nutrients

MDPI

Article

Development of an Unified Food Composition Database for the European Project "Stance4Health"

Daniel Hinojosa-Nogueira [1], Sergio Pérez-Burillo [1,2], Beatriz Navajas-Porras [1], Bartolomé Ortiz-Viso [3], Silvia Pastoriza de la Cueva [1], Fabio Lauria [4], Alexandra Fatouros [5], Kostas N. Priftis [6], Verónica González-Vigil [7] and José Ángel Rufián-Henares [1,8,*]

1 Centro de Investigación Biomédica, Departamento de Nutrición y Bromatología, Instituto de Nutrición y Tecnología de los Alimentos, Universidad de Granada, 18071 Granada, Spain; dhinojosa@ugr.es (D.H.-N.); spburillo@ugr.es (S.P.-B.); beatriznavajas@ugr.es (B.N.-P.); spdelacueva@ugr.es (S.P.d.l.C.)
2 Department of Biochemistry and Molecular Biology, Boonshoft School of Medicine, Wright State University, Dayton, OH 45435, USA
3 Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, 18071 Granada, Spain; bortiz@ugr.es
4 Institute of Food Sciences, National Research Council, 83100 Avellino, Italy; fabio.lauria@isa.cnr.it
5 Department of Food Chemistry and Analysis, Institute of Food Technology and Food Chemistry, Technische Universität Berlin, 13355 Berlin, Germany; a.urbisch@tu-berlin.de
6 Allergology and Pulmonology Unit, 3rd Paediatric Department, National and Kapodistrian University of Athens, 12462 Athens, Greece; kpriftis@otenet.gr
7 Gestión de Salud y Nutrición S.L., 33003 Oviedo, Spain; vgonzalez@gsnsoft.es
8 Instituto de Investigación Biosanitaria ibs.GRANADA, Universidad de Granada, 18071 Granada, Spain
* Correspondence: jarufian@ugr.es

Abstract: The European Commission funded project Stance4Health (S4H) aims to develop a complete personalised nutrition service. In order to succeed, sources of information on nutritional composition and other characteristics of foods need to be as comprehensive as possible. Food composition tables or databases (FCT/FCDB) are the most commonly used tools for this purpose. The aim of this study is to describe the harmonisation efforts carried out to obtain the Stance4Health FCDB. A total of 10 FCT/FCDB were selected from different countries and organizations. Data were classified using FoodEx2 and INFOODS tagnames to harmonise the information. Hazard analysis and critical control points analysis was applied as the quality control method. Data were processed by spreadsheets and MySQL. S4H's FCDB is composed of 880 elements, including nutrients and bioactive compounds. A total of 2648 unified foods were used to complete the missing values of the national FCDB used. Recipes and dishes were estimated following EuroFIR standards via linked tables. S4H's FCDB will be part of the smartphone app developed in the framework of the Stance4Health European project, which will be used in different personalized nutrition intervention studies. S4H FCDB has great perspectives, being one of the most complete in terms of number of harmonized foods, nutrients and bioactive compounds included.

Keywords: food composition database; food standardization; food data; nutrients; bioactive compounds; public health; personalized nutrition

check for updates

Citation: Hinojosa-Nogueira, D.; Pérez-Burillo, S.; Navajas-Porras, B.; Ortiz-Viso, B.; de la Cueva, S.P.; Lauria, F.; Fatouros, A.; Priftis, K.N.; González-Vigil, V.; Rufián-Henares, J.Á. Development of an Unified Food Composition Database for the European Project "Stance4Health". *Nutrients* **2021**, *13*, 4206. https://doi.org/10.3390/nu13124206

Academic Editors: Alessandra Durazzo, Rosa Casas and Massimo Lucarini

Received: 27 October 2021
Accepted: 22 November 2021
Published: 24 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

There is a close relationship between eating habits, nutrition and health [1]. Many efforts have been made to investigate the nutrient composition of foods consumed by the population [2]. Food composition data describe the content in terms of energy, macronutrients and micronutrients, as well as other compounds such as phytochemicals, antinutrients, bioactive compounds or toxic compounds in foods [3]. Generally, food composition data are published via food composition tables (FCT) and more recently as food composition databases (FCDB) [4–7].

The FCT/FCDB provide data of foods and beverages consumed by the largest portion of a population [8–10]. Currently, there are new agents to take into account, such as climate change [11–13] or the loss of biodiversity [12,14]. Add to this the constant change in consumer preferences [15,16], such as the increased consumption of processed products [17], novel foods [12,18], and an increase in global trade [5,19]. Due to these factors, FCDB are increasingly trying to collect a greater number of nutrients, bioactive compounds and foods.

FCDB data can come from: (i) original analytical data; (ii) published or imputed values of a specific or similar food; and (iii) calculated values or data provided by other FCDB [6,9,10]. On the other hand, food composition can be influenced by different factors [5,6,11,15,20–22] as depicted in Figure 1. All these factors can result in a somewhat different food composition between countries, and even between regions from the same country, thus requiring the development of more detailed and higher quality FCDB [14,15].

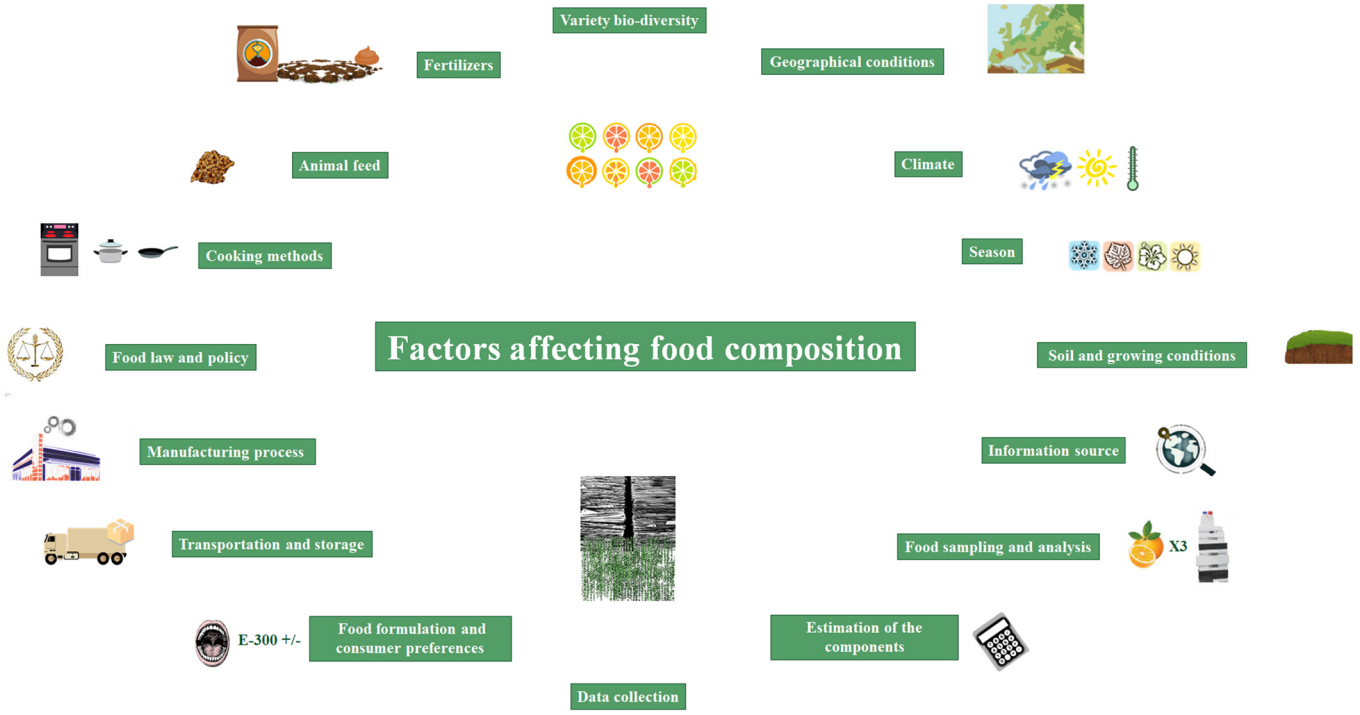

Figure 1. Main factors affecting the nutrient content and food composition. Numbers correspond to appropriate references.

