



Published in final edited form as:

J Food Compost Anal. 2017 October ; 63: 79–88. doi:10.1016/j.jfca.2017.07.037.

Development of a harmonized food grouping system for between-country comparisons in the TEDDY Study

Gesa Joslowski^a, Jimin Yang^b, Carin Andrén Aronsson^c, Suvi Ahonen^d, Martha Butterworth^e, Jenna Rautanen^f, Jill M. Norris^g, Suvi M. Virtanen^h, Ulla Uusitaloⁱ, and the TEDDY Study Group

^aInstitute of Diabetes Research, Helmholtz Zentrum München and Forschergruppe Diabetes, Klinikum rechts der Isar, Technische Universität München and Forschergruppe Diabetes e.V., Koelner Platz 1, Munich, 80804, Germany ^bHealth Informatics Institute, Department of Pediatrics, Morsani College of Medicine, University of South Florida, 3650 Spectrum Blvd, Tampa, FL, 33612, USA ^cDepartment of Clinical Sciences, Lund University, CRC 60:11, Jan Waldenströms gata 35, Malmö, SE-20502, Sweden ^dNational Institute for Health and Welfare, Department of Public Health Solutions, Nutrition Unit. University of Tampere, Faculty of Social Sciences. Science Centre, Tampere University Hospital. Mannerheimintie 166, Helsinki, 00300, Finland ^eHealth Informatics Institute, Department of Pediatrics, Morsani College of Medicine, University of South Florida, 3650 Spectrum Blvd, Tampa, FL, 33612, USA ^fNational Institute for Health and Welfare, Department of Public Health Solutions, Nutrition Unit. Mannerheimintie 166, Helsinki, 00300, Finland ^gDepartment of Epidemiology, University of Colorado Denver, Colorado School of Public Health, 13001 East 17th Place, Aurora, CO, 80045, USA ^hNational Institute for Health and Welfare, Department of Public Health Solutions, Nutrition Unit. University of Tampere, Faculty of Social Sciences. Science Centre, Tampere University Hospital. Center for Child Health Research, University of Tampere and Tampere University Hospital. Mannerheimintie 166, Helsinki, 00300, Finland ⁱHealth Informatics Institute, Department of Pediatrics, Morsani College of Medicine, University of South Florida, 3650 Spectrum Blvd, Tampa, FL, 33612, USA

Abstract

The Environmental Determinants of Diabetes in the Young (TEDDY) is an international study aiming to investigate associations between dietary and other environmental factors and the risk of developing islet autoimmunity and type 1 diabetes. Dietary intake was assessed using a 24-hour recall and repeated 3-day food records and analyzed using country-specific food composition databases (FCDBs) in Finland, Germany, Sweden, and the U.S. with respective in-house calculation programs. A food grouping harmonization process between four country-specific FCDBs was conducted to evaluate and achieve comparability on food group definitions and

Corresponding author: Ulla Uusitalo, Ulla.Uusitalo@epi.usf.edu, Health Informatics Institute, Department of Pediatrics, University of South Florida Morsani College of Medicine, 3650 Spectrum Blvd, Tampa, FL 33612, USA, Phone: 813-396-9516, Fax: 813-910-5916.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

quantification of food consumption across the countries. Systematic review revealed that the majority of existing food groups of the TEDDY FCDBs were not comparable. Therefore, a completely new classification system of 15 mutually exclusive main food groups (e.g. vegetables) and 89 subgroups (e.g. root vegetables, leafy vegetables) was developed. Foods and beverages were categorized into basic foods (single ingredient) and composite dishes (multiple ingredients). Composite dishes were broken down to ingredients using food composition data available in the FCDBs or generic recipes created for the harmonization effort. The daily consumption of every food group across FCDBs was quantified consistently as either raw or prepared weight depending on the food group to achieve maximal comparability.

Keywords

diet; food analysis; food composition; food consumption; food exposure; food grouping; food composition database; food grouping harmonization; TEDDY Study; Type 1 diabetes

1. Introduction

The Environmental Determinants of Diabetes in the Young (TEDDY) study is a prospective, multi-center, international study in which 8676 children with increased genetic susceptibility to type 1 diabetes and celiac disease are followed across six study centers worldwide (one each in Finland, Germany, and Sweden; and three in the United States (U.S.)). The participants are monitored for islet autoantibodies and tissue transglutaminase autoantibodies until 15 years of age or until they develop type 1 diabetes. The study aims to examine associations between islet autoimmunity and various environmental exposures such as dietary intake (TEDDY Study Group, 2008).

Infant and childhood diet plays an important role in the etiology of type 1 diabetes and islet autoimmunity (Virtanen, 2016) and celiac disease (Andrén Aronsson et al., 2016a; Chmielewska et al., 2015). In order to examine nutrient intakes in the TEDDY study, considerable effort was made to estimate the comparability of the nutrients between the country specific food composition databases (FCDBs) linked to the study (Uusitalo et al., 2011). Besides the role of nutrients there is a growing emphasis on the importance of food level exposures to better understand diet and disease relationships (Norris, 2010; Virtanen, 2016). One of the aims of the TEDDY study is to examine the quantitative exposure to foods such as cow's milk, root vegetables and gluten-containing cereals which have been associated with increased risk of islet autoimmunity, type 1 diabetes and/or celiac disease in infancy and later in childhood (Knip et al., 2010; Virtanen et al., 2000; Virtanen et al., 2012; Lamb et al., 2015a; Lamb et al., 2015b; Virtanen, 2016; Szajewska et al., 2016). It is critical to quantify the consumption of foods consistently across the participating countries when studying the putative associations between dietary intake and outcome.

Two approaches to group or categorize foods are described in the literature: the behavioral and the epidemiological approach (Cullen et al., 1999). The approaches differ in the contribution of specific foods to the food groups as well as in the application for the approach. The behavioral approach is more practical and (mostly) based on household servings. For example, the coding does not take small amounts of vegetables into account

when coding a pizza because vegetables within a pizza do not contribute enough to be included in the vegetable food group (Mitchel, 2001). Nonetheless, the pizza itself is counted a suitable food group as a main dish. The epidemiological approach, on the other hand, is more “detailed” and includes all ingredients of food items that contribute to dietary intake, no matter how small the contribution is. Weights are usually counted as gram amounts. Composite dishes such as pizza are broken down into their ingredients and then grouped under the respective food groups (Cullen et al., 1999; Mitchel, 2001).

Difficulties associated with categorizing foods start with the definition of food groups and/or subgroups, the use of raw or prepared weight, as well as flexibility and limitations of databases to assign foods into respective food groups (Mitchel, 2001; Ireland et al., 2002). Decision processes are complex and international food grouping is even more complicated because of diverse food nomenclature, terminology or even cultural differences which may all lead to different food grouping (Truswell et al., 1991). The use of standardized systems enables linking and describing food composition data across country specific databases (Ireland & Møller, 2013a). To overcome food categorization difficulties, the Langua aLimentaria (LanguaL)TM method can be used to provide a standardized language for describing foods in a systematic way (Ireland & Møller, 2013b). There also exist inconsistencies in the food grouping algorithms between countries in national FCDBs (Ireland et al., 2002). For successful comparisons, single food items and ingredients of composite dishes need to be allocated into suitable and comparable categories. Furthermore, the form of the foods should be considered. It needs to be defined whether the weight of a food group is given as raw or prepared weight; dry, fresh or liquid weight. If one FCDB provides only raw and another only prepared weights the results are not comparable. Thus, in order to produce comparable results, the FCDBs must contain mutually comparable food grouping data, including an agreement on food groups and reported weights. Preferably raw weights should be reported, using national conversion factors reflecting national differences in food preparation (Ireland et al., 2002). One example of an international standardized system for classifying foods is that of the European Food Safety Authority (EFSA) (European Food Safety Authority (EFSA), 2015).

In the TEDDY study, food grouping harmonization was conducted to evaluate and achieve comparability on food group definitions and the quantification of food intakes across FCDBs used in Finland (FINELI), Germany (LEBTAB), Sweden (TEDDY-SWE), and the U.S. (Nutrition Data System for Research). The aim of this paper is to describe the harmonization efforts and the TEDDY food grouping system, which was developed primarily based on existing hypotheses in the etiology of type 1 diabetes and celiac disease.

2. Material and methods

In the TEDDY Study, dietary intake was assessed by interviewing parents using one 24-hour recall on the first visit at child's age of 3 months. Thereafter, all foods and beverages consumed are recorded using 3-day food records, carried out at the ages of 6, 9, and 12 months and from there biannually until the child is 15 years old. The food consumption data are entered and analyzed using the country-specific FCDB, i.e. the national food composition database FINELI in Finland, LEbensmittelTAbelle (LEBTAB) in Germany,

NFA FCDB (The TEDDY Malmö version of the NFA Database; TEDDY-SWE), and Nutrition Data System for Research (NDSR) in the U.S. and their respective in-house dietary intake data processing software's. Details on the FCDBs and nutrient harmonization have been previously published (Uusitalo et al., 2011). The TEDDY Data Coordinating Center in Tampa (FL, U.S.) gathers, stores, and processes the nutrient analysis data and food grouping output files from the national FCDBs. A brief description of each FCDB is given below.

