

APPROVED: 28 June 2016

Mapping from FoodEx2 to FoodEx1

European Food Safety Authority; Sofia Ioannidou, Francesco Vernazza and Ilaria Magliano

Abstract

The mapping between FoodEx2 codes and FoodEx1 codes was carried out in order to convert to FoodEx1 occurrence and consumption data provided to EFSA using FoodEx2. This conversion is necessary during the transitional period between the two systems, when data are still analysed with FoodEx1. The FoodEx2 codes considered in the mapping are mainly the reportable terms of the Exposure and Reporting hierarchies. A perfect matching between the two systems was possible for many FoodEx2 terms, but not for all. For the other terms, a series of rules for converting FoodEx2 codes to FoodEx1 codes were identified. When necessary, product treatment descriptors were added to the FoodEx1 matching code in order to keep as much as possible of the information included in the FoodEx2 codes. The resulting mapping was integrated in the field 'foodexOldCode' in the FoodEx2 database.

© European Food Safety Authority, 2016

Key words: food classification, mapping, FoodEx1, FoodEx2

Requestor: EFSA

Question number: EFSA-Q-2016-00367

Correspondence: data.collection@efsa.europa.eu

Acknowledgements: EFSA wishes to thank the EFSA DATA unit for the support provided to this report.

Suggested citation: EFSA (European Food Safety Authority); Ioannidou S, Vernazza F and Magliano I, 2016. Mapping from FoodEx2 to FoodEx1. EFSA supporting publication 2016:EN-1056. 13 pp.

© European Food Safety Authority, 2016

Reproduction is authorised provided the source is acknowledged.

Summary

The European Food Safety Authority (EFSA) and the EU Member States are facing a transitional period where occurrence and food consumption data are provided using one of two food classification and description systems: FoodEx1, which is the first detailed system developed by EFSA and FoodEx2, which is the new system progressively being introduced to replace FoodEx1. Currently, most of the data are still processed and analysed by EFSA using the FoodEx1 system. Therefore tools to facilitate the conversion of data already reported using FoodEx2 codes back to FoodEx1 codes are needed.

The mapping from FoodEx2 codes to FoodEx1 codes was carried out focusing on a total of 4,357 reportable terms belonging to both the Exposure and Reporting hierarchies and 134 hierarchy terms belonging to the Exposure hierarchy. The reportable terms are those identified in the FoodEx2 terminology by the 'Corex' attributes C (Core term), E (Extended term), M (Generic term) or P (Non-specific term) while the hierarchy terms are identified by the 'Corex' attribute H (hierarchy term).

A perfect matching between the two systems was possible only for some FoodEx2 terms. For the remaining codes, a series of rules was defined aiming at linking FoodEx2 terms to the best possible FoodEx1 terms. These rules are described in the present report. When necessary, elements of the PRODTR catalogue, used to code the product treatment in the first version (still in use) of the Standard Sample Description data model (SSD1), were added to the FoodEx1 code representing the best compromise in order to keep as much as possible of the information included in the FoodEx2 codes.

Table of contents

Abstract.....	1
Summary	3
1. Introduction.....	5
1.1. Background and Terms of Reference as provided by the requestor	5
2. Data and Methodologies	6
2.1. Data.....	6
2.2. Methodologies	6
2.2.1. Tools.....	6
2.2.2. Mapping process	6
3. Development of the mapping.....	7
4. Conclusions	11
References.....	12
Glossary and Abbreviations	13

1. Introduction

1.1. Background and Terms of Reference as provided by the requestor

In 2010, the first version of the Guidance on Standard Sample Description (SSD1) (EFSA, 2010) was published by EFSA aiming at harmonising the collection of analytical data for the occurrence of chemical substances in food and feed. In 2013, the data model was extended to cover additional data collection domains such as zoonotic agents in food and animals, antimicrobial resistance and food additives (EFSA, 2013). The revised model (Standard Sample Description ver. 2.0, SSD2) specifies the data elements and the data structure to describe the characteristics of several samples and analytical results transmitted from data providers to EFSA in support to its risk assessments.

Data on chemical substances and microbiological agents are collected in Europe from different matrices (e.g. food, feed, animals, water, environmental samples and food contact materials) and are provided to EFSA according to these data models.