FCT/FCDB are an essential tool in a wide range of areas. For example, in the field of public health [10], health programs and clinical practice [10,23], nutritional epidemiology [24], in research and food safety [18,20,25], in the food industry [20], and in agricultural programs and policies [7,23].

Currently, a growing number of countries are updating their FCT/FCDB, for example, McCance and Widdowson's food composition table [26], the Dutch Food Composition Database (NEVO) [27] and Frida Food Data [28] which include a wide range of foods and compounds, making them a reference at an international level [12]. However, several countries still lack their own data sets [10,15,20,22,29] so they often resort to foreign FCT/FCDB [9,24] such as the Food and Agriculture Organization (FAO) [30]; or the United States Department of Agriculture (USDA) FCDB [31], among others. Nevertheless, most food composition data are based on fresh foods, while information on processed foods, recipes or fortified foods is usually missing or not up to date [9,15,16,29]. Organizations such as the International Network of Food Data Systems (INFOODS) [23,32,33] are making great efforts to provide information about different FCT/FCDB, promoting the reliability and up-to-date nature of the data [34,35].

Therefore, by having a harmonized and standardized FCT/FCDB, comparisons between countries would be possible and nutritional data would be more accurate and com-

prehensive [10,15,34]. In order to standardize terms between data bases, different ontologies are used in nutritional research [9,36,37]. The most common is LanguaL™ [38,39]. LanguaL™ is based on the concept that any food (or food product) may be described systematically by a combination of characteristics [39]. There are other descriptors such as those developed by the European Food Safety Authority (EFSA), and FoodEx2 [40]. FoodEx2 is a standardized food classification that consists of individual food items aggregated into food groups and food categories in a hierarchical structure [15,37,40–42]. Recently, efforts have been made to map FoodEx2 facet descriptors with LanguaL codes [39,42].

In order to overcome this challenge, in the last decades great progress has been made to develop standards and guidelines focused on the harmonization and standardization of FCDB [10,15,23]. Among the most important is FAO/INFOODS that coordinates food composition activities at the international level [6,32]. FAO/INFOODS has developed different strategies in an attempt to harmonize data and make it comparable across countries [4,6,32,33,43–46].

Additionally, there have also been numerous EU-funded initiatives to standardize and harmonize food compositional data [15] such EUROFOODS, COST99 or NORFOODS [16]. More recently, they continued via the European Food Information Resource Network (EuroFIR), now known as EuroFIR AISBL [2,10,16,47]. The main objective of EuroFIR is to contribute to the harmonization of high-quality food composition data in Europe [47–49]. For this purpose, the EuroFIR project has developed different tools like its own LanguaL™ descriptors; EuroFIR Theasauri or FoodEXplorer. FoodEXplorer is a query tool that includes food composition data across more than 30 countries [20,47] and is updated regularly [2,47,50]. In Europe, these networks allowed the development of large multicenter nutritional studies. For example, the European Prospective Investigation into Cancer and Nutrition (EPIC) [51]. Notably, food composition analysis is very expensive and can be time consuming [22,46]. However, an increasing number of FCDB are introducing as many nutrients and bioactive compounds available as possible [30,52–54]. FooDB (https://www.foodb.ca, accessed on 27 October 2021) represents the most comprehensive effort to integrate food composition data [24] and a large amount of different compounds [55].

Stance4Health (Smart Technologies for personalized Nutrition and Consumer Engagement) (S4H) is a project funded by European Union's Horizon 2020 research and innovation program, aimed at evaluating the benefits of a novel smart personalized nutrition service in a large clinical study [56]. One of the main tasks of the project is to build a nutritional database (with as many foods and nutrients as possible) to complete the national FCDBs from the countries involved in the project. As the FCT/FCDB of the countries is completed, a more accurate approximation of the users' diets will be achieved.

The aim of the present study is to describe all the harmonization efforts and introduce this novel and unified Stance4Health's FCDB (S4H FCDB). This database will be part of the app developed in the framework of the European project, which will be used in different personalized nutritional intervention studies (Trial ID: ISRCTN63745549).

2. Materials and Methods

2.1. Working Group Organization and Training

A working team composed of two coordinators and a committee (including researchers, computer scientists and compilers, all of whom were dietitians and nutritionists) was established for the preparation of the S4H FCDB. Both the coordinators and the compilers completed the e-learning course offered free of charge by FAO/INFOODS [57]. The e-FoodComp course on food composition was designed by experts to be used by different professional users. The course consisted of 14 lessons structured in five units, for a total of approximately 10 h. The course offers a large number of examples and exercises suitable for on-the-job training. In addition, different guides and research were chosen to be used as a reference for the standardization and harmonization processes [3,25,32,33,41–51,57–68]. The coordinators established the general guidelines, and also helped choosing and obtain-

ing the FCDBs used. In addition, they were subsequently responsible for checking and assessing the quality of the harmonized procedures and data. The remaining committee members performed the rest of the tasks.

2.2. Data Collection, Harmonization and Standardization Methods

A personalized nutrition intervention for different populations in Spain, Germany and Greece will be carried out within the S4H Project [56]. For this reason, the three national FCT/FCDB of the intervention countries were used as references [69–71]. These FCT/FCDB were completed with values of nutrients, bioactive compounds, such as polyphenols, and foods from different databases [14,26,27,30,31,72–75] (Table 1). All FCT/FCDB were either free of charge or permissions were granted when needed. The original FCDB data, such as original name or food identifier, were kept for the purpose of future checks or updates. In addition, quality and traceability of the documented data was guaranteed. However, the data needed to undergo some conversions before being added to our FCDB. All data were harmonized in order to obtain standardized foods and nutrients. Subsequently, all the information was entered into dynamic spread-sheets that related the data and characteristics to each other. As all foods were not in one single language, names and recipes were translated into English. All foods were uniquely identified using the standardized food classification and description system proposed by EFSA FoodEx2 [40,42]. The coding was carried out by qualified compilers and the last version of FoodEx2 system was used [40]. FoodEx2 allowed coding of all foods and beverages present in the FCDB into 20 main food categories, divided into subgroups up to a maximum of four levels [68]. Fortified foods, dietary supplements, food commercial brands, recipes or prepared dishes were discarded from the FCDB. Cooked foods were included, and the cooking method was extracted as an additional data element. Generic unbranded processed foods (such as canned foods, pickles, processed meats or pastries, among others) were also included.

Table 1. Compilation of FCDBs used in the construction of the S4H FCDB.

Name of the FCDB	Last Update	Nº FoodsT [1]	Nº Foods [2]	Nº Items [3]	References
Tabla de Composición de Alimentos de Martin Peña actualizada de la version original por i-Diet (Spain)	2019	726	711	90	[69]
Composition tables of foods and Greek dishes (Greece)	2007	88	84	18	[71]
Bundeslebensmittelschlüssel (BLS) (Germany)	2014	936	715	146	[70]
Banca Dati di Composizione degli Alimenti per Studi Epidemiologici in Italia (Italy)	2015	978	976	97	[72]
Dutch Food Composition Database (NEVO) (Netherlands)	2019	2152	949	144	[27]
McCance and Widdowson's 'composition of foods (United Kingdom)	2019	2910	1208	280	[26]
Food and Nutrient Database for Dietary Studies (FNDDS) (United States)	2018	7083	609	69	[31]
FAO/INFOODS Analytical food composition database version 2.0 (AnFooD2. 0)	2017	2953	346	378	[30]
FAO/INFOODS food composition database for biodiversity (BioFoodComp4.0)	2017	7953	355	538	[14]
Phenol-Explorer 3.6 database on polyphenol content in food	2015	458	457	520	[73–75]

[1] Number of foods used out of total. [2] Number of foods included. [3] Number of items collected, including information on food nutrients and other compounds and data.

