EXTERNAL SCIENTIFIC REPORT

Updated food composition database for nutrient intake [1]

M.A. Roe, S. Bell, M. Oseredczuk, T. Christensen, S. Westenbrink, H. Pakkala, K. Presser and P.M. Finglas

On behalf of the project consortium:

Institute of Food Research, Norwich, UK; EuroFIR AISBL, Brussels, Belgium; ANSES, Paris, France; National Food Institute, Technical University of Denmark, Copenhagen, Denmark; Institute of Medical Research, Centre of Research Excellence in Nutrition and Metabolism, University of Belgrade, Serbia; INRAN, Rome, Italy; Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisbon, Portugal; Jožef Stefan Institute, Ljubljana, Slovenia; Matis ohf – Icelandic Food and Biotech R&D, Iceland; Max Rubner-Institut, Karhlsruhe, Germany; RIVM, Bilthoven, Netherlands; THL, Helsinki, Finland; Hellenic Health Foundation (HHF), Athens, Greece; NFA, Uppsala, Sweden; TUBITAK, Turkey; Presser Informatik GMbH, Andelfingen, Switzerland (sub-contracted to Institute of Food Research).

ABSTRACT

This report is the final report of the EFSA contract CFT/EFSA/DCM/2011/03. The aim of the project was to provide EFSA with an updated food composition database covering approximately 1750 foods in combination with additional FoodEx2 facet descriptors included in the EFSA FoodEx2 classification system, and to expand the dataset to include harmonised information on the most common composite recipes of European countries and harmonised information on food supplements. Proposals for models of nutrient composition, composite dishes and food supplement databases were developed as well as guidelines for mapping food data in national datasets to EFSA FoodEx2 codes and facet descriptors. A model for data transfer, compatible with the EuroFIR technical annex and CEN Food Data Standard, and the EFSA data structure was developed and tested. Fourteen national food database compiler organisations supplied initial food lists mapped to the EFSA food list. Information on commonly consumed composite dishes was provided by compilers from thirteen countries and information on food supplements was provided by compilers from eight countries. Datasets compatible with EFSA's data structure were produced based on the models. In addition, guidelines developed and limitations of the data produced are discussed.

KEY WORDS

Food composition, nutrient, FoodEx2, composite dishes, food supplement, dietary intake and EuroFIR.

DISCLAIMER

[1] Question No EFSA-Q-2011-00916.

Any enquiries related to this output should be addressed to DCM@efsa.europa.eu

Suggested citation: M.A. Roe, S. Bell, M. Oseredczuk, T. Christensen, S. Westenbrink, H. Pakkala, K. Presser and P.M. Finglas, 2013. Updated food composition database for nutrient intake. EFSA supporting publication 2013:EN-355, 21 pp.

Available online: www.efsa.europa.eu/publications

SUMMARY

This report is the final report of the EFSA contract CFT/EFSA/DCM/2011/03, taking into account comments and discussions at the final project meeting. The aim of the project was to provide EFSA with an updated food composition database covering approximately 1750 foods in combination with additional FoodEx2 facet descriptors included in the EFSA FoodEx2 classification system, and to expand the dataset to include harmonised information on the most common composite recipes of European countries and harmonised information on food supplements.

An initial food list was provided by EFSA and consisted of 1727 foods from the FoodEx2 food list. The food list was refined to include facet descriptors in combination with FoodEx2 codes to allow an increased level of food description and a total of 3468 FoodEx2 codes and facet descriptors. Proposals for models of nutrient composition, composite dishes and supplements databases were developed as well as guidelines for mapping food data in national datasets to EFSA FoodEx2 codes and facet descriptors.

National food composition database compilers from 14 countries mapped food codes in their published food composition datasets to EFSA FoodEx2 codes and facet descriptors using guidelines agreed by the project Steering Group and EFSA. Codes were matched exactly where possible and close matches were used where no exact match was available. To ensure a complete nutrient dataset was provided, where data was not available in a national database, values were borrowed from another country. The percentage of values borrowed for each dataset ranged from 5% to 90% depending on the range of information available in each national dataset. Countries with extensive datasets borrowed approximately 40% of the values in their final dataset because of the requirement to provide data for all relevant FoodEx2 codes and because of the comprehensive nature of the FoodEx2 food list that contains a large number of foods that would be rarely consumed. Values were borrowed from countries that could provide the data and, in the opinion of each compiler, consume similar types of foods. The large number of borrowed values were not assessed for applicability at the level of each value but national datasets contain the best available data for each country and the main focus is on providing information for foods that are commonly consumed and important to the national diet, therefore borrowed values should not have a significant impact on nutrient intake assessments. Nutrient information was provided for over 100 prioritized nutrients for 1320 of the list of 1724 FoodEx2 codes and 1258 combinations of FoodEx2 codes and facet descriptors. Nutrient information was provided according to nutrient data available in each national food composition database and varied depending on relevance to each food. All datasets provided data for energy, protein, total fat and carbohydrates. Values for fatty acids were provided in all datasets but coverage across the range of foods was limited in some datasets and coverage of specific isomers varied. Data for vitamins B1, B2, B3, B6, B12, A, C, D, E and folate were provided by a minimum of 10 countries. Biotin and pantothenic acid data were provided by 6 countries. All countries provided values for minerals and trace elements but coverage varied by country and food. Amino acid data was available from 2 countries but only for a limited selection of foods. There were 887 (26%) of the 3468 FoodEx2 codes and facet descriptor combinations that did not have a code assigned by any FCDB. There was no data available from any country for 53 of the EFSA FoodEx2 codes (excluding facet descriptors) and these foods were therefore considered not relevant to food consumption.

Compilers from 10 countries provided information for at least 200 commonly consumed composite dishes. Information for a smaller number of composite dishes was also provided by four other countries. All ingredients of the composite dishes were disaggregated and were linked to an EFSA FoodEx2 code to enable linking to the Comprehensive European Food Consumption Database.

Nine countries provided lists of food supplements based on consumption recorded in national intake surveys. Supplement types included in the dataset included minerals, vitamins, combinations of vitamins and minerals, plant extracts, protein and amino acids supplements and fatty acid supplements. Eight of the supplement lists also included information on supplement composition (including fatty acids) held in their national datasets or compiled from product information sources.

Data was collected by compilers in a format compatible with EuroFIR specifications and thesauri and was managed using the FoodCASE food composition database management system. This database structure was amended with the addition of fields and tables that were required for compatibility with the EFSA specification for standard sample description for food and feed. EuroFIR thesauri for data documentation were mapped to EFSA controlled vocabularies to ensure compatibility with the EFSA Data Collection Framework. A model for data transfer, compatible with the EFSA data structure was developed and tested and the compiled databases were supplied to EFSA in accordance with a standard data model and through the EFSA Data Collection Framework.

Limitations of the data were discussed and recommendations for further development of harmonised and standardized datasets were proposed.

- The FoodEx2 food list and facet descriptors should be reviewed for use with food composition data and further clarifications, scope notes, and examples provided to facilitate improved accuracy and consistency of matching food composition data to FoodEx2 codes.