In order to estimate the consumer's exposure to hazards in food, EFSA needs to combine occurrence data with data on food consumption across the EU Member States. The food consumption data have been gathered by EFSA in the Comprehensive European Food Consumption Database; the data in this database have been collected with different methodologies in different countries. Since 2011, EFSA has launched the EU Menu project, which, based on the guidance on the EU Menu methodology (EFSA, 2014), aims at harmonizing the European Food consumption survey methodology across all European countries, thus improving the comparability of the data and allowing a more accurate exposure assessment.

A key element needed to collect, collate, compare and combine data for food safety and to allow dietary exposure assessment is a precise and harmonised description of the food matrix analysed or consumed. The first standardised detailed Food Classification and Description System (initially referred to as FoodEx (EFSA, 2011a), now as FoodEx1) was developed by EFSA at the end of 2008 within the SSD1 framework. It consists of about 1,700 terms (food groups) organized in a hierarchical system based on 20 main food categories that are further divided into subgroups up to a maximum of 4 levels; each term is identified by a unique code. FoodEx1 is the classification used in the 'Comprehensive European food consumption database' project started by EFSA in 2008. In the following years, based on the experience gathered in the use of FoodEx1, a more detailed Food Classification and Description System (named FoodEx2 revision1) was developed and released in 2011 (EFSA, 2011b,c). The new system was tested during three years, leading to highlight further areas for improvement. Considering the outcome of the testing phase, in 2015 FoodEx2 revision 2 was published by EFSA (EFSA, 2015).

FoodEx2 is a comprehensive 'Food classification and description system for exposure assessment' applicable across different food safety domains including food consumption, chemical contaminants, pesticide residues, zoonoses and food composition. The system is based on detailed food groups and different facets. Facets are collections of descriptors, each providing different options to describe a particular aspect of a food group (e.g. treatment, production method, fat content, packaging material, etc.). FoodEx2 is organized in different hierarchies for each relevant food safety domain such as the Exposure hierarchy, the Zoonoses hierarchy or the Pesticides hierarchy. In particular, the Exposure hierarchy was developed for the collection of food consumption data and the estimation of dietary exposure. The Exposure hierarchy consists (as of end of May 2016) of 4,491 terms, including hierarchy terms (134 terms) and reportable terms (4,357 terms). The reportable terms present in the Exposure hierarchy are the same as the reportable terms of the Food section of the Reporting hierarchy, but are aggregated differently in broader categories. The reportable terms are those identified in the FoodEx2 terminology by the 'Corex' attributes C (Core term), E (Extended term), M (Generic term) or P (Non-specific term), while the broader aggregation terms are identified by the 'Corex' attribute H (Hierarchy term).

Currently, EFSA is facing a transitional period, during which consumption data are already collected with FoodEx2 while occurrence data are transmitted by data providers using different standards (SSD1-FoodEx1, SSD2-FoodEx2 format and eventually SSD1-FoodEx2). In the meanwhile, until the switch to SSD2-FoodEx2 is completed, dietary exposure is still calculated using FoodEx1 categories. For this reason the occurrence and food consumption data provided using the FoodEx2 classification

need to be re-coded to FoodEx1 for the purpose of exposure calculation. This transitional period will last until the process of moving the entire data collection and the data analysis to SSD2-FoodEx2 is completed.

Subsequently, the need for tools to facilitate the conversion of FoodEx2 codes to FoodEx1 codes has been identified. The present report describes the development of a mapping of the FoodEx2 Exposure hierarchy terms to FoodEx1 terms or FoodEx1 term combined with product treatment information; the information on product treatment is reported using the catalogue 'PRODTR' of the SSD1. This conversion was foreseen in the structure of FoodEx2, where a field named 'foodexOldCode' was already included and filled to a limited extent in the first revision of the system. However, the rapid evolution of FoodEx2 required launching a specific project to completely amend the initial mapping.

2. Data and Methodologies

2.1. Data

As starting point, a preliminary mapping exercise created for internal use by EFSA in 2012 within the Comprehensive Food Consumption Database framework (using at that time FoodEx2 revision1) was used. Additionally, FoodEx2 revision 2 was now available, with a largely extended set of terms. The final dataset also included the results of the 2015 annual maintenance of FoodEx2 (EFSA, 2016).