The complete dataset was examined and converted into standard units [3,43]. The tagnames for food components developed by INFOODS were used for this purpose [33,60]. In order to ensure harmonization, standard tagnames were designed for each compound. The original FCDB compounds that were in different units or did not correspond to those described in the INFOODS tagnames, were transformed and recalculated to match the one expressed in the standard tagname (i.e., change of units from grams to milligrams) [33,43].

Only in specific cases were tagnames not modified (as in the case of some polyphenols) where the coordinators decided that it was more functional to leave all compounds with the same units. Those compounds that did not have labels were assigned one that was proposed by compilers. The labels and units can be found in Supplementary Material S1 (Excel sheet). All compounds were expressed in amount per 100 g or 100 mL of food and edible portion values were extracted for further calculations as recommended [3,33]. All changes were made manually or semi-automatically in spreadsheets. All changes were monitored and subsequently validated as described in Section 2.5.

2.3. Mapping and Unification Process

Once the data were harmonized, a single FCDB was created. The data were differentiated by origin, but organized in a homogeneous structure. The mapping process involved matching foods based on the FoodEx2 identification code. The data were cleaned by eliminating 0 values and treated as missing to eliminate possible errors in the matching. Standard rounding values were taken [43]. Statistical parameters (mean, median, standard deviation) were calculated for each compound whenever a food had the same code. After all the data were evaluated, the coordinators decided to use the median as the final value. Unification was applied to foods with the same codes. The median was used in order to unify and complete the values of a food as long as the matchings were identical. The results were filtered using different filters as values to locate the values of the outer layers. Afterwards, the quality of the data was evaluated. All changes were made in spreadsheets, and Python 3.0 was used for unification and statistical calculations. The scripts used are shown in Supplementary Material S2. For the S4H FCDB, energy was recalculated using the Atwater factors [62]. Once the values were obtained, they could be inputted in the national FCDB for those foods that are not yet included, or for those nutrients or compounds that were missing.

2.4. Recipe Calculation and Additional Factors

Recipes or prepared dishes will be introduced as part of another database. Recipes will be linked to the S4H FCDB in order to obtain all the necessary information. For the calculations, the edible portion, cooking method and those factors that can generate changes in the nutrient content (such as retention factors (RF) and yield factors (YF)) will be taken into account. In addition, allergen data and preparation methods will be implemented.

For the harmonized calculation of recipes, a mixed model was used, since it is the most widely used and accepted [3,76]. This method was proposed as standard by EuroFIR, and consist of applying YF at the recipe level and the RF to each individual ingredient [48,77]. This procedure requires incorporating beforehand the standardized YF and RF based on the food group classification system [25,78]. YF and RF values were obtained from different sources in order to cover the largest number of foods and cooking methods [26,50,76–80]. For the RF of polyphenols, in addition to those given in Phenol-Explorer 3.6 [75], the values retrieved from the EPIC study [61] were also used. The calculation method involved the following steps: first, weights of the raw ingredients were collected. Second, nutrient and compound levels were corrected for edible portions, if applicable. Next, ingredients were modified to account for the effects of cooking by using yield factors to adjust the raw weights. In addition, retention factors were also applied for nutrient losses or gains during cooking. Finally, the ingredient values were summed to obtain recipe values. Final values were expressed per 100 g of recipe and per total recipe weight. The estimates were performed automatically and entered as recipes in the database.

2.5. Information Management and Data Quality

Tables and FCDB were implemented in MySQL open-source software. MySQL is a cross-platform relational database management system. A total of eight tables were implemented and interrelated. Tables were disaggregated to provide more versatility and security. All values were subjected to a variation range. Organizations such as INFOODS or

EUROFIR propose different methodologies to ensure and validate data quality [25,33,50]. However, in this case, the coordinators decided to follow a system of hazard analysis and critical control points (HACCP) [50]. For each data input, an original document and a working document identified with the same code were stored. For each step identified as HACCP, a series of validation tests were performed. These tests were based on different recommendations [3,25,33,50,57]. The validation procedure was followed by corrections, if necessary. The corrections of the conflicting foods were checked by data traceability extending to the original FCDB. The verifications performed are shown in Table 2. Those processes were applied at each stage of quality control, trying to minimize systematic and random errors. All tests were performed manually or semi-automatically by the coordinators, except for the recipes, which were automated.

Table 2. Steps identified as HACCP and validation testing.

Validation Testing	Step as HACCP
Verification of food name and description, possible misspellings or translation errors.	Harmonization
Classification and consistency verification of food name and grouping.	Harmonization
Verification of FoodEx2 coding and INFOODS compound tagging.	Harmonization
Sum of (water + protein + fat + total carbohydrates + alcohol + ash) is within the range: 95–105 g.	Unification
Implausible values, such total fat value, is = 0, Fatty acids = 0 and Cholesterol = 0, or fiber in fish.	Unification
Outlier values within each nutrient or compound.	Unification
Spreadsheet to MySQL data transfer checking.	Management
Model recipe testing.	Recipes Calculation

3. Results

Around 26,200 foods were collected from different FCDBs. Branded foods, recipes or ready-to-eat products, among others, were excluded and a total of 6410 foods were obtained. The Netherlands, the Italian and the United Kingdom's FCDB were the ones that contributed the largest number of foods in the unification process. A large number of foods were excluded from the FAO FCDB due to incomplete information. Subsequent to unification, filtering and quality validation, 2648 foods were obtained for the S4H FCDB and 47% of them had an equivalent food in another FCDB, so that achieved unified values. The foods were grouped by food groups and shown in Supplementary Material S1 (Excel sheet).

Regarding nutrients, bioactive compounds and other information, 880 items were collected. About 95% of the items corresponded to nutrients or other food compounds. Only 5% corresponded to other items such as the food group, its code or some additional factors. During harmonization and standardization, 78.7% of the tagnames were kept with the recommended INFOOD standards units [33,60], without taking into account the polyphenol tagnames. However, the majority of the polyphenols did not have standard tagnames and represented 55.7% of the total of items. Only 5.3% of other compounds did not have standard tagnames. The standard units of 8.4% of the total number of compounds was modified to more functional units.

Germany contributed the highest percentage (15%) of total nutrients, Spain 9% and Greece 2%. It should be noted that 65.5% of the nutrients included in the database were polyphenols from Phenol-Explorer 3.6. If we do not take this into account, the percentages are tripled, as shown in Figure 2. For example, Spain and Germany had around 88% of the 40 most used nutrients in epidemiology, while Greece had only 40%. After Phenol-Explorer, the FAO FCDB is the one with the highest percentage of compounds, around 28.2%. However, the English and Italian FCDBs were the ones with the highest percentage of nutrient values used in epidemiology, with more than 95%.

Figure 2. Absence or presence of different compounds and nutrients in the FCDB. The FAO/INFOODS tagnames are expressed with the S4H IDs, listed in the Supplementary Material Table S1. Not all tagnames are shown; the Phenol-Explorer Database is not included and the complete figure is depicted as Supplementary Material Table S3.

Figure 3 shows an example of the values of the unification process for the item A00MH Spinaches, raw. Raw spinach was selected because it was included in most FCT/FCDB. The value of total proteins is quite similar, which confirms a correct classification of the food. However, micronutrient values were more heterogeneous among the different FCT/FCDB. With the unification, the S4H FCDB obtained intermediate values considering the possible variability and also, in the case of Selenium, it retrieved values similar to those of the national FCT/FCDB.