2.1. Food composition databases - their original structure and food grouping systems

The Finnish FCDB **FINELI** is maintained by the National Institute for Health and Welfare, Finland consisting of more than 7,000 food items (ingredients and dishes) and more than 900 supplements. Composite dishes can be automatically broken down into ingredients except for the following core foods: crisp breads, breakfast cereal and cookie recipes created before 2010 (cookie recipes have been replaced by noncore ones gradually along the years), pasta (also with mixed flours), cereal mixtures, and nachos. Recipes can be used within recipes and there is no restriction of recipe levels within FINELI. For both recipes and basic foods, yield factors are applied at ingredient level. The weights given by FINELI are expressed as eaten if counted at the food use level or given as raw if counted at ingredient level. The food items are structured by food use class (FUCLASS) and ingredient class (IGCLASS), relating to both main food groups and subgroups captured on both food use level and the ingredient level (National Institute for Health and Welfare, 2016). While the food use class (FUCLASS) is based on the international classification "COST Action 99" (Working Group on Food Data Management and Interchange, 2000), which has been modified to better serve the Finnish needs, the ingredient class (IGCLASS) was developed in-house. Overall, there are 17 main food groups for the FUCLASS with 104 subgroups and 18 main food groups for the IGCLASS with 89 subgroups.

The **LEBTAB** FCDB including its food grouping system was developed in-house for the longitudinal DONALD (Dortmund Nutritional and Anthropometric Longitudinally Designed) study (Kroke et al., 2004). Currently, the TEDDY version of LEBTAB contains >20,300 food items including basic foods, composite dishes, additives, supplements and medicine. Of these approx. 2,200 are basic food items and 17,300 are and composite dishes. Composite dishes can be broken down to their ingredients according to the underlying recipes. An ingredient within a recipe may be a composite dish itself, which can be broken down to its ingredients in a second step. In LEBTAB, the use of recipes within a recipe is restricted to one level, i.e. the second ingredient breakdown contains only basic food items. Within the recipe simulation the yield factors are applied on recipe level, e.g. a commercial baby food containing carrots, potato, and beef is assigned with one yield factor. For single foods, the yield factor is applied on ingredient level, e.g. lunch containing: cooked carrots, baked potato, and fried beef. The weights given by LEBTAB are raw except for canned foods, cooked ham, some pasta, rice, and commercial products, which are given as prepared. LEBTAB contains 23 main food groups and no subgroups, as previously described (Sichert-Hellert et al., 2007).

The National Food Administration (Livsmedelverket) maintains the National Food Database (NFA FCDB) in Sweden. The NFA FCDB contains more than 2,000 food items (ingredients and dishes) (National Food Administration, 2014) and is the foundation of the TEDDY-SWE food database. The NFA FCDB did not contain any nutrient information on commercial baby foods (infant formulas, infant porridge, gruel), but since the start of TEDDY, all nutrient data on commercial baby foods have been created/compiled from information based on ingredient information and manufacturer data. Moreover, recipes on dishes cooked by the TEDDY families were added to the TEDDY-SWE database. To date, the TEDDY-SWE database contains over 12 000 food codes (food items, core foods, recipes) in addition to the over 2000 food items provided by NFA FCDB. Composite dishes can be broken down into their ingredients except for the following core foods: Crisp breads, cookies/crackers, breakfast cereals, commercial sweet bakery products, mixed vegetables (canned or frozen), and apple sauce. Within TEDDY-SWE there is no restriction of recipe levels, however, only one level is used in practice, i.e. one recipe is used as an ingredient within another recipe. Within the recipe calculation, food items are entered raw and yield factors applied at ingredient level.

The use of recipes within a recipe is allowed and all recipes break down to single ingredients. For single foods, TEDDY-SWE database provides raw and prepared basic food items and core foods, respectively. In order to convert prepared basic food items to their raw weight recipe simulation is used, i.e. application of yield factors at ingredient level. TEDDY-SWE database contains 41 main food groups and 240 subgroups.

The NDSR FCDB is maintained by the University of Minnesota Nutrition Coordinating Center (NCC) in the U.S., including over 18,000 food items (ingredients and dishes) (Nutrition Coordinating Center (NCC) University of Minnesota, 2015). With the availability of many ingredient choices as well as preparation methods NDSR provides more than 160,000 food variants. By default, NCC provides two levels of ingredient information. Composite dishes are broken down to their ingredients according to the underlying recipe. Recipes can be used within recipes and there is no restriction of recipe levels within NDSR. Nevertheless, many ingredients within a recipe are composite dishes which are not broken down automatically, e.g. breakfast cereals, breads, bakery sweets and savory, pasta, baby foods, fast foods such as chicken nuggets or other commercial dishes. To create composite dishes, NDSR distinguishes between recipes, formulations, and core foods. The yield factor is taken into account on recipe or ingredient level. In general, the main goal of the recipes and formulations used in NDSR is to match the nutrient profile of a food as closely as possible. The weights given by NDSR are all as consumed, i.e. either raw or prepared. The only raw meat or fish within NDSR would be carpaccio or sashimi/sushi. The NCC Food Group Serving Count System, which includes 9 major food categories and 168 sub-categories, was designed to quantify foods by the number of servings. The reference serving sizes used by the NCC are based on the recommendations made by the 2000 Dietary Guidelines for Americans and the Food and Drug Administration.

2.2. Systematic comparison

Systematic comparison of food groups was carried out between the four country-specific FCDBs within the TEDDY study. The process involved evaluation and comparison of existing food grouping systems and types of available variables (e.g. ingredients vs. recipes/ composite dishes, default ingredient breakdown, possibility of regrouping ingredients manually). Furthermore, the handling of foods weights, i.e. raw vs. prepared and the application of yield factors was discussed. FCDB experts outside the TEDDY study were also consulted.

3. Results

Systematic review revealed that the majority of existing food groups in the FCDBs were not comparable. Table 1 illustrates the challenges to find common ground in food classification with the food group “cereals and cereal products” as an example. For cereal FINELI contains on food use level one main group and under it 14 food use subgroups and on ingredient level one main group and under it 9 ingredient subgroups, while LEPTAB contains 3 main food groups to capture cereals and cereal products. TEDDY-SWE comprises 9 main food groups with 53 subgroups and NDSR includes a total of 35 subgroups within the NCC Food Group Serving Count System. This example shows that the groups did not compare very well across the countries. Therefore, an alternative and completely new classification system was developed in-house for all food items and food groups that could be made comparable, to a large extent based on TEDDY hypotheses, i.e. for cereals to distinguish seven food groups on an ingredient level, i.e. rice; wheat; rye; oats; barley; corn; other gluten free flours and starches, to improve the comparability of cereal consumption data between the TEDDY countries. All TEDDY food groups are described in Table 2. They were classified in 15 mutually exclusive main food groups and 89 subgroups. They were designed to examine the quantitative exposure to specific food groups such as wheat, cow's milk, berries, or root vegetables across countries. The harmonization process is described in the following:

3.1 Classification of foods in the TEDDY food grouping system

All foods and beverages consumed by TEDDY participants are categorized into basic foods (single ingredient) or composite dishes (multiple ingredients). Within the TEDDY food grouping each food item undergoes an assignment process – automatically or manually – in which it is determined whether the food item is a basic food or a composite dish and whether composite dishes can or cannot be broken down into its ingredients. Figure 1 depicts the disaggregation procedure of basic food items and composite dishes which serves the food grouping process. While basic foods such as apples could be directly assigned to the corresponding food group, composite dishes, e.g. apple pie, were broken down to ingredients using recipes available in the country-specific FCDBs. There were differences in the breakdown procedures between the TEDDY countries (see also 3.3). While some software linked to the FCDBs broke down most of the dishes in the database others had only few automated complete break-down processes. Generic recipes were created for dishes that did not break down into their ingredients and programmed to be used in the food grouping procedure. Core dishes are in general terms often regarded as their own food groups for

which the ingredient information is not on demand, however, in TEDDY this ingredient information is of interest. Examples of these include wheat pasta and crisp bread. Eventually ingredients were allocated to one of the 89 mutually exclusive subgroups with a few exceptions: e.g. spices, vinegars, artificial sweeteners, xylitol containing products, HP Sauce, water, as well as the food group containing sugar, syrups, honey, jams and jellies. The total amount of daily consumption of every food group was consistently expressed as either raw, prepared, dry or liquid gram weight. Powdered products were converted to liquids when applicable, to achieve the maximal comparability.