2.2. Methodologies

The FoodEx2 terms included in the Exposure hierarchy were matched to the most suitable FoodEx1 codes or to the combination of FoodEx1 code and PRODTR code, using expert judgement based on the general guidance for coding included in the report on development of FoodEx2 revision 2 (EFSA, 2015).

2.2.1. Tools

The tools used for managing the codes were Microsoft Excel® and the FoodEx2 Browser, a user friendly tool developed in Java by staff of EFSA (DATA Unit) to help creating FoodEx2 codes¹.

2.2.2. Mapping process

The structures of the two Food Classification Systems (FoodEx1 and FoodEx2) have a different level of complexity (i.e. level of detail). Therefore, different approaches depending on the case were used in order to identify the most appropriate FoodEx1 code matching each FoodEx2 code.

In some cases a perfect consistency between the two systems was found. In these cases the FoodEx1 code identifying or including the food item described in FoodEx2 was attributed to the FoodEx2 term. However, there were cases when a direct link was not possible: for example, when food groups included in FoodEx2 were not present in FoodEx1; or a FoodEx2 subgroup contained more detailed food descriptors than the equivalent FoodEx1 group. A series of rules for defining the best approximate mapping were set and subsequently applied depending on a case-by-case basis. More details about the mapping process and the rules applied are further explained in the following chapter.

¹ Available online at <http://www.efsa.europa.eu/en/datex/datexfoodclass.htm>

3. Development of the mapping

Since the FoodEx2 system covers a wider range of food items in comparison to the FoodEx1 catalogue, a perfect matching between the two codes was possible only for some food descriptors. Table 1 shows some cases of perfect matching.

Table 1: Examples of perfect matching between a FoodEx2 code and a FoodEx1 code

FoodEx2		FoodEx1	
code	name	code	name
A001H	Rice grain, red	A.01.000035	Rice, red
A002G	Buckwheat flour	A.01.000068	Buckwheat flour
A046P	Aspartame	A.01.001291	Aspartame
A03PZ	Infant formulae, powder	A.01.001716	Infant formulae, powder

However, for other FoodEx2 codes identifying a logic link between FoodEx2 and FoodEx1 terms was more complex. The following rules for mapping two not perfectly matching codes were defined:

1. In the case where there are two or more FoodEx2 terms, one of which has a reasonably precise map to a FoodEx1 term and the other is similar, but more generic (e.g. food 'X' and similar), the FoodEx1 code mapping the most precise FoodEx2 term was used also for the more generic term(s). Examples are shown in Table 2.

Table 2: Examples of matching between FoodEx2 parent terms and its extended term(s) to the same FoodEx1 code.

FoodEx2		FoodEx1	
code	name	code	name
A0D9Y	Barley and similar-	A.01.000019	Barley grain
A000P	Barley grains	A.01.000019	Barley grain
A0D9E	Chamomile and similar-	A.01.000407	Camomile flowers (<i>Matricaria recutita</i>)
A03JC	Chamomile	A.01.000407	Camomile flowers (<i>Matricaria recutita</i>)
A0D9D	Common chamomile	A.01.000407	Camomile flowers (<i>Matricaria recutita</i>)

2. In some cases, FoodEx2 codes are more detailed than the descriptors available in FoodEx1; in these cases the nearest more generic FoodEx1 term was selected. For example, several *Zea mays* subspecies are present (under A000S code 'Maize and similar-') in the Exposure hierarchy, but only one term identifying generically 'Corn grain' is available in FoodEx1 (Table 3).

Table 3: Example of matching when FoodEx2 terms are more detailed than the descriptors available in FoodEx1

FoodEx2		FoodEx1	
code	name	code	name
A000S	Maize and similar-	A.01.000022	Corn grain
A000V	Popcorn kernels	A.01.000022	Corn grain
A000X	Teosinte grain	A.01.000022	Corn grain
A000T	Maize grain	A.01.000022	Corn grain
A0D9S	Indian corn grain	A.01.000022	Corn grain
A048Z	Yeast cultures	A.01.001705	Yeast
A049A	Baking yeast	A.01.001705	Yeast

3. In rare cases, FoodEx1 terms are highly detailed while their parent codes are too generic. Thus, the FoodEx2 extended terms (the most detailed) are perfectly mapped, but their parent terms would be linked to an excessively generic term leading to a loss of information. In these cases, it

was decided to choose the highly detailed FoodEx1 code corresponding to the food item more frequently consumed in Europe. Data from the EFSA Comprehensive European Food Consumption Database² were used to determine which food item is more commonly consumed among the European population. Table 4 shows two cases for which this approach was used: sweet almonds and cultivated fungi are more frequently consumed than bitter almond and wild edible fungi, respectively. Therefore, these were used for the generic FoodEx2 terms (Almonds and Fungi).

