GUIDANCE OF EFSA

Standard sample description for food and feed[1]

European Food Safety Authority[2, 3]

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

Data collection is an important task of EFSA[4] and a fundamental component of risk assessment (Articles 22 and 23 of Regulation EC No 178/2002). The Technical Working Group on Data Collection (TWG-DC) developed a guideline on the standard description of samples and analytical results (Standard Sample Description). The document provides specifications aimed at harmonising the collection from Member States of analytical measurement data for the presence of harmful or beneficial chemical substances in food, feed and water. The standard sample description includes a list of standardised data elements (items describing characteristics of samples or analytical results such as country of origin, product, analytical method, limit of detection, result, etc…), controlled terminologies and validation rules to enhance data quality. These can be used both by data providers and data recipients to accurately describe analytical samples for evaluation purposes. This work intends to develop a generalised model to harmonise the collection of a wide range of measurements in the area of food and feed safety assessment.

KEY WORDS

Standard, Format, Chemical, Occurrence, Residues, transmission, XML, Data, Pesticides

---

1 On request of EFSA, Question No EFSA-Q-2009-00698, issued on 18 January 2010.

2 Correspondence: datex@efsa.europa.eu

3 The European Food Safety Authority wishes to thank the members of the Working Group on Data Collection for the preparation of this guidance of EFSA: Jens Hinge Andersen (National Food Institute, Technical University of Denmark - Denmark), Eileen O'Dea (Food Safety Authority of Ireland - Ireland), Luisa Oliviera (Instituto Nacional de Saúde Dr. Ricardo Jorge - Portugal), Jean Cédric Reninger (Agence Française de Sécurité Sanitaire des Aliments - France), Renata Del Rosario (ESTAT), Stijn Saevels (FASFC - Belgium), Lars Wiehle (Bundesamt für Verbraucherschutz und Lebensmittelsicherheit), Josef Wolf (Austrian Agency for Health and Food Safety - Austria), Roger Wood (Food Standards Agency – United Kingdom). In addition, EFSA wishes to thank EFSA's staff members Fabrizio Abbinante, Daniela Brocca, Stefano Cappè, Jane Richardson for their contributions to this guidance of EFSA.

4 REGULATION (EC) No 178/2002 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety OJ L 31, 1.2.2002, p. 1–24.

For citation purposes: European Food Safety Authority; Standard sample description for food and feed. EFSA Journal 2010;8(1):1457 [54 pp.]. doi:10.2903/j.efsa.2010.1457. Available online: www.efsa.europa.eu

SUMMARY

Data collection is an important task of EFSA[5] and a fundamental component of risk assessment (Articles 22 and 23 of Regulation EC No 178/2002). EFSA is receiving from different providers (Member States, the European Commission, industry etc…) an increasing volume of data in support of its scientific activities.

The Technical Working Group on Data Collection (TWG-DC) was mandated by EFSA to develop a proposal

to harmonise the collection of analytical measurement data for the presence of harmful or beneficial chemical substances in food, feed and water.

The TWG-DC was requested to produce guidance documents on:

- the harmonised description of data on analytical measurements in food and feed samples *(Guidance on Standard Sample Description for Food and Feed)*: including a list of standardised data elements (items describing characteristics of samples or analytical results such as country of origin, product, analytical method, limit of detection, result, etc…), controlled terminologies and validation rules to enhance data quality.;
- the procedures to efficiently transmit and exchange data between Member States and EFSA (Guidance on data exchange) taking care of selecting specific file formats for data transmission (e.g. XML, Microsoft Excel etc…) and specific data transmission protocols to support electronic data exchange..

The Guidance on Standard Sample Description for Food and Feed, specifies the data elements and the data structure of the samples and the analytical results for chemical contaminants and residues in food and feed included in monitoring and control programmes (e.g. sample description, analytical methods and the analytical results). The TWG-DC aimed to build a description as general as possible to facilitate its application to a wide range of measurements taken for food and feed safety assessment.

The TWG-DC agreed that a Standard Description, defining variables and terminologies, and a standard data transmission format and mechanism are only the first step to support the harmonisation of the data transmissions between EFSA and Member States. The key element for a successful implementation of the Standard Sample Description is to set up a maintenance and evolution programme of the Standard Sample Description. This process will be needed to continue the process of adapting and improving the Standard to areas not currently in the scope of this version of the Standard Sample Description. Whilst this version of Standard Sample Description has been developed specifically to address transmission of Chemical Occurrence and Pesticides data, to date it has only been piloted in the Pesticides domain (2008 Annual Data Collection). Feedback from this experience has been incorporated into the standard data model. Further experience in this and other domains will contribute to the enhancement of the data model over time. Further, the TWG-DC highlights that an evaluation of the Standard Sample Description's applicability to collection of data in the Zoonoses domain will also be needed.

[5] REGULATION (EC) No 178/2002 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety OJ L 31, 1.2.2002, p. 1–24.

The group recognises that the ability of each Member State to transmit data to EFSA according to the standard data model will vary. Therefore it should be also viewed as guidance for Member States to be used when planning future developments and evolution of local, regional and national systems with the objective of harmonising data transmissions.

The harmonisation of data collection is recognised as a fundamental step for the development of an effective EFSA Data Warehouse.

1. TABLE OF CONTENTS

2.

BACKGROUND AS PROVIDED BY EFSA

Data collection is an important task of EFSA[6] and a fundamental component of risk assessment (Articles 22 and 23 of Regulation EC No 178/2002). EFSA is receiving from different providers an increasing volume of data in support of its scientific activities. Data may be submitted to EFSA by a variety of providers: national competent authorities, local authorities, laboratories, universities and others.

The increasing volume of data transmitted to EFSA involves a number of challenges:

Efficient use of human resources in data collection processes, both on the side of the data providers and on the EFSA side where data have to be collected, collated and analysed.

Quality of transmitted data, which may differ in origin, language, description and codification. A standard data quality facilitates the task of collating and preparing data for the analysis.

Capacity of managing high volumes of data; many data collections contain a huge quantity of analytical results, such as data from control monitoring programs and targeted surveys. The quantity of these data makes their manual processing unfeasible and introduces the need for automated methods of formatting, transmitting, processing and analysing the data.

Capacity to analyse the data and to produce valuable reports for the different stakeholders; the increased volume of data requires more advanced techniques for storing and analysing the data. Data should be as much as possible readily available for standardised analyses functions to harmonise processing among the different food data stakeholders.

EFSA is requesting the Data Collection and Exposure Unit, through a self-tasking mandate, to harmonise the collection of analytical measurement data for the presence of harmful or beneficial chemical substances in food, feed and water.

The Data Collection and Exposure Unit in coordination and cooperation with the Pesticide Risk Assessment (PRAPeR) Unit should establish, for this purpose, a working group of experts (Technical Working Group on Data Collection).

The working group should report to the Chemical Occurrence Expert Group and to the Networking Group on Pesticide Residues. These networking groups will review and approve all deliverables of the working group.

[6] REGULATION (EC) No 178/2002 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety OJ L 31, 1.2.2002, p. 1–24.

3. TERMS OF REFERENCE AS PROVIDED BY EFSA

The working group should focus on both chemical occurrence and pesticide residue data, due to the close similarities between these areas. The working group should build on, expand further and finalise the effort of harmonising the analytical measurement data descriptions already started by the EFSA internal working group on controlled terminology and define the data exchange model format to be used.

The working group is requested to produce guidance documents on:

- the harmonised description of data on analytical measurements in food and feed samples *(Guidance on standard sample description for food and feed)*;
- the methods to efficiently transmit and exchange data between Member States and EFSA (*Guidance on data exchange*)[7].

The guidance documents will be presented for approval to the EFSA Chemical Occurrence Expert Group and to the EFSA Networking Group on Pesticide Residues.

Standard sample description

The standard sample description should contain a list of data elements that are standardised and can be conveniently used by both data providers and data recipients to fully describe samples and analytical parameters for evaluation purposes. This standard should provide:

- name and structure of the defined data elements to be used;
- controlled terminologies, where needed, with exclusion of the food dictionary which will be addressed by a specific working group;
- validation rules to assess the validity of the information supplied, to ensure an adequate level of data quality in data export, transmission and storage.

The proposal for standardised sample description should be based on the initiative developed by the EFSA internal Working Group on Controlled Terminology. The group elaborated a draft standard for the description of chemical contaminants and pesticide residues data (Annex II). This document has been presented at the IT Expert Meeting on 17-18 February 2009 and will be the basis for a pilot project on pesticide residues data transmission and for an EFSA grant on chemical contaminant data.

Methods for efficient transmission and exchange of data between Member States and EFSA[7]

The working group should propose a "Guidance on Data Exchange" to define methods for the transmission and the exchange of data which would be suitable for the data collected with the standard sample description for food and feed. The guidance should take into account the software tools already set up by the EFSA IT Unit and provide inputs for their enhancement. The guidance should promote the use of semi-automated or automated methods for the data transmission to lower the costs of manual human intervention.

[7] The "Guidance on data exchange" is part of the mandate of the technical working group on data collection but will be addressed on a separate document.

4. ACKNOWLEDGEMENTS

The European Food Safety Authority wishes to thank the members of the Working Group on Data Collection for the preparation of this guidance of EFSA: Jens Hinge Andersen (National Food Institute, Technical University of Denmark - Denmark), Eileen O'Dea (Food Safety Authority of Ireland - Ireland), Luisa Oliviera (Instituto Nacional de Saúde Dr. Ricardo Jorge - Portugal), Jean Cédric Reninger (Agence Française de Sécurité Sanitaire des Aliments - France), Renata Del Rosario (ESTAT), Stijn Saevels (FASFC - Belgium), Lars Wiehle (Bundesamt für Verbraucherschutz und Lebensmittelsicherheit), Josef Wolf (Austrian Agency for Health and Food Safety - Austria), Roger Wood (Food Standards Agency – United Kingdom). In addition, EFSA wishes to thank EFSA's staff members Fabrizio Abbinante, Daniela Brocca, Stefano Cappè, Jane Richardson for their technical and scientific contributions to this guidance of EFSA.

5. INTRODUCTION

The *Standard Sample Description for Food and Feed* (hereinafter referred to as 'Standard Description') specifies the data elements and the data structure to describe samples and results coming from analytical measurements in food and feed. The guidance focuses on the definition of a logical model which is independent from the file format. Therefore data providers and receivers can use different file formats – e.g. Microsoft Excel, Comma Separated Values (CSV), Extensible Markup Language (XML), etc… – to submit transmissions depending on their technological constraints. The definition of the supported file formats and the actual implementation of the standard logical model will be discussed in a separate guidance of the Technical Working Group on Data Collection: "Guidance on Data Exchange".
The specification of a logical model for Standard Description is composed of:

a) data elements definition and structure,

b) controlled terminologies,

c) validation rules to assess the validity of the information supplied.

a. DATA ELEMENTS DEFINITION AND STRUCTURE

The data elements are referenced by a sequential alphanumeric code. A unique element name is provided; this is to be used for column names, field names and tags depending on the software programs, files or databases implementing the Standard Description. The unique element name is composed only of characters from the Roman alphabet and does not include spaces or any other special characters. The unique element names should be considered case sensitive[8] to ensure compatibility with information systems especially XML standards. The data elements are described also by a label to be used in reports, print outs or in the graphical interfaces of the software programs that will manage the Standard Description. A data type is associated to each data element and it

[8] Distinguishing upper- and lower-case letters. Often used in computer science to indicate a distinction is made in comparison or equality of letters based on case. For example, a case-sensitive password will not recognize "Password" and "password" as the same, but a case insensitive comparison would

defines the values that it can contain. Data types will be defined using the W3C XML schemas data types specification[9].

The Standard Description also defines the data structure that is the sequence and relationships between the elements for effective information exchange.

b. CONTROLLED TERMINOLOGIES

The Standard Description includes controlled terminologies. A controlled terminology is a finite, enumerated set of terms intended to convey information unambiguously. The use of controlled terminologies facilitates the aggregation of data during analysis and ensures comparability between datasets. Controlled terminologies are language independent, only the code needs to be returned, the description of the code in any language can be linked to the code. To ensure that the data are of sufficient quality to allow analysis at EU level, controlled terminologies have been applied extensively in the Standard Description. For some elements with controlled terminologies an additional companion free text element is included. This allows the provision of relevant information when the controlled terminology is insufficient to fully characterise the item described.

c. VALIDATION RULES

Validation rules can either check the validity of a value reported in a individual data element (single data element validation) or they can check inter-dependent values reported in more data elements (inter dependent data element validation).

- *Single data element validation*: checks on specific data elements e.g. the verification whether the code reported in a field constrained to a controlled terminology is correct, or if the value reported is within a certain acceptance range.

- *Inter-dependent data element validation*: checks on two or more data elements for or instance if the result is reported to be below the LOQ that the LOQ has been supplied.

Certain validation rules are specific to the legislation or the project under which the data is collected and should be defined data collection specifically. Only those business rules which apply to all data collections will be described in this document.

[9] W3C recommendation, "XML Schema Part 2: Datatypes Second Edition", 28 October 2004, http://www.w3.org/TR/xmlschema-2/ (last access 22/10/2009)

6. BACKGROUND TO THE DEVELOPMENT OF THE STANDARD DESCRIPTION

The Standard Description was originally based on proposals received within the Article 36 grant "Development of a Standard Food Classification and Sample Description System for Chemical Occurrence Data Storage" (CFP/EFSA/DATEX/2007/02) awarded to the Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL). An initial draft Standard Description was then defined by an EFSA Internal Working involving the Assessment Methodology unit (AMU), the Data Collection and Exposure unit (DATEX), and the Pesticide Risk Assessment Peer Review unit (PRAPeR) and Zoonoses unit, with purpose of extending it particularly to the requirements for pesticide residues transmissions. The draft data model was presented for comments at the IT Expert Meeting on 18-19 February, 2009. A revised model was published on 3 April 2009 after consideration of received comments. To test the data model six Member States participated to a pilot submitting the 2008 pesticide monitoring data using the draft Standard Description and an XML file submission format. The participating Member States were Austria, Denmark, Germany, Ireland, The Netherlands and Slovenia. The Member States submitted data in July 2009 and the pilot project was reviewed in autumn 2009. Recommendations for improving the draft Standard Description, which were part of the outcome of the pilot, were addressed by the Technical Working Group on Data Collection as described in paragraph 4.

7. PESTICIDE PILOT OUTCOME

On the 1 Sept 2008 under EC Regulation 396/2005 the responsibility for the collation of pesticides monitoring data and the Annual Report on Pesticides Residues was transferred from the Food and Veterinary Office (FVO) to the European Food Safety Agency (EFSA).

Currently the required information is submitted using formatted Excel workbooks. These workbooks are complex to complete, difficult to validate and provide insufficient information to perform a robust exposure assessment. The development of the Standard Description provided the opportunity to explore methods to improve the data collection process.

In the IT Meeting on Data Collection for Pesticide Residues (18 February 2009, Parma) six Member States agreed to participate in a pilot project to submit the 2008 monitoring data via the Data Collection Framework (DCF) according to the standard data model for analytical measurements in food and feed. The participating reporting organisations were Department of Agriculture, Fisheries and Food (Ireland), Food and Consumer Product Safety Authority, VWA (The Netherlands), Austrian Agency for Health and Food Safety, Ministry of Health, Health Inspectorate of the Republic of Slovenia, National Food Institute, DTU (Denmark) and Federal Office of Consumer Protection and Food Safety (Germany).

All participants were able to submit the 2008 pesticide monitoring data according to the standard data model in XML format via the DCF. The following elements in the data model were identified for amendments or clarification of the completion instructions:

- "Product treatment": this field should be a mandatory field at least for pesticide transmissions and a dictionary should be developed according to OECD and CODEX recommendations for the application of processing factors. Please note that a harmonised food classification for processed products is not available yet.

- "Parameter code": To ensure comparability at EU level results of the EU coordinated programme results should be reported according to the EC MRL residue definition (with conversion factors applied for summed residue definitions). Clear instructions need to be given in the correct PARAMCLS codes to be used, specifically that codes from "Residue definition" substance category should be used when reporting the results of the EU coordinated programme.

- "Possible reason for non compliance": reporting of this data element was poor. The requirement for this information is specified in Reg. 396/2005 Article 32. Reporting organisations need to explore methods for improving completion of this element.

- "Result evaluation": This element field should be mandatory to enable verification of the rate of compliance and ensure compatibility with published national results.

The size of the XML transmissions presented a challenge during the pilot project. Splitting the results into smaller manageable portions (batching e.g. by month or number of records) resolved some issues, however improvements to the EFSA information system are required to improve processing of these transmissions. In addition the use of short element names could significantly reduce file size.

8. SCOPE OF THE STANDARD SAMPLE DESCRIPTION

The Standard Description is targeted to support the data collection and the data transmission of the samples and the results of analytical measurement to support exposure assessments for food and feed safety. The legislation taken into consideration for the design and specification of the standard logical model were the chemical contaminants (e.g. Chemicals included in Regulation (EC) No. 1881/2006 and its amendments) and the pesticide residues (Reporting of results of the monitoring of pesticide residues in food according to Regulation (EC) No. 396/2005). The Standard Description is primarily designed for control and monitoring programs run by the Member States or by other private organisations. While some provision of specific variables were made for ad-hoc studies (i.e. product packaging, product ingredients, etc…), additional model validations and considerations should be performed before applying the data model outside the scope it was designed.

The standard logical model is designed to support the Standard Description in data transmissions and it cannot be directly converted into a database for data storage or analysis, since the Standard Description is not optimised for these tasks.

The guidance focuses on the definition of a logical model which is independent from the file format. Therefore data providers and receivers can use different file formats – e.g. Microsoft Excel, Comma Separated Values (CSV), Extensible Markup Language (XML), etc… – to submit transmissions depending on their technological constraints. The definition of the supported file formats and the actual implementation of the standard logical model will be discussed in the guidance of the Technical Working Group on Data Collection: "Guidance on Data Exchange".

The guidance describes additional data elements which connect the Standard Description with local systems of data providers and receivers (e.g. sending organisation, receiving organisation, the data collection title and the risk assessment area to which the data collection belongs to).These data elements are described in the paragraph "Context information of data transmissions". They are dependent on needs of data providers and receivers and, therefore, they are not to be considered part of the Standard Description.

The Standard Description was not designed to report data on zoonotic agents or similar biological entities.

The Standard Description does not support aggregated data but only individual results from the measurement of concentration of residues/substances in food and feed samples (i.e. sample level data).

Finally the Standard Description is intended to support the following analyses:

1) Assessment of acute and/or chronic consumer exposure

2) The number of samples for specific product / parameter combinations below and above detection limits

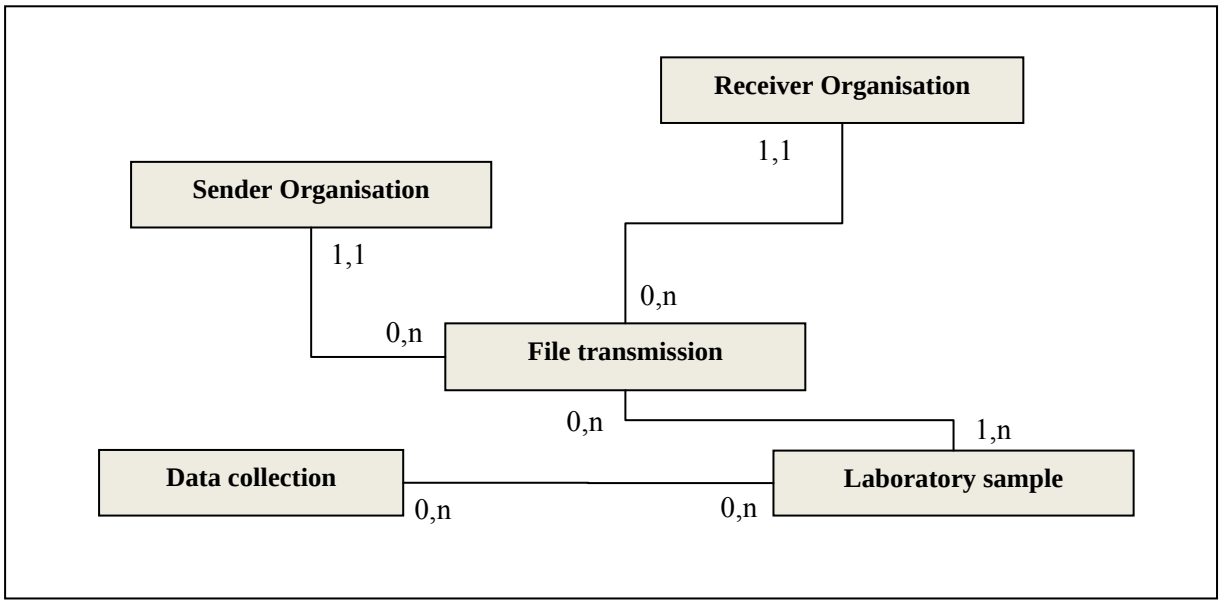
3) The number of samples for specific product / parameter combinations above legal limits (legal compliance)

9. CONTEXT INFORMATION OF DATA TRANSMISSIONS

Together with the data elements described by the Standard Description, the local systems should also keep records of some contextual information for data transmission. Although this contextual information mainly depends on the set up of the local systems, the following entities should be recorded in the local databases:

1. Sender organisation: Entity describing the organisation transmitting the data;
2. Receiver organisation: Entity describing the organisation receiving the data;
3. File transmission: Entity linking all laboratory samples submitted or received in a single file transmission. The entity should be described by some additional attributes such as the transmission date, the receipt date and other additional logging dates that may be needed by the transmission or receiver systems. A link to the physical original copy of the transmitted or received file should be maintained, as well.
4. Data collection: Entity linking all laboratory samples included in a single collection of data on specific risk assessment areas, year of sampling etc... In general terms data collections will be defined on ad-hoc basis by the data receiver: e.g. Heavy metal data collection, Pesticide residues 2009, etc…

Figure 1 – Structure of the main entities of the context in formation of data transmissions



10. DATA STRUCTURE OF THE STANDARD DESCRIPTION

In order to collect data on analytical measurements in food and feed the Standard Description needs to describe and maintain the relationships between four key entities (described below).

1. *Laboratory sample/Laboratory sub-sample*: the portion of material which is obtained by the specified sampling procedure, and which is sent to the laboratory for testing. Information relating to the laboratory sample includes product analysed, sampling plan, sampling dates, origin etc. These variables are coded S.01 – S.39.

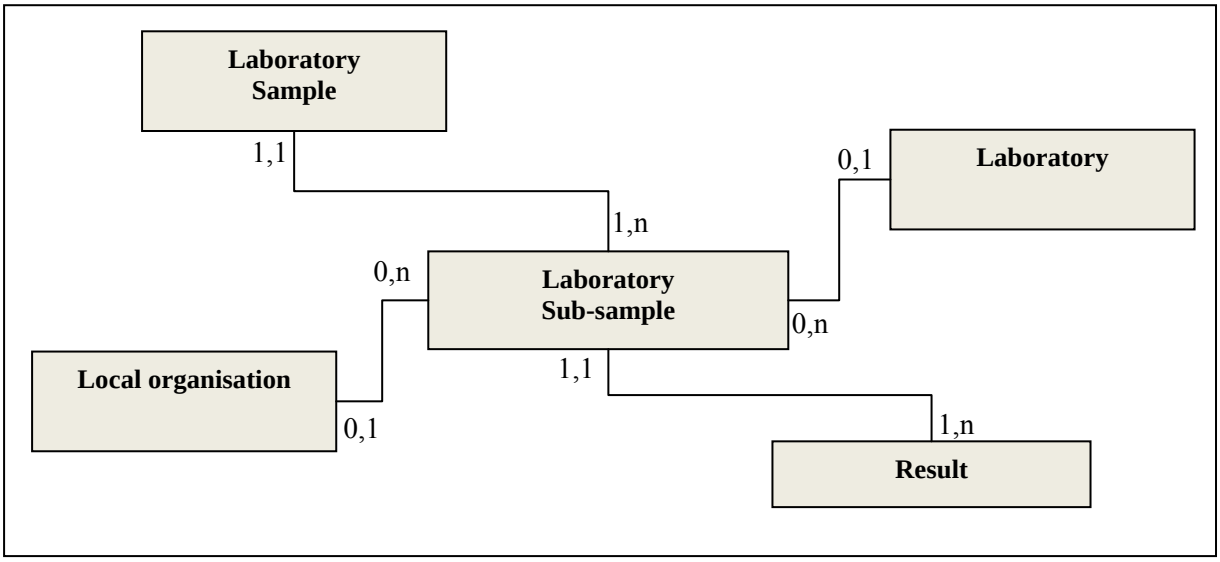
2. *Laboratory*: the laboratory that performed the analysis. These variables are coded L.01 – L.03.

3. *Local Organisation*: the organisation (local competent authority or other type of local organisation) who requested initially the analysis. These variable are coded O.01, O.02.

4. *Result*: the result of the laboratory tests, a quantitative or qualitative outcome value. Information relating to the result includes the parameter analysed, method of analysis, limits of detection, and for control programmes the legal limit used for compliance testing and the evaluation of the result. These variables are coded R.01 – R.32

The diagram in Figure 2 represents the entity-relationship diagrams between the represented entities.

Figure 2 Structure of the main Standard Description entities



The data structure reported in Figure 2 can be implemented in the standard logical model in two ways: normalised or denormalised. If the structure is normalised then the entity "Laboratory sub-sample" is a nested element of the entity "Laboratory Sample" and the entity "Result" is a nested element of the entity "Laboratory sub-sample". This approach creates a structure similar to a tree. Alternatively it is possible to have a denormalised structure avoiding the use of nested elements, repeating the value of each sample or sub-sample for each analytical result belonging to the same sample. This approach creates a structure similar to a table. The Technical Working Group on Data Collection opted for the denormalised approach for the Standard Description. The pros and cons of the denormalised approach are summarised below:

Pros:

- Simplified generation of the data files, since no nested elements are included;
- Flexibility in handling a variety of cases where the individual variables may belong to different levels;
- Similarity with an Microsoft Excel table, into/from which the denormalised structure can be easily imported/exported;

Cons:

- The normalised approach may avoid errors such as not repeating correctly the code for the high level entities;
- The file size is smaller because the information is not repeated.

The preference to the denormalised structure was given because the simpler generation was judged to deserve the higher priority. The issue of the larger file size was seen as the major fall back. In this case the use of compressed archives (such as zip files) for the transmitted files to mitigate the file size issue should be considered. The larger file size could be controlled also introducing a limit for the number of records that can be transmitted in a single file, but at this stage no file size limits were decided, and this issue should be reconsidered during the specification of the actual transmission format in the "Guidance on Data Exchange".

11. DATA ELEMENTS

Table 1 contains the list of the data element for the Standard Description. In the context of the data model, the term product does not refer only to the industrially produced products but more in general to all the matrices that can be chemically analysed.

Table 1. Data element list

Element Code	Element Name	Element Label	Type[9]	Controlled terminology	Description	Mandatory
S.01	**labSampCode**	**Laboratory sample code**	**xs:string (20)**		**Alphanumeric code of the analysed sample.**	**Yes**
S.02	**labSubSampCode**	**Laboratory sub-sample code**	**xs:decimal (4,0)**		**Numeric sequence number reflecting a subgroup of the analysed sample. The default value is 1.**	No
S.03	**lang**	**Language**	**xs:string (2)**	**LANG**	**Language used to fill in the free text fields (ISO-639-1).**	**Yes**
S.04	**sampCountry**	**Country of sampling**	**xs:string (2)**	**COUNTRY**	**Country where the sample was collected. (ISO 3166-1-alpha-2).**	**Yes**
S.05	**sampArea**	**Area of sampling**	**xs:string(5)**	**NUTS**	**Area where the sample was collected (Nomenclature of territorial units for statistics – NUTS – coding system valid only for EEA and Switzerland).**	No
S.06	**origCountry**	**Country of origin of the product**	**xs:string (2)**	**COUNTRY**	**Country of origin of the product (ISO 3166-1-alpha-2 country code).**	**Yes**
S.07	**origArea**	**Area of origin of the product**	**xs:string (5)**	**NUTS**	**Area of origin of the product (Nomenclature of territorial units for statistics – NUTS – coding system valid only for EEA and Switzerland).**	No
S.08	**origFishAreaCode**	**Area of origin for fisheries or aquaculture**	**xs:string (10)**	**FAREA**	**Fisheries or aquaculture area specifying the origin of the sample (FAO Fisheries areas).**	No

Element Code	Element Name	Element Label	Type[9]	Controlled terminology	Description	Mandatory
		activities code				
S.09	origFishAreaText	Area of origin for fisheries or aquaculture activities text	xs:string (250)		Fisheries or aquaculture area specified in free text	No
S.10	procCountry	Country of processing	xs:string (2)	COUNTRY	Country where the food was processed (ISO 3166-1-alpha-2).	No
S.11	procArea	Area of processing	xs:string (5)	NUTS	Area of product processing (Nomenclature of territorial units for statistics – NUTS – coding system valid only for EEA and Switzerland).	No
S.12	EFSAProdCode	EFSA Product Code	xs:string (250)	To be defined	Product under analysis described according to the EFSA Food Classification and Description System, currently under development.	No
S.13	prodCode	Product code	xs:string (20)	MATRIX	Product under analysis described according to the MATRIX catalogue, currently available.	Yes
S.14	prodText	Product full text description	xs:string (250)		Free text to describe in detail the product sampled. The text should provide additional information in respect to S.13. This element becomes mandatory if "product code" is 'XXXXXXA' (Not in list).	No
S.15	prodProdMeth	Method of production	xs:string (5)	PRODMD	Code providing additional information on the type of production for the food under analysis	No
S.16	prodPack	Packaging	xs:string (5)	PRODPAC	Describe container or wrapper that holds the product. Common type of packaging: paper or plastic bags, boxes, tinplate or aluminium cans, plastic trays, plastic bottles, glass bottles or jars.	No
S.17	prodTreat	Product treatment	xs:string(5)	PRODTR	Used to characterise a food product based on the treatment or processes	Yes

Element Code	Element Name	Element Label	Type[9]	Controlled terminology	Description	Mandatory
					applied to the product or any indexed ingredient.	
S.18	**prodBrandName**	**Brand name**	**xs:string(250)**		**Brand name of the product under analysis**	No
S.19	**prodManuf**	**Manufacturer**	**xs:string (250)**		**Company manufacturer of the product**	No
S.20	**prodIngred**	**Ingredients**	**xs:string(250)**		**List of ingredients, separated by "$", for the product under analysis. Use to provide further information on composite product.**	No
S.21	**prodCom**	**Product comment**	**xs:string (250)**		**Additional information on the product, particularly home preparation details if available.**	No
S.22	**prodY**	**Year of production**	**xs:decimal (4,0)**		**Year of production**	No
S.23	**prodM**	**Month of production**	**xs:decimal(2,0)**		**Month of production**	No
S.24	**prodD**	**Day of production**	**xs:decimal (2,0)**		**Day of production**	No
S.25	**expiryY**	**Year of expiry**	**xs:decimal (4,0)**		**Best before year or use by year or other indication of the expiry year.**	No
S.26	**expiryM**	**Month of expiry**	**xs:decimal (2,0)**		**Best before month or use by month or other indication of expiry month.**	No
S.27	**expiryD**	**Day of expiry**	**xs:decimal (2,0)**		**Best before day or use by day or other indication of the expiry day.**	No
S.28	**sampY**	**Year of sampling**	**xs:decimal (4,0)**		**Year of sampling. If the measure is the result of a sampling over a period of time, this field should contain the year when the first sample was collected.**	**Yes**
S.29	**sampM**	**Month of sampling**	**xs:decimal (2,0)**		**Month of sampling. If the measure is the result of a sampling over a period of time, this field should contain the month when the first sample was collected.**	No
S.30	**sampD**	**Day of sampling**	**xs:decimal (2,0)**		**Day of sampling. If the measure is the result of a sampling over a period of time, this field should contain the day when the first sample was collected.**	No
S.31	**progCode**	**Sampling**	**xs:string (20)**		**Sender's unique identification code of**	No

Element Code	Element Name	Element Label	Type[9]	Controlled terminology	Description	Mandatory
		programme code			the programme or project for which the sample analysed was taken.	
S.32	progLegalRef	Programme legal reference	xs:string (100)		Reference to the legislation for the programme defined by programme number.	No
S.33	progSampStrategy	Sampling strategy	xs:string (5)	SAMPSTR	Sampling strategy (ref. EUROSTAT - Typology of sampling strategy, version of July 2009) performed in the programme or project identified by programme code.	Yes
S.34	progType	Programme type	xs:string (5)	SRCTYP	Indicate the type programme for which the samples have been collected.	Yes
S.35	sampMethod	Sampling method	xs:string (5)	SAMPMD	Code describing the sampling method	Yes
S.36	sampleNum	Number of samples	xs:integer		Number of food samples analysed, only if composite samples were used.	No
S.37	lotSize	Lot size	xs:double		Size of the lot the sample belongs to	No
S.38	lotSizeUnit	Lot size unit	xs:string (5)	UNIT	Unit in which the lot size is expressed.	No
S.39	sampPoint	Sampling point	xs:string (10)	SAMPNT	Point in the food chain where the sample was taken. (Doc. ESTAT/F5/ES/155 “Data dictionary of activities of the establishments”).	Yes
L.01	labCode	Laboratory	xs:string (100)		Laboratory code (National laboratory code if available). This code should be unique and consistent through the transmissions.	No
L.02	labAccred	Laboratory accreditation	xs:string (1)	LABACC	The laboratory accreditation to ISO/IEC 17025.	Yes
L.03	labCountry	Laboratory country	xs:string (2)	COUNTRY	Country where the laboratory is placed. (ISO 3166-1-alpha-2).	No
O.01	localOrg	Local organisation	xs:string (100)		Local or regional organisation (Competent authority or company affiliate) who requested initially the analysis.	No
O.02	localOrgCountry	Local organisation country	xs:string (2)	COUNTRY	Country where the local organisation is placed. (ISO 3166-1-alpha-2).	No

Element Code	Element Name	Element Label	Type[9]	Controlled terminology	Description	Mandatory
R.01	resultCode	Result code	xs:string (40)		Unique identification number of an analytical result (a row of the data table) in the transmitted file. The result code must be maintained at organisation level and it will be used in further updated/deletion operation from the senders.	Yes
R.02	analysisY	Year of analysis	xs:decimal (4,0)		Year when the analysis was completed.	Yes
R.03	analysisM	Month of analysis	xs:decimal (2,0)		Month when the analysis was completed.	No
R.04	analysisD	Day of analysis	xs:decimal (2,0)		Day when the analysis was completed.	No
R.05	EFSAParamCode	EFSA Parameter Code	xs:string (250)	To be defined	Parameter/analyte of the analysis described according to the EFSA Parameters System, currently under development.	No
R.06	paramCode	Parameter code	xs:string (20)	PARAM	Parameter/analyte of the analysis described according to the Substance Code of the PARAM catalogue	Yes
R.07	paramText	Parameter text	xs:string (250)		Parameter subject of the analysis described according to the PARAM catalogue This element becomes a mandatory free text field if “Parameter code” is ’ RF-XXXX-XXX-XXX’ (Not in list).	No
R.08	paramType	Type of parameter	xs:string (5)	PARTYP	Define if the parameter reported is an individual residue/analyte, a summed residue definition or part of a sum a summed residue definition.	Yes
R.09	anMethRefCode	Analytical method reference code	xs:string(500)		Identifier for the method used. When validated methods are used, the official reference code should be provided.	No
R.10	anMethCode	Analytical method code	xs:string (5)	ANLYMD	Code describing the instrument used in the method.	No
R.11	anMethText	Analytical method text	xs:string (250)		Free text describing the analytical instrument used, particularly if “other”	No

Element Code	Element Name	Element Label	Type[9]	Controlled terminology	Description	Mandatory
					was reported for "Analytical method code".	
R.12	**accredProc**	**Accreditation procedure for the analytical method**	**xs:string (5)**	**MDSTAT**	**Accreditation procedure for the analytical method used**	No
R.13	**resUnit**	**Result unit**	**xs:string (5)**	**UNIT**	**Unit of measurement for the values reported in "Result LOD", "result LOQ, "CC alpha", "CC beta", "Result value", "Result value uncertainty standard deviation, "Result value uncertainty" and "Result legal limit".**	**Yes/No**[10]
R.14	**resLOD**	**Result LOD**	**xs:double**[11]		**Limit of detection reported in the unit specified by the variable "Result unit".**	No
R.15	**resLOQ**	**Result LOQ**	**xs:double**		**Limit of quantification reported in the unit specified by the variable "Result unit"**	No
R.16	**Ccalpha**	**CC alpha**	**xs:double**		**CC alpha value (decision limit) reported in the unit specified by the variable "Result unit"**	No
R.17	**Ccbeta**	**CC beta**	**xs:double**		**CC beta value (detection capability) reported in the unit specified by the variable "Result unit"**	No
R.18	**resVal**	**Result value**	**xs:double**		**The result of the analytical measure reported in the unit specified by the variable "Result unit",**	No
R.19	**resValRec**	**Result value recovery**	**xs:double**		**Recovery value associated with the concentration measurement expressed as a percentage (%). i.e. report 100 for 100%.**	No
R.20	**resValRecCorr**	**Result value**	**xs:string (1)**	**YESNO**	**Define if the result value has been**	No

[10] The unit of measurement should always be provided. Results without unit of measurement are accepted only if qualitative results are provided without any other numeric fields (e.g. LOD, LOQ, etc…)
[11] The data type xs:double and the other numeric data types which allow decimal separator requires the decimal separator to be a "." while the decimal separator "," is not allowed.

Element Code	Element Name	Element Label	Type[9]	Controlled terminology	Description	Mandatory
		corrected for recovery			corrected by calculation for recovery.	
R.21	resValUncertSD	Result value uncertainty Standard deviation	xs:double		Standard deviation for the uncertainty measure	No
R.22	resValUncert	Result value uncertainty	xs:double		Indicate the expanded uncertainty (usually 95% confidence interval) value associated with the measurement expressed in the unit reported in the field "Result unit".	No
R.23	moistPerc	Percentage of moisture in the original sample	xs:double		Percentage of moisture in the original sample	No
R.24	fatPerc	Percentage of fat in the original sample	xs:double		Percentage of fat in the original sample	No
R.25	exprRes	Expression of result	xs:string (5)	EXRES	Code to describe how the result has been expressed: Whole weight, fat weight, dry weight, etc…	No
R.26	resQualValue	Result qualitative value	xs:string (3)	POSNEG	This field should be completed only if the result value is qualitative e.g. Positive / Negative. In this case the element "Result value" should be left blank	No
R.27	resType	Type of result	xs:string (3)	VALTYP	Indicate the type of result, whether it could be quantified/determined or not.	Yes
R.28	resLegalLimit	Legal Limit for the result	xs:double		Report the legal limit for the analyte in the product sampled.	No
R.29	resLegalLimitType	Type of legal limit	xs:string(5)	LMTTYP	Type of legal limit applied for the evaluation of the result. ML, MRPL, MRL, action limit etc..	No
R.30	resEvaluation	Evaluation of the	xs:string (5)	RESEVAL	Indicate if the result exceeds a legal	Yes/No[12]

[12] This field is mandatory for all parameters where a legal limit exists e.g. for pesticide residues(parameter codes ending in –PPP), for veterinary drug residues (parameter codes ending in –VET) and for other control programmes where legal limits must be assessed

Element Code	Element Name	Element Label	Type[9]	Controlled terminology	Description	Mandatory
		result			limit.	
R.31	actTakenCode	Action Taken	xs:string (5)	ACTION	Describe any follow-up actions taken as a result of the exceeding a legal limit.	No
R.32	resComm	Comment of the result	xs:string (250)		Additional comments for this analytical result	No

12. SAMPLE DESCRIPTION

Several variables described below make use of controlled terminologies. The lists of all values belonging to the controlled terminology are reported in the document standardSampleDescription.xls.

Laboratory sample code (S.01)

As a consequence of the denormalised structure the laboratory sample should be identified by a unique sample identification number. No specification is provided to the data providers on the format of the sample identification number but data providers must ensure that the sample identification number is unique at data provider level. Where multiple analytical results are reported for a sample the unique sample identification number must be maintained for that sample in all transmissions. Additionally the information contained in elements S.01 – S.39 should be identical for each result returned.

Laboratory sub-sample code (S.02)

For some contaminants it is required to report multiple identical sub-samples belonging to the same original sample. In these cases the sub-sample code should be used. The sub-sample code should be a sequence number starting from 1 to n sub-samples analysed from the original sample e.g. aflatoxins in dried fruits where three sub-samples must be analysed. In most of cases the entity of interest will be only the laboratory sample and the result. In this cases where the laboratory sub-sample code is not explicitly reported it will be assumed to be equal to 1. The sub-sample should not be used if the sub-samples are not identical: in this case they should be reported as different samples. The sample can be analysed for the same parameter more than once to perform a counteranalysis to confirm a positive sample. In these cases the only result to report is the final result and the sub-sample code should not be used.

The specification of sub-samples is required in some cases by the legislation e.g. aflatoxins in dried fruits (Regulation (EC) No. 401/2006)

Language (S.03)

The language element records the language used to complete the free text fields. All free text fields should be completed with the same language. It is recommended that the free text fields are completed in English whenever possible. The language must be specified using the ISO-639-1 language code list.

Country of sampling (S.04), origin (S.06), processing (S.10)

Countries should be encoded using the standard ISO-3166-1-alpha-2 coding system. An extract is reported in the COUNTRY catalogue.

The country of sampling is the country where the commodity was selected for laboratory testing. As a result, only Member States or EEA country codes should be used.

In addition to the ISO standard codes, the codes EU, AA, XC, XD and XX have been added according to the provisions of the ISO-3166-1-alpha-2 for user-assigned code elements. When the country is unknown these options must be used in the provided order, being as specific as possible.

EU – Unspecified country that is part of European Union (EU)

AA – Unspecified country that is part of the European Economic Area (EEA) including EU

XC – Unspecified third country non EEA

XD – Country not domestic, import

XX – Unknown (i.e. nothing is known about the country).

The EEA is made up by the 27 EU Member States and the three EEA States (Iceland, Liechtenstein and Norway).

The country or area of origin of the commodity should be considered the place where the main commodity was grown, raised etc. Additional details on the source of the item samples if available or requested can be supplied in the "Product comment" (S.21) element.

The country or area of processing is the location where the processed commodity was manufactured. This element should be used for processed commodities only.

Area of sampling (S.05), origin (S.07), processing (S.11)

The area of sampling, origin and production provides more detailed geographical information on locations according to the definitions described in the section above (6.4). Use the Nomenclature of territorial units for statistics (NUTS) code as described in NUTS catalogue. This coding system only covered regions within countries in EEA.

Area of origin for fisheries or aquaculture activities code and text (S.08, S.09)

Fishing areas are coded using the FAO fishing area coding system, prefixed with the letter "M". Additional codes have been added in case details on the part of the ocean are unknown or if the fishing area is unknown. More detail on the fishing place (e.g. ICES codes, name of river or lake, place of catch…) can be reported in the element "Origin fish area text" (S.09).

EFSA Product Code (S.12)

EFSA is currently developing a food classification and description system for exposure assessment. This element will be used for reporting the EFSA product code once the classification scheme has been validated and implemented.

Product code (S.13), product full text description (S.14), brand name (S.18), manufacturer (S.19) and product comment (S.21)

The product code includes the coding systems currently in use in the PRAPeR unit for pesticide residues data collection as defined in Reg. EC 178/2006. Baby food samples (baby-food, infant formulae/follow-up formulae and processed cereal-based foods) are also listed as "REG. EC 178/2006" in the "codingSystem" column of the MATRIX catalogue even though the appropriate pieces of EC legislation that describes these food products are Directive 2006/141/EC and Directive 2006/125/EC.

A food classification applicable to all food risk assessment areas is under development by the Food Classification Working Group[13]. Data collection specific guidelines should be consulted to determine the correct coding system to be used. If information is available to further characterise the product this should be supplied in the free text elements "Product full text description", ""Brand name", "Manufacturer", "Product comment". In particular information on home preparation, if relevant, should be supplied in the "Product comment" (S.21) element. Data providers should be aware that if "Product code" is equal to 'XXXXXXA' (Not in list) then the "Product text" becomes a mandatory field.

[13] M-2009-0135 Development of a food classification system for exposure assessment. For further information on the progresses of this work contact datex@efsa.europa.eu.

Method of production (S.15)

The method of production contains a list of codes to describe the agricultural production method used to produce the food or feed sampled. This coding list is a combination of the coding list provided by Federal Agency for the Safety of the Food Chain (Belgium) and LANGUAL.
The reporting of multiple codes is accepted for this data element and the codes reported should be reported separated by '$'. The order of the codes is not important.

Packaging (S.16)

The packaging element describes the container or wrapper of the product, for example paper or plastic bags, boxes, tinplate or aluminium cans plastic trays, plastic bottles, glass bottles or jars. Where the exact type of packaging is unknown simply state if the product was unpacked, wrapped or packed.

Product treatment (S.17)

This element is used to discriminate between processed, unprocessed commodities and in this case the type of processing used. Processing indicates the product has been changed from its original form as a raw agricultural commodity by the application of physical, chemical or biological treatments. The controlled terminology (PRODTR) to describe the treatment is taken from the OECD and FAO documents (FAO: Further consideration of processing as related to the establishment of MRLs for processed foods: Recommendations on principles and practices OECD:http://www.oecd.org/dataoecd/508/12/39736351.pdf) which discuss processing factors which should be applied in consumer exposure assessments. For pesticide residues a final classification of processed commodities is not yet available.

Ingredients (S.20)

The data element describes the list of ingredients of a product. The ingredients should be separated by the character "$". Ingredients should be reported in decreasing order of content. The data element is used to provide further information on composite products.

Year, month and day of production/expiry (S.22 – S.27)

This data element is describing additional information for the product under analysis. It is possible to report the date of production and the expiry date. In some cases the date of expiry may not be available and it is replaced by a best-before date. In these case the best-before date should be reported.

Year, month and day of sampling (S.28 – S.30)

Sampling date divided into the year, month and day elements. It is mandatory to report the year while reporting the month or the day is optional. It is possible to report the day only if the month is also reported.

In case the sampling has been performed over a period of time the start date of sampling should be reported.

Sampling programme code (S.31)

The data element should contain the laboratory or data provider organisation's unique identification code for the sampling programme or project for which the sample described by the Sample Description was taken. Each code will identify a specific group of samples analysed for a programme or project. Further description on the reason for sampling can be provided in the element "Programme legal reference" (see below S.32).

Programme legal reference (S.32)

This data element is a free text description of the reason for initiating the sampling programme identified by the code specified in the Sampling programme code. Where the sampling programme is described in legislation the regulation or directive should be referenced.

In cases where EC regulations are quoted the following format should be used: Regulation (EC) No XXX/CCYY. In cases where EC directives are quoted the following format should be used Directive CCYY/XXX/EC.

Sampling strategy (S.33)

Sampling strategy describes how the sample was selected from the population being monitored or surveyed. A list was defined starting from the previous typology defined by Eurostat[14]. In order to make reference to the existent Eurostat controlled terminologies for the sampling strategy the following definitions will apply for this section of the document. Eurostat's definitions refer to reporting of aggregate data.

a) The **control objective**: this will define why the control is undertaken, what it is the aspect to control and the population to control;
b) **population**: set of homogeneous units with respect to certain characteristics (population of units of a certain kind of fruit, population of containers of imported vegetables, population of milk producing holdings, population of slaughtered bovines, population of bovine herds, …) in a certain space and time;

[14] Previous typology defined by Eurostat in "Typology of sampling strategies used in control and monitoring activities" doc. ESTAT/F5/ES/104 Rev 4 – Part 1 (now under revision) in http://circa.europa.eu/Public/irc/dsis/foodsafertystats/library?l=/documents_statistics/statistics_monitoring&vm=detailed&sb=Title

c) **sample**: subset of the population on which the control activity is conducted; each unit of the sample is a "sample unit";
d) **sample unit taken**: the material physically extracted from the sample unit to meet the objective of the control within the sample. The sample unit taken depends on the control to be undertaken and its capability of representing the sampling unit.

The term "sample unit taken" in the Eurostat definition corresponds to the "laboratory sample" used in the data model.

Code	Description	Definition
ST10A	Objective sampling	Strategy based on the selection of a random sample from a population on which the data are reported. It includes also other random samplings as: "stratified" in subpopulations and sampling with proportional criterion, multistage sampling, …
ST20A	Selective sampling	Strategy based on the selection of a random sample from a subpopulation (or more frequently from subpopulations) of a population on which the data are reported. The subpopulations are determined on a risk basis or not. The sampling from each subpopulation is not proportional: the sample size is proportionally bigger for instance in subpopulations considered at high risk. This sampling includes also the case when the data reported refer themselves to censuses on subpopulations.
ST30A	Suspect sampling	Selection of an individual product or establishment in order to confirm or reject a suspicion of non-conformity. It's a not random sampling. The data reported refer themselves to suspect units of the population.
ST40A	Convenient sampling	Strategy based on the selection of a sample for which units are selected only on the basis of feasibility or ease of data collection. It's a not random sampling. The data reported refer themselves to units selected according to this strategy.
ST50A	Census	When the totality of a population, on which the data are reported, is controlled.
ST90A	Other	
STXXA	Not specified	

As the data model reports data on single sample (according to data model terminology), and not on aggregate data as Eurostat's sampling strategies refer to, the sampling strategy to assign is the one related to the programme specified in "Sampling programme code (S.31)"

Some examples:

Example 1: a sample of milk checked for the presence of melamine belongs to a programme on controls on all milk imported from China; the sampling strategy to assign to this sample is "Census". The population of reference is all milk imported from China.

Example 2: a commodity included in the EU coordinated programme (e.g. Orange) sampled randomly from retail outlets in a Member State and tested for the pesticides specified in the EU coordinated programme; the sampling strategy to assign to this sample is "Objective sampling". The population of reference is all oranges available to the consumer in the Member State.

Example 3: a sample of orange checked for the presence of a pesticide residue belongs to a programme on random controls on oranges imported from any non-EU Countries; according to the programme the sampling is higher for Countries considered at high risk. The sampling strategy to assign to this

sample is "Selective sampling". The population of reference is the set of all "units" (single fruits, lots, consignments …) of oranges imported from extra-EU Countries.

Example 4: a sample of orange checked for the presence of a pesticide residue available from a precise wholesaler, as a consequence of previous non-conformity results. The sampling strategy to assign to this sample is "Suspect sampling".

Programme type (S.34)

The programme type should be reported to indicate the type of control programme or other type of source to which the sample belongs. It is important to determine whether the programme was designed to assess consumer exposure at the EU level or at National level as the commodities sampled may differ depending on the dietary habits of the population under study.

Sampling method (S.35) and number of samples (S.36)

The sampling method defines the way the samples have been collected for analysis. If the sampling method is described in official legislation the code to identify this legislation should be provided. The data structure supports individual sample transmissions where each sample provides a single analytical result. Pooled samples, where multiple samples may be analysed together to provide a single result, are also allowed. The number of individual samples that was pooled should be entered in S.36 Number of samples. The default value for this field is "1".

Lot size (S.37), Lot size unit (S.38)

Data elements providing the size and unit of the lot from which the sample was taken.

A lot is a definite quantity of some commodity manufactured or produced under conditions, which are presumed uniform for the purpose of these Guidelines[15].
For the goods presumed heterogeneous, sampling can only be achieved on each homogeneous part of this heterogeneous lot. In that case, the final sample is called a stratified sample.

NOTE: A continuous series of lots is a series of lots produced, manufactured or commercialised on a continuous manner, under conditions presumed uniform. The inspection of a continuous series of lots can only be achieved at the production or processing stage.

Sampling point (S.39)

This element defines the point of the food chain where the sample was taken.

The controlled terminology to be used in the data element is based on the data dictionary of activities. (Reference document "Data dictionary of activities of the establishments", doc. ESTAT/F5/ES/155[16]).

[15] "General Guidelines on Sampling", CAC/GL 50-2004, FAO/WHO

The list details the activities of establishments at different points in the food chain. The activities are described at different levels of detail, and data providers are requested to report at the most detailed level available. This list of activities is intended to indicate the type of establishment from which the sample was taken.

[16] "Data dictionary of activities of the establishments", doc. ESTAT/F5/ES/155 available here: http://circa.europa.eu/Public/irc/dsis/foodsafertystats/library?l=/documents_statistics/statistics_monitoring&vm=detailed&sb=Title

13. LABORATORY

Laboratory code (L.01)

A unique code to identify each laboratory providing laboratory results should be reported here. If a national laboratory coding system exists this code should be reported. This code should be reported consistently for all transmissions of data. Further information may be requested separately, for instance participation in proficiency tests, the unique code should also be reported in these cases.

Laboratory accreditation (L.02)

This element indicates whether accreditation of the laboratories performing the analysis has been achieved. In accordance with Art 12 of Regulation 882/2004, laboratories designated for official controls must be accredited to ISO/IEC 17025, or avail of the derogation in Art 18 of Regulation 2076/2005.

Laboratory country (L.03)

Report the country where the laboratory performing the testing is located in this element.

14. LOCAL ORGANISATION

Local organisation, local organisation country (O.01, O.02)

This element indicates that the sender received the data in the transmission from a regional organisation or a regional competent authority or an affiliate for a commercial organisation, who requested initially the analysis. The country the organisation is located in should also be reported. This information is important in countries where the data collection is decentralised to local organisations. If the data collection is decentralised it is possible that duplicated samples could be transmitted to EFSA or other data receivers. In this case the information on the local organisation can help in the detection of duplicate samples.

15. ANALYTICAL RESULT DESCRIPTION

Result code (R.01)

This element should contain the unique identification number of an analytical result (a row of the data table) in the transmitted file. This code is mandatory, as it will be used as reference for operation of deleting or updating an individual result if this procedure is supported by the data collection protocol.

Year, month and day of analysis (R.02 – R.04)

Analysis date divided into the year, month and day elements. It is mandatory to report the year while reporting the month or the day is optional. It is possible to report the day only if the month is also reported.

If the analysis has been performed over a period of time the completion date of analysis should be stated.

EFSA Parameter Code (R.05)

This data element is included to support possible future developments in coding systems to describe the parameter under analysis.

Parameter code, parameter text and parameter type (R.06, R.07, R.08)

The "Parameter code" includes the coding system currently in use in the DATEX, PRAPeR and Zoonoses Units to describe parameters under analysis, particularly chemical contaminants and pesticide residues. Data collection specific guidelines should be consulted to determine the correct section of the coding system to be used. In case the parameter is not included in the PARAM catalogue the code "Not in list" should be reported and the name of the parameter should be specified in the "Parameter Text" variable. The "Parameter text" element should be completed using the International Union of Pure and Applied Chemistry (IUPAC) name. CAS number can be included for clarity using the "$" sign to separate the chemical name from the CAS number.

In order to facilitate the reporting of parameters according to complex residue definitions and ensure that the assessment of multiple residues in a product is accurate, the "Parameter type" data element must be completed to indicate whether parameter reported is summed according to a parameter sum or residue definition, a part of a parameter sum or residue definition or an individual parameter or residue.This is necessary because in some cases it is diffucult to reproduce on the receiver side which individual parameters are included in summed parameters (such as residues definition, dioxin TEQ), making the calculation of summed parameters more transparent.

Analytical method reference code (R.09)

This data element identifies the method used. The identifier used should enable the laboratory to uniquely identify the actual method procedure used for the procuring the result.
When validated methods are used, the official reference code should be provided otherwise an internal sender identifier can be provided.

Analytical method code, analytical method text (R.10, R.11)

This data element is the code describing the type of method applied (sometimes represented by the main instrument) used in the method. If the method is not in the list or it is necessary to provide more details with respect to the method available in the list, the associated free text should be used. The free text field must be used if "Classification not possible" was reported for "Analytical Method Code".

Accreditation procedure for the analytical method (R.12)

It is essential that any data which is submitted to EFSA is of sufficient quality such that it is fit-for-purpose. There are internationally accepted data quality standards for analytical laboratories supplying such data. As described in Article 12 of Regulation (EC) No 882/2004 on "Official controls performed to ensure the verification of compliance with feed and food law, animal health and welfare rules" data provider should state which quality assurance procedures (e.g. ISO/IEC17025, other third party quality assessment, internal validation) were used. Multiple options are not possible; the data provider must choose the option which is most appropriate.

Result unit (R.13)

This data element indicates the unit of measurement for the values reported in "Result LOD", "Result LOQ", "Result Value", "Result uncertainty", "Result uncertainty standard deviation", "CC alpha", "CC beta" or "Result legal limit". This should be consistent for all elements. This field is by default not mandatory, but it becomes mandatory if at least one of the following fields is provided: "Result LOD", "Result LOQ", "Result Value", "Result uncertainty", "Result Uncertainty Standard Deviation", "CC alpha", "CC beta" or "Result legal limit". By convention, 'per' is used at the beginning of an element value and the '/' symbol is used to mean per when it appears within an element value without any difference in meaning intended.

Result LOD (R.14)

The Limit of Detection[17] (LOD) is the lowest concentration level that can be determined to be statistically different from a blank (Keith, et al., 1983). Usually a confidence level of 95% or 99% is used. This value must be expressed in the unit reported in the data element "Result unit".

Result LOQ (R.15)

The limit of quantification[17] (LOQ) is the level above which quantitative results may be obtained with a specified degree of confidence (Keith, et al., 1983). This value must be expressed in the unit reported in the data element "Result unit".

[17] Many definitions of LOD and LOQ have been suggested throughout the years and in different analytical areas. For recent international definitions see e.g. Report Of The Thirtieth Session Of The Codex Committee On Methods Of Analysis And Sampling (Balatonalmádi, Hungary, 9 - 13 March 2009), Method validation and quality control procedures for pesticide residue analysis in food and feed (Document N° SANCO/2007/3131), Commission Directive 2009/90/EC laying down, pursuant to Directive 2000/60/EC of the European Parliament and of the Council, technical specifications for chemical analysis and monitoring of water status, IUPAC Compendium of Chemical Terminology (Gold Book).

CC alpha (R.16)

Decision limit[18] (CC α) means the limit at and above which it can be concluded with an error probability of α that a sample is non-compliant.
If no permitted limit has been established for a substance, the decision limit is the lowest concentration level at which a method can discriminate with a statistical certainty of 1 – α that the particular analyte is present.
If a permitted limit has been established for a substance, the decision limit is the concentration above which it can be decided with a statistical certainty of 1 – α that the permitted limit has been truly exceeded.
Alpha (α) error means the probability that the tested sample is compliant, even though a non-compliant measurement has been obtained (false non-compliant decision).
CC alpha must be expressed in the unit reported in the data element "Result unit".

CC beta (R.17)

Detection capability[19] (CCβ) means the smallest content of the substance that may be detected, identified and/or quantified in a sample with an error probability of β.
In the case of substances for which no permitted limit has been established, the detection capability is the lowest concentration at which a method is able to detect truly contaminated samples with a statistical certainty of 1 – β.
In the case of substances with an established permitted limit, this means that the detection capability is the concentration at which the method is able to detect permitted limit concentrations with a statistical certainty of 1 – β.
Beta (β) error means the probability that the tested sample is truly non-compliant, even though a compliant measurement has been obtained (false compliant decision).
CC beta must be expressed in the unit reported in the data element "Result unit".

Result value (R.18), Result qualitative value (R.26) and type of result (R.27)

The data elements "Result value", "Result qualitative value" and "type of result" are used to describe different type of results of an analysis.

"Result value" stores the final value reported for a measured of computed quantity. If the result is numeric, the data element "Result value" must be completed and the "Type of result" should be set to "VAL". In this case the result of the analytical measure must be reported in the unit specified by the data element "Result unit".

The "Result value" should not be corrected for uncertainty. The uncertainty associated with the measure must be expressed using the data elements "Result uncertainty standard deviation" and "Result uncertainty". The uncertainty and the uncertainty standard deviation are always provided in the same unit as the result value.

[18] 2002/657/EC: Commission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (Text with EEA relevance) (notified under document number C(2002) 3044)

[19] 2002/657/EC: Commission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (Text with EEA relevance) (notified under document number C(2002) 3044)

If not otherwise specified, the result is considered expressed in “whole weight”. In case the data provider has to report the result expressed in other way (e.g. “fat weight”, “dry weight” etc…), the data element “Expression of result” must be populated and the fat and/or moisture content provided.

If not otherwise specified, the result expressed is considered not corrected for recovery. If the data provider reports a result corrected for recovery, the data element “Result value corrected for recovery” must be set to “yes”. The recovery for the result must be provided as percentage (the value 120 indicates 120% recovery, the value 1.20 indicate 1.20%) using the data element “Result value recovery”. It should be noted that for specific data collections (e.g. pesticide residues) the result reported should not be corrected for recovery.”

In cases when the “Result value” cannot be determined the result value could be left empty but the “Type of result” must indicate which type of limit was used (e.g. Non detected, non quantified, under CCalpha, under CCbeta.

In cases when the result value is below LOQ (or lower than CC alpha), the data providers are invited to report the result value, stating “Type of result” equal to “LOQ” (or “CCA”).

In some cases the analytical method may not have a LOQ defined; in these cases, if the analytical method is taking into account a limit of reporting, this information should be reported in the LOQ data element.

If the result of the analysis is qualitative i.e. positive or negative then the data element “Result qualitative value” should be populated and the “Type of result” should be set to “BIN”.

There are a number of validation rules that apply to the relationships between the result elements. These are described in section 17: validation rules.

Result value recovery (R.19) and result value corrected for recovery (R.20)

Recovery value associated with the concentration measurement expressed as a percentage (%). This element must be completed if "Result value corrected for recovery" (R.20) has been set to "YES".
Setting "Result value corrected for recovery" (R.20) to "No" and supplying a value for "Result value recovery" (R.19) enables the data receiver to correct "Result value" for recovery by calculation.
If the R.20 is set to "No", in case the value needs to be adjusted for recovery, the result value will be multiplied for the result value recovery (R.19) by the receiver. The recovery for the result must be provided as percentage (the value 120 indicates 120% recovery, the value 1.20 indicate 1.20%).

Result value uncertainty standard deviation (R.21)

This field contains the standard deviation for the uncertainty measure expressed in the unit reported in the data element result unit.

Result value uncertainty (R.22)

Indicate the expanded uncertainty (normally 95% confidence interval) value associated with the measurement expressed in the unit reported in the data element result unit.

Percent moisture (R.23), percent fat (R.24) and expression of result (R.25)

The result should be expressed on a "whole" weight basis. Only when the legislation explicitly requires the expression of results in dry weight or in fat basis the information should be reported in this manner. If the "Result value" is not expressed on a whole weight basis the type of expression of result should be indicated using the EXRES terminology in the "Expression of result" element. If the expression of the result is on a fat weight basis the "Percentage of fat in the original sample" should be completed and if the expression of the result in on a dry weight basics "Percentage of moisture in the original sample" should be completed. This is described in the business rules section.

Legal Limit for the result (R.28)

Report the legal limit for the residue/chemical and commodity. The "Result value" should be reported according to the legal limit definition and in the same units as the legal limit. The legal limit reported should be the one applicable at the time of compliance assessment. Where the official EU legal limit has not been used to evaluate the result, then the legal limit which was used should be provided.

Type of the legal limit (R.29)

Report the type of legal limit used to assess the result value. If a legal limit is provided in R.28 then the type of legal limit used should be also provided in this field (R.29).

Evaluation of the result (R.30)

This element should contain the result of the laboratory or risk assessor's evaluation of the measured parameter in a commodity. If the laboratory considers the result has failed in internal quality control procedures the result should not be transmitted. The result should be compared with the legal limit applicable at the time of sampling and the correct code from the RESEVAL terminology selected to indicate whether the result was complaint with the legal limit or not. Since the value reported here will

be used in reports assessing the rate of compliance for commodities in the EU market place, this data element is mandatory for all control programmes for all parameters where a legal limit exists.

Action taken (R.31)

Describe any follow-up actions taken as a result of the exceeding a legal limit. More than one action may be reported using the "$" sign as the separator character.

Comment of the result (R.32)

A free text comment for each analytical result is provided. Ideally, free text fields should be transmitted in English (see also 1.3 Language (S.03)).

16. VALIDATION RULES

Validation rules describe the constraints used to accept data from receiver side, during the data transmissions. The receivers must always reply back to the sender through an acknowledgment transmission whether a Sample Description transmission is accepted or not. In case of failure of the transmission, the receiver should also describe the reasons which determined the failure of the message transmission. The specific protocol for exchanging Sample Description and acknowledgment transmissions to business rules validation will be described in the "Guidance for data exchange".

In order to reduce the possible occurrence of transmission failures, the sender is recommended to implement the business rules on the sender side and to validate all the files before submission to EFSA.

The validation rules can generate an error or a warning message. Transmissions containing only warnings are still accepted by the receiver. Transmissions containing at least one error are rejected by the receiver.

The following general validation rules produce errors upon validation of the data file:

- If the data element is mandatory, the element must be present and the value must be reported;
- Check if the value reported is compliant with the type defined in the Standard Description;
- If the type requires a controlled terminology then the code reported must be included in the controlled terminology specified.

In addition, the following validations must be performed. The type column specifies whether the check will produce an error (E) or a warning (W).

In cases where the data element is not mandatory, the rules apply only if the relevant element is provided. In cases where the rule involves a comparison of one or more data elements, the rule can only be applied if the data elements which are the subject of comparison are provided or can be built from the data provided (e.g. the rule for the data element S.05 "Area of sampling" will be applied only if the element S.05 is provided).

:Element Code	Element Name	Element Label	Validation	Type
S.01	**labSampCode**	**Laboratory sample code**	For s.03-s.39: all values in each data elements must be equal for all records with same labSampCode	**E**
S.02	**labSubSampCode**	**Laboratory sub-sample code**		
S.03	**Lang**	**Language**		
S.04	**sampCountry**	**Country of sampling**		
S.05	**sampArea**	**Area of sampling**	The "area of sampling" reported must be included in the country reported in "Country of sampling"	**E**
S.06	**origCountry**	**Country of origin of the product**		
S.07	**origArea**	**Area of origin of the product**	The "area of origin" reported must be included in the country reported in "Country of origin of the product"	**E**
S.08	**origFishAreaCode**	**Area of origin for fisheries or aquaculture activities**		
S.09	**origFishAreaText**	**Area of origin of the product expressed as free**		

:Element Code	Element Name	Element Label	Validation	Type
		text		
S.10	**procCountry**	**Country of processing**		
S.11	**procArea**	**Area of processing**	The "Area of processing" reported must be included in the country reported in "Country of processing"	**E**
S.12	**EFSAProductCode**	**EFSA Product Code**		
S.13	**prodCode**	**Product code**		
S.14	**prodText**	**Product full text description**	If "Product code" is equal to "XXXXXXA" (not in list) then prodText must be provided	**E**
S.15	**prodProdMeth**	**Method of production**		
S.16	**prodPack**	**Packaging**		
S.17	**prodTreat**	**Product treatment**		
S.18	**prodBrandName**	**Brand name**		
S.19	**prodManuf**	**Manufacturer**		
S.20	**prodIngred**	**Ingredients**		
S.21	**prodCom**	**Product comment**		
S.22	**prodY**	**Year of production**	"Year of production" has to be less or equal to the current year	**E**
			"Year of production" must be less than or equal to the "Year of expiry"	**E**
			"Year of production must be less than or equal to "Year of sampling	**E**
			"Year of production must be less than or equal to "Year of analysis"	**E**
S.23	**prodM**	**Month of production**	"Month of production" has to be between 1 and 12	**E**
			"Month of Production" has to be filled in if "Day of production" is filled in	**E**
			The partial date prodM/prodY must be less or equal to the current partial date M/Y.	**E**
			The partial date prodM/prodY must be less or equal to expiryM/expiryY.	**E**
			The partial date prodM/prodY must be less or equal to sampM/sampY.	**E**
			The partial date prodM/prodY must be less than or equal to analysisM/analysisY	**E**
S.24	**prodD**	**Day of production**	"Day of production" has to be between 1 and 31	**E**
			The date prodD/prodM/prodY must be a valid date	**E**
			The date prodD/prodM/prodY must be less or equal to the current date D/M/Y.	**E**

:Element Code	Element Name	Element Label	Validation	Type
			The date prodD/prodM/prodY must be less than expiryD/expiryM/expiryY.	**E**
			The date prodD/prodM/prodY must be less than sampD/sampM/sampY.	**E**
			The date prodD/prodM/prodY must be less than analysisD/analysisM/analysisY	**E**
S.25	**expiryY**	**Year of expiry**		
S.26	**expiryM**	**Month of expiry**	"Month of expiry" has to be between 1 and 12	**E**
			"Month of expiry" has to be filled in if "Day of expiry" is filled in	**E**
S.27	**expiryD**	**Day of expiry**	"Day of expiry" has to be between 1 and 31	**E**
			The date prodD/prodM/prodY must be a valid date	**E**
S.28	**sampY**	**Year of sampling**	"Year of sampling" must be less or equal to the current year	**E**
			"year of sampling" must be less than or equal to "year of analysis"	**E**
S.29	**sampM**	**Month of sampling**	"Month of sampling" has to be between 1 and 12	**E**
			"Month of sampling" has to be filled in if "Day of sampling" is filled in	**E**
			The partial date sampM/sampY must be less or equal to the current partial date M/Y	**E**
			The partial date sampM/sampY must be less or equal to the partial date analysisM/analysisY	**E**
S.30	**sampD**	**Day of sampling**	"Day of sampling" has to be between 1 and 31	**E**
			The date sampD/sampM/sampY must be a valid date	**E**
			The date sampD/sampM/sampY must be less or equal to the current date D/M/Y	**E**
			The date sampD/sampM/sampY must be less or equal to the date analysisD/analysisM/analysisY	**E**
S.31	**progCode**	**Sampling programme code**		
S.32	**progLegalRef**	**Programme legal reference**		
S.33	**progSampStrategy**	**Sampling strategy**		
S.34	**progType**	**Type of sampling programme**		
S.35	**sampMethod**	**Sampling method**		
S.36	**sampleNum**	**Number of samples**		

:Element Code	Element Name	Element Label	Validation	Type
S.37	**lotSize**	**Lot size**		
S.38	**lotSizeUnit**	**Lot size unit**	If the lot size is provided the lot size unit must be provided	**E**
S.39	**sampPoint**	**Sampling point**		
L.01	**LabCode**	**Laboratory**		
L.02	**labAccred**	**Laboratory accreditation**		
L.03	**labCountry**	**Laboratory country**		
O.01	**localOrg**	**Local organisation**		
O.02	**localOrgCountry**	**Local organisation country**		
R.01	**resultCode**	**Result code**		
R.02	**analysisY**	**Year of analysis**	"Year of analysis" has to be less than or equal to the current year	**E**
R.03	**analysisM**	**Month of analysis**	"Month of analysis" has to be between 1 and 12	**E**
			"Month of analysis" has to be filled in if "Day of analysis" is filled in	**E**
			The partial date analysisM/analysisY must be less or equal to the current partial date M/Y	**E**
R.04	**analysisD**	**Day of analysis**	"Day of analysis" has to be between 1 and 31	**E**
			The date analysisD/analysisM/analysisY must be a valid date	**E**
			The date analysisD/analysisM/analysisY must be less or equal to the current date D/M/Y	**E**
R.05	**EFSAParamCode**	**EFSA Parameter Code**		
R.06	**paramCode**	**Parameter code**	Where paramCode <> "RF-XXXX-XXX-XXX" (Not in list) then (paramCode, labSampCode, labSubSampCode) must be unique for a data provider;	**E**
R.07	**paramText**	**Parameter text**	Where paramCode = "RF-XXXX-XXX-XXX" (Not in list) then paramText must be provided	**E**
R.08	**paramType**	**Type of parameter**		**E**
R.09	**anMethRefCode**	**Analytical method reference code**		
R.10	**anMethCode**	**Analytical method code**		
R.11	**anMethText**	**Analytical method text**	If anMethCode is "F001A" (Classification not possible) then anMethText must be provided	**E**
R.12	**accredProc**	**Accreditation procedure for the analytical method**		
R.13	**resUnit**	**Result unit**	If "Type of result" is different from "BIN" then "Result unit" must be specified If at least one of resLOD, resLOQ, CCalpha, CCbeta, resVal, resValUncertSD, resValUncert, resLegalLimit is provided then resUnit must be provided.	**E**
R.14	**resLOD**	**Result LOD**	"Result LOD" has to be filled in if "Type of result" is equal to "LOD".	**E**
			"Result LOD" must be greater than 0	**W**
			"Result LOD" must be lower than or equal to "Result LOQ"	**E**

:Element Code	Element Name	Element Label	Validation	Type
R.15	resLOQ	Result LOQ	"Result LOQ" has to be filled in if "Type of result" is equal to "LOQ".	E
			"Result LOQ" must be greater than 0	W
R.16	CCalpha	CC alpha	"CC alpha" has to be filled in if "Type of result" is equal to "CCA".	E
			"CC alpha" must be greater than 0	W
			"CC alpha" must be lower than "CC beta"	E
R.17	CCbeta	CC beta	"CC beta" has to be filled in if "Type of result" is equal to "CCB".	E
			"CC beta" must be greater than 0	W
R.18	resVal	Result value	ResVal must be filled in if ResType is equal to "VAL".	E
			ResVal must be greater than 0	W
			ResVal has to be missing when resType is LOD.	E
R.19	resValRec	Result value recovery	resValRec must be greater than 0	W
R.20	resValRecCorr	Result value corrected for recovery		
R.21	resValUncertSD	Result value uncertainty Standard deviation	resValUncertSD must be greater than 0	W
R.22	resValUncert	Result value uncertainty	resValUncert must be greater than 0	W
R.23	moistPerc	Percentage of moisture in the original sample	MoistPerc has to be between 0 and 100.	E
			MoistPerc must be provided if "Expression of result" is 'B002' "dry weight"	E
R.24	fatPerc	Percentage of fat in the original sample	FatPerc has to be between 0 and 100.	E
			FatPerc must be provided if "Expression of result" is 'B003' "fat weight"	E
R.25	exprRes	Expression of result		
R.26	resQualValue	Result qualitative value	ResQualValue has to be filled in if ResType="BIN".	E
R.27	resType	Type of result	If resType = "LOD" then resLOD must be completed If resType = "LOQ" then resLOQ must be completed If resType = "VAL" then resVal must be completed If resType = "BIN" then resQualValue must be completed If resType = "CCA" then CCalpha must be completed If resType = "CCB" then CCbeta must be completed	E

:Element Code	Element Name	Element Label	Validation	Type
R.28	**resLegalLimit**	**Legal Limit for the result**		
R.29	**resLegalLimitType**	**Type of legal limit**	If resLegalLimit is provided then the resLegalLimitType must be provided	
R.30	**resEvaluation**	**Evaluation of the result**	Where resVal greater than resLegalLimit then resEvaluation must be different from "J002A" (≤ maximum permissible quantities (Compliant result))	**E**
R.31	**actTakenCode**	**Action Taken**	Where resEvaluation = "J003A" – (> maximum permissible quantities (Non compliant result)) than actTakenCode should be provided	**W**
R.32	**resComm**	**Comment of the result**		

17. DESCRIPTION FOR THE TERMINOLOGIES AND VERSIONING

As already pointed out in the previous chapters, controlled terminologies are paramount to the Standard Description. In a similar way to the Standard Description, terminologies need to have a standard set of data elements describing the terms, so that they can be easily imported into the data provider databases.
These data elements include not only the 'term code', 'term name' and the 'term definition' but also additional information used to describe the relationships between the terms and the terms version control.

The maintenance of the controlled terminology is one of the major challenge in order to have a fully working Sample Description. For this reasons, provisions are made to add some data elements to the controlled terminology intended to describe the life cycle of the terms.
For this reasons, provisions are to describe the logical model (described in table 1) and the life cycle of the terms.

The life cycle of the term can be summarised as follows:

- After the term is published and distributed, the '*term code'* cannot be changed.
- Correction for spelling mistakes in the '*term name*', '*term short name*' and '*term definition*' can be considered minor changes and will be amended without publishing a new term.
- When terminology changes are required (replacement of a term, new term addition,…) new terms are added to the controlled terminology and if needed some terms may be flagged as "outdated".

In the validation of the Sample Description, the terms reported, will be checked against the controlled terminology active at the time of the data transmission ("Valid from" <= transmission date < "Valid to" or transmission date > "Valid from" if the "Valid to" date is missing.
The amount of information stored in the terminologies, due to history management, although very useful for IT system management, can be misleading for users.
In the first version of the controlled terminology "valid from" date will be set to "1 Dec 2009".
In order to support a better user understanding of the Standard Description, terminologies will also be distributed in a user friendly format, which will report only the terms active at the moment they were published and only the data element subset which is useful for user. The user friendly format will be a Microsoft Excel workbook containing all the terminologies used in the Standard Description.
The full list of terms and data elements will then be distributed in machine friendly format, to be discussed and described in the "Guidance on data interchange".

The list of data elements for the description of terms is presented in the table below:

Table 3: Data elements for terminologies description

Variable Code	Element Name	Element Label	Type[9]	Description	Mandatory
T.01	**code**	**Term code**	**xs:string (20)**	**Unique code for the term. This is the only code that should be reported in the Standard Description.**	**Yes**
T.02	**name**	**Term name**	**xs:string (50)**	**Term name**	**Yes**
T.03	**short**	**Short name/ acronym/ symbol**	**xs:string (20)**	**Short name, acronym or symbol mainly for concise display**	**No**
T.04	**def**	**Term definition/ description**	**xs:string (250)**	**Definition for the term where available**	**No**
T.05	**ref**	**Term references**	**xs:string (250)**	**Reference for the term**	**No**
T.06	**parentCode**	**Parent code**	**xs:string(20)**	**Link to the parent node in case of hierarchies**	**No**
T.07	**hierarchyCode**	**Hierarchy code**	**xs:string (20)**	**Code representing the hierarchy, for displaying and sorting purposes. This code should not be used for reporting.**	**No**
T.XX	*Additional element name*	*Additional element label*	*Additional element type*	*Additional element description*	*Yes/No*
T.97	**validFrom**	**Valid from**	**xs:date**	**Start date of the validity interval**	**Yes**
T.98	**validTo**	**Valid to**	**xs:date**	**End date of the validity interval. If null the term is valid**	**Yes**
T.99	**comm**	**Comment**	**xs:string (250)**	**Comment to describe the term of reason of changes update.**	**No**

18. CONCLUSIONS AND RECOMMENDATIONS

a. CONCLUSIONS

The Technical Working Group on Data Collection indicates that the Standard Sample Description, defining variables and terminologies, is the first step to harmonising the transmission of sample level data from Member States to EFSA. Furthermore, it is essential to plan and put in place a process for further maintenance and evolution of the standard to allow enhancing and extending the Standard Sample Description for Risk Assessment purposes.

Whilst the Standard Sample Description has been developed specifically to address transmission of Chemical Occurrence and Pesticides data, to date it has only been piloted in the Pesticides domain (2008 Annual Data Collection). Feedback from this experience has been incorporated into the standard data model. Further experience in this and other areas will contribute to enhancing and extending of the data model over time. Moreover, it should be noted that an evaluation of the model's applicability to collection of data in the Zoonoses domain will also be needed.

The group recognises that the ability of each member state to transmit data to EFSA according to the standard data model will vary. Therefore it should be also intended as guidance for Member States to use in planning future developments and evolution of local, regional and national systems with the objective of harmonising data transmissions.

Harmonisation of data collections is recognised as fundamental step to the development of an effective EFSA Data Warehouse. The establishment of the EFSA Data Warehouse is seen as a resource for Europe-wide risk assessment by EFSA and – with appropriate access policies- for Member States.

b. RECOMMENDATIONS

The Technical Working Group on Data Collection, after examining the details of the data elements and the controlled terminologies, makes the following recommendations which are aimed to harmonising the format and mechanism of transmission of data to EFSA for the Chemical Contaminants and Pesticides Residues domains. Common requirements for these domains are addressed in the model. In addition, during the development of this harmonised Sample Description, attention has been paid to commonality with the requirements for the collection of Zoonoses data, although domain specific adjustments have not been addressed.

The following recommendations are addressed to the European Food Safety Authority as the leading organisation and it is anticipated that the work necessary to achieve these will be undertaken in conjunction with Member States and the Commission.

Recommendation 1: Usage and maintenance of the Standard Description

Different reporting templates are currently used and defined in the legislation for the reporting of chemical occurrence and pesticide residues. Efforts should be made for rationalise the reporting requirements of these data and for endorsing the usage of the Standard Description. A maintenance process should be set up to enhance the Sample Description to add new data elements that may be required to cover additional scenarios not taken into account in this version. The maintenance process should foresee a review committee that should meet at regular intervals to discuss changes to the Standard Description. Contact points in the different Member States should be established to coordinate the feedback on the implementation of the Standard Description. A first review is recommended to be organised at the end of 2010.

Recommendation 2: Usage and maintenance of Terminologies

Wherever possible, existing terminologies have been adopted to control the values which will be transmitted for the variables. These may be from international or national sources and have been specified in this document either exactly as per the source terminology or in an adapted form. Where the terminology has been adapted, this has been noted.

It is recommended that maintenance procedures be developed to ensure that amendments to the source terminologies are identified and assessed. An impact analysis on the Sample Description and its user community should be performed before amendments will be published. It is also recommended that the implementation schedule for such changes should be annual and take account of the ability of Member States to implement such changes. EFSA should also look at the development of a software system to maintain and make accessible the terminology to the user community.

Recommendation 3: Food Classification

The development of a matrix catalogue specifying the terminology for food description is paramount to the implementation of the Sample Description. Data providers, in many cases, will use in their local system and their local food classification.

Currently data providers, when reporting data, have to use different food classification for different risk assessment areas. With the new matrix catalogue data providers should be able to use a unique code for reporting, instead of separate codes for the different risk assessment areas. The new food classification is expected to include facets covering a number of product characteristics defined in this document including Product Code, Product Treatment, Method of Production, Packaging and Ingredients. Through this code the receiver should be able to link the required food classification for its specific risk assessment area. In the interim a very basic food sample classification terminology has been defined in this document and provision to accommodate a more comprehensive terminology has been included.

In addition it is recommended that the outcome of the Working Group on Food Classification be implemented in the Standard according to the maintenance procedure of the Standard Description (Recommendation 1).

Recommendation 4: Policies for access and use of data

The group recognises that data transmitted to EFSA using the Standard Sample Description contains a much deeper level of details than previously designed summary data transmission. As such, the data will provide EFSA with much greater flexibility of data analysis and risk assessment but also a corresponding responsibility for maintaining secure access controls to prevent breaches of data confidentiality.

It is recommended that clear policies for access and use of the data are defined and published to establish clarity regarding what are appropriate uses of the transmitted data.

Recommendation 5: Business Rules and Validations

Business rules and validations have been defined in this document which addresses the general validation rules for the domains of Pesticides Residues and Chemical Contaminants. The group identified the existence of different business rules and validations applicable to specific data risk assessment areas. It is recommended that the general validation rules will be made available to those defining specific data collection so that domain specific validations be defined as required. Specific validations should be then communicated to the data providers so that they will not impact on the successful transmission of data.

Recommendation 6: Languages

At present, and partially due to time constraints on this Technical Working Group, the variable and terminologies have been defined in English only.

It is recommended that translations of the variables and terminologies be accommodated as synonyms of the English values defined and further, that the development of such translations be co-ordinated with the relevant member states. It is also recommended that, wherever possible, data in free text fields should be transmitted in English.

It should be noted that throughout the data model, the use of free text fields has been minimised both to reduce the amount of non-standard data and to reduce the difficulty of multiple languages in the free text.

Recommendation 7: Multiple Character sets

Due to language variations across Member States, a number of characters are in common usage in one or more languages which are not standard in all languages e.g.Greek and Cyrillic characters.

It is recommended that the file formats and further, the EFSA Data Warehouse system will support the UNICODE character set, so to support all the character sets currently in usage in Europe.

Recommendation 8: LOD/LOQ definitions can be data collection specific

The group notes that the definition of Limit of Detection (LOD) and Limit of Quantification (LOQ) can vary according to legislation applicable to different data collections. It is recommended that when the data collections are launched, the definitions and the applicable terminologies are reviewed and adjusted as applicable.

Recommendation 9: Review definitions in light of transmissions to evolve data model

The group anticipates that the transmission of data for specific data collections will produce feedback regarding the completeness and accuracy of the terminologies defined in this document. It is recommended that such feedback be incorporated in the next revisions of the lists of variables and terminologies to ensure the data model evolves in parallel with the needs of data senders and EFSA.

Recommendation 10: Harmonise data reporting requirements in legislation for chemicals

Currently, different reporting formats are defined in the legislation for chemical occurrence data. It is recommended that efforts should be made to rationalise and unify, at legislative level, the reporting formats, endorsing the Standard Description.

Recommendation 11: Consider the extension of the Data Model for future Zoonoses sample based data collection

In many national databases, controls on biological data are hosted, on individual sample basis, in the same database as data on chemicals. It should be therefore desirable to extend the Standard Description to include also Zoonoses sample basis data.

Zoonoses data are, at the moment, collected by EFSA at aggregated level. It is anticipated that zoonoses data collection may be extended to implement individual sample based data. In that case, this group recommends also extending the Standard Sample Description defined in this document to cater for transmissions of Zoonoses sample basis data.

Recommendation 12: Possible extension to include Residues of Veterinary Medicines.

As for biological data, in many national databases, veterinary drug residues data are hosted in the same database as chemical contaminants data. It is recommended that consideration be given to the use of this Standard Sample Description for the transmission of data in the domain of residues of veterinary medicines including any necessary adjustments to terminologies to accommodate this data.

Recommendation 13: Compatibility with CEN

The group recognised the ongoing work of the CEN/TC 387 Project entitled “Food Data” in the area of content and interchange format of data about food. It is recommended that this document be shared with that group to identify compatibility challenges.

19. REFERENCES

Keith L. H. et al., "Principles of environmental analysis", *Analytical Chemistry* **55** (14), 2210-2218 (1983).

"General Requirements for the Competence and Testing of Calibration Laboratories", *EN ISO/IEC 17025:2005*

Horwitz W., "Protocol for the design, conduct and interpretation of method performance studies", *Pure Appl. Chem*. **67**, 331–343 (1995).

"Report to the Standing Committee on the Food Chain and Animal Health on the relationship between analytical results, the measurement uncertainty, recovery factors and the provisions in EU food and feed legislation with particular focus on the community legislation concerning contaminants in food and undesirable substances in feed", *EU Website :http://ec.europa.eu/comm/food/food/chemicalsafety/contaminants/report-sampling_analysis_2004_en.pdf*

Thompson M. and Wood R. "The international harmonised protocol for the proficiency testing of (chemical) analytical laboratories", *Pure Appl. Chem*. **65**, 2123–2144 (1993). (Also published in *J. AOAC Int*. **76**, 926–940 (1993).

Thompson M. and Wood R. "Harmonised guidelines for internal quality control in analytical chemistry laboratories", *Pure Appl. Chem*. **67** (4), 49–56 (1995).

Thompson M., Ellison S., Fajgelj A., Willetts P and Wood R. "Harmonised guidelines for the use of recovery information in analytical measurement", *Pure Appl. Chem*. **71** (2), 337–348 (1999).

"Precision of Test Methods", *Geneva, ISO 5725 (1994).*

"Guide to the Expression of Uncertainty in Measurement", *ISO, Geneva (1993)*.

Williams A., Ellison S. L. R., Roesslein, M (Eds.). Quantifying Uncertainty in Analytical Measurement*, 2nd ed. (English) (2000), ISBN 0-948926-1595,* http://www.eurachem.org/.

Thompson M., Ellison S. L. R. and Wood R "Harmonized Guidelines for Single-Laboratory Validation of Methods of Analysis", *Pure Appl. Chem.*, Vol. 74, No. 5, pp. 835–855, 2002

Procedural Manual of the Codex Alimentarius Commission, 18th Edition, *ISBN 978-92-5-106147-3, FAO/WTO, (2009)*

"Guidelines For The Assessment Of The Competence Of Testing Laboratories Involved In The Import And Export Control Of Food", *CAC/GL 27-1997, FAO/WHO*

"Harmonized IUPAC Guidelines For The Use Of Recovery Information In Analytical Measurement" *CAC/GL 37-2001, FAO/WHO*

"Harmonized IUPAC Guidelines for Single-Laboratory Validation of Methods of Analysis," *CAC/GL 49-2003, FAO/WHO*

Guidelines on Measurement Uncertainty" *CAC/GL 54-2004, FAO/WHO*

Guidelines on Measurement Uncertainty", *CAC/GL 54-2004, FAO/WHO*

Harmonized Guidelines for Internal Quality Control in Analytical Chemistry Laboratories", *CAC/GL 65-1997, FAO/WHO*

"Commission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results", *Official Journal L 221, 17.8.2002*

"Regulation (EC) no 882/2004 of the European Parliament and of the Council of 29 April 2004 on official controls performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules" *Official Journal L 165, 30.4.2004*

APPENDICES

The list of all controlled terminologies is available in the Microsoft Excel workbook "standardSampleDescription.xls"