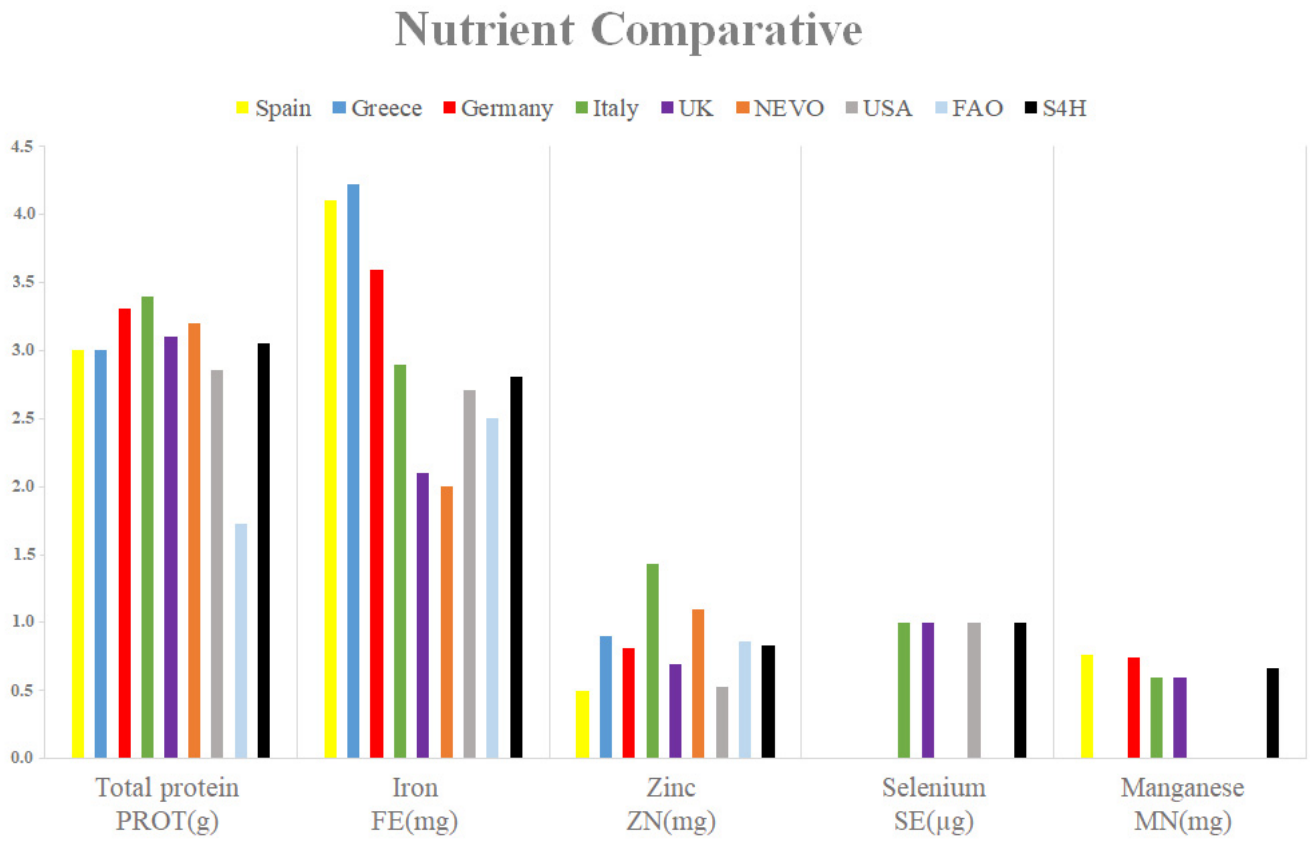

Figure 3. Protein, zinc, selenium and manganese content of the food categorized as spinach in different FCDBs and the unified values corresponding to the S4H FCDB.

Regarding recipes, tables and interrelations for energy, protein, carbohydrate, fat, sodium, calcium, riboflavin, Vitamin C, the flavonols group and (-)-epicatechin were checked for correctness. A set of recipes was selected from the database to perform manual

and automatic calculations; the results were identical in 80% of the cases, and when they were not, mismatches came from the compilers' failure to choose performance or retention factors. This problem disappears when automated.

After data validation, no errors were detected in the transformation of units because there were no systemic deviations detected in any specific nutrient or compound. In 1.8% of the foods, some nutrients showed extreme standard deviations, most likely coming from the original FCDB. In addition, 7.5% of the foods had high deviations in some nutrients; all of these values, coming from the harmonization and coding phase, were reviewed and corrected. No differences were detected when using either mean or median values, except in some specific cases, such as unified foods with more than six FCDB. Nevertheless, the median value gave estimates closer to the overall computation of the data. In addition, 4.9% of foods had macronutrients that did not meet the established quality limits; the same happened with the sum of total fats, where 2.8% presented mismatches. Therefore, 17% of the food products had some type of error. Of this percentage, about 88% could be resolved by excluding 54 food items, resulting in a total of 2648 foods. The data were transferred properly and all MySQL interrelations were checked.

4. Discussion

The aim of this study was to develop a FCDB as complete as possible in terms of food, nutrients and other compounds. This is especially important because globally there is a large nutritional data gap [33]. This trend is changing, since according to Finglas et al. [49] many countries are making efforts to create or update their FCDB. Epidemiological studies where several countries are involved, such as the EPIC study, are becoming more and more common. According to Slimani et al. [51] during the EPIC study a total of 26 nutrients for more than 550 foods per country were selected; after appropriate standardization they were used for cross-country comparisons.

Since the S4H project involves several countries, we used the EPIC study as our reference [51]. All the databases were chosen by agreement between coordinators and researchers. The three countries involved in the intervention were used as the main sources for FCT/FCDB [69–71]. The national FCT/FCDB were selected assuming that the most reliable values were available at the local level. Some nutrients were missing from one or more FCDBs. Therefore, we decided to include three more to make ours more representative of European foods [10]. These widely recognized databases were from Italy [72], the Netherlands [26,27] and the United Kingdom, since these countries had more updated versions [81,82]. Finally, four more international FCT/FCDB were included to enrich nutritional composition: the USDA FCDB, since it is widely used [31], the INFOODS/FAO—FCT/FCDB [14,30] to increase the number of nutrients and to take into account the biodiversity of some foods, and, finally, Phenol explorer 3.6 was chosen [73–75] due to the great implication that polyphenols have on diet and health [61]; this allowed for the enhancement of national FCT/CBDT through the addition of more foods and the inclusion of more than 600 bioactive compounds.. We discarded 75% of the foods since quality issues were reported in the estimates when introducing new commercial foods [83], emerging dietary components [61], fortified foods or dietary supplements [84], since these are specific to each country. Recipes and prepared dishes were also not incorporated due to the great variability of preparations in each country [85]. Recipes will be linked from another interconnected database under construction. The national FCDB will input those foods or missing values from the S4H FCDB.

4.1. Standardization and Unification

The use of food composition data from different countries needs a high level of harmonization of both food values and the nutrients that are included [48]. Data processing requires precise nomenclature and standardized methods, such as the use of ontologies or tags that allow correct classification and description [86]. Nutrients from the TEDDY study were compared between four countries. According to Uusitalo et al. [67], harmonizing

datasets before calculations generally made the results comparable, as systematic and random errors were minimized. This approach was previously used for ten European countries in the EPIC study, producing similar results [51,66].

Due to the large amount of available food items, the implementation of artificial intelligence and computational approaches is recommended [87]. Currently there are many automatic and semi-automatic tools that are extensively used to classify FCDBs [9,41,87,88]. A clear example is the ASA24 system that uses automated methods for several databases [16,87]. Another example is StandFood, a semi-automated system that obtained an overall result accuracy of 79% [41]. New techniques of natural language processing [88], machine learning, and statistical models, such as Monte Carlo simulations [12] or extraction of 'big data' [20], make the process faster than manual work [16]. However, due to the complex work, a manual *post hoc* review is always required [82,87]. After a first approach using different methodologies, manual and semi-automated harmonization and standardization work was decided to be performed in the S4H FDCB. Although human errors are still possible, this work guarantees a higher accuracy when comparing the same foods from different FCDB than automated predictions [41].

