Name laboratory that performed the confirmation test
				Confirmation Date: date the laboratory performed the confirmation test
				Instrument Name: analytical instrument used to identify analyte ex: Whole-Genome Sequencing (WGS), platforms, test kits, etc.

|

5.7 Conclusions

FOSCOLLAB currently provides users with dashboards integrating various databases available internally to WHO namely the JECFA and JMPR databases on chemical risk assessments, the WHO database on Collaborating Centres and the GEMS Food databases on food consumption and chemical occurrence in food. In parallel, a number of external databases were identified as fitting the needs of food safety professionals and are planned for integration in the near future. The objective of WHO is both to be inclusive of all relevant data available worldwide and to maintain a high level of quality for the benefit of data users.

The challenge for the working group was to include a pool of potential data elements as broad as possible without compromising its usefulness for food safety professionals. The hierarchy of data elements described above answers part of this need. However, consideration also needs to be given to how FOSCOLLAB itself will operate.

The group also recognised that much contaminant data is collected in academic journals and these (which may be written in many languages), not in international databases, so there is a significant risk of excluding useful data that cannot be directly integrated into FOSCOLLAB. However, FOSCOLLAB is intended to be a portal through which food-safety professionals can readily access existing food-safety data. In no way will it act as a replacement for the sources it links together. Linking to large numbers of data sources of varying quality could undermine the value of its user-friendly interface and increase the needs for its maintenance. This in turn could reduce the value of FOSCOLLAB itself for users.

> *Consequently, it is desirable for FOSCOLLAB to remain a portal to a limited selection of quality internationally-accessible data sources, and to encourage users to visit the individual source programmes for further details. Those international data providers could be encouraged to facilitate the acquisition and consolidation of data with broad geographic and demographic spread.*

In the longer term, FOSCOLLAB should also be an international platform to improve harmonization of data reporting between data producers and to promote and share best practices to generate, compile and exchange analytical data. In addition, FOSCOLLAB should serve as a platform to promote best practice regarding data generation and sharing, as well as capacity building on a global level.

6 Data and information from other sources

6.1 Introduction

Data and information from other sources are defined in the context of FOSCOLLAB as structured or unstructured data not generated by a laboratory. Such data can be useful for food safety professionals but cannot fulfil the criteria described in chapter 5. We need a good example, other than lab data that show the added value of FOSCOLLAB. This should be included in the non-structured/structured data. This example may also describe how aggregate data can enhance decision making processes.

6.2   Examples

6.2.1 Structured data

The amounts of foodstuff either traded or consumed can be considered as an example of structured non-analytical data. The International Trade Centre (ITC) is providing trade statistics[9] by countries and by product group. Similarly the Food and Agriculture Organization of the United Nations (FAO) is collecting, analysing and making available data on food production and consumption[10]. This type of summary statistics should be integrated into FOSCOLLAB. They should be related to the four main parameters of FOSCOLLAB (Food name, hazard name, date, country name).

The origin of the raw data and the way they were obtained should be documented. In some cases a laboratory can propose to report summary statistics for analysis rather than individual analytical results: in such a case the majority of data elements applying to individual data results also apply here.

6.2.2 Non-structured data

The conclusion from an expert Committee on risk assessment (e.g. JECFA) is an example of non-structured data. When JECFA is producing a risk characterization for a chemical contaminant, the text includes ad'hoc considerations on groups at risk as well as on inherent uncertainty of the assessment. Such information cannot be translated in a quantified value but may be useful for risk managers.

Another example of relevant non-structured data is the INFAL network information system among member laboratories in American region. This network includes 63 labs from 30 countries to:

♦   Facilitate availability of reference materials and participation in inter-laboratory testing

♦   Organize and promote programs for training and education promoting the sharing of experiences and resources available

♦   Promote and strengthen the inter-sectoral participation in building and operating national networks

♦   Promote and strengthen integration of member laboratories food safety programs and epidemiological surveillance

For its international dimension as well as for its overall goal to promote scientific exchanges and to improve efficiency and quality of analytical laboratories, the INFAL data also represents a useful source of non-analytical data.

The reputation of an institution producing the information is a consideration for assessing data quality. The data should be associated with the text from which it is extracted or/and with information on the context around its production.

---

9 http://www.intracen.org/exporters/statistics-import-product-country/

10 http://faostat3.fao.org/faostat-gateway/go/to/home/E

6.3   Criteria to include other data and information into FOSCOLLAB

Data sources for non-analytical data should be internationally recognized, provide an international coverage and the method for data collection should be documented. Moreover the data source should be sustainable and be regularly updated.

- ♦ International recognition of the source

- ♦ International coverage of the data

- ♦ Availability of the methodology for data collection

- ♦ Availability of quality assurance system

- ♦ Linkage with one or more of the parameters used in FOSCOLLAB (i.e. Food name, hazard name, date, country name).

- ♦ Availability of the methodology for data collection, analysis and review

- ♦ Availability of quality control protocols

7 Recommendations

7.1 General recommendations

In order to improve the efficiency of FOSCOLLAB the working group recommends:

- ♦ To form a standing advisory committee to advise WHO on the future developments of FOSCOLLAB.

- ♦ To develop a Taxonomy of data sources to help delineate the types of data users are working with

- ♦ To perform a review of the food classification system and ensure it is compatible with others

- ♦ To include products beyond food (e.g. Animal reservoirs, OIE datasetsetc…) which could be also useful for food safety professionals

- ♦ To highlight gaps in data availability in specific disciplines and/or regions

- ♦ To track the use of data for decision-making.

7.2 Analytical data

- ♦ The table elaborated by the working group and reported in chapter 5 should be used to evaluate the presence of various data elements and to classify data sources in three tiers.

- ♦ Tier 1 is a minimum requirement for linking relevant datasets to to FOSCOLLAB

♦   WHO should develop guidance on the use of shared data to avoid the potential misuse of analytical data

♦   WHO should further explore standardized data interchange formats (for example, autofill XML files that can be instantly uploaded into iRisk or another tool)

♦   Future developments of FOSCOLLAB must take into account emerging technologies and work with relevant communities e.g. Genomics Data

♦   The integration of data on predictive microbiology should be discussed to avoid duplication of other tools already available

♦   FDA and PAHO should develop a pilot to leverage systems including INFAL and eLEXNET to demonstrate how aggregated, quality data can support FOSCOLLAB and users.

7.3 Non-analytical data

♦   WHO should increase the integration of non-analytical data and explore the possibility for connecting with INFAL

♦   In addition to risk assessment and risk management goals FOSCOLLAB should further explore its use for capacity development to improve sampling, mechanism for reporting and addressing gaps

8   Appendices

8.1 Members of the Working Group in attendance

Dr Paul DANSTED – Chair
Manager Food Assurance
Market Assurance Directorate, Standards Branch, Ministry for Primary Industries
New Zealand

Dr Mary GILSENAN
Head of Evidence Management Unit
European Food Safety Authority (EFSA)
Italy

Dr Karen KEDDY
Centre for Enteric Diseases, National Institute for Communicable Diseases
Division of the National Health Laboratory Service
South Africa

Mr Greg PAOLI
Principal Risk Scientist, COO - Risk Sciences International
Canada

Dr Pisan PONGSAPITCH
Director of Office of Standard Development
National Bureau of Agricultural Commodity and Food Standards - Ministry of Agriculture and Coopera-
tive
Thailand

Mr Carl SCIACCHITANO
Senior Advisor For Scientific International Affairs
Office of the Commissioner, U.S. Food and Drug Administration
United States of America

Dr Ying XIAO
Food Safety Officer (Risk Assessment)
Center for Food Safety, Hong Kong SAR
People's Republic of China

8.2 WHO Secretariat

Mr Tim CORRIGAN
Project Officer
Department of Food Safety and Zoonoses, World Health Organization
Switzerland

Dr Philippe VERGER
Scientist
Department of Food Safety and Zoonoses, World Health Organization
Switzerland

8.3 Observers

Ms Rachel Fefer
Office of the Commissioner, U.S. Food and Drug Administration
United States of America

Ms Kanchana Iyer
Office of the Commissioner, U.S. Food and Drug Administration
United States of America

Dr Ben Tall
U.S. Food and Drug Administration
United States of America

WHO/HSE/FOS/2015.2