- Procedures for providing data when values for a food are not available for a particular country should be reviewed. There are limitations to the approach of borrowing values from another dataset because it cannot be guaranteed that borrowed values are directly comparable to foods consumed within a country. Published food composition datasets do not always take into account recent changes in food production that may impact on nutrient content. Further additional evaluation is possible now that the final dataset is complete and borrowed data can be easily identified. In particular guidelines should be considered for foods where nutrient content may differ due to processing practices, e.g. fortification.

- Further consideration should be given to the use of retention factors in recipe calculations. Recommendations exist but practices are very different between countries and some countries do not use them. The applicability of retention factors could be assessed by reviewing differences between calculated and analysed values.

- Information on food supplements is not consistently collected or maintained. Collection of data from manufacturers on a regular systematic basis should be investigated to allow maintenance of a more comprehensive database.

- Member states and candidate countries should be encouraged to provide additional resources for generation of new analytical data on food to improve the quality of national datasets and avoid the necessity to calculate values or borrow values from another country.

TABLE OF CONTENTS

BACKGROUND

The Dietary and Chemical Monitoring Unit of EFSA collects and collates consumption and chemical occurrence and nutrient composition data for risk assessment purposes and other scientific tasks of EFSA. The Comprehensive European Food Consumption Database consists of detailed food consumption information collected in 2008-2009 from 20 countries for adults and 12 countries for children, and it was updated with the latest food consumption data from 8 countries during 2012. The FoodEx food description and classification system was developed by EFSA to link food consumption data to the collection of chemical occurrence data. FoodEx is being refined by the Working Group on "Development of a Food Classification and Description System for exposure assessment" with the aim of serving a broad range of needs, including nutrient intake calculations. FoodEx2 was used in the current project.

Food composition data is essential for calculation of nutrient intake from consumption data and is available in most countries (Finglas et al., 2009). There is a need for composition data to be collected and presented in a standardised way to enable comparison between countries and across Europe (Becker, 2010). This problem has been addressed by several European projects, most recently by the completed FP6 EuroFIR (European Food Information Resource) Network of Excellence project, which has led to the development of a draft European standard for food data, and the establishment of the EuroFIR Food Data Platform in Europe. The work is ongoing through the EuroFIR Nexus project (FP7; 2011-13), led by the Institute of Food Research, and supporting activities coordinated through EuroFIR AISBL, a non-profit International Association of food database compilers, users and stakeholders, based in Belgium.

EFSA requires an updated food composition dataset to cover both foods and food supplements and to link the data to the European Comprehensive Food Consumption Database for calculation of nutrient intake and exposure. This requirement will enable EFSA to evaluate nutrient composition of European foods and estimate nutrient intakes and patterns in Europe, in response to information requests from the Commission and member states. The current proposal brings together the expertise and resources of 14 national food database compiler organisations and EuroFIR AISBL.

TERMS OF REFERENCE

Overall objective

The purpose was to provide EFSA with an updated food composition database covering approximately 2,000 foods included in the EFSA FoodEx2 food classification system and to expand the EFSA nutrient dataset to include harmonised information on the most common composite recipes of European countries and harmonised information on food supplements.

Specific objectives

There were 8 specific objectives:

(1) To organise a 1-day kick-off project meeting to be held in Parma at EFSA's premises within 1 month of the contract start date.

(2) To organise an interim web/teleconference to be held within 5 months of the contract start date.

(3) To organise a 1 day final project meeting to be held in Parma at EFSA's premises within 9 months of the contract start date.

(4) To plan the harmonised data compilation and evaluate the availability and quality of data.

(5) To assemble and provide in a database the most recent (published in a national food composition dataset since 2005) information on macro-nutrients, micronutrients and other dietary compounds required for use with FoodEx2 (published in December 2011). This will cover approximately 2,000 foods based on entry codes contained in the EFSA FoodEx2 food classification system. The type of data and other background information related to nutrient value should also be provided.

(6) To compile a database of the most common European composite dishes to include food name, ingredient list, quantification of raw ingredients and recipe handling procedures. A minimum of 200 composite dishes should be included for each country. Data compilation must be harmonised between each country and ingredient information must be related to the EFSA food classification system entry codes.

(7) To compile a food supplement database containing information on at least 100 of the most commonly sold vitamins, minerals and other food supplements.

(8) To propose (and/or further develop) harmonised and standardised rules for the compilation of each of the databases (nutrient composition, composite dishes and food supplements).

This contract was awarded by EFSA to the consortium consisting of Institute of Food Research, Norwich Research Park, Colney, Norwich, United Kingdom; EuroFIR AISBL, Brussels, Belgium.; ANSES, Paris, France.; National Food Institute, Technical University of Denmark, Copenhagen, Denmark.; Institute of Medical Research, Centre of Research Excellence in Nutrition and Metabolism, University of Belgrade, Serbia.; INRAN, Rome, Italy.; Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisbon, Portugal.; Jožef Stefan Institute, Ljubljana, Slovenia.; Matis ohf – Icelandic Food and Biotech R&D, Iceland.; Max Rubner-Institut, Karhlsruhe, Germany.; RIVM, Bilthoven, Netherlands.; THL, Helsinki, Finland.; Hellenic Health Foundation, Athens, Greece.; NFA, Uppsala, Sweden.; TUBITAK, Turkey with Institute of Food Research having the leader role.

Contract title: Updated food composition database for nutrient intake

Contract number: CFT/EFSA/DCM/2011/03

INTRODUCTION AND OBJECTIVES

Introduction

The aim of this project was to provide EFSA with an updated food composition database covering approximately 2,000 foods included in the EFSA FoodEx2 food classification system and to expand the dataset to include harmonised information on the most common composite recipes of European countries and harmonised information on food supplements. This report is the final report (project deliverable 8) describing the results of the project, including the methods used for preparing nutrient, composite dish and food supplement datasets. Development of a standard method for data transfer is also described.

MATERIALS AND METHODS

1. PREPARATION OF THE PROJECT PLAN

An agenda for the first project kick off meeting (EFSA, Parma, Italy, 9-10th February, 2012) was agreed with members of the Project Steering Group and EFSA. The meeting agenda included:

- An overview of EFSA's requirements;
- An update on FoodEx2 and discussion on the nutrient database food list;
- Revision of the project schedule and deliverable list to take into account the Christmas and New Year period falling between the contract signature and the kick off meeting;
- Project consortium presentations and discussion on:
    - Outline proposals for models of the nutrient composition, composite dishes and food supplement databases
    - Outline rules for harmonising data compilation and evaluation of data availability (including mapping foods in national datasets to those included in FoodEx2) and data quality)
    - Outline model for data transfer, based on the EuroFIR Food Data Transfer Protocol.

A report of the meeting, a revised action plan and deliverables were included in the agreed minutes of the meeting.

An interim web/teleconference to discuss progress and plans for project completion was held on 19th June 2012. Minutes of the meeting were agreed and reported as Deliverable 4.

An interim report describing the progress of the project, coverage and quality of the data to be included in the databases and describing the harmonised procedures to be used to reach the goals of the database compilation was submitted as Deliverable 5.

Databases on i) micro and macronutrients and other dietary compounds, ii) names, list of ingredients and quantitative breakdown to raw ingredients of composite dishes, and iii) food supplements were provided as Deliverable 6.

2. NUTRIENT COMPOSITION DATA

An initial food list was provided by EFSA and consisted of 1727 foods from the FoodEx2 food list. A list of facet descriptors that would provide additional food description and that would be relevant to foods that were reported in the EFSA Comprehensive food consumption database was also included.

Draft guidelines for data compilation, including rules for nutrient data harmonisation and rules for compilation of composite dish and supplement data were agreed with the project Steering Group. The guidelines outlined the procedure for compilers to map codes from their national food composition datasets to FoodEx2 codes, and gave examples of possible problems related to food description and coding. A final version, taking into account comments from the project Steering Group and EFSA was agreed on 27/3/12 (Appendix A).

Based on the guidelines for data compilation, the food list was revised to produce a list that could be used by both EFSA and food composition data compilers. EFSA confirmed that combinations of EFSA food code and facet descriptors should be treated as different foods, and that if a particular facet descriptor would not apply to a code then that code/facet descriptor combination could be excluded from the food list. EFSA also provided a mapping of FoodEx1 codes towards FoodEx2 codes to help compilers who already had FoodEx1 codes in their national datasets.

The food list was finalised (Appendix B) following EFSA clarifications and consisted of 3468 combinations of food/facet descriptor with EFSA FoodEx2 codes and facet descriptors combined as a single string. A template containing the food list was sent to compilers for adding matched food codes from national datasets to combinations of EFSA FoodEx2 code and facet descriptors.

Compilers provided feedback on the code mapping process and problems with mapping national codes with combinations of EFSA FoodEx2 code and facet descriptors were identified. Possible approaches to include additional codes, based on close matches were proposed by compilers and agreed by the project Steering Group (Appendix C).

3. COMPOSITE DISH DATA

Compilers were asked to provide information on 200 commonly consumed composite dishes as outlined in the guidelines for data compilation (Appendix A).

The aim of this dataset was to provide the names and quantitative breakdown to raw ingredients, mapped according to FoodEx2, of the most commonly consumed composite dishes in different European countries. The recipe data included clear descriptions of the recipe handling procedures with yield and retentions factors used. For this project compilers considered a composite dish as described in the scope note for composite foods in FoodEx2:

- ‘foods containing more than one distinct ingredient and having a recipe behind them, where the ingredients are so diverse that no relevant main category can be identified’.

Composite dishes were identified based on intake frequency from consumption surveys wherever possible or from composite dishes included in national datasets. A template (Appendix D) for data compilation was produced that is compatible with the EuroFIR technical annex and CEN Food Data Standard.

4. SUPPLEMENT DATA

Compilers were asked to provide information on commonly consumed food supplements based on information from national consumption surveys/research projects and market share data, where available. The aim was to provide information on a minimum of 100 vitamin, mineral, fatty acid and other supplements. A template (Appendix E) to collect this information in a standard format, compatible with the EuroFIR technical annex was produced.

5. DATA TRANSFER

The compiled databases were supplied to EFSA in accordance with a standard data model and through the EFSA Data Collection Framework. EuroFIR uses a standardised relational database structure, compatible with the EuroFIR technical annex, to host data for the EuroFIR e-Search tool. This database structure was amended with the addition of fields and tables that were required for compatibility with the EFSA specification for standard sample description for food and feed. EuroFIR thesauri for data documentation were mapped to EFSA controlled vocabularies to ensure compatibility with the EFSA Data Collection Framework.

EuroFIR AISBL has an agreement in principle for use of the FoodCASE food data management software among its full members. This software has been developed by ETH Zurich and may be further modified in the future subject to agreement with ETH. The project team compiled data into Excel format templates. FoodCASE was then used to access nutrient composition values for national codes that had been matched to EFSA FoodEx2 codes and facet descriptors. Details of EuroFIR standards and proposed data formats were supplied to EFSA on 17/4/12 and 10/5/12. EFSA provided an Excel format schema (Appendix F) that was mapped to a subset of the EuroFIR technical annex on 22/5/12. The EFSA schema described details of elements and controlled terminologies for food and nutrient data and composite dishes, including yield and retention factors. Datasets produced in FoodCASE were exported in Excel formats and sent to EFSA for initial testing in July 2012. As the exports were not compatible with the proposed EFSA schema further necessary changes were identified and Excel outputs compatible with the EFSA schema were produced.

Retention factors used for calculation of micronutrients in composite dishes are handled in different ways by national compilers. Although standard factors have been proposed for use by EuroFIR, many national compilers use variations of these standard factors, depending on past practice, to retain comparability of their data within their own datasets. Retention factors are not associated with all values in national datasets and therefore it was only possible to provide information at the composite dish, ingredient and component level for data from 7 countries. Four countries (France, Netherlands, Serbia and Greece) do not use retention factors but use values for cooked ingredients in recipe calculations.

This data transfer model is sufficiently flexible to allow future updates to reflect changes in national food composition data and to allow for additional datasets to be added.

RESULTS

1 NUTRIENT COMPOSITION DATA

Details of national codes mapped to EFSA FoodEx2 codes/facet descriptor combinations are presented in Table 1.

Table 1: National food codes mapped to EFSA FoodEx2 codes/facet descriptors[1]

Country	Food codes in national dataset	National food code matched to EFSA code	% of national codes mapped to EFSA codes	EFSA food list codes with no exact or close match[2]
Denmark	1049	626	59.7	2842
Finland	2093	804	38.4	2664
France	1351	893	66.1	2575
Germany	14814	2120	14.3	1348
Greece	305	217	71.1	3251
Iceland	1265	547	43.2	2921
Italy	790	678	85.8	2789
Netherlands	2080	1024	49.2	2444
Portugal	962	511	53.1	2957
Serbia[3]	1198	496	41.4	2962
Slovenia	437	390	89.2	3070
Sweden	1500	628	41.9	2840
Turkey	500	149	29.8	3319
UK	3423	1074	31.4	2394
eBasis/PlantLibra	391	189	48.3	-

[1]Total number of EFSA codes/facet descriptors in the food list used within this project was 3468

[2]Foods where there was not an exact match to the EFSA code/facet descriptor in the national dataset

[3]The Serbian national dataset includes EFSA food codes but does not currently use facet descriptors so the level of food matching may be different to other compilers who all used facet descriptors

National food composition database (FCDB) compiler organisations mapped existing national codes to EFSA FoodEx2 codes and facet descriptors as closely as possible. The number of codes matched to an EFSA FoodEx2 code/facet descriptor is for foods that are exact or very close matches based on food names and descriptions. The proportion of 'close' matches used was low (<10) and mainly related to cooking method facet descriptors e.g. using a code for 'cooked in water' to match 'steamed'. Where a FoodEx2 code could not be matched to a food within a national dataset a code was borrowed from another country's dataset. 189 Codes used in the EuroFIR e-BASIS/ PlantLibra database were also mapped to EFSA FoodEx2 codes to enable values for bioactive compounds to be obtained, however e-BASIS contains many values for the same compound and plant food and it is not possible to identify a single 'best' value.