The first step was to achieve harmonization to classify foods. Durazzo et al. [37] classified foods based on different criteria. One of the classifications used is the FoodEx2 classification implemented by EFSA [40]. We selected these classification criteria due to its hierarchical nature and its widespread use. All foods were harmonized and linked between the different FCDBs. This classification provides the possibility to match foods, although full comparability is not guaranteed [2,59]. Secondly, since all nutrients and compounds have to be made comparable, they were defined in the same way, according to measurement units [51,67]. The tagnames proposed by INFOODS [43], indicating the name of the component, units and analytical method [60], have been implemented in different FCDBs around the world [45]. INFOODS tagnames allowed us to normalize variables from all databases to reference units (such as µg or mg) with faster results. Also, when unifying two nutrients, it allowed us to ensure that they were expressed in the same way and could be comparable. We modified 8.4% of the units of the tagnames to obtain a more functional FCDB. Most of the individual phenolic compounds did not have a tagname, and a new tagname was created to facilitate their integration into S4H FCDB. After standardizing both foods and nutrients, we had the opportunity to unify those foods that were categorized as identical. This would allow the inputting of those missing data and foods in the national FCDB.

Several studies claim that for research purposes in nutritional epidemiology, it is better to approximate nutrient values than to leave them as missing [51]. Not imputing data could lead to systematic underestimations of nutrient intake [18]. Although authors and institutions recognize this as a reliable method [33], others are critical, arguing that food composition changes considerably from one country to another [2]. S4H FCDB inputted the values of a weighted estimate of several FCDBs, making the values of high quality and taking into account the biodiversity of foods, thereby improving the estimations [14]. The inputting of missing values are frequent mechanisms that are performed when using FCDBs with recognized data quality [88]; typically, the data come from FCDBs from the United States, Europe, or other countries in the same region [5,12]. An example is the FCDBs from countries in sub-Saharan Africa, which import up to 88% of data about animal-source food [22]. Another example is the Middle East FCDBs, which inputted food from the United Kingdom FCDB [81]. The S4H FCDB uses an ad hoc approach to standardize the FCDB, as was done in the EPIC study [66]. This approach will make it possible to add foods or replace the value of a missing compound from other FCDBs with comparable estimated and weighted values [51,66].

During the first unification tests between foods, large standard deviations were identified in some macronutrient or micronutrient, largely coming from beverages and spices. The reason was that most of the 0 values for a compound or nutrient were not of the 'logical zero' type. Authors such as Pérez Grana or Westenbrink et al. [1,25] recommend that miss-

ing values should never be replaced by 0 and even modify the 'logical zero' values so as to avoid affecting the estimations. Before unifying the values, all the 0 values were removed. Then, by unifying the values, most of the data were homogenized, thus improving the results. The loss of the 'logical zero' values would not affect the calculations since they should remain at 0 and can be incorporated later.

On the other hand, although the mean and median were calculated, median values were chosen as the reference value after unification. Although some authors choose the mean [9,25], the median value is, in some cases, a better measure of central tendency [88], especially for extreme values from national FCDB. This ensures homogenization of the data and prevents wrong estimations. The unification allowed the inclusion of many foods and compounds. Figure 3 shows how the unification guarantees the homogenization of values. Once the values were unified and cleaned, as recommended by FAO or EuroFIR [2,33], estimated energy was recalculated using the Atwater coefficients [62].

Organizations such as FAO work with spreadsheets due to their simplicity, wide availability and familiarity to users [44,89]. Our work started out using spreadsheets, although the amount of data quickly became rather difficult to handle [89]. Therefore, the software MySQL was used, which allowed us to send and retrieve data through its interrelated tables [45,90]. This ensured traceability and quality controls, and also facilitated the relationship of S4H FCDB with the recipe tables for subsequent calculations.

4.2. Data Quality and Recipe Calculation

High quality data are essential for nutritional studies [48]. The use of the HACCP system [50] allowed us to quickly and sensibly evaluate data quality at different stages. In addition, the FAO guidelines served as a reference in the detection of critical points at any stage of the process [33]. Initial training was essential to successfully complete all the tasks, while guaranteeing the highest possible accuracy and quality.

For S4H FCDB, name verification and food description, as well as translations, were corrected thanks to the collaboration of researchers whose first language was mostly the language of the FCDB.

An FCDB should be frequently updated. For example, in the TEDDY study, the FCDB was updated at least once a year [67]. The incorporation of the original food IDs to guarantee the traceability of the food was a critical control point. Original food IDs allowed us to identify and correct errors and even to retrieve or update the information. Failures in the classification and verification of food grouping and compound labeling were detected due to outliers or manual coding by using standard deviations. Three different checking approaches were used: (i) Checking that the sum of macronutrients was within the range or the presence of implausible values detected semi-automatically in the spreadsheets; (ii) Checking for data transfer from spreadsheet to MySQL by direct verifications between versions and table relationships; and (iii) Checking the model recipe by manual verifications by compilers and automatic verifications by interconnecting the different databases. These verifications made it possible to ensure the comparability and reliability of the data.

Performing chemical analyses for all recipes and complex food matrices is not achievable. Calculations are performed indirectly using each ingredient's nutritional information [10,91]. In order to properly calculate a recipe, different parameters must be taken into account, such RF or YF. One of them is that values should not be missing, since these may lead to a biased underestimation of nutrient intake [43]. During unification and the inputting of values, this problem was solved to a large extent. The EuroFIR recipe calculation procedure was selected as a reference because it is one of the most commonly used [76]. There are several studies that use an app or software to estimate or perform interventions in nutrition and health [61,92,93]. Accordingly, the S4H FCDB will be interlinked with a recipe database. It will therefore make possible the automatic calculation of recipe values, taking into account all necessary parameters, such as edible portion, retention factors and

yield factors or even allergens. Thus, the recipes will be as adequate and representative as possible to cover the needs of the population.

4.3. Strengths and Limitations of the S4H FCDB

With the continuous expansion of food trade worldwide [10], climate change or innovation in agriculture [13], international FCDBs are essential. For this reason, S4H FCDB wants to be a reference in the creation of a unified FCDB. Much effort has been made to overcome the common drawbacks that are generally associated with the FCDB's construction. The variability in food composition (when using different FCDBs) is one of the most detected limitations [7,20]. S4H FCDB attempts to address this limitation by using the median value as the reference estimation. Additionally, there is no guarantee that national FCDB data are free of errors [2]. However, all national FCDBs are used in their own country. The unification gave us a global view of possible wrong values, allowing them to be corrected. Another limitation was represented by missing foods and nutrients from the national FCDB [47]. The S4H FCDB inputs those missing foods and compounds giving coverage and completing those values in the national FCDB. Discrepancies may exist between the tagnames proposed by FAO/INFOODS and their units [15]. However, the decisions to change units were consensual and made to improve their functionality. Moreover, inputted values from other datasets, especially dishes and recipes, did not guarantee directly related values [10,65]; for this reason, recipes and ready-to-eat products were removed. Recipes will be calculated thanks to the interconnection between the S4H FCDB and a recipe database.

The work was complex, and although the compilers were experts in nutrition, mistakes may have been made when choosing codes for harmonization [15,23]. However, the use of guidelines and data validation throughout the whole process allowed for the verification and correction of possible mistakes. The preparation of this material required a long time, and perhaps with automated methods and a subsequent exhaustive check, similar results could have been obtained [88]. There may have been failures during the translation of some foods [47], especially regional foods, although if no reliable translation was found, foods were discarded. Even so, our results are encouraging. Misspellings and translation mistakes were detected while manually identifying and classifying. Thus, one of the limitations may have actually been a strength.