3.2 Automatized and manual assignment of food items and food groups

The existing food groups in every FCDB were either integrated verbatim to the TEDDY food groups if the definitions matched or they were manually re-grouped into the respective TEDDY food groups. All foods in **FINELI** were manually assigned to the respective TEDDY food groups. Existing **FINELI** food classifications were used to help this manual assignment. The main food groups in **LEBTAB** could not easily convert to the TEDDY food groups. Therefore, recorded food items were manually assigned to TEDDY food groups. Within the **TEDDY-SWE** FCDB most food groups were created to match the TEDDY food grouping system when the FCDB was updated in 2007, i.e. existing main food groups and subgroups were aggregated. Manual assignment of food items into the correct TEDDY food group was necessary for those food items whose original food group did not match the TEDDY food group. This was the case for approximately 200-250 food items. Within **NDSR** only a few groups of the NCC Food Group Serving Count System matched the TEDDY food groups. This applied to 75 of 168 groups such as citrus juice, vegetable juice, yogurt, and cold cuts and sausage. Other groups of the NCC Food Group Serving Count System were automatically assigned to break down into their ingredients. This was the case for 45 groups such as grains, flours and dry mixes of pancakes, breads, muffins, etc.; loaf-type bread and plain rolls; other Breads (quick breads, corn muffins, tortillas). Food items of the remaining NCC Food Group Serving Count System groups were manually assigned to corresponding TEDDY food groups. **NDSR** annually updates its database version, which results in the manual assignment of approximately 1000 new food items to the appropriate TEDDY food groups per year. Similarly, new food items are added to **FINELI**, **LEBTAB**, and **TEDDY-SWE** on a regular basis, which are then assigned to a corresponding TEDDY food group.

3.3 Ingredient breakdown and core dishes

Among all FCDBs **FINELI**, **LEBTAB**, **TEDDY-SWE**, and **NDSR**, composite dishes can be broken down into their respective ingredients. However, **FINELI** and the **TEDDY-SWE** FCDB contain core dishes, which are not automatically broken down. Country-specific generic recipes were used to break down these core dishes into their ingredients. The ingredient breakdown of core dishes was done manually within the countries FCDB or at the TEDDY Data Coordinating Center, in order to provide an estimation of the ingredient proportions and to estimate the ingredient amount consumed. The development of country-specific generic recipes was done according to local food manufacturing guidelines. If the ingredient breakdown was not performed by software linked to the FCDB, manually created recipes were gathered in a separate file and applied to the core dishes using the SAS

environment after food calculation process and to assign them to their respective TEDDY food group (e.g. mixed vegetables).

NDSR also contains core dishes, which are not automatically broken down. Therefore, the NCC created a special ingredient breakdown file for the TEDDY study, which includes the ingredient information of the remaining core dishes. Unlike the structure in LEBTAB, NDSR does not restrict the extent of recipes being used as ingredients within another recipe, so multiple layers of breakdown are often needed to reach the ingredient level. This ingredient breakdown for the TEDDY study is conducted on a yearly basis together with the release of the new database version. After this ingredient breakdown was provided by the NCC only pasta and applesauce remained as core dishes. Generic recipes were then created by the NCC to obtain an estimate of the ingredients of the core dish.

3.4 Conversions

In order to express the amount of every food group consistently as either raw or prepared gram weight, conversions needed to be applied. Most of those conversions could be made with the countries' FCDBs or the corresponding data processing software using national conversion factors to convert e.g. cooked vegetables into raw ingredients as these reflect national differences in food preparation. For some food items such as cooked or fried potatoes, French fries, potato chips, canned vegetables and sprouts the conversions are done at the TEDDY Data Coordinating Center. Moreover, conversion factors needed to be employed to allow the aggregations of processed food items into TEDDY food grouping subgroups. For instance, to estimate the overall quantity of fat-free, low-fat, high-fat milk, milk powder had to be multiplied by country specific conversion factors to be transformed into liquid equivalents of fresh milk which could be added to the corresponding subgroups. Finland (Valio, 2016) and the U.S. (Nutrition Coordinating Center (NCC) University of Minnesota, 2015) use the conversion factors 11 and 10.4, respectively. For Germany and Sweden, the Finnish conversion factor was borrowed to convert milk powder into liquid equivalents of fresh milk. Other examples are cream powder, whey concentrate, infant formula, dried potatoes, and potato flakes.

To quantify the amount of breast milk consumed, algorithms developed by the Institute of Medicine were used (Institute of Medicine, 2005). The estimated energy requirement (EER) based on a child's age and weight is determined for every visit as well as the energy from complementary feeding reported on the food record (FR Kcal, harmonized). Then the energy of breast milk (BM Kcal) can be calculated ($BM\ Kcal = EER - FR\ Kcal$). If the EER exceeds the FR Kcal, the caloric difference is assumed to be from breast milk. The amount of breast milk consumed by a child is calculated as follows: $breast\ milk\ ingested\ (g) = BM\ Kcal / energy\ density\ per\ 100\ g$, where the energy density is 65 kcal in Finland, 69 kcal in Germany, 75 kcal in Sweden, and 70 kcal in the U.S. per 100g.

3.5 Additional variable

In Germany, an additional variable was introduced to capture whether a food item contains added pro-, pre- or symbiotics on the level of food grouping. Finland, Sweden, and the U.S.

flag probiotic foods on the data entry level, i.e. per food item and provide this information in the nutrient intake file separate from the food grouping file.

3.6 Food grouping database

All food grouping data are transferred to the TEDDY Data Coordinating Center where country-specific files are stored in a centralized food grouping database. Within the food grouping database generic recipes are applied to the core foods and other necessary conversions are executed. For analyses purposes, food level intake data can be summarized at various levels as needed, e.g. total intake per day per food group for every child at every age of data collection.

3.7 Food grouping exercises

Several food grouping exercises were carried out across all participating countries to see whether the countries were able to classify recorded food items in their FCDBs according to the TEDDY food grouping system. The result was that all the countries were able to use the new classification system. However, it also became clear that a lot of manual work would be needed to fully implement the TEDDY food grouping system and all countries expressed the need for further directions to assign food items to the different food groups in a comparable way and to agree on how the gram amounts within a food group should be reported (i.e. raw / prepared / liquid / dry weight). A detailed manual with clear definitions of every food group was developed to ensure that food items were grouped in the correct TEDDY food group using the correct gram weight. The food group assignment of every food item was made by dietitians or nutritionists who were from the same country and were familiar with the food items.

4. Discussion

To the best of our knowledge this is the first FCDB harmonization at food group level conducted across country-specific FCDBs in an international multicenter cohort study. This harmonization enables assessment of the quantitative exposure to specific food groups such as wheat, cow's milk, berries or root vegetables across the TEDDY countries.

Previous approaches to harmonize food grouping data have shown that food items need to be aggregated to the lowest and least detailed level of information in order to make data among countries comparable. This is because the level of detail in which food data were collected ranges from very detailed food groups to more aggregated ones (Lagiou & Trichopoulou, 2001; Ireland et al., 2002; Slimani et al., 2000). The European Prospective Investigation into Cancer and Nutrition (EPIC) study conducted a comparison to evaluate the nature and magnitude of differences between national FCDBs (Slimani et al., 2000). At food level, not only the number of food items reported but also the level of detail reported (e.g. cooking or preservation method) varied substantially between FCDBs (Slimani et al., 2000). Within the EPIC study, detailed food lists derived from existing standardized 24-hour recalls could be used for food grouping. A standardized database was built where EPIC food lists and nutrient lists were matched to those available in national databases or, alternatively, defined how to calculate or adjust them (Slimani et al., 2000; Slimani et al., 2007). Because numbers

of food items varied from 3500 in Greece to 15 000 in France, initial EPIC food occurrences were aggregated to obtain 547 – 1537 foods per country using a calibration approach requiring good estimates of mean population intakes (Slimani et al., 2007). The differences among TEDDY FCDBs presented similar challenges, however, within TEDDY there was no option to lose the level of detail and subsequently the opportunity to analyze exposures to specific food groups.

To date, a universal food classification system is lacking. Different approaches to classify food items result from different objectives or may even reflect different legislations. Sometimes classification systems may be contradictory and their very existence shows that there can be no single international classification system that serves all needs of every food composition database compiler (Ireland & Møller, 2000). The Euro Food group (EFG) classification system with 33 main food groups was created to attempt a comparison of food consumption data collected using different food classification systems across European countries (Food and Agriculture Organization of the United Nations (FAO) balance sheet, World Health Organization Global Environment Monitoring System (WHO GEMS) / FOOD, Data Food Networking (DAFNE), Eurocode 2, French survey, Dutch survey, and EPIC Soft). Verger et al. concluded that the EFG system is the best compromise between different classification systems (Verger et al., 2002), even though further work would be needed to establish an acceptable level of comparability where discrepancies between different reporting methods of food consumption data (“as consumed level”, “raw ingredient level” or both) are solved (Ireland et al., 2002; Verger et al., 2002). Eventually, foods can only be made comparable at the “raw ingredient” level (Ireland et al., 2002).

The EFG classification system would not fit the TEDDY purpose and/or the FCDBs used in TEDDY. The goal of the TEDDY food grouping is to quantify exposure to very specific food groups which are hypothesized in the etiology of type 1 diabetes and celiac disease and the EFG system would not have been specific enough. For example, exposures to gluten and cow's milk are important factors to be investigated in TEDDY, but the EFG cereal and cereal products groups (“bread and rolls”; “breakfast cereal”; “flour”; “pasta”; “bakery products”; “rice and other cereal products”) do not distinguish by the amount and existence of gluten as needed in TEDDY (rice, wheat, rye, oats, barley, corn or other gluten free flours and starches). The EFG milk and milk product groups (“milk”; “cheese”; “other milk products”) do not specifically capture cow's milk vs. other kinds of milks either. It was necessary to conduct a harmonization between the FCDBs used in TEDDY on the food level, to achieve a harmonized food grouping system serving the needs of TEDDY. TEDDY collects nutritional epidemiologic data to identify risk factors for type 1 diabetes and celiac disease. Therefore, it was logical to employ the epidemiological approach to avoid losing details even though this approach requires a huge amount of manual work (Mitchel, 2001).

In order to make food groups comparable, we needed to make a decision on how to report the weight. Since foods can only be made comparable at the “raw ingredient” level most TEDDY food groups are given as raw and/or dry weight. A food grouping harmonization on raw ingredient level was achieved for most food groups, such as cereals (except for “rice milk” and “oat milk”), fruits and berries (except for “dried fruits” and “fruit juices”), vegetables (except for “vegetable juices”), soy beans, nuts and seeds (except for the

respective “milks”), various kinds of milk, meat (except for “processed meats and sausages”), fish (except for “processed fish” and “canned fish”), and eggs.

A challenge within the food grouping was that not all composite food items break down fully to ingredients. To overcome this issue, generic recipes were used within the TEDDY food grouping to estimate ingredient proportions. It is common practice across FCDBs to estimate mean nutrient values of mixed dishes or average weighted food items using generic recipes (Slimani et al., 2000). One example is a commercial product such as a “gluten free chocolate cupcake, bought in a store” where no other information is given. In addition, this method can be used for core foods such as pasta which cannot be broken down into its ingredients. Recipe information can be taken from product packages, cookbooks or the internet. Within the TEDDY study, the process of recipe simulation is conducted by trained dietitians and nutritionists. The addition of new food items into a FCDB either with known ingredients and exact recipes or generic recipes offers flexibility. If a FCDB does not contain a specific food item e.g. wheat corn pasta, it can be added to the database, eventually breaking it down into its ingredients and thus providing the much needed ingredient information for the TEDDY food grouping. The use of generic recipes can be monitored throughout time to see how often generic recipes are used and to see whether data differences between countries may stem from these estimations.

The TEDDY food grouping harmonization was not accomplished for all food items across the countries. As described, examples of exceptions were spices, vinegars, water, as well as the food group containing sugar, syrups, honey, jams and jellies because we were not able to comparably quantify the weight of those food groups across the countries. Therefore, they were excluded in the exposure analysis. However, the amount of sugar from these sources are captured by the nutrient variables “total sugars” and/or “added sugars”. Overall, we were able to harmonize all 89 TEDDY food groups providing comparable weight ingredient information across participating TEDDY countries. All participating TEDDY countries submit the food grouping data in the same format, either as an Excel or CSV file, to the TEDDY Data Coordinating Center where country-specific files are retained and stored in a centralized food grouping database. Data files can then be further processed, e.g. by applying generic recipes to core foods or other necessary conversions. The final food grouping data will be combined with nutrient data and used for statistical analysis.

Within the TEDDY food grouping no harmonization of conversion factors took place. Most of the conversions are done within individual FCDBs or the corresponding data processing software using national conversion factors that reflect national differences in food preparation. Given not all countries have their own yield factors for every single food item, conversion factors used in one FCDB (e.g. dry milk powder to liquid milk) may be referenced in another FCDB.

The TEDDY food grouping comprises a lot of continuous, manual work across all FCDBs to re-group food items into TEDDY food groups or to create new food items / recipes within or outside each FCDBs, thus being very flexible and dynamic. When new foods are introduced into FCDBs which are not yet captured by the TEDDY food grouping system, the FCDBs can easily be adjusted to accommodate the new foods. For future attempts to harmonize food

grouping between country specific FCDBs it would be highly beneficial if all FCDBs had appropriate documentation with references available e.g. definitions of food groups and all terms used, and conversion factors for converting prepared weights to original mutually comparable raw weights whenever applicable. This would very much facilitate harmonization and probably make the whole process much faster.

5. Conclusions

The development of a harmonized food grouping system for between-country comparisons in the TEDDY study was completed to produce comparable quantification of food exposures. The harmonized TEDDY food grouping data can be used to conduct descriptive analyses, to analyze association between food exposure and islet autoimmunity, type 1 diabetes or to identify risk factors for celiac disease such as gluten exposure across TEDDY countries. So far, this has only been possible using Swedish data (Andrén Aronsson et al., 2016b). Moreover, exposure data at food level will provide the opportunity for dietary pattern analyses.

Acknowledgments

The authors gratefully acknowledge the contributions from Janet Petite, Mary Stevens, Sally Schakel, Irene Mattisson, Stefanie Schoen, Kristina Foterek, Nicole Janz, Sara Sibthorpe, Anne Riikonen, Mirva Koreasalo, Mari Åkerlund, Carina Kronberg-Kippilä, Maryouri Avendano, Wendy McLeod, and Lori Ballard.

Funding disclosure: Funding disclosure: This work was supported by U01 DK63829, U01 DK63861, U01 DK63821, U01 DK63865, U01 DK63863, U01 DK63836, U01 DK63790, UC4 DK63829, UC4 DK63861, UC4 DK63821, UC4 DK63865, UC4 DK63863, UC4 DK63836, UC4 DK95300, UC4 DK100238, UC4 DK106955, and Contract No. HHSN267200700014C from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Child Health and Human Development (NICHD), National Institute of Environmental Health Sciences (NIEHS), Juvenile Diabetes Research Foundation (JDRF), and Centers for Disease Control and Prevention (CDC). This work supported in part by the NIH/NCATS Clinical and Translational Science Awards to the University of Florida (UL1 TR000064) and the University of Colorado (UL1 TR001082).

Appendix: The TEDDY study group

Colorado Clinical Center: Marian Rewers, M.D., Ph.D., ^{pi1,4,5,6,10,11}, Kimberly Bautista¹², Judith Baxter^{9,10,12,15}, Ruth Bedoy², Daniel Felipe-Morales, Kimberly Driscoll, Ph.D.⁹, Brigitte I. Frohnert, M.D.^{2,14}, Patricia Gesualdo^{2,6,12,14,15}, Michelle Hoffman^{12,13,14}, Rachel Karban¹², Edwin Liu, M.D.¹³, Jill Norris, Ph.D.^{2,3,12}, Adela Samper-Imaz, Andrea Steck, M.D.^{3,14}, Kathleen Waugh^{6,7,12,15}, Hali Wright¹². University of Colorado, Anschutz Medical Campus, Barbara Davis Center for Childhood Diabetes.

Finland Clinical Center: Jorma Toppari, M.D., Ph.D., ^{PI[¥]1,4,11,14}, Olli g. Simell, M.D., Ph.D.^{¥1,4,11,13}, Annika Adamsson, Ph.D.^{^12}, Suvi Ahonen^{*±§}, Heikki Hyöty, M.D., Ph.D.^{*±6}, Jorma Ilonen, M.D., Ph.D.^{¥13}, Sanna Jokipuu[^], Tiina Kallio[^], Leena Karlsson[^], Miia Kähönen^{μ±}, Mikael Knip, M.D., Ph.D.^{*±5}, Lea Kovanen^{*±§}, Mirva Koreasalo^{*±§2}, Kalle Kurppa, M.D., Ph.D.^{*±13}, Tiina Latva-aho^{μ±}, Maria Lönnrot, M.D., Ph.D.^{*±6}, Elina Mäntymäki[^], Katja Multasuo^{μ±}, Juha Mykkänen, Ph.D.^{¥3}, Tiina Niininen^{±*12}, Sari Niinistö^{±§}, Mia Nyblom^{*±}, Petra Rajala[^], Jenna Rautanen^{±§}, Anne Riikonen^{*±§}, Mika Riikonen[^], Jenni Rouhiainen[^], Minna Romo[^], Tuula Simell, Ph.D., Ville Simell^{^¥13}, Maij a

Sjöberg^{¥12,14}, Aino Stenius^{µ12}, Maria Leppänen[^], Sini Vainionpää[^], Eeva Varjonen^{¥12}, Riitta Veijola, M.D., Ph.D.^{µ14}, Suvi M. Virtanen, M.D., Ph.D.^{*±§2}, Mari Vähä-Mäkilä[^], Mari Åkerlund^{*±§}, Katri Lindfors, Ph.D.^{*13} University of Turku, *University of Tampere, µUniversity of Oulu, ^Turku University Hospital, Hospital District of Southwest Finland, ±Tampere University Hospital, ºOulu University Hospital, §National Institute for Health and Welfare, Finland, ¶University of Kuopio.

Georgia/Florida Clinical Center: Jin-Xiong She, Ph.D., PI^{1,3,4,11}, Desmond Schatz, M.D.^{*4,5,7,8}, Diane Hopkins¹², Leigh Steed^{12,13,14,15}, Jamie Thomas^{*6,12}, Janey Adams^{*12}, Katherine Silvis², Michael Haller, M.D.^{*14}, Melissa Gardiner, Richard McIndoe, Ph.D., Ashok Sharma, Joshua Williams, Gabriela Young, Stephen W. Anderson, M.D.[^], Laura Jacobsen, M.D.^{*14} Center for Biotechnology and Genomic Medicine, Augusta University. *University of Florida, ^Pediatric Endocrine Associates, Atlanta.

Germany Clinical Center: Anette G. Ziegler, M.D., PI^{1,3,4,11}, Andreas Beyerlein, Ph.D.², Ezio Bonifacio Ph.D.^{*5}, Michael Hummel, M.D.¹³, Sandra Hummel, Ph.D.², Kristina Foterek^{¥2}, Nicole Janz, Mathilde Kersting, Ph.D.^{¥2}, Annette Knopff⁷, Sibylle Koletzko, M.D.^{¶13}, Claudia Peplow¹², Roswith Roth, Ph.D.⁹, Marlon Scholz, Joanna Stock^{9,12,14}, Katharina Warncke, M.D.¹⁴, Lorena Wendel, Christiane Winkler, Ph.D.^{2,12,15}. Forschergruppe Diabetes e.V. and Institute of Diabetes Research, Helmholtz Zentrum München, and Klinikum rechts der Isar, Technische Universität München. *Center for Regenerative Therapies, TU Dresden, ¶Dr. von Hauner Children's Hospital, Department of Gastroenterology, Ludwig Maximilians University Munich, ¥Research Institute for Child Nutrition, Dortmund.

Sweden Clinical Center: Åke Lernmark, Ph.D., PI^{1,3,4,5,6,8,10,11,15}, Daniel Agardh, M.D., Ph.D.¹³, Carin Andrén Aronsson^{2,12,13}, Maria Ask, Jenny Bremer, Ulla-Marie Carlsson, Corrado Cilio, Ph.D., M.D.⁵, Emelie Ericson-Hallström, Lina Fransson, Thomas Gard, Joanna Gerardsson, Rasmus Bennet, Monica Hansen, Gertie Hansson, Susanne Hyberg, Fredrik Johansen, Berglind Jonsdottir, M.D., Helena Elding Larsson, M.D., Ph.D.^{6,14}, Marielle Lindström, Markus Lundgren, M.D.¹⁴, Maria Månsson-Martinez, Maria Markan, Jessica Melin¹², Zeliha Mestan, Karin Ottosson, Kobra Rahmati, Anita Ramelius, Falastin Salami, Sara Sibthorpe, Birgitta Sjöberg, Ulrica Swartling, Ph.D.^{9,12}, Evelyn Tekum Amboh, Carina Törn, Ph.D.^{3,15}, Anne Wallin, Åsa Wimar^{12,14}, Sofie Åberg. Lund University.

Washington Clinical Center: William A. Hagopian, M.D., Ph.D., PI^{1,3,4,5,6,7,11,13,14}, Michael Killian^{6,7,12,13}, Claire Cowen Crouch^{12,14,15}, Jennifer Skidmore², Josephine Carson, Maria Dalzell, Kayleen Dunson, Rachel Hervey, Corbin Johnson, Rachel Lyons, Arlene Meyer, Denise Mulenga, Alexander Tarr, Morgan Uland, John Willis. Pacific Northwest Diabetes Research Institute.

Pennsylvania Satellite Center: Dorothy Becker, M.D., Margaret Franciscus, MaryEllen Dalmagro-Elias Smith², Ashi Daftary, M.D., Mary Beth Klein, Chrystal Yates. Children's Hospital of Pittsburgh of UPMC.

Data Coordinating Center: Jeffrey P. Krischer, Ph.D., PI^{1,4,5,10,11}, Michael Abbondandolo, Sarah Austin-Gonzalez, Maryouri Avendano, Sandra Baethke, Rasheedah Brown^{12,15}, Brant Burkhardt, Ph.D.^{5,6}, Martha Butterworth², Joanna Clasen, David Cuthbertson, Christopher Eberhard, Steven Fiske⁹, Dena Garcia, Jennifer Garmeson, Veena Gowda, Kathleen Heyman, Francisco Perez Laras, Hye-Seung Lee, Ph.D.^{1,2,13,15}, Shu Liu, Xiang Liu, Ph.D.^{2,3,9,14}, Kristian Lynch, Ph.D.^{5,6,9,15}, Jamie Malloy, Cristina McCarthy^{12,15}, Steven Meulemans, Hemang Parikh, Ph.D.³, Chris Shaffer, Laura Smith, Ph.D.^{9,12}, Susan Smith^{12,15}, Noah Sulman, Ph.D., Roy Tamura, Ph.D.^{1,2,13}, Ulla Uusitalo, Ph.D.^{2,15}, Kendra Vehik, Ph.D.^{4,5,6,14,15}, Ponni Vijayakandipan, Keith Wood, Jimin Yang, Ph.D., R.D.^{2,15}. *Past staff: Lori Ballard, David Hadley, Ph.D., Wendy McLeod.* University of South Florida.

Project scientist: Beena Akolkar, Ph.D.^{1,3,4,5,6,7,10,11}. National Institutes of Diabetes and Digestive and Kidney Diseases.

Other contributors: Kasia Bourcier, Ph.D.⁵, National Institutes of Allergy and Infectious Diseases. Thomas Briese, Ph.D.^{6,15}, Columbia University. Suzanne Bennett Johnson, Ph.D.^{9,12}, Florida State University. Eric Triplett, Ph.D.⁶, University of Florida.

¹Ancillary Studies, ²Diet, ³Genetics, ⁴Human Subjects/Publicity/Publications, ⁵Immune Markers, ⁶Infectious Agents, ⁷Laboratory Implementation, ⁸Maternal Studies, ⁹Psychosocial, ¹⁰Quality Assurance, ¹¹Steering, ¹²Study Coordinators, ¹³Celiac Disease, ¹⁴Clinical Implementation, ¹⁵Quality Assurance Subcommittee on Data Quality.

References

- Andrén Aronsson C, Kurppa K, Agardh D. Gluten in infants and celiac disease risk. Expert review of gastroenterology & hepatology. 2016a; 10:669–670. [PubMed: 27027529]
- Andrén Aronsson C, Lee HS, Koletzko S, Uusitalo U, Yang J, Virtanen SM, Liu E, Lernmark Å, Norris JM, Agardh D. Effects of Gluten Intake on Risk of Celiac Disease: A Case-Control Study on a Swedish Birth Cohort. Clinical Gastroenterology And Hepatology: The Official Clinical Practice Journal of the American Gastroenterological Association. 2016b; 14:403–409.e3. [PubMed: 26453955]
- Chmielewska A, Pie cik-Lech M, Szajewska H, Shamir R. Primary Prevention of Celiac Disease: Environmental Factors with a Focus on Early Nutrition. Annals of nutrition & metabolism. 2015; 67(2):43–50. [PubMed: 26605913]
- Cullen KW, Baranowski T, Baranowski J, Hebert D, Moor Cde. Behavioral or epidemiologic coding of fruit and vegetable consumption from 24-hour dietary recalls: research question guides choice. Journal of the American Dietetic Association. 1999; 99:849–851. [PubMed: 10405684]
- European Food Safety Authority (EFSA). Data standardisation. 2015. Retrieved February 2016 from: <http://www.efsa.europa.eu/en/data/data-standardisation>
- Institute of Medicine. Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids (Macronutrients). Washington, DC, USA: The National Academies Press; 2005.
- Ireland J, Møller A. Review of International Food Classification and Description. Journal of Food Composition and Analysis. 2000; 13:529–538.
- Ireland J, Møller A. US, Canadian, Australian, and New Zealand datasets seen through foreign eyes. Procedia Food Science. 2013a; 2:195–202.
- Ireland J, Møller A. What's new in LanguaL™? Procedia Food Science. 2013b; 2:117–121.

- Ireland J, van Erp-Baart AMJ, Charrondière UR, Møller A, Smithers G, Trichopoulou A. Selection of a food classification system and a food composition database for future food consumption surveys. *European journal of clinical nutrition*. 2002; 56(2):S33–45. [PubMed: 12082516]
- Knip M, Virtanen SM, Akerblom HK. Infant feeding and the risk of type 1 diabetes. *The American journal of clinical nutrition*. 2010; 91:1506S–1513S. [PubMed: 20335552]
- Kroke A, Manz F, Kersting M, Remer T, Sichert-Hellert W, Alexy U, Lentze MJ. The DONALD Study. History, current status and future perspectives. *European journal of nutrition*. 2004; 43:45–54. [PubMed: 14991269]
- Lagiou P, Trichopoulou A. The DAFNE initiative: the methodology for assessing dietary patterns across Europe using household budget survey data. *Public health nutrition*. 2001; 4:1135–1141. [PubMed: 11924937]
- Lamb MM, Frederiksen B, Seifert JA, Kroehl M, Rewers M, Norris JM. Sugar intake is associated with progression from islet autoimmunity to type 1 diabetes: the Diabetes Autoimmunity Study in the Young. *Diabetologia*. 2015a; 58:2027–2034. [PubMed: 26048237]
- Lamb MM, Miller M, Seifert JA, Frederiksen B, Kroehl M, Rewers M, Norris JM. The effect of childhood cow's milk intake and HLA-DR genotype on risk of islet autoimmunity and type 1 diabetes: the Diabetes Autoimmunity Study in the Young. *Pediatric diabetes*. 2015b; 16:31–38. [PubMed: 24444005]
- Mitchel D. Database Requirements to Meet Food Grouping Needs. *Journal of Food Composition and Analysis*. 2001:279–285.
- National Food Administration. Livsmedelsverket. 2014 Retrieved August 2014.
- National Institute for Health and Welfare. Fineli ® - Finnish Food Composition Database. 2016. Retrieved March 2016 from: <http://www.fineli.fi/index.php?lang=en>
- Norris JM. Infant and childhood diet and type 1 diabetes risk: recent advances and prospects. *Current diabetes reports*. 2010; 10:345–349. [PubMed: 20640941]
- Nutrition Coordinating Center (NCC) University of Minnesota. Nutrition Data System for Research (NDSR) Manual. 2015. from: <http://www.ncc.umn.edu/>
- Reinivuo, H., Laitinen, K. Proposal for the harmonisation of recipe calculation procedures: WP2.2 Composite foods - Harmonisation of recipe calculation procedures (D2.2.12/M2.2.4). 2007. from: http://www.fao.org/uploads/media/reinivuo_laitinen_2007_Eurofir_recipe_harmonisation_D2.2.12_M2.2.4_02.pdf
- Sichert-Hellert W, Kersting M, Chahda C, Schäfer R, Kroke A. German food composition database for dietary evaluations in children and adolescents. *Journal of Food Composition and Analysis*. 2007; 20:63–70.
- Slimani N, Charrondière UR, van Staveren W, Riboli E. Standardization of Food Composition Databases for the European Prospective Investigation into Cancer and Nutrition (EPIC): General Theoretical Concept. *Journal of Food Composition and Analysis*. 2000; 13:567–584.
- Slimani N, Deharveng G, Unwin I, Southgate DAT, Vignat J, Skeie G, Salvini S, Parpinel M, Møller A, Ireland J, Becker W, Farran A, Westenbrink S, Vasilopoulou E, Unwin J, Borgejordet A, Rohrmann S, Church S, Gnagnarella P, Casagrande C, van Bakel M, Niravong M, Boutron-Ruault MC, Stripp C, Tjønneland A, Trichopoulou A, Georga K, Nilsson S, Mattisson I, Ray J, Boeing H, Ocké M, Peeters PHM, Jakszyn P, Amiano P, Engeset D, Lund E, de Magistris MSantucci, Sacerdote C, Welch A, Bingham S, Subar AF, Riboli E. The EPIC nutrient database project (ENDB): a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. *European journal of clinical nutrition*. 2007; 61:1037–1056. [PubMed: 17375121]
- Szajewska H, Shamir R, Mearin L, Ribes-Koninckx C, Catassi C, Domellöf M, Fewtrell MS, Husby S, Papadopoulou A, Vandenplas Y, Castillejo G, Kolacek S, Koletzko S, Korponay-Szabó IR, Lionetti E, Polanco I, Troncone R. Gluten Introduction and the Risk of Coeliac Disease: A Position Paper by the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition. *Journal of pediatric gastroenterology and nutrition*. 2016; 62:507–513. [PubMed: 26815017]
- TEDDY Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY) Study. *Annals of the New York Academy of Sciences*. 2008; 1150:1–13.

- Truswell A, Bateson DJ, Madafoglio KC, Pennington JA, Rand WM, Klensin JC. INFOODS Guidelines for Describing Foods: A Systematic Approach to Describing Foods to Facilitate International Exchange of Food Composition Data. *Journal of Food Composition and Analysis*. 1991; 4:18–38.
- Uusitalo U, Kronberg-Kippila C, Aronsson CA, Schakel S, Schoen S, Mattisson I, Reinivuo H, Silvis K, Sichert-Hellert W, Stevens M, Norris JM, Virtanen SM. Food composition database harmonization for between-country comparisons of nutrient data in the TEDDY Study. *Journal of Food Composition and Analysis*. 2011; 24:494–505. [PubMed: 22058606]
- Valio. Mainion monipuolinen maitojauhe. 2016. Retrieved January 2017 from: <https://www.valio.fi/tuotteet/artikkeli/mainio-maitojauhe/>
- Verger P, Ireland J, Møller A, Abravicius JA, Henauf Sde, Naska A. Improvement of comparability of dietary intake assessment using currently available individual food consumption surveys. *European journal of clinical nutrition*. 2002; 56(2):S18–24. [PubMed: 12082514]
- Virtanen SM. Dietary factors in the development of type 1 diabetes. *Pediatric diabetes*. 2016; 17(22): 49–55. [PubMed: 27411437]
- Virtanen SM, Läärä E, Hyppönen E, Reijonen H, Räsänen L, Aro A, Knip M, Ilonen J, Akerblom HK. Cow's milk consumption, HLA-DQB1 genotype, and type 1 diabetes: a nested case-control study of siblings of children with diabetes. Childhood diabetes in Finland study group. *Diabetes*. 2000; 49:912–917. [PubMed: 10866042]
- Virtanen SM, Nevalainen J, Kronberg-Kippilä C, Ahonen S, Tapanainen H, Uusitalo L, Takkinen HM, Niinistö S, Ovaskainen ML, Kenward MG, Veijola R, Ilonen J, Simell O, Knip M. Food consumption and advanced β cell autoimmunity in young children with HLA-conferred susceptibility to type 1 diabetes: a nested case-control design. *The American journal of clinical nutrition*. 2012; 95:471–478. [PubMed: 22237062]
- Working Group on Food Data Management and Interchange. Eurofoods recommendations for food composition database management and data interchange. Cost Action 99, Research action on food consumption and composition data. 2000

Definitions of terms used in the manuscript / TEDDY terminology

Basic food

Food comprising a single **ingredient** (see also: Ingredient; e.g. carrot, beef), which can be raw or prepared.

Break down

Process to separate out the information about **ingredients** (see also: Ingredient) of a **composite dish** or a **mixed dish** (see also: Dish / composite dish / mixed dish), i.e. 'breaking down' a dish to its ingredients.

Core dish / core food

A **core dish** (e.g. pasta sauce, veggie burger) does not break down (see also: Break down) into ingredients (see also: Ingredient) automatically in the food grouping calculation process. It is sometimes called a **core food** especially if referring to a common commercial **product** (see also: Product; e.g. crisp bread, breakfast cereal).

Dish / composite dish / mixed dish

A **food item** (see also: Food item) that has been prepared at home or by industrial or catering processes (Reinivuo & Laitinen, 2007). A **dish** has one or more **ingredients** (see also: Ingredient). If more than one **ingredient**, then it is called a **composite dish** or a **mixed dish**. A dish breaks down (see also: Break down) into **ingredients** in food grouping process.

Dry weight

The weight of the **food items** (see also: Food item) in a food group is given as dry weight, e.g. flour, rice, wheat, starches, baby formula powder.

Food item

General term used if spoken about **basic foods** (see also: Basic food) and **dishes** (see also: Dish / composite dish / mixed dish).

Form of food

General term used to describe the appearance of a food item (see also: Food item), e.g. raw or prepared; fresh, dry, liquid or powder.

Formulation

If a **recipe** (see also: Recipe) includes nutrients as **ingredient** (see also: Ingredient) to match the Nutrition Facts-label information of a **dish** (see also: Dish / composite dish / mixed dish) the **recipe** is called a **formulation**. The term is only used in the U.S. food composition database (FCDB).

Generic recipe / standard recipe

Used to estimate average weighted **food items** (see also: Food item) (Slimani et al., 2000) among several **recipe** (see also: Recipe) variations of the same **dish** (see also: Dish / composite dish / mixed dish).

Ingredient

A **food item** (see also: Food item) included in a **recipe** (see also: Recipe) (Reinivuo & Laitinen, 2007).

Liquid weight

The weight of the **food items** (see also: Food item) in a food group is given as **liquid weight**. This can be liquid itself (e.g. coffee, juice, soft drinks or infant formula) or it is possible that conversions need to be made from **dry weight** (see also: Dry weight) to **liquid weight** (e.g. baby formula as powder will be converted to the ready-to-drink form).

Prepared weight

The weight of the **food items** (see also: Food item) in a food group is given as prepared which resulted from cooking e.g. steaming, baking or frying.

Product

A **dish** (see also: Dish / composite dish / mixed dish) that has been industrially prepared.

Raw weight

The weight of the **food items** (see also: Food item) in a food group is given as raw food. In this case, it is possible that conversions need to be made from **prepared weight** (see also: Prepared weight) to **raw weight** within that food group, e.g. conversion of the weight of cooked peas to their **raw weight**.

Recipe

A list of **ingredients** (see also: Ingredient), including the amounts, which are needed to prepare a selected **dish** (see also: Dish / composite dish / mixed dish) (Reinivuo & Laitinen, 2007). Note: A **recipe** can be included in a **recipe**, e.g. in filled pasta recipe, pasta has its own recipe.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

3-5 bullet points to include core findings need to be submitted in a separate file, with a maximum of 85 characters (including spaces) per bullet point

- TEDDY study investigates the environmental determinants of diabetes in the young (82 characters incl. space)
- Food grouping definitions were harmonized across four food composition databases (82 characters incl. spaces)
- The harmonized food groups comprise 15 main groups and 89 subgroups (69 characters incl. spaces)
- Food grouping harmonization enables studying food intakes across TEDDY countries (*80 characters incl. spaces*)

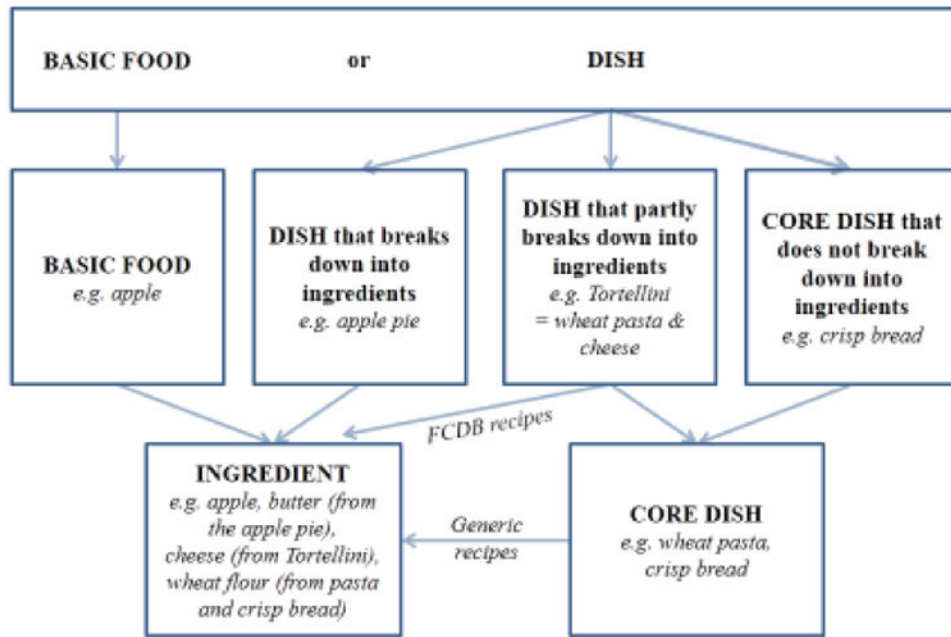


Figure 1. Overview of the disaggregation procedure of basic food items and composite dishes
 FCDB – Food Composition Database; Both FCDB and generic recipes are country-specific.

Table 1
Comparison of original food grouping systems across the food composition databases used in The Environmental Determinants of Diabetes in the Young (TEDDY) study (with cereals and cereal products as an example)

FINELI^a (Finland)	LEBTAB^b (Germany)	TEDDY-SWE^c (Sweden)	NDSR^d (U.S.)
Food use classes (Cereals and cereal products) Biscuits Bread, mixed flour Breakfast cereals Buns Pasta dishes Pies and other cereal products Pizza Porridge Rice as a side dish Rye bread Sandwiches and burgers Savory bakery Sweet bakery Wheat bread	Cereal grains, breakfast cereals, pasta Baked products: breads, bread rolls, etc. Baked products: cakes, biscuits, etc.	Bread Crisp bread; Soft white bread; Rye bread, coarse rye bread; Soft whole meal bread; Soft and crispy flatbread; Gluten free bread and rice crackers; Dishes based on bread Porridge; gruel Porridge with gluten; Porridge gluten free; Gruel with gluten; Gruel gluten free Cereals; muesli Cereals, muesli low/unsweetened; Cereals, muesli sweetened; Commercial baby foods based on fruit purée and cereals Pancakes; waffles; crêpes Pancakes, waffles, crêpes; Dishes incl. meat/sausages/poultry; Dishes incl. fish and seafood; Dishes incl. vegetables; Gluten free dishes; Unspecified Pizza; pie; pirogues Dough; Dishes incl. meat/sausage/poultry; Dishes incl. fish and seafood; Dishes incl. vegetables; Dishes incl. cheese and egg Rice and grains Rice; Rice flour; Rice milk; Grains and seeds, cooked; Dishes incl. meat/sausage/poultry; Dishes incl. fish and seafood; Dishes incl. vegetables; Rice based desserts Pasta Pasta; Pasta gluten free; Pasta dishes – simple; gluten free; Dishes incl. meat/sausage/poultry; Dishes incl. fish and seafood; Dishes incl. vegetables; Dishes incl. cheese and egg Bakery sweet Buns, crust; Biscuits, cookies, crackers; Cakes, pastries, swiss roll; Sponge cake w/o filling; Gluten free sweet bakery Flour; starch Wheat; Rye; Oats; Barley; Other cereals; Flour blends; Gluten	Grains, Flour and Dry Mixes - Whole Grain; Some Whole Grain; Refined Grain Loaf-type Bread and Plain Rolls - Whole Grain; Some Whole Grain; Refined Grain Other Breads (quick breads, corn muffins, tortillas) - Whole Grain; Some Whole Grain; Refined Grain Crackers - Whole Grain; Some Whole Grain; Refined Grain Pasta - Whole Grain; Some Whole Grain; Refined Grain Ready-to-eat Cereal (not presweetened) - Whole Grain; Some Whole Grain; Refined Grain Ready-to-eat Cereal (presweetened) - Whole Cakes, Cookies, Pies, Pastries, Danish, Doughnuts and Cobblers - Whole Grain; Some Whole Grain; Refined Grain Snack Bars - Whole Grain; Some Whole Grain; Refined Grain Snack Chips - Whole Grain; Some Whole Grain; Refined Grain Popcorn Flavored Popcorn Baby Food Grain Mixtures - Whole Grain; Some Whole Grain; Refined Grain
Ingredient Classes (Cereals) Rice Wheat Rye Oats and barley Other grains Starches Pasta and macaroni Crispbread			

^aFinnish food composition database;

^bLebensmittelTabelle;

^cThe national Swedish food composition database (FCDB) is called NFA FCDB which does not include information about commercial baby foods. After the TEDDY Study added the baby food information to a copy of the NFA FCDB the new version of the database was named TEDDY-SWE to distinguish it from the original version.;

^dNutrition Data System for Research

Table 2
Definitions and contents of The Environmental Determinants of Diabetes in the Young (TEDDY) food (sub-) groups

Food groups	Definition and examples of content	WEIGHT
CEREALS	Various types, e.g. cracked, flour, flakes, grit	
Rice	Various types of rice; Two groups will be separated: Rice milk Dry or parboiled rice	Liquid Dry
Wheat	Various types of wheat (e.g. wheat flour, wheat flakes)	Dry
Rye	Various types of rye (e.g. rye flour, rye flakes)	Dry
Oats	Various types of oats; Two groups will be separated: Dry oats Oat milk	Dry Liquid
Barley	Various types of barley (e.g. barley flour, barley flakes)	Dry
Corn	Corn meal, polenta	Dry
Other gluten free flours and starches	Non-gluten cereals that cannot be categorized to any of the previous categories, e.g. millet, buckwheat, quinoa, tapioca as well as potato flour, starch, wheat starch	Dry
FRUIT AND BERRIES	Various types, e.g. fresh, cooked, canned, juice, purees (Fruit purees are counted as such under their corresponding fruit group, i.e. prepared)	
Citrus fruit	Clementine, grapefruit, lemon, lime, mandarin, orange, tangerine	Raw
Apple	Apple	Raw
Berries	Strawberry, raspberry, blueberry, lingonberry, bilberry, cloudberry, blackberry, rose hip, acai, cranberry, arctic brambleberry	Raw
Other fruits	Banana, nectarine, peach, pineapple, pear, plum, melons, grapes, pomegranate, coconut, kiwi, passion fruit, cherry	Raw
Canned fruits	All canned fruits are grouped together	Raw
Dried fruits and berries	Apricot, cranberries, raisins, currants	Dry
Citrus juice	Orange-, grapefruit-, tangerine juice (100% juice)	Liquid
Apple juice	Apple juice (100% juice), apple cider	Liquid
Berry juices	All 100% berry juices	Liquid
Other fruit juices	Cherry-, pineapple juice (100% juice)	Liquid
Mixed Juices	Mixtures of berry and apple juices	Liquid
Juices, mixture of vegetable and fruit	ACE-juice, i.e. a mixture of carrot and apple/orange	Liquid

Food groups	Definition and examples of content	WEIGHT
VEGETABLES	Various types, e.g. fresh, cooked, canned, juice	
Potatoes	Various types of potatoes including cooked, baked, boiled, roasted or pan-fried potatoes, French fries, dried potatoes, potato flakes, and potato chips	Raw
Roots, sweet potatoes	Sweet potato, carrot, rutabaga, radish, cassava, yams, celery root, turnip, radish, lotus root, water chestnut	Raw
Leafy vegetables	Lettuce, arugula, celery stem, mustard green, spinach, asparagus, artichoke, Swiss chard, beet leaves, fresh basil, fresh coriander, Alfalfa	Raw
Cabbages	Kale, red cabbage, Chinese cabbage, Savoy, bok choy, broccoli, brussels sprouts, cauliflower, collard green, Sauerkraut	Raw
Fruit vegetables	Tomato, eggplant, peppers, zucchini and other squashes, cucumber, plantain, avocado, pumpkin, fresh corn, okra, heart of palm, bamboo sprouts	Raw
Onions	All types of onion (e.g. white, yellow, red), spring onion, leek, chives, garlic	Raw
Mushrooms	All wild and cultivated mushrooms	Raw
Legumes, beans, peas	Lentils, red, kidney, white, lima, black, green beans, peas Note: Dried peas, lentils are not converted to raw! This applies for FIN and GER, not SWE and US	Raw
Vegetables, dried	All types of dried vegetables (except legumes, beans and peas)	Dry
Vegetables, canned	All types of canned vegetables	Prepared
Vegetable juices	Carrot juice, beet juice, vegetable juices like V8, all 100% vegetable juices	Liquid
SOY		
Soy beans	Soy nuts, flour, germs, roasted bean	Raw
Soy products	Soy tofu and cheese, curd	Prepared
Soy milk	Soy milk and cream	Liquid
Soy dessert	Soy pudding, soy yoghurt, soy ice cream	Prepared
Soy sausage		Prepared
Soy sauce	All varieties, e.g. Chinese, Japanese, Vietnamese soy sauce	Liquid
NUTS, SEEDS	Various types, e.g. grounded, crushed, flakes, milk	
Nuts, seeds	Various types of nuts; Three groups will be separated:	
	Nut/seed milks, e.g. almond milk, coconut milk	Liquid
	Chestnut, peanut, walnut, pistachio, pecan, cashew, sunflower seed, pumpkin seed, sesame, pine nut, macadamia	Raw
	Nut spreads	Prepared
FATS AND OILS		
Vegetable oils	Canola-, olive-, sunflower-, corn oils, various nut oils (= liquid at room temperature)	Prepared
Solid vegetable fat	Vegetable fats such as palm oil, coconut oil, palm kernel oil, and shortening (= naturally solid at room temperature)	Prepared

Food groups	Definition and examples of content	WEIGHT
Fish oil	All fish oils	Prepared
Animal fats	Lard, tallow	Prepared
Butter	Two groups will be separated, i.e. butter with a fat content <50% and 50%	Prepared
Margarines	Two groups will be separated, i.e. margarine with a fat content <50% and 50%	Prepared
Butter-margarine mixtures	Two groups will be separated, i.e. butter-margarine mixtures with a fat content <50% and 50%	Prepared
MILK AND MILK PRODUCTS	Various types, e.g. slices, shredded, cubes form is taken into account: powder conversion to liquid form	
Breast milk	Human breast milk	Raw
Fat-free milk	Cow's milk or sweetened milk drinks such as cocoa <1%	Raw
Low-fat milk	Cow's milk or sweetened milk drinks such as cocoa 1-2%	Raw
High-fat milk	Cow's milk or sweetened milk drinks such as cocoa >2% and 5%	Raw
Other animal milk	Horse-, goat-, sheep milk	Raw
Creams	Creams >5% fat	Prepared
Ice cream	All dairy ice creams	Prepared
Sour milk and sour milk products	Sour milk, yogurt, kefir, cultured milk, butter milk, sour cream, crème fraîche	Prepared
Cheese	Various types of cheese; Two groups will be separated: Fresh cheeses, e.g. Cottage cheese, cream cheese and other similar cheese spreads, homemade cheese, mozzarella Aged cheeses, e.g. Gouda, Emmentaler, Edam, Gruyere, Camembert, Brie, cheese spreads (e.g. in tubes, contain often aged cheese, check the ingredients first before categorization), blue cheese like Roquefort, Finnish "sulatejuusto" (made out of Edam)	Prepared Prepared
Whey	Liquid whey, whey concentrate, whey isolate (powders/concentrates are converted to liquid)	Liquid
NON-DAIRY PRODUCTS		
Non-dairy products	Non-dairy yoghurts, ice cream/desserts, kefir, creams, e.g. almond, rice, oat yoghurt and ice cream	Prepared
MEAT AND MEAT PRODUCTS	Various types, e.g. fillet, minced, cold-cuts	
Pork	Pork, including non-cured bacon	Raw
Beef	Beef, veal	Raw
Poultry	Various poultry meat, e.g. goose, duck, game bird, pheasant	Raw
Lamb, goat, horse	Lamb, goat meat, horse	Raw
Game	Various game meat, e.g. rabbit, venison, deer, elk	Raw
Processed meats and sausages	Cold-cuts, sausages, brawn, bologna, spam, liver sausage, smoked meat, cured bacon	Prepared
Organ meats/offal	Offal from all kind of meat types	Raw

Food groups	Definition and examples of content	WEIGHT
FISH AND FISH PRODUCTS	Various types, e.g. fillet, whole fish (edible part)	
Fresh and frozen fish	All types of fresh and frozen fish	Raw
Processed fish	Pickled, fermented, dried or smoked fish	Prepared
Canned fish	All types of canned fish e.g. canned tuna, salmon	Prepared
Shellfish, other seafood	Shrimp, clam, crawfish, oysters, crab, caviar, lobster, crab cake as well as canned shell fish	Raw
EGGS	Various types, e.g. cooked, poached, fried form is taken into account: powder, conversion to raw liquid form	
Eggs	All type of eggs, e.g. chicken, duck, goose; Egg powder will be converted back to raw	Raw
BEVERAGES	Various types; form is taken into account: powder conversion to liquid form	
Coffee	Coffee, consider powders	Liquid
Tea	Teas from tea plant, unsweetened ice tea, powders e.g. lemon tea; not including herbal or fruit teas	Liquid
Light Beverages including cola	Unsweetened or artificially sweetened / low-calorie (e.g. Stevia) sweetened soft drinks, fruit and berry drinks, ice, and nectar as well as instant teas (< 20kcal/100g)	Liquid
Sugar sweetened beverages including cola	Sugar sweetened soft drinks, fruit and berry drinks, ice, and nectar as well as regular ice tea, instant teas (>20kcal/100g)	Liquid
Alcohol	All types, e.g. wine, beer, cognac, brandy, liquor	Liquid
CONFECTIONARY		
Sweets	Total sugar candies, non-chocolate power- and energy bars, candy mixtures with chocolate as minor component	Prepared
Chocolate	High-fat sweets, chocolate, M&M, raisins with chocolate, and chocolate power- and energy bars, cocoa powder in a cocoa drink will not be counted, cocoa as liquid is grouped under the corresponding milk group (see above)	Prepared
INFANT FORMULAS	Various types; form is taken into account: powder conversion to liquid form	
Regular cow's milk based infant formulas	All types of regular cow's milk based infant formulas	Liquid
Partially hydrolyzed cow's milk based infant formulas	All types of partially hydrolyzed cow's milk based infant formulas	Liquid
Fully hydrolyzed cow's milk based infant formulas and amino acid (elemental) formulas	All types of fully hydrolyzed cow's milk based infant formulas and amino acid infant formulas	Liquid
Soy base infant formulas	All types of soy based infant formulas	Liquid
Other non-dairy infant formulas	All types of non-dairy infant formulas	Liquid
Other animal (dairy) infant formulas	All types of infant formulas from goat, sheep or horse milk	Liquid
MISC.		
Ketchup	Various types	Prepared