2 GAP FILLING NUTRIENT DATASETS

To ensure a complete nutrient dataset was provided, where data was not available in a national database, values were borrowed from another country based on priorities assigned by the national compiler. A script was produced to complete datasets by searching for codes available from other countries and assigning the code from the most appropriate country with available data. A similar programme was used to complete datasets where values for a particular nutrient were not available.

The percentage of values borrowed for each dataset ranged from 5% to 90% depending on the range of information available in each national dataset. Countries with extensive datasets borrowed approximately 40% of the values in their final dataset because of the requirement to provide data for all relevant FoodEx2 codes included in the EFSA food list and because of the comprehensive nature of the FoodEx2 food list that contains a large number of foods that would be rarely consumed and would not be nationally relevant. Values were borrowed from countries that could provide the data and, in the opinion of each compiler, consume similar types of foods. The large number of borrowed values were not assessed for applicability at the level of each value but national datasets contain the best available data for each country, in the opinion of the national compiler, and the main focus is on providing information for foods that are commonly consumed and important to the national diet. Borrowed values mainly related to foods that would not be relevant nationally and therefore should not have a significant impact on nutrient intake assessments. Recipe calculations were not used because the vast majority of FoodEx2 codes that could not be matched were for single food items rather than composite dishes and recipe calculations.

There were 887 (26%) of the 3468 FoodEx2 codes and facet descriptor combinations that did not have a code assigned by any FCDB. There were 53 EFSA FoodEx2 codes (excluding facet descriptors) that did not have a national code assigned to any FoodEx2/facet descriptor combination by any FCDB. The fact that no data was available from any country indicates that these combinations of code and facet descriptor could be considered much less relevant to food consumption.

3 LIMITATIONS OF NUTRIENT DATASET

During the process of mapping national food codes to EFSA codes and facet descriptors, several problems were identified and potentially may affect the usability and applicability of the dataset as follows:

- **No direct match in national dataset:** In these cases codes were borrowed from another country's dataset based on prioritized datasets assigned by each compiler, e.g. Denmark borrowed codes from Germany, Sweden or Finland wherever possible.

- **More than one national code could be matched to the FoodEx2 code/facet descriptor:** In these cases, compilers chose the code that would most commonly be consumed or would be the best match based on their judgement.

- **A national code could be mapped to more than one FoodEx2 code/facet descriptor:** In many cases the same national code was mapped to more than one FoodEx2 code/facet descriptor, particularly for FoodEx2 code/facet descriptor combinations within the same FoodEx2 code and also where both generic and more specific versions of a food are included in the EFSA food list.

- **FoodEx2 code and facet descriptor did not accurately define food, e.g. A005K Bread and rolls with special ingredients added:** In these cases the compiler mapped a national code based on what would most commonly be consumed, e.g. bread with tomato, bread with fibre, bread with seeds etc. The food used by the national compiler will be specified in the original food name and additional information is linked to Langual codes associated with the national dataset.

- **FoodEx2 code was well defined but foods mapped in different national datasets or within a dataset could be quite different in type and composition:** Compilers used a national code that would best fit the description and would match what could most commonly be reported as being consumed in a dietary survey, e.g. A00BC Muffins, which were coded as American style muffins in the UK dataset but could also be mapped to English style muffins.

- **Nutrient values may not represent most up to date food composition:** Nutrient data is from the most recent published national datasets but may not reflect recent changes in composition arising from changes in product reformulation or legislation, e.g. the UK dataset does not take into account recent reductions in trans fat and saturated fat content of some processed foods. The data has been collected but not yet published in the national dataset. Following UK publication this information will be available for any updated versions of the nutrient dataset. For all national datasets, the data provided was as published in July 2012.

- **Where a food code was borrowed from another country, it was not possible, at this stage, to assess whether the published values for that food had already been borrowed from another dataset:** In these cases an evaluation would already have been made by the national compiler and documentation associated with the values would provide information on what kind of publication the data came from with a reference where known.

- **Translation inaccuracies:** Each FCDB was responsible for any translation and mapping of foods. In some cases, food descriptions had to be translated from the national language into English and some mapping inconsistencies and errors may have occurred at that stage. Each country was responsible for checking their own food code mapping prior to submission to the coordinator. Mapping consistency was checked centrally (by IFR/EuroFIR AISBL) where possible; however, without detailed knowledge of the local language some food translation errors may not have been identified. This is estimated to be <5% based on a central evaluation of the consistency of both the mapping and translation by each country.

- **Nutrient gap filling:** Where food codes did not have values for a particular nutrient, values were borrowed from another country using the same method as for borrowing codes. The proportion of nutrient values borrowed depends on the nutrient coverage of each national dataset. Where a national food code was mapped to a FoodEx2 code, nutrient values were available for all major macronutrients, minerals and vitamins. However for lower priority nutrients identified by EFSA, a large percentage of the values was borrowed and came from the few countries where that data was available. Availability of data for required components is described in Table 2, although it should be noted that not all components are relevant for all foods and coverage of some components was limited within datasets. In some cases it is possible that a food in one country was not well matched to a food in another country and values used may not necessarily reflect the food as consumed. Since the dataset contains >4 million values (3468 foods and >100 nutrients (including fatty acids and summation of fatty acid isomers) from 14 countries), it is not possible to check the overall robustness of each value in the final datasets, although the compilers of each FCDB has already checked all values in their national published datasets. Where data has been borrowed the source of data is 'tracked' in FoodCASE and any value can be investigated in comparison with values from another country. The borrowed values published by other compilers had all been evaluated prior to national publication and were considered to be the best values available. Therefore no distinction was made between values that were calculated and values that were directly analysed. Calculated values are usually derived from analysed values, although conversion factors are used for some nutrients, and all values are subject to a degree of variation depending on the food and nutrient. A check of sum of proximates highlighted a very small number of foods (approx. 75) where the sum of proximates was <90 or > 105 and could not be explained or corrected. In all cases the data were as published in the national dataset. It will be possible to identify any outliers or incorrect published data by reviewing documentation included in the EuroFIR e-Search dataset.

Table 2: Component coverage in foods included in the EFSA nutrient dataset[1]

Nutrient/Country	DK	FI	FR	DE	GR	IS	IT	NL	PT	RS	SI	SE	TR	UK
Energy (kJ, kcal)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Protein	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fat, total	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Carbohydrates, total[2]	✓	x	x	✓	✓	✓	x	x	x	✓	✓	x	x	x
Carbohydrates, glycaemic[2]	✓	✓	✓	x	✓	x	✓	✓	✓	✓	✓	✓	✓	✓
Starch	✓	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	x	x	✓
Dietary fibre (AOAC)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sugars, sum	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	x	x	✓
Sugar, added[3]	✓	x	x	x	x	✓	x	x	x	x	x	✓	x	x
Cholesterol	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Alcohol	✓	✓	✓	✓	x	✓	✓	✓	✓	x	✓	✓	x	✓
Ash	✓	✓	x	✓	✓	✓	x	✓	✓	✓	✓	✓	✓	x
Water	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fatty acids[4]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Amino acids[5]	x	x	x	✓	x	x	x	x	x	x	✓	x	x	x
Mono and disaccharides[6]	✓	x	x	✓	✓	x	✓	✓	✓	✓	✓	✓	x	✓
Biotin[7]	✓	x	x	✓	x	x	✓	x	x	✓	✓	x	x	✓
Folate	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	✓	✓	x	✓
Niacin	✓	✓	✓	✓	x	✓	✓	✓	✓	x	✓	✓	x	✓
Pantothenic acid[7]	✓	x	✓	✓	x	x	✓	x	x	✓	x	x	x	✓
Pyrodoxine (B6)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	x	✓
Retinol	✓	✓	✓	✓	✓	✓	x	✓	✓	x	x	✓	x	✓
Riboflavin	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Thiamin	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Vitamin A (ret equiv)	✓	✓	x	✓	x	✓	✓	✓	✓	✓	x	✓	✓	✓
Vitamin B12	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	✓	✓	x	✓
Vitamin C	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	x	✓	✓

Nutrient/Country	DK	FI	FR	DE	GR	IS	IT	NL	PT	RS	SI	SE	TR	UK
Vitamin D	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	✓	✓	x	✓
Vitamin E	✓	✓	✓	✓	✓	✓	✓	✓	x	✓	✓	✓	✓	✓
Vitamin K (Phylloquinone) [8]	✓	x	x	✓	x	x	x	x	x	✓	✓	✓	x	✓
Calcium	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Chromium[7]	✓	x	x	x	x	✓	x	x	x	✓	✓	x	✓	x
Copper	✓	✓	✓	✓	✓	✓	✓	✓	x	✓	✓	x	✓	✓
Fluoride	x	x	x	✓	x	✓	x	x	x	x	✓	x	x	x
Iodine	✓	✓	✓	✓	x	✓	✓	✓	x	✓	✓	✓	x	✓
Iron	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Magnesium	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Manganese	✓	x	✓	✓	x	✓	✓	x	x	✓	✓	x	✓	✓
Phosphorus	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Potassium	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Selenium	✓	✓	✓	x	✓	✓	✓	✓	x	✓	✓	✓	x	✓
Sodium	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sodium chloride (Salt)	x	x	x	✓	x	x	x	x	x	x	✓	x	x	x
Zinc	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Beta-carotene	✓	✓	✓	✓	✓	✓	x	✓	x	✓	✓	✓	✓	✓
Carotenoids[1]	x	x	x	✓	x	x	x	✓	x	x	✓	x	x	x
Choline	x	x	x	x	x	x	x	x	x	✓	✓	x	x	x
Isoflavones[7]	x	x	x	x	✓	x	x	x	x	x	x	x	x	x
Phytosterols (sum) [1]	x	✓	x	x	✓	x	x	x	x	x	✓	x	x	✓

[1]Data may be available from some countries for some components for a limited range of foods that are not included in the EFSA food list.
[2]Data is provided as total or glycaemic carbohydrates depending on the calculation method used by each country.
[3]Data for added sugars is only available from three countries.
[4]Values for fatty acids are available from all countries, however coverage is limited in some datasets and coverage of specific isomers will vary between countries and between foods.
[5]Amino acid data is only available from two countries for foods included in the EFSA dataset. Older data is available for a limited range of foods from six countries but those food codes are not mapped to the EFSA nutrient dataset. Coverage of individual amino acids is variable and is likely to include data for cystine (the oxidised disulfide derivative) rather than cysteine.
[6]Individual sugars are available in 10 datasets but coverage is often limited to relevant foods. Total monosaccharides and disaccharides are not included as components.

[7]Finland has data for a very limited range of foods that are not included in the EFSA food list.
[8]Included as total vitamin K in some datasets. Vitamin K_2 is not routinely measured.

4 COMPOSITE DISH DATA

Composite dish data was compiled as reported in Table 3.

Table 3: Composite dish data provided by compilers

Country	Number of dishes identified
Denmark	246
Finland	208
France	299
Germany	200
Greece	233
Iceland	67[1]
Italy	200
Netherlands	212
Portugal	190
Serbia	67
Slovenia	201
Sweden	559
Turkey	243
UK	197

[1]There are a limited number of composite dishes consumed in Iceland and the total represents a significant proportion of dishes likely to be consumed

Compilers from 10 countries provided information for at least 200 composite dishes. Information for a smaller number of composite dishes was also provided by Iceland, Portugal, Serbia and UK. All ingredients are disaggregated and are linked to an EFSA FoodEx2 code to enable linking to the Comprehensive European Food Consumption Database. Where countries provided recipe ingredient quantities as percentages, they were converted to grams. Details of yield factors, applied at food level to take into account weight change during cooking, used in recipe calculations were provided for each recipe. Country specific retention factors were provided for six countries (Germany, France, UK, Netherlands, Serbia and Sweden). Seven countries were unable to provide information related to specific values but used EuroFIR recommended factors (Bognar, 2002) applied at food level or

calculated recipes based on cooked ingredients. Greece does not use retention factors in recipe calculations.

5 SUPPLEMENT DATA

Supplement data provided by compilers is given in Table 4.

Table 4: Supplement data provided by compilers

Country	Supplement data available	Number of supplements listed	Composition information provided
Denmark	No	0	-
Finland	Yes	92	Yes
France	Yes	449	Yes
Germany	No	0	-
Greece	Yes	76	No[1]
Iceland	No	0	-
Italy	Yes	126	Yes
Netherlands	Yes	90[2]	Yes
Portugal	No	0	-
Serbia	Yes	8	Yes
Slovenia	Yes	95	Yes
Sweden	Yes	122	Yes
Turkey	No	0	-
UK	Yes	103	Yes

[1]Not currently available from Greece.

[2]Does not include supplements for children.

National FCDB from nine countries provided lists of food supplements based on consumption recorded in national intake surveys. Eight of the food supplement lists also included information on food supplement composition (including fatty acids) held in their national datasets or compiled from product information sources.

CONCLUSIONS

The project objectives were challenging, both in terms of timescale and delivery of the quantity of nutrient data to meet EFSA's exact requirements via the Data Collection Framework.

The following conclusions can be drawn from the project:

- National food composition data has been provided for 1320 single FoodEx2 codes and 1258 combinations of FoodEx2 codes and facet descriptors. Foods as described in food composition datasets did not match well for many FoodEx2 codes without the addition of further description provided by facet descriptors.

- FoodEx2 codes, even with the addition of a limited range of facet descriptors, was not always adequate to allow unambiguous matching of foods in national datasets to foods require by EFSA and may have led to some inconsistencies in matching foods in national datasets to FoodEx2 codes.

- Nutrient data was delivered from 14 countries with data for over 100 components. Countries were not able to provide data for all foods and nutrients because of the limited range of data currently available in national datasets. Nutrient coverage by each country matched the data currently available in national datasets and matched the component coverage descriptions and limitations provided in the technical offer.

- Procedures for providing data when values for combinations of foods and nutrients were not available for a particular country were developed based on borrowing data from countries with available data, where food consumption would be similar. There are some limitations to this approach because it cannot be guaranteed that borrowed values are directly comparable to foods consumed within a country. However national compilers already estimate (using recipe calculations or estimates based on content in similar foods) missing values for important foods and nutrients in their published FCDBs, so borrowed values mainly relate to foods which are less important or not relevant to national consumption. In many cases borrowed values are from countries with the largest datasets. Values were borrowed from countries that could provide the data and, in the opinion of each compiler, consume similar types of foods. The large number of borrowed values were not assessed for applicability at the level of each value but national datasets contain the best available data for each country and the main focus is on providing information for foods that are commonly consumed and important to the national diet, therefore borrowed values should not have a significant impact on nutrient intake assessments.

- Data on composite dishes with quantified information on ingredients and yield factors was provided by 13 countries. Country specific retention factors were provided for six countries (Germany, France, UK, Netherlands, Serbia, Sweden). Seven countries were unable to provide information related to specific values but used EuroFIR recommended factors (Bognar, 2002) applied at food level or calculated recipes based on cooked ingredients. Greece does not use retention factors in recipe calculations.

- Information on food supplements was provided by eight countries, including composition information. A list of commonly consumed supplements was also provided by Greece but

composition data was not provided so it was not included in the final dataset. The food supplement dataset provides an important central resource for information that has not been consistently compiled and maintained by European food data compilers.

- The data for the three datasets was compiled according to the EuroFIR standard, which is compatible with the draft CEN standard for food composition data. The datasets were managed within the FoodCASE database management system and the datasets were exported in a format mapped to the EFSA data collection framework and EFSA controlled vocabularies. EuroFIR thesauri developed for use by food composition data compilers were successfully mapped to EFSA controlled vocabularies.

RECOMMENDATIONS

- The FoodEx2 food list and facet descriptors should be reviewed for use with food composition data and further clarifications, scope notes, and examples provided to facilitate improved accuracy and consistency of matching food composition data to FoodEx2 codes.

- Training on use of the FoodEx2 classification system for compilers of nutrient composition data could be considered.

- Consideration should be given to updating and maintaining national food composition datasets using procedures developed within this project and extending to other member state and candidate countries.

- Member state and candidate countries should be encouraged to provide additional resources for generation of new analytical data on food to improve the quality of food composition databases and avoid the necessity to calculate values or borrow values from another country.

REFERENCES

Finglas P., Weichselbaum E., Buttriss JL. (2009). The 3rd International EuroFIR Congress 2009: European food composition data for better diet, nutrition and food quality. European Journal of Clinical Nutrition 64(3) s1-s3.

Becker W. (2010). Towards a CEN standard on food data. European Journal of Clinical Nutrition 64(3) s49-s52.

Bognár, A. (2002). Tables on weight yield of food and retention factors of food constituents for the calculation of nutrient composition of cooked foods (dishes). BFE, Karlsruhe, Germany, p. 7-11; p. 41-43; p. 95-97

APPENDICES

APPENDIX A: GUIDELINES FOR DATA COMPILATION

EFSA PROJECT CFT/EFSA/DCM/2011/03

Guidelines for data compilation, including rules for nutrient data harmonisation and rules for compilation of composite dish and supplement data

1. Nutrient Data

1.1 Mapping the EFSA food list to national food composition data (each partner)

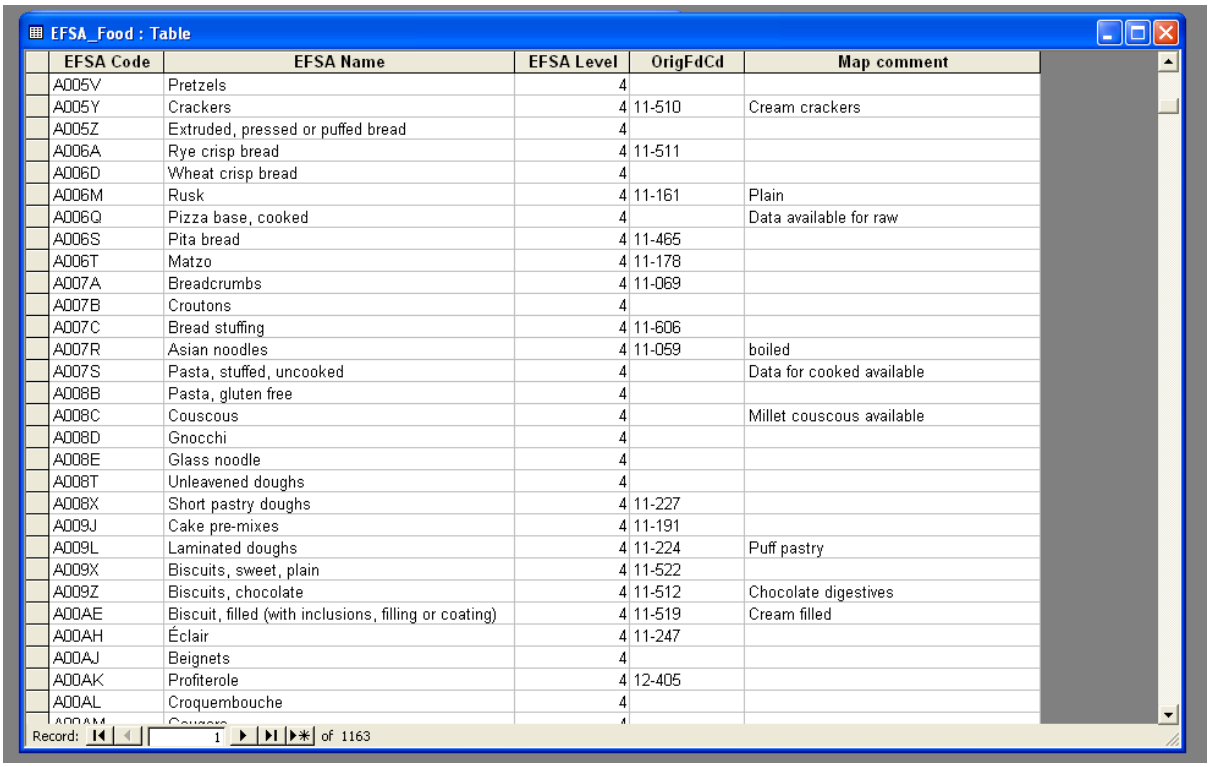
EFSA has produced a draft list of foods (based on FoodEx2) that are required for the nutrient dataset (see 'EFSANutrient databaseFoodListDraft1.xls'). The food list (worksheet 'EFSA_DCMFoodListDraftINCLUDED' consists of foods from levels 3, 4, 5 and 6 in combination with facet descriptors that describe facets including, fat content, fat type, sugar content, salt content, fortification, cooking process, preservation process and key ingredients.

There are 1727 'basic' foods listed and approximately 2000 combinations of basic food + facet descriptor, e.g. oat porridge (A00EQ) made up with water, oat porridge made up with full fat milk and oat porridge made up with skimmed milk. Each partner should map all combinations of basic food and facet descriptor to codes in their eSearch nutrient dataset.

The mapping exercise should focus on the food composition data related to foods that are consumed nationally in order to provide data that is applicable to consumption in each country. Some foods in the EFSA food list are mainly consumed in a limited number of countries and where a compiler does not have relevant data the 'OrigFdCd' field should be marked as N/A (not available).

Each partner will be provided with an Access database table (or Excel spread sheet if preferred) to map national dataset codes to the EFSA food list.

An example is given below



EFSA_Food : Table

EFSA Code	EFSA Name	EFSA Level	OrigFdCd	Map comment
A005V	Pretzels	4		
A005Y	Crackers	4	11-510	Cream crackers
A005Z	Extruded, pressed or puffed bread	4		
A006A	Rye crisp bread	4	11-511	
A006D	Wheat crisp bread	4		
A006M	Rusk	4	11-161	Plain
A006Q	Pizza base, cooked	4		Data available for raw
A006S	Pita bread	4	11-465	
A006T	Matzo	4	11-178	
A007A	Breadcrumbs	4	11-069	
A007B	Croutons	4		
A007C	Bread stuffing	4	11-606	
A007R	Asian noodles	4	11-059	boiled
A007S	Pasta, stuffed, uncooked	4		Data for cooked available
A008B	Pasta, gluten free	4		
A008C	Couscous	4		Millet couscous available
A008D	Gnocchi	4		
A008E	Glass noodle	4		
A008T	Unleavened doughs	4		
A008X	Short pastry doughs	4	11-227	
A009J	Cake pre-mixes	4	11-191	
A009L	Laminated doughs	4	11-224	Puff pastry
A009X	Biscuits, sweet, plain	4	11-522	
A009Z	Biscuits, chocolate	4	11-512	Chocolate digestives
A00AE	Biscuit, filled (with inclusions, filling or coating)	4	11-519	Cream filled
A00AH	Éclair	4	11-247	
A00AJ	Beignets	4		
A00AK	Profiterole	4	12-405	
A00AL	Croquembouche	4		

Record: 1 of 1163

The EFSA code and EFSA food name, with facet descriptors added where necessary, will be provided and the Original Food Code should be mapped as closely as possible to national data in the eSearch national data set. Where mapping is not clear comments should be added to explain any mapping limitations (see examples in 1.2).

Some partners have already mapped codes to the FoodEx1 food list and it is expected that EFSA will provide a mapping translation from FoodEx1 to FoodEx2 so that partners who have mapped to FoodEx1 can make use of existing data.

In many cases it is likely that there will not be a direct match between the level of detail provided by EFSA and that in national datasets meaning that national compilers will need to map based on their own knowledge and interpretation of foods available in their dataset. In some cases a single EFSA code could be mapped to more than one national code. All decisions related to code mapping should be documented in the comments field.

1.2 Possible food description and coding issues

- EFSA code may refer to an aggregate of national codes

  Example - A004Y Wheat bread and rolls, white (refined) flour
  Datasets may contain separate codes for bread and rolls and there may be a number of different codes for white bread and rolls. In this case the mapped code should relate to the most commonly consumed type or to a code where values are close to an 'average' for the type of food. Compilers should comment on which code is used and on limitations related to its use and include EFSA facet descriptors with comments where possible.

- EFSA code is not well defined

Example - A005K Bread and rolls with special ingredients added
This could be bread with dried fruit, sundried tomatoes, garlic, herbs, cheese or seeds with nutrient composition being quite different in each case. Where possible, compilers should base their code on consumption data so that the data represents the type or the combination of types that is most likely to be consumed.

- EFSA code is for raw food rather than food as consumed

  Example - A007S Pasta, stuffed, uncooked
  Data is available for cooked but not uncooked. This will be common for many datasets because data is usually available for foods as consumed. In this case compilers should comment that data is available for cooked but not raw. Values for the raw food should then be borrowed from another country.

- EFSA code is defined but foods mapped in different national datasets may be quite different

  Example - A00AR Cheesecake
  In the UK this is a frozen dessert, often with fruit topping. In other countries it would probably refer to traditional baked cheesecake. Nutrient compositions will be quite different but the codes used should match what would commonly be consumed in each country

  Example - A00BC Muffins
  In the UK these could be English muffins or American muffins. Composition could be very different but consumption of American muffins is probably much higher so that code should be used.

In some cases these problems would be best resolved by recalculating and adding additional foods to national datasets but this may not be possible for many compilers (see 1.3 Updating national e-Search datasets).

1.3 Updating national e-Search datasets

Some compilers may have published new data since providing their data in eSearch and in that case, to provide the most up to date nutrient data, national eSearch datasets should be updated.

eSearch datasets are either hosted by EuroFIR or by compiler organisations using web services and partners who wish to update datasets can either update by FDTP or could use an Access/Excel template generated from the current eSearch data format (single format for all partners) and provide that data format to EuroFIR for import into the EFSA dataset.

Where compilers produce new data from the EFSA mapping process, the eSearch dataset should be updated to include that new data.

1.4 Adding data to mapped EFSA food list

Tables of mapped codes, provided by each partner, will be queried against the eSearch dataset to return values for all mapped codes. Datasets, in an eSearch format database, will then be returned to each partner to check and fill gaps for missing nutrients where possible.

1.5 Adding missing data

Data for missing foods and nutrients should be completed using data from the eSearch platform, using the most appropriate data available. Data used should be applicable to the food as consumed in each partner country and the source of data used should be recorded and commented on where necessary. Once values for mapped codes have been returned from the eSearch dataset, missing data will be compiled centrally with each compiler being asked to complete data for a proportion of the foods and to comment on selection of data. Each compiler will then be asked to check a proportion of data for missing foods added by another partner.

1.6 Compiling national datasets into a single EFSA dataset

When each partner has completed their dataset, the data should be returned to IFR for merging into a single EFSA dataset. The dataset will be managed in FoodCASE and a single format will be delivered to EFSA in an agreed format.

2. Composite Dish Data

For this project a composite dish will be defined according to the EFSA definition agreed following the 14th Scientific Colloquium, i.e. 'a composite food is a food item, which consists of more than one ingredient. It can be either homemade or manufactured and either raw or prepared'. The scope note for composite foods in FoodEx2 further defines composite dishes as 'foods containing more than one distinct ingredient and having a recipe behind them, where the ingredients are so diverse that no relevant main category can be identified'.

The aim of this dataset is to provide a quantified list of ingredients for the most commonly consumed composite dishes so that intake of individual ingredients, e.g. vegetables, meat, fish, can be disaggregated from composite dishes. The focus should be more on composite dishes that are likely to be home prepared than on composite dishes that would be ready prepared or consumed outside the home. Foods such as bread, cheese, biscuits are composite foods as defined by the 14th Scientific Colloquium but fall outside the FoodEx2 scope note description so for the purposes of this study should not be included because disaggregation would not provide an additional level of useful information, i.e. fat content, sugar content would already be available as nutrient composition data and other ingredients would not be the most important for disaggregation.

A database structure for composite dish information will be produced based on the EuroFIR technical annex (Recipe and Ingredient entities) and will be supplied to each partner for data compilation.

2.1 Composite dish consumption analysis

Where possible the choice of 200 composite dishes should be supported by consumption survey data to ensure that the dishes reflect consumption frequency. Each compiler should produce a short description of how the selected composite dishes were identified.

2.2 Composite dish information

Information to be compiled for each composite dish includes:

Composite dish name
Original food code (where the composite dish is included with nutrient data in the national dataset)
Recipe procedure
Yield factor
Reference for retention factors
Portion size (as consumed)
Comments
Ingredients
Quantity of each ingredient
EFSA food code (where there is a good match to the recipe)

Each ingredient should be mapped, by each compiler, to an EFSA food code wherever possible.
Where an ingredient is itself a composite dish, e.g. tomato sauce, pastry, the ingredients of that composite should be recorded.

2.3 Collect composite dish information

Information for each dish should be collected from reliable sources and either a reference should be provided or the process of formulating the ingredient composition should be documented. Ingredients included in composite dishes should be generic and should reflect the dish as it would commonly be consumed.

2.4 Compilation of composite dish data

Each partner will return completed composite dish datasets to IFR and the datasets will be merged into a single dataset for delivery to EFSA.

3. Supplement Data

3.1 Analysis of available supplement data

Where information on food supplement data is available, e.g. from consumption surveys, each partner should compile a list of food supplements consumed in their country.

Information is known to be available in

- France
- Netherlands
- Greece
- Slovenia
- Italy
- UK
- Finland

Partners who do not hold information on supplement consumption should obtain data from other sources e.g. manufacturers and retailers.

3.2 Compilation of supplement data

A database or spreadsheet structure to input information will be provided to all partners and will require information on;

- Product name
- Product type
- Distributor
- Producer
- Brand name
- Package sizes
- Unit weights
- Nutrient composition (amount and weight)
- Data source

Information will be collated by IFR/EuroFIR and will be compiled into a single dataset. Products will be grouped into types for data handling purposes based on the USDA structured vocabulary for indexing food supplements (Saldanha et al., J. Fd. Comp. Analy. 25(2012), 226-233) e.g.

- Vitamin
- Mineral
- Herbal/botanical, e.g. yeast, algae, fungus
- Amino acid/protein
- Other dietary substance, e.g. fibre, enzymes
- Metabolite, constituent, extract, isolate or combination of these, e.g. hormone precursor, steroid precursor, 7-dehydrocholesterol, lutein, omega-3s
- Combinations of types listed above e.g.
    - Multivitamin/mineral (MVM)
    - Mineral & botanical
    - MVM & amino acids
    - MVM & botanical
    - MVM & fatty acids
    - Vitamin(s) & botanicals
    - Vitamins(s) & fatty acids

The LanguaL food supplement structured vocabulary will also be used as the basis for systematically capturing additional information related to supplements.

APPENDIX B: FINAL NUTRIENT DATABASE FOOD LIST FOR COMPILERS

MICROSOFT EXCEL SPREADSHEET - EFSANutrient database_CompilerFoodList010512.xlsx

Enclosed document

APPENDIX C: GUIDELINES FOR GAP FILLING CODES AT THE FOOD CODE LEVEL

EFSA PROJECT CFT/EFSA/DCM/2011/03

Guidelines for gap filling codes at the food code level

Problem: No code for food that is rarely consumed

Solution
1) Mark as 'N/R'
2) Use code from another country

Problem: No code for uncooked food

Solution
1) Use code from another country after checking that data would be comparable

Problem: No code for 'cooked in water' and/or 'steamed' foods

Solutions
1) Use available 'cooked in water' or 'steamed' code for missing EFSA codes. It most cases, assuming water contents are not significantly different, there will only be a small impact on micronutrient content.
2) Use a suitable code from another country after checking that data would be comparable.
3) Consider 'N/R' if food is very unlikely to be cooked in water or steamed e.g. dough

Note: Do not use a code for 'cooked with fat' for any 'cooked without fat' EFSA code

Problem: No code for 'baked' foods

Solutions
1) Use a suitable code from another country after checking that data would be comparable.
2) If no difference in water content, a code for 'cooked in water' or 'steamed' could be considered.
3) Consider 'N/R' if food is very unlikely to be baked, e.g. rice

Problem: No code for food 'cooked with fat'

Solution
1) Use code from another country after checking that data would be comparable

Note: Do not use a code for 'cooked without fat' for any 'cooked with fat' EFSA code

Problem: No code for one or more fortification facets

Solution
1) Use code from another country after checking that data would be comparable
2) Use code from unfortified food and gap fill fortified nutrients from another countries dataset

Problem: No code for food with fructose as a characteristic ingredient (products made with fructose corn syrup)

Solution

1) Use default code and gap fill sugars from another countries dataset or similar food within own dataset

Problem: No code for food made up with characteristic ingredient (water/whole milk/skimmed milk/half fat milk) or packed with characteristic ingredient (juice/syrup)

Solution

1) Use code from another country after checking that data would be comparable

Problem: No code for higher level food when code for lower level is available
Example: Onion and similar (A00HB, level 4) and Onion bulb (A00HC, level 5)

Solution

1) Use lower level code

Problem: No code for facet related to fat/salt/sugar content

Solution

1) Use code from another country after checking that data would be comparable
2) Use code from 'standard' food and gap fill nutrients from another countries dataset

Problem: No code for pickled food

Solution

1) Use code from another country after checking that data would be comparable

APPENDIX D: TEMPLATE FOR COMPOSITE DISH INFORMATION

MICROSOFT EXCEL SPREADSHEET - Recipes_Template010612.xlsx

Enclosed document

APPENDIX E: TEMPLATE FOR SUPPLEMENT INFORMATION

MICROSOFT EXCEL SPREADSHEET - Supplement Template.xlsx

Enclosed document

APPENDIX F: EFSA DATA TRANSFER SCHEMA

MICROSOFT EXCEL DOCUMENT - EFSA_Schema.xls

Enclosed document

ABBREVIATIONS

EuroFIR AISBL	Association Internationale Sans But Lucratif
ANSES	Agence Nationale de Sécurité Sanitaire de l'Alimentation, de l'Environnement et du Travail (French Agency for Food, Environment and Occupational Health Safety)
EFSA	European Food Safety Authority
EuroFIR	European Food Information Resource
FCDB	Food Composition Database
INRAN	The National Research Institute on Food and Nutrition
INSA	Instituto Nacional de Saúde Doutor Ricardo Jorge
NFA	The National Food Administration
RIVM	National Institute for Public Health and the Environment
THL	National Institute for Health and Welfare
TUBITAK	The Scientific and Technological Research Council of Turkey
XML	Extensible Markup Language