In most nutritional epidemiological studies, results are similarly interpreted regardless of how they make estimations or which FCDB is used. This generates an unrealistic relationship of nutrient intakes and their impact on health [94]. An increasingly large number of epidemiological studies attempt to make their data comparable [51,67,95,96]. One of the strengths of the S4H FCDB is that with unified values, data from different countries could be compared, as it would take biodiversity and different parameters affecting the same kind of food into account. Another option is to use national FCDB data and only fill in the missing nutrients and compounds to avoid underestimations [6,18]. Organizations such as EUROFIR have the potential to create a standardized FCDB which should be free to use [48]. EFSA already has a tool as a first step towards the unification of nutrients [97]. The S4H FCDB is one of the most comprehensive FCDB regarding the number of foods and nutrients, being able to collect more than 800 compounds from each foodstuff. Thus, to date it is only surpassed by the https://foodb.ca (accessed on 27 October 2021) project supported by the Canadian Institutes of Health Research and by The Metabolomics Innovation Centre. This Database includes not only nutritional information, but also a large amount of bioactive compounds [24,55,98]. However, it must be noted that the S4H FCDB uses different FCT/FCDB, giving much more homogeneous and comparable nutritional values.

4.4. S4H FCDB's Future Perspective

The S4H FCDB consists of interlinked tables that make a complete nutritional information system. S4H FCDB not only allows accurate calculations, but also provides the user with information on different aspects integrated in the personalized nutrition system. The purpose of the S4H FCDB was for it to be used in epidemiological studies, in particular precision nutritional studies. This S4H FCDB will be connected to an app that will be used during the nutritional intervention of the project. The study aims to generate personalized nutritional recommendations to different populations, more specifically adults and children [56]. The app derived from the S4H Project will be an automated diet evaluator and generator used from smartphones. A set of more than 10,000 recipes from all countries is expected to be available. All recipes will be implemented in a mobile app for future nutritional intervention. An example is depicted in Figure 4. Other similar apps have a smaller number of foods and were developed from a smaller number of food data sources [82].

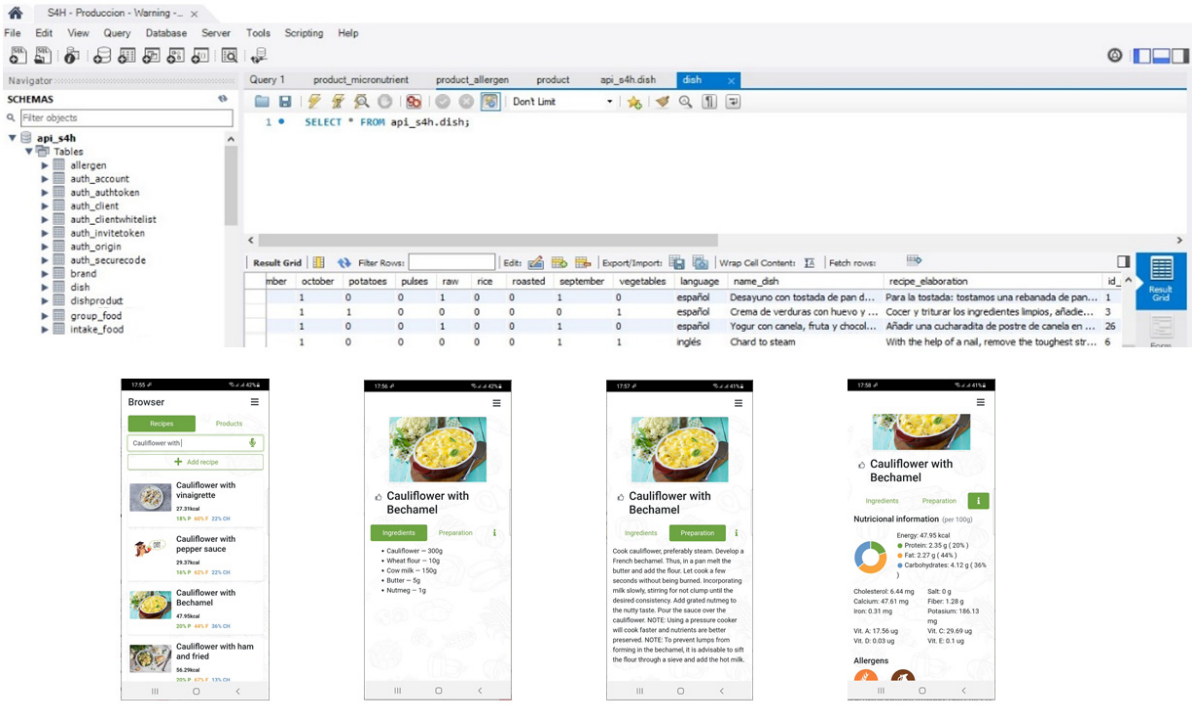

Figure 4. Shows the relationship between MySQL tables and how the recipe is generated in the APP.

The S4H FCDB will be able to connect to other tools. One of the milestones of personalized nutrition is to understand the health level of the gut microbiota of a given patient. The S4H FCDB generated data will also be completed with the use of AGREDA [99], an extended reconstruction of diet metabolism by the human gut microbiota. The S4H FCDB will also introduce commercial products, incorporating allergens and different scores as used in Open Food Facts [16]. These products and fast foods from the different countries of the project will make S4H FCDB more comprehensive and representative [38]. The data will also be updated periodically to avoid obsolescence [5], which will be possible thanks to traceability.

In the future it is expected that the S4H FCDB will be extended by implementing toxic substances, such as food processing contaminants, as few FCDBs contain these components [15,88,100–102]. Due to the importance of climate change in nutrition, sustainability parameters and different markers of climate change would be an added value to be included [13]. Finally, in order to identify food-disease associations [55], food biomarkers could be introduced by linking them to FOBI (Food-Biomarker Ontology) [36], or extending

the compounds related to https://foodb.ca (accessed on 27 October 2021) or other big data sources [24,98].

5. Conclusions

S4H FCDB was built through a huge scientific work to collect and harmonize all the nutritional data. S4H FCDB is one of the most comprehensive FCDB with more than ten FCDBs used, which is one of its main unique characteristics. This food database is comparable to that used in other relevant studies, such as EPIC. A large number of harmonized foods (over 2000) and more than 800 nutrients and bioactive compounds (such as polyphenols) have been included, the inclusion of such a large number of bioactive compounds being another unique strength of the paper. S4H FCDB attempts to mitigate the usual limitations, such as variability in food composition, errors, and missing values in the national FCT/FCDB databases. Trained personnel following the guidelines of official agencies were able to homogenize the information. This made it possible to unify foods, their nutrients and bioactive compounds among the FCT/FCDBs using the median value as the reference value. The values obtained were less extreme and made it possible to complete the national FCT/FCDB. The S4H FCDB has many perspectives, not only the implementation in nutritional studies through an application. But it is also capable of being part of other tools and has the versatility to be continuously enhanced with much more information. Thus, S4H FCDB becomes a solid and indispensable tool to approach the age of personalized nutrition.

Supplementary Materials: The following are available online at http://www.mdpi.com/xxx/s1, Table S1: INFOODS Tagnames used in S4H FCDB, Code S2: Code used in the S4H FCDB unification stage. Table S3: Figure with the absence or presence of different compounds and nutrients in the FCDB. All compounds are included and expressed with FAO/INFOODS names.

Author Contributions: Conceptualization, V.G.-V. and J.Á.R.-H.; Methodology, D.H.-N., B.N.-P. and S.P.d.l.C.; Software, B.O.-V. and V.G.-V.; Validation, D.H.-N. and S.P.-B.; Formal analysis, D.H.-N., S.P.-B., B.N.-P. and B.O.-V.; Investigation, A.F. and S.P.d.l.C.; Resources, S.P.-B., F.L. and V.G.-V.; Data curation, D.H.-N. and B.O.-V.; Writing—Original draft preparation, D.H.-N. and S.P.-B.; Writing—Review and editing, A.F., F.L., K.N.P., S.P.-B. and J.Á.R.-H.; Supervision, V.G.-V. and J.Á.R.-H.; Funding acquisition, J.Á.R.-H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Research Commission (Research Executive Agency) under the research project Stance4Health (Grant Contract No. 816303) to José Á. Rufián-Henares. This work was supported by the "Plan propio de Investigación y Transferencia" of the University of Granada under the programs "Intensificación de la Investigación, modalidad B", granted to José Á. Rufián-Henares.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The S4H FCDB is not freely available at the present moment, since it will be used along the research project. Those researchers aiming at using the S4H FCDB should request a license to José A. Rufián-Henares, who will transfer the request to the relevant S4H committee.

Acknowledgments: This work will be a part of the doctoral thesis of Daniel Hinojosa-Nogueira conducted within the context of the 'Program of Nutrition and Food Sciences' at the University of Granada.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pérez Grana, R. Exactitud de las tablas de composición de alimentos en la determinación de nutrientes. *Sanid. Mil.* **2013**, *69*, 102–111. [CrossRef]
2. Westenbrink, S.; Kadvan, A.; Roe, M.; Koroušić Seljak, B.; Mantur-Vierendeel, A.; Finglas, P. 12th IFDC 2017 special issue—Evaluation of harmonized eurofir documentation for macronutrient values in 26 European food composition databases. *J. Food Compos. Anal.* **2019**, *80*, 40–50. [CrossRef]
3. Greenfield, H.; Southgate, D.A. *Food Composition Data: Production, Management, and Use*; Food and Agriculture Organization: Roma, Italy, 2003; ISBN 92-5-104949-1.
4. Charrondiere, U.; Stadlmayr, B.; Haytowitz, D.; Oseredczuk, M.; Ireland, J.; Wolmarans, P.; Rittenschober, D.; Selley, B.; Puwastien, P.; Reykdal, Ó. *FAO/INFOODS Guidelines for Checking Food Composition Data Prior to the Publication of a User Table/Database Version 1.0*; FAO: Roma, Italy, 2012.
5. Wafa, S.; Colaiezzi, B.; Some, J.; Bell, W.; Prata, C.; Perlick, H.; Rogers, B.; Coates, J. INDDEX24: An innovative global dietary assessment platform to scale up the availability, access, and use of dietary data. *Curr. Dev. Nutr.* **2020**, *4*, 1180. [CrossRef]
6. Ene-Obong, H.; Schönfeldt, H.C.; Campaore, E.; Kimani, A.; Mwaisaka, R.; Vincent, A.; El Ati, J.; Kouebou, P.; Presser, K.; Finglas, P.; et al. Importance and use of reliable food composition data generation by nutrition/dietetic professionals towards solving africa's nutrition problem: Constraints and the role of FAO/INFOODS/AFROFOODS and other stakeholders in future initiatives. *Proc. Nutr. Soc.* **2019**, *78*, 496–505. [CrossRef]
7. Lupiañez-Barbero, A.; González Blanco, C.; de Leiva Hidalgo, A. Tablas y bases de datos de composición de alimentos españolas: Necesidad de un referente para los profesionales de la salud. *Endocrinol. Diabetes Nutr.* **2018**, *65*, 361–373. [CrossRef]
8. Emmett, P.M.; Jones, L.R.; Northstone, K.; Pounis, G.; Taylor, C.M. Collection and management of dietary data. In *Analysis in Nutrition Research*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 43–73. ISBN 978-0-12-814556-2.
9. Ispirova, G.; Eftimov, T.; Koroušić Seljak, B.; Korošec, P. Mapping food composition data from various data sources to a domain-specific ontology. In Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, Funchal, Portugal, 1–3 November 2017; SCITEPRESS—Science and Technology Publications: Funchal, Portugal, 2017; pp. 203–210.
10. Marconi, S.; Durazzo, A.; Camilli, E.; Lisciani, S.; Gabrielli, P.; Aguzzi, A.; Gambelli, L.; Lucarini, M.; Marletta, L. Food composition databases: Considerations about complex food matrices. *Foods* **2018**, *7*, 2. [CrossRef] [PubMed]
11. Ortiz, J.; Astudillo, G.; Castro, M.; Castro, C.; Astudillo, S.; Donoso, S. 12th IFDC 2017 special issue—Seasonal variations in nutrient composition of plant-based foods produced at the southern highlands of ecuador. *J. Food Compos. Anal.* **2019**, *83*, 103284. [CrossRef]
12. Smith, M.R.; Micha, R.; Golden, C.D.; Mozaffarian, D.; Myers, S.S. Global expanded nutrient supply (genus) model: A new method for estimating the global dietary supply of nutrients. *PLoS ONE* **2016**, *11*, e0146976. [CrossRef] [PubMed]
13. Drabo, A. Climate change mitigation and agricultural development models: Primary commodity exports or local consumption production? *Ecol. Econ.* **2017**, *137*, 110–125. [CrossRef]
14. Charrondière, U.R.; Stadlmayr, B.; Rittenschober, D.; Mouille, B.; Nilsson, E.; Medhammar, E.; Olango, T.; Eisenwagen, S.; Persijn, D.; Ebanks, K. FAO/INFOODS food composition database for biodiversity. *Food Chem.* **2013**, *140*, 408–412. [CrossRef] [PubMed]
15. Kapsokefalou, M.; Roe, M.; Turrini, A.; Costa, H.S.; Martinez-Victoria, E.; Marletta, L.; Berry, R.; Finglas, P. Food composition at present: New challenges. *Nutrients* **2019**, *11*, 1714. [CrossRef]
16. Zeb, A.; Soininen, J.-P.; Sozer, N. Data harmonisation as a key to enable digitalisation of the food sector: A review. *Food Bioprod. Process.* **2021**, *127*, 360–370. [CrossRef]
17. Lauria, F.; Russo, M.D.; Formisano, A.; De Henauw, S.; Hebestreit, A.; Hunsberger, M.; Kroghe, V.; Intemann, T.; Lissnerd, L.; Molnar, D.; et al. Ultra-processed foods consumption and diet quality of European children, adolescents and adults: Results from the I.Family study. *Nutr. Metab. Cardiovasc. Dis.* **2021**, *31*, 3031–3043. [CrossRef] [PubMed]
18. Ocké, M.C.; Westenbrink, S.; van Rossum, C.T.M.; Temme, E.H.M.; van der Vossen-Wijmenga, W.; Verkaik-Kloosterman, J. The essential role of food composition databases for public health nutrition—Experiences from the Netherlands. *J. Food Compos. Anal.* **2021**, *101*, 103967. [CrossRef]
19. Egan, M.B.; Fragodt, A.; Raats, M.M.; Hodgkins, C.; Lumbers, M. The importance of harmonizing food composition data across Europe. *Eur. J. Clin. Nutr.* **2007**, *61*, 813–821. [CrossRef] [PubMed]
20. Traka, M.H.; Plumb, J.; Berry, R.; Pinchen, H.; Finglas, P.M. Maintaining and updating food composition datasets for multiple users and novel technologies: Current challenges from a UK perspective. *Nutr. Bull.* **2020**, *45*, 230–240. [CrossRef]
21. Delgado-Andrade, C.; Rufián-Henares, J.A.; Morales, F.J. Lysine availability is diminished in commercial fibre-enriched breakfast cereals. *Food Chem.* **2007**, *100*, 725–731. [CrossRef]
22. Grande, F.; Vincent, A. The importance of food composition data for estimating micronutrient intake: What do we know now and into the future? In *Global Landscape of Nutrition Challenges in Infants and Children*; Michaelsen, K.F., Neufeld, L.M., Prentice, A.M., Eds.; Nestlé Nutrition Institute Workshop Series; S. Karger AG: Basel, Switzerland, 2020; Volume 93, pp. 39–50. ISBN 978-3-318-06648-7.
23. Micha, R.; Coates, J.; Leclercq, C.; Charrondiere, U.R.; Mozaffarian, D. Global dietary surveillance: Data gaps and challenges. *Food Nutr. Bull.* **2018**, *39*, 175–205. [CrossRef]

24. Barabási, A.-L.; Menichetti, G.; Loscalzo, J. The unmapped chemical complexity of our diet. *Nat. Food* **2020**, *1*, 33–37. [CrossRef]
25. Westenbrink, S.; Presser, K.; Roe, M.; Ireland, J.; Finglas, P. Documentation of aggregated/compiled values in food composition databases; EuroFIR default to improve harmonization. *J. Food Compos. Anal.* **2021**, *101*, 103968. [CrossRef]
26. Roe, M.; Pinchen, H.; Church, S.; Finglas, P. McC ance and widdowson's the composition of foods seventh summary edition and updated composition of foods integrated dataset. *Nutr. Bull.* **2015**, *40*, 36–39. [CrossRef]
27. Westenbrink, S.; Jansen-van der Vliet, M.; van Rossum, C. Updated folate data in the dutch food composition database and implications for intake estimates. *Food Nutr. Res.* **2012**, *56*, 5449. [CrossRef]
28. Nielsen, C.W. *Fødevaredatabanken August 2019*; DTU: Lyngby, Denmark, 2019.
29. Delgado, A.; Issaoui, M.; Vieira, M.C.; Saraiva de Carvalho, I.; Fardet, A. Food composition databases: Does it matter to human health? *Nutrients* **2021**, *13*, 2816. [CrossRef] [PubMed]
30. Charrondière, U.; Rittenschober, D.; Grande, F.; Nowak, V.; Stadlmayr, B. *FAO/INFOODS Analytical Food Composition Database Version 2.0*; FAO: Roma, Italy, 2017.
31. Montville, J.B.; Ahuja, J.K.; Martin, C.L.; Heendeniya, K.Y.; Omolewa-Tomobi, G.; Steinfeldt, L.C.; Anand, J.; Adler, M.E.; LaComb, R.P.; Moshfegh, A. USDA food and nutrient database for dietary studies (FNDDS), 5.0. *Procedia Food Sci.* **2013**, *2*, 99–112. [CrossRef]
32. Murphy, S.P.; Charrondiere, U.R.; Burlingame, B. Thirty years of progress in harmonizing and compiling food data as a result of the establishment of INFOODS. *Food Chem.* **2016**, *193*, 2–5. [CrossRef] [PubMed]
33. Charrondiere, U.R.; Rittenschober, D.; Nowak, V.; Stadlmayr, B.; Wijesinha-Bettoni, R.; Haytowitz, D. Improving food composition data quality: Three new FAO/INFOODS guidelines on conversions, data evaluation and food matching. *Food Chem.* **2016**, *193*, 75–81. [CrossRef] [PubMed]
34. Apekey, T.A.; Copeman, J.; Kime, N.H.; Tashani, O.A.; Kittaneh, M.; Walsh, D.; Maynard, M.J. Nutrient composition of popularly consumed african and caribbean foods in the UK. *Foods* **2019**, *8*, 500. [CrossRef]
35. Hulshof, P.; Doets, E.; Seyha, S.; Bunthang, T.; Vonglokham, M.; Kounnavong, S.; Famida, U.; Muslimatun, S.; Santika, O.; Prihatini, S.; et al. Food composition tables in southeast asia: The contribution of the SMILING project. *Matern. Child Health J.* **2019**, *23*, 46–54. [CrossRef]
36. Castellano-Escuder, P.; González-Domínguez, R.; Wishart, D.S.; Andrés-Lacueva, C.; Sánchez-Pla, A. FOBI: An ontology to represent food intake data and associate it with metabolomic data. *Database* **2020**, *2020*, baaa033. [CrossRef]
37. Durazzo, A.; Camilli, E.; D'Addezio, L.; Sette, S.; Marconi, S.; Piccinelli, R.; Le Donne, C.; Turrini, A.; Marletta, L. Italian composite dishes: Description and classification by LanguaL^{TM} and FoodEx2. *Eur. Food Res. Technol.* **2020**, *246*, 287–295. [CrossRef]
38. Katidi, A.; Vlassopoulos, A.; Kapsokefalou, M. Development of the hellenic food thesaurus (HelTH), a branded food composition database: Aims, design and preliminary findings. *Food Chem.* **2021**, *347*, 129010. [CrossRef] [PubMed]
39. Møller, A.; Ireland, J. LanguaL^{TM} 2017—The LanguaL^{TM} thesaurus technical report. *Dan. Food Inform.* **2018**. [CrossRef]
40. European Food Safety Authority (EFSA). *The Food Classification and Description System FoodEx 2 (Revision 2)*; EFSA: Palma, Italy, 2015; Volume 12. [CrossRef]
41. Eftimov, T.; Korošec, P.; Koroušić Seljak, B. StandFood: Standardization of foods using a semi-automatic system for classifying and describing foods according to FoodEx2. *Nutrients* **2017**, *9*, 542. [CrossRef] [PubMed]
42. European Food Safety Authority (EFSA); Ioannidou, S. *EFSA FoodEx2 Interpreting and Checking Tool User Guide*; EFSA: Palma, Italy, 2019; Volume 16. [CrossRef]
43. Charrondière, U.R.; Rittenschober, D.; Nowak, V.; Wijesinha-Bettoni, R.; Stadlmayr, B.; Haytowitz, D.; Persijn, D. *FAO/INFOODS Guidelines for Converting Units, Denominators and Expressions, Version 1.0*; FAO: Roma, Italy, 2012.
44. Charrondiere, U.R.; Burlingame, B. Report on the FAO/INFOODS compilation tool: A simple system to manage food composition data. *J. Food Compos. Anal.* **2011**, *24*, 711–715. [CrossRef]
45. Schlotke, F.; Becker, W.; Ireland, J.; Møller, A.; Ovaskainen, M.-L.; Monspart, J.; Unwin, I. EUROFOODS Recommendations for Food Composition Database Management and Data Interchange. *J. Food Compos. Anal.* **2000**, *13*, 709–744. [CrossRef]
46. Merchant, A.T.; Dehghan, M. Food Composition Database Development for between Country Comparisons. *Nutr. J.* **2006**, *5*, 2. [CrossRef] [PubMed]
47. Finglas, P.M.; Berry, R.; Astley, S. Assessing and Improving the Quality of Food Composition Databases for Nutrition and Health Applications in Europe: The Contribution of EuroFIR. *Adv. Nutr.* **2014**, *5*, 608S–614S. [CrossRef]
48. Westenbrink, S.; Roe, M.; Oseredczuk, M.; Castanheira, I.; Finglas, P. EuroFIR Quality Approach for Managing Food Composition Data; Where Are We in 2014? *Food Chem.* **2016**, *193*, 69–74. [CrossRef] [PubMed]
49. Finglas, P.; Roe, M.; Pinchen, H.; Astley, S. The Contribution of Food Composition Resources to Nutrition Science Methodology. *Nutr. Bull.* **2017**, *42*, 198–206. [CrossRef]
50. Westenbrink, S.; Oseredczuk, M.; Castanheira, I.; Roe, M. Food Composition Databases: The EuroFIR Approach to Develop Tools to Assure the Quality of the Data Compilation Process. *Food Chem.* **2009**, *113*, 759–767. [CrossRef]
51. Slimani, N.; Deharveng, G.; Unwin, I.; Southgate, D.A.T.; Vignat, J.; Skeie, G.; Salvini, S.; Parpinel, M.; Møller, A.; Ireland, J.; et al. The EPIC Nutrient Database Project (ENDB): A First Attempt to Standardize Nutrