CODEX ALIMENTARIUS COMMISSION **E**

Food and Agriculture Organization of the United Nations

World Health Organization

Viale delle Terme di Caracalla, 00153 Rome, Italy - Tel: (+39) 06 57051 - Fax: (+39) 06 5705 4593 - E-mail: codex@fao.org - www.codexalimentarius.net

JOINT FAO/WHO FOOD STANDARDS PROGRAMME
CODEX COMMITTEE ON CONTAMINANTS IN FOODS
9th Session
New Delhi, India, 16 – 20 March 2015
DISCUSSION PAPER ON SUBMISSION AND USE OF DATA FROM GEMS/Food
(including Guidance document for extraction of data on chemical contaminants in food and the diet)
(Prepared by WHO
with inputs from FAO and the Codex Secretariat)

Background

1.   The purpose of the GEMS/Food database is to allow a web-based submission of data on food contamination from different countries and institutions, and to inform the Codex Alimentarius Commission and other interested parties on the levels and trends of contaminants in food and their contribution to the total human exposure.

2.   Risk management decisions e.g. the establishment of food safety standards such as the establishment of maximum levels for contaminants in food and feed, are highly dependent on comparable and reliable data on chemical levels in foods and estimates of dietary exposure. Data can be extracted from the GEMS/Food database in an excel format, which allows data users to analyze the distribution of occurrence and to calculate statistical descriptors (e.g. mean, median, percentiles, maximum).

3.   There have been cases where electronic working groups (EWGs) worked in consultation with the JECFA Secretariat to develop a call for data inviting Codex members to submit data to GEMS/Foods while other EWGs developed their own template(s) and requested data from members of the EWG for the purpose of developing proposals for MLs for contaminants in foods for consideration by CCCF. There have also been overlapping situations by which two requests for data have been issued for the same purpose, one developed by the EWG with a specific template and another issued by JECFA with the GEMS/Foods template which may duplicate work and create confusion among Codex members.

4.   During the 8th Session of the Committee on Contaminants in Foods, the Representative of WHO informed the Committee on new achievements of the GEMS/Food program and emphasized the importance of linking this to the work of CCCF. Data collected by EWGs needed to be included in the GEMS/food database and EWGs need to be able to extract and analyze data from this global database. In order to assure best linkage and assure consistent use of data, the Representative of WHO recommended the elaboration of a guidance document for EWGs on submission and use of data from GEMS/Food.

5.   The Committee agreed with this proposal and requested the GEMS/Food Secretariat to develop such a guidance paper, in collaboration with FAO and EWG chairs whose work includes data collection and analysis for consideration at its next session.[1]

Conclusions and Recommendations

6.   The GEMS/Food Secretariat prepared a guidance document attached as Appendix I.

7.   This document will assist Codex members to submit and use data from the GEMS/Food database. In doing so, the guidance document will ensure a consistent approach to data collection by using the GEMS/Food as the primary reference for data collection. The guidance document can be made available to EWGs when developing work that requires data collection and analysis to facilitate uploading and downloading of data to/from GEMS/Foods.

8.   When additional data/information are needed to fulfill the task of EWGs, separate templates for specific information required by the EWG could be developed in consultation with the JECFA Secretariat.

9.   The CCCF is invited to agree that when development of MLs required data collection, such request should be prepared in accordance with the GEMS/Food template using the guidance document. Additional data may be gathered through supplementary templates developed in consultation with the JECFA Secretariat. Such templates should not request information that can be collected through the GEMS/Food template.

---

[1] REP14/CF, paras 13-14

GUIDANCE DOCUMENT FOR EXTRACTION OF DATA ON CHEMICAL CONTAMINANTS IN FOOD AND THE DIET

Table of Contents

1 Preamble: Access to OPAL web[2]

- To create a WHO account, please complete the following online registration form: https://extranet.who.int/ads/adswebinterface/create.aspx

- Once you have successfully registered and activated your account please send your new username by e.mail to vergerp@who.int in order to complete the creation of your new GEMS/Food database account.

- Once you registered you can access the GEMS/Food website: https://extranet.who.int/gemsfood/ and you should log in to the database using your new WHO account created in step 1 by clicking the **Login** link in the top right of the home page.

- Once you have logged in to the database, you will have access to the **Search** tab for access to the data.

2 Introduction

2.1. Purpose of the GEMS/Food Database

The GEMS/Food database is the main outcome of the International GEMS/Food Programme administered by the World Health Organization Head Quarter (Geneva) and collecting contaminants concentration in food. The purpose of the GEMS/Food database is to allow a web-based submission of data on food contamination from different countries and institutions, and to inform the Codex Alimentarius Commission and other interested parties on the levels and trends of contaminants in food and their contribution to total human exposure. Food contamination monitoring is an essential component of assuring the safety of food supplies and managing health risks at the international level. The GEMS/Food database is open to Competent Authorities[3] to submit and share their data on food surveillance and monitoring.

Risk management decisions e.g. the establishment of food standards such as maximum levels (MLs) in Codex Alimentarius, are highly dependent on comparable and reliable data on contaminant levels in foods and estimates of dietary exposure. Data can be extracted from the GEMS/Food database in an Excel format, which allows data users to analyse the distribution of occurrence and to calculate statistical descriptors (e.g. mean, median, percentiles, maximum).

---

[2] Important note: it is possible to perform a search for data without login into the system but in that case a part of the information classified as confidential will not appear in the extracted data e.g. name of country generating the data.
[3] Competent Authorities in that case are Institutions recognized by their National Ministry of Health to collect data on food contamination and to submit these data to International Organizations.

2.2. Objective of the document

In recent years the amount of data submitted by Member States to the Codex Committee on Contaminants in Food (CCCF) has increased, as has the need for robust data analysis to support the Standard Setting process. The GEMS/Food database provides a unique tool to allow a first level of harmonization of international data and to facilitate data sharing and data analysis in a transparent and sustainable way. A complete harmonization of data submitted (e.g. unique categorization system, similar Limits of Detection…) is not fully achievable between national Authorities, however it is important to develop good practices for combining and analysing data in a consistent way across different working groups. The overall objective of the current document is to propose a harmonized way to extract and to analyse data on contaminants in food from the GEMS database for the purpose of CCCF and related bodies.

3   Protocols for the extraction of contaminant concentration data

Experts authorised by WHO should access the OPAL website https://extranet.who.int/gemsfood/ and login using the user name and password provided by the WHO GEMS/Food manager.

The home page of the GEMS/Food website provides useful information about the programme as well as access to published reports.

3.1. Data search

Data should be extracted from the GEMS/Food database from the SEARCH tab.

The search can be designed by the selection of several criteria:

- WHO Region(s)

- Contaminant(s)

- Country(s)

- Food Category(s)

- Food Name(s)

Regarding the other criteria. i.e. "batch number" and/or "sampling period", their use for an initial search is not recommended.

3.1.1.   The region(s) of interest

It is possible to choose all "*WHO Region(s)*", several WHO Regions or one WHO Region according to your goal. The extraction can also be focused on one or several "*Country(s)*".

3.1.2.   The contaminant

It is possible to select one or several "*Contaminant(s)*". However separate searches for each contaminant are recommended.

If a hazard consists of several congeners or is known in different name(s), a broad selection and a check for all the congeners or related substances in the database is recommended. Example: <u>if the search is about</u> "**Aflatoxin**", all relevant names such as "**Aflatoxin (total)**", "**Aflatoxin B1**", "**Aflatoxin B1 and B2**"… should be selected.

3.1.3.   The food category (s)

Foods are classified based on food commodity descriptors used in the *Codex Alimentarius* Committees. When submitting data, data providers convert the local food name into a standardised food name. Because some providers may misclassify foods, a first search at a broad level i.e. "*Food Category(s)*" is recommended rather than a more focused search by "food name". Then, it's possible to refine the initial search within the Excel sheet (the procedure is explained below: part 3.3.1).

3.1.4.   The results of the search

As a first result of the search, **a summary of the number of records** available is provided. This allows to determine the number of records for each possible combination of contaminant and food and the future possibilities for data analysis.

Summary: For the initial search, the use of broad criteria are recommended to include all relevant data. Subsequently, it is possible to refine the search in a second step (see 3.3). An initial search that is too narrow may exclude relevant data.

3.2. Data extraction

The search results can be printed and/or saved to an Excel file. The recommendations for data extraction and analysis are described below.

The results can be exported to an Excel file under 2 formats: xls (version 97-2003) or xlsx (version 2010). The xlsx version is mandatory for searches that result in more than 50,000 records. If Excel 2010 is not available on your computer, the search should be restricted to get fewer than 50,000 lines. This can be done by un-selecting region(s), food(s) and/or contaminant(s).

Contaminant	Food Category	# of record(s)
Aflatoxin (total)	Cereals and cereal-based products	7936
Aflatoxin B1	Cereals and cereal-based products	12260
Aflatoxin B1 and B2	Cereals and cereal-based products	669
Aflatoxin B2	Cereals and cereal-based products	7796
Aflatoxin G1	Cereals and cereal-based products	7811
Aflatoxin G2	Cereals and cereal-based products	7818
Aflatoxin M1	Cereals and cereal-based products	8

Search Results   Export to excel | Export to excel (xlsx) | Print | Get link to this search | Email this search

Record Type	Region	Country	Contaminant	Food Group		WHO Food Identifier	WHO Food Code
Individual	WHO Western Pacific Region	Japan	Aflatoxin B1	Cereals and cereal-based products	Maize		GC 0645
Individual	WHO Western Pacific Region	Japan	Aflatoxin B1	Cereals and cereal-based products	Maize		GC 0645
Individual	WHO Western Pacific Region	Japan	Aflatoxin B1	Cereals and cereal-based products	Maize		GC 0645

Summary: The Excel worksheet containing data extracted from GEMS should be saved before making any additional selections or edits.

3.3. Data analysis

3.3.1. Detecting mapping errors

Each record extracted by food category in the excel sheet are classified with a CODEX code ("*FoodCode*": column G) and a CODEX name ("*FoodName*": column F). This classification corresponds to the detailed food commodity descriptors used in the *Codex Alimentarius* Committees[4].

As the Codex classification system does not cover all kinds of food, different food items or produces may use the same generic name. Therefore, the local food name ("*LocalFoodName*": column H) may indicate more detailed information. The local food name is the sample description provided by data submitters from their own countries' national databases in the original language and mapped with Codex name into the templates for submission. The local food name is therefore not harmonized between countries but represents the best description of the food by the data submitter/national data manager. When reviewing lines of data in the Excel file, the data user should check the "*LocalFoodName*" field to confirm that the results belong in the desired dataset.

It should also be noted that the "*Remarks*" field (Column AR) can also contain useful information related to the sample description.

Important: In case of doubt on possible misclassification for certain samples, the manager of the GEMS/Food database can contact the data provider for clarification.

---

[4] If the local food name is a generic name e.g. "Fish" it could have been mapped with a "*Food Name*", in capital-letter and corresponding to the food category i.e. FISHES. In such a case no refinement is possible.

	A	B	C	D	E	F	G	H	I	J	K	L
1	RecordType	RegionName	CountryName	ContaminantName	FoodCategory	FoodName	FoodCode	LocalFoodName	FoodStateName	ResultValue	ResultText	UnitName
2	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Maize	GC 0645	toumorokoshi	Raw	0	ND	ug/kg
3	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Maize	GC 0645	toumorokoshi	Raw	0	ND	ug/kg
4	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Maize	GC 0645	toumorokoshi	Raw	0	ND	ug/kg
5	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Cereals and	A.01	korngurutenmiru	Raw	0	ND	ug/kg
6	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Barley	GC 0640	oomugi	Raw	0	ND	ug/kg
7	Individual	WHO Western	Japan	Aflatoxin B1	Cereals and ce	Bran, unpro	CM 0081	fusuma	Raw	0	ND	ug/kg
8	Individual	WHO Western	Japan	Aflatoxin B1	Cereals and ce	Cereals and	A.01	korngurutenmiru	Raw	0	ND	ug/kg
9	Individual	WHO Western	Japan	Aflatoxin B1	Cereals and ce	Barley	GC 0640	oomugi	Raw	0	ND	ug/kg
10	Individual	WHO Western	Japan	Aflatoxin B1	Cereals and ce	Maize	GC 0645	toumorokoshi	Raw	0	ND	ug/kg
11	Individual	WHO Western	Japan	Aflatoxin B1	Cereals and ce	Maize	GC 0645	toumorokoshi	Raw	0	ND	ug/kg
12	Individual	WHO Western	Japan	Aflatoxin B1	Cereals and ce	Maize	GC 0645	toumorokoshi	Raw	0	ND	ug/kg
13	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Bran, unpro	CM 0081	fusuma	Raw	0	ND	ug/kg
14	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Cereals and	A.01	korngurutenmiru	Raw	5	5	ug/kg
15	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Cereals and	A.01	toumorokoshijis	Raw	2.3	2.3	ug/kg
16	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Cereals and	A.01	korngurutenfido	Raw	4	4	ug/kg
17	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Cereals and	A.01	korngurutenfido	Raw	6	6	ug/kg
18	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Cereals and	A.01	korngurutenfido	Raw	0	ND	ug/kg
19	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Maize	GC 0645	toumorokoshi	Raw	3	3	ug/kg
20	Individual	WHO Western	Japan	Aflatoxin (total)	Cereals and ce	Maize	GC 0645	toumorokoshi	Raw	0.5	0.5	ug/kg

3.3.2.   Refining the dataset

Some data should be excluded if they don't fit for the purpose of the assessment: It's possible to sort and filter the "*FoodName*" column to identify one food name in particular (e.g. in the food category "*Milk*" the food name "*milk powder"* can be out of the scope of the analysis).

It is also possible to refine the first search, by returning to the SEARCH tab and refining the dataset based on "*Food Name*" (column F).

It is also possible to refine the dataset based on "*FoodStateName*" (column I) and on other characteristics of the samples (e.g. "s*amplingPeriod*": column Y). However, keep in mind that different countries may have interpreted "*results based on*" (Column AL) and "*state of food analysed*" (Column I) differently, so be cautious when interpreting the entries in these columns to mean that results are raw/powdered/diluted/cooked/etc. (Refer to the Annex for a list of the fields and their correspondence with column's numbering).

Important: Data extracted by the search and considered not relevant for the analysis should be kept in a separate tab in order to allow traceability of the process.

> *Summary: After an initial extraction based on broad categories i.e. "Food Category(s)" (column E), it is important to check the data for consistency and fit for purpose: check the "FoodName" column (column F), "LocalFoodName" column (column H) and "Remark" (column AR).*
>
> *Refinement of the dataset can also be based on other parameters of the samples (see annex).*

3.3.3.   Analysing results reported to be below LOD or LOQ

Analytical results below the Limit of Determination (LOD) and sometimes also results below the Limit of Quantification (LOQ) are reported as "0" in the column "*ResultValue*" (column J).

It is important to evaluate and to report in the results of the assessment, the proportion of results below the LOD and LOQ.

When using data to evaluate proposed MLs, the limit of quantification should be compared with the proposed MLs: analytical results based on LOQs above the current or proposed MLs should be withdrawn from the analysis.

3.3.4.   Individual vs Aggregated Contaminant Concentration data

Datasets extracted from GEMS include both individual (1 record - 1 sample) and aggregated (1 record - several samples) data. Depending upon the purpose of the data analysis, the aggregated results can be integrated with individual results, analysed separately or excluded.

In column A, filter and visualise the number of aggregated records in order to help your decision.

If the decision is to analyse separately or to exclude these data, select the rows corresponding to aggregated data, then cut and paste them in a separate tab.

If the decision is to integrate these data together with the individual data, follow the steps below:

- For "aggregated" data, the columns "*ResultValue*" (column J) is empty and the values in <u>column S</u> (called "mean") should be used as the best estimate of the contamination for the aggregated samples.

- The data should be selected in <u>column S</u>, copied and pasted in <u>column J</u>.

In certain cases aggregated results can be weighted by the number of individual samples into the pooled sample (e.g. for dietary exposure assessment). In other cases they could/should be excluded of the dataset (e.g. for probabilistic assessment).

3.3.5. Harmonising the units

Analytical results (<u>column L</u>) as well as LODs and LOQs, can be expressed in different units (e.g., mg/kg or µg/kg). For the analysis, figures need to be converted into a single unit. Converting the whole dataset to the smallest unit is recommended to avoid using decimal numbers as much as possible.

- The "*UnitName*" column (column L), once selected, should be "Sort A to Z" to visualise each of the units appearing in the dataset

- A new column should be inserted just before the"*UnitName*" column and named "*ResultValue - CONVERTED*" (new column L).

- In the new column "*ResultValue- CONVERTED*", insert all the results to be converted multiplied by the relevant factor (For example: results in *mg/kg* should be multiplied by 1000 to be converted to *µg/kg* (L2=J2*1000) in the new column)

- In the new column "*ResultValue- CONVERTED*", insert all results already in the desired unit without using any conversion factor (Lx=Jx).

Important: It is recommended to use different colours to highlight the results reported in different units in order to facilitate a quality check.

> *Summary: A uniform analysis should apply to harmonize the dataset:*
>
> *1/ to quantify the proportion of results below the LOD/LOQ in the considered dataset.*
>
> *2/ to identify results obtained on pool samples vs individual samples and*
>
> *3/ to harmonize the units for chemical concentration.*

Harmonization of concentration units

Step 1:

Step 2:

Step 3:

Step 4:

Step 5:

3   Conclusions

The resulting dataset can be used to provide a description of the occurrence of contaminants in various foodstuffs and to facilitate establishment of standards and/or MLs. It can also be used to simulate the impact of various MLs on the number of lots to be removed from the International market.

In parallel, WHO should continue to improve the harmonization of data submitted to GEMS/Food in close collaboration with Collaborating Institutions. The guidance on submitting data to GEMS should be updated based on the experience of data users.

ANNEX: GEMS/Database – Data Extraction File Format

Following is the list of columns and field names in the data files extracted from the GEMS/Food database.

Column	Field name in extracted file	Field related to individual, aggregated or both record types	Convert all to same unit (e.g., mg/kg)	Fields with food descriptors	Check to ensure consistency of combined data
A	Record Type	Both			
B	Region Name	Both			
C	Country Name	Both			
D	Contaminant Name	Both			
E	Food Category	Both		√	
F	Food Name	Both		√	
G	Food Code	Both		√	
H	Local Food Name	Both		√	
I	Food State Name	Both			√
J	Result Value	Individual	√		
K	Result Text	Individual	√		
L	Unit Name	Both			
M	LOD	Individual	√		
N	LOQ	Individual	√		
O	LOD (Min)	Aggregated	√		
P	LOD (Max)	Aggregated	√		
Q	LOQ (Min)	Aggregated	√		
R	LOQ (Max)	Aggregated	√		
S	Mean	Aggregated	√		
T	Mean Lower	Aggregated	√		
U	Mean Upper	Aggregated	√		
V	Median	Aggregated	√		
W	90th Percentile	Aggregated	√		
X	Standard Deviation	Aggregated	√		
Y	Sampling Period Start	Both			
Z	Sampling Period Start Text	Both			
AA	Sampling Period End	Both			
AB	Sampling Period End Text	Both			
AC	Representativeness Name	Both			

Column	Field name in extracted file	Field related to individual, aggregated or both record types	Convert all to same unit (e.g., mg/kg)	Fields with food descriptors	Check to ensure consistency of combined data
AD	Lab Count	Both			
AE	Lab Number	Both			
AF	Food Origin Name	Both			
AG	Analytical QA Name	Both			
AH	Sample Count	Both			
AI	Samples Below LOQ	Both			
AJ	Range (Min)	Aggregated	√		
AK	Range (Max)	Aggregated	√		
AL	Result Basis Name	Both			√
AM	Portion Type Name	Both			√
AN	Serial Number	Both			
AO	Batch ID	Both			
AP	Is Confidential	Both			
AQ	Submission Date	Both			
AR	Remarks	Both			
AS	Row Num	Both			