Table 4: Examples of matching based on frequency of consumption (in absence of a proper detailed match)

FoodEx2		FoodEx1	
code	name	code	name
A0DYP	Almonds and similar-	A.01.000514	Almond, sweet (<i>Prunus amygalus dulcis</i>)
A014D	Almonds	A.01.000514	Almond, sweet (<i>Prunus amygalus dulcis</i>)
A014F	Almonds sweet	A.01.000514	Almond, sweet (<i>Prunus amygalus dulcis</i>)
A014E	Bitter almonds	A.01.000515	Almond, bitter (<i>Prunus amygalus amara</i>)
A0ETG	Fungi	A.01.000453	Fungi, cultivated
A00TP	Cultivated fungi and similar-	A.01.000453	Fungi, cultivated
A00TD	Wild fungi and similar-	A.01.000458	Fungi, wild, edible

4. In some cases, FoodEx1 terms identifying 'other' not listed items within a group were used to map FoodEx2 codes not present in the FoodEx1 system. Table 5 presents some examples.

Table 5: Examples of matching to FoodEx1 terms identifying 'other' not listed items

FoodEx2		FoodEx1	
code	name	code	name
A0D9X	Foxtail amaranth grain	A.01.000038	Other grains
A0DPL	Blue taros	A.01.000480	Other starchy roots and tubers

5. In the case of FoodEx2 codes for moderately concentrated forms of foods, when both a diluted and a very concentrated form are available in FoodEx1, the approach of 'minimizing the over- or under-estimation in the exposure assessment' was followed. When a concentrated or powdered food is missing in FoodEx1, then the mapping is done to the concentrated or diluted form being nearer in terms of concentration/dilution factor. As an example, 'Liquid coffee concentrate' (A0DLJ) was matched to 'Instant coffee, powder' (A.01.000429) being this the nearest form of coffee in terms of concentration (Table 6).

Table 6: Examples of matching based on the nearest concentration approach

FoodEx2		FoodEx1	
code	name	code	name
A0DLJ	Liquid coffee concentrate	A.01.000429	Instant coffee, powder

6. For some FoodEx2 codes the preferred solution was to select a FoodEx1 generic parent group term since any other more detailed matching, according to the previously described rules, was not suitable (examples shown in Table 7).

Table 7: Examples of matching to FoodEx1 generic parent group term

FoodEx2		FoodEx1	
code	name	code	name
A0DCV	Mosses and lichens	A.01.000317	Vegetables and vegetable products (including fungi)

² <http://www.efsa.europa.eu/en/food-consumption/comprehensive-database>

FoodEx2		FoodEx1	
code	name	code	name
A0C6M	Baumkuchen and similar	A.01.000253	Pastries and cakes
A041V	Meat and vegetable soup	A.01.001856	Ready to eat soups

7. When an entire category of food items was included in FoodEx2 but not in FoodEx1, a 'family' similarity criterion was applied. While applying the similarity criterion, in some cases it was possible to keep at least part of the detail available in the FoodEx2 term (a), though in other cases it was not possible (b). Examples are shown in Table 8.
- In FoodEx1 a subgroup of rolled cereals within the breakfast cereal group does not exist. If by hypothesis all the codes identifying 'rolled cereals' would have been mapped to the generic FoodEx1 term A.01.000184 'Breakfast cereals' the information on the source commodities (i.e. the specific type of grain) would have been lost. In the FoodEx1 system, within the 'Breakfast cereals' group the available subgroups are: 'Cereal flakes', 'Muesli', 'Cereal bars', 'Popped cereals', 'Mixed breakfast cereals', 'Grits' and 'Porridge'. Among these subgroups, 'Cereal flakes' was found to be the most similar food group. Using the terms in this group would allow keeping the information on the source commodity; therefore, the codes belonging to the 'Cereal rolled grains' group of FoodEx2 were matched to the 'Cereal flakes' codes of FoodEx1. This was also the case of e.g. the terms belonging to the 'Animal carcase' group, 'Animal mechanically separated meat (MSM)' group and 'Animal fresh fat tissues' group, which were linked to the FoodEx1 terms included in A.01.000727 'Meat and meat products (including edible offal)' group in order to avoid a loss of information on the animal source.
 - For other FoodEx2 categories not included in FoodEx1, it was only possible to link them to a generic term identifying a similar (or related in the use) food matrix. For example, most of the codes belonging to the group A00SF 'Sprouts, shoots and similar' were matched to the code A.01.000359 'Leaf vegetables' since no appropriate FoodEx1 term was present, and the choice done represented the best approximation even if the 'sprout' concept was lost. The same mapping was done also for A0ESZ 'Flowers used as vegetables'. Similarly, all aromatic flowers were linked to A.01.001581 'Herbs'.

Table 8: Examples of matching based on a similarity criterion

FoodEx2		FoodEx1	
code	name	code	name
A00DF	Millet rolled grains	A.01.000192	Millet flakes
A00DH	Oat rolled grains	A.01.000196	Oat flakes
A00DJ	Oat rolled grains, instant	A.01.000198	Oat flakes, instant
A04AA	Pig carcase	A.01.000731	Pork / piglet meat (Sus scrofa)
A04BZ	Rabbit msm ^(a)	A.01.000735	Rabbit meat (Lepus cuniculus)
A01TX	Bovine fat tissue	A.01.000729	Beef meat (Bos spp.)
A0DKT	Daikon sprouts	A.01.000359	Leaf vegetables
A0DGR	Courgette (edible flowers)	A.01.000359	Leaf vegetables

(a): msm: mechanically separated meat

8. For certain food categories, it was not possible to discriminate whether a food item in FoodEx1 is considered processed or not. In the case of processed food products, it would therefore not have been possible to keep information about the processing by simply using a FoodEx1 code. In these cases, it was decided to add, when possible, the information about the process using terms from the PRODTR catalogue (a specific processing catalogue present in SSD1) in addition to the FoodEx1 term with 'unspecified' treatment. PRODTR is used to describe the treatment applied to a food within the SSD1 framework. This SSD1 element is part of FoodEx2 (as facet) in the SSD2 framework. For the mapping purpose, adding the PRODTR code allows to discriminate e.g. between the dried and the liquid form of certain food items or between a raw primary commodity and its derivative otherwise not distinguishable using the available FoodEx1 codes. In this case the FoodEx1 code and the PRODTR descriptor were joined in a single string as shown

below. Examples of this approach are presented in Table 9. When a PRODTR code was added to a FoodEx1 code, the mapping is composed by the FoodEx1 code followed by a hash character '#', 'PRODTR.', and the PRODTR descriptor code:

A.01.000246#PRODTR.T127A

Table 9: Examples of mapping based on the FoodEx1 list plus the PRODTR catalogue

FoodEx2 code	FoodEx2 name	Mapping code	FoodEx1 name	PRODTR name
A00EN	Porridge (in dry form, to be diluted)	A.01.000246	Porridge	-
A0F0V	Porridge (ready to eat)	A.01.000246#PRODTR.T127A	Porridge	Cooking
A01RN	Horse fresh meat	A.01.000734	Horse, asses, mules or hinnies meat (Equus spp.)	-
A022N	Horse meat, dried	A.01.000734#PRODTR.T131A	Horse, asses, mules or hinnies meat (Equus spp.)	Dehydration
A0C75	Salmons	A.01.000883	Salmon and trout (Salmo spp.)	-
A02KF	Smoked salmon	A.01.000883#PRODTR.T135A	Salmon and trout (Salmo spp.)	Smoking
A041X	Fish soup	A.01.001861	Fish soup	-
A0B9N	Fish soup, dry	A.01.001861#PRODTR.T131A	Fish soup	Dehydration

—: no PRODTR

By applying the rules described above, a total of 4,357 reportable terms belonging to both the Exposure and Reporting hierarchies and 134 hierarchy terms belonging to the Exposure hierarchy of FoodEx2 were mapped to FoodEx1 codes.

The complete mapping list was integrated in the FoodEx2 catalogue in the Correlated Codes section in the field 'foodexOldCode'. It will be therefore available in the FoodEx2 browser tool and in the DCF Catalogue Management System together with the revisions applied with the 2015 annual maintenance (EFSA, 2016).

4. Conclusions

- The full mapping from FoodEx2 terms to FoodEx1 terms was a challenging task because FoodEx2 has considerably increased number of food groups with respect to FoodEx1. Nonetheless, using different mapping rules it was possible to provide a match for all the terms suitable to be used in exposure assessment.
- This mapping list can be used for the conversion of FoodEx2 codes to FoodEx1 codes in order to allow the analysis according to FoodEx1 groups of chemical occurrence and food consumption data collected by EFSA during the transitional period of switch between the FoodEx1 and FoodEx2 standards.
- A perfect match between the two systems was possible only for part of the FoodEx2 terms.
- For the remaining codes, the match was achieved by applying a series of similarity rules for linking FoodEx2 terms to FoodEx1 terms.
- When necessary, codes describing specific treatments from the SSD1 'PRODTR' catalogue were added to the FoodEx1 matching code in order to keep as much as possible of the information included in the FoodEx2 codes.
- The mapping was fully integrated in the FoodEx2 catalogue, via the 'foodexOldCode' attribute.

References

- EFSA (European Food Safety Authority), 2010. EFSA Guidance on Standard sample description for food and feed. EFSA Journal 2010; 8(1):1457, 54 pp., doi:10.2903/j.efsa.2010.1457
- EFSA (European Food Safety Authority), 2011a. Evaluation of the FoodEx, the food classification system applied to the development of the EFSA Comprehensive European Food Consumption Database. EFSA Journal 2011; 9(3):1970, 27 pp. doi:10.2903/j.efsa.2011.1970
- EFSA (European Food Safety Authority), 2011b. Report on the development of a food classification and description system for exposure assessment and guidance on its implementation and use. EFSA Journal 2011; 9(12):2489, 84 pp., doi:10.2903/j.efsa.2011.2489
- EFSA (European Food Safety Authority), 2011c. The food classification and description system FoodEx 2 (draft revision 1). Supporting Publications 2011:215, 438 pp. Available online: <http://www.efsa.europa.eu/en/supporting/doc/215e.pdf>
- EFSA (European Food Safety Authority), 2013. Standard Sample Description ver. 2.0. EFSA Journal 2013; 11(10):3424, 114 pp. doi:10.2903/j.efsa.2013.3424
- EFSA (European Food Safety Authority), 2014. Guidance on the EU Menu methodology. EFSA Journal 2014;12(12):3944, 77 pp. doi:10.2903/j.efsa.2014.3944
- EFSA (European Food Safety Authority), 2015. The food classification and description system FoodEx2 (revision 2). EFSA supporting publication 2015:EN-804. 90 pp.
- EFSA (European Food Safety Authority); Vernazza F and Magliano I, 2016. FoodEx2 annual maintenance 2015. EFSA supporting publication 2016:EN-1049, 24 pp.

Glossary and Abbreviations

Corex	Corex is an attribute of each FoodEx2 term indicating the level of aggregation of the term in the basic food list (core term, extended term, generic term, non-specific term, hierarchy term).
DATA Unit	Evidence Management Unit of EFSA.
DCF	Data Collection Framework.
EFSA	European Food Safety Authority.
EU	European Union.
EU Menu	Fully harmonised European Food consumption survey methodology.
Exposure hierarchy	Food hierarchy included in FoodEx2 arbitrarily defined based on experiences in exposure assessment of chemical contaminants. It particularly focuses on the needs of data analysis and dietary exposure calculation.
Facets	Collections of (food) descriptors from specific points of view included in FoodEx2
FoodEx1	Preliminary detailed food classification developed by EFSA in 2008.
FoodEx2	Multi-purpose food classification and description system developed by EFSA. Revision 1 of FoodEx2 was published in 2011, while revision 2 was released in 2015.
PRODTR	Element of the Standard sample description (SSD1) which allows reporting the product treatment using a specific catalogue.
SSD	Standard sample description.
SSD1	Standard Sample Description rev.1.
SSD2	Standard sample description rev.2 (extended to different food safety domains, including Zoonoses, residues of veterinary medicines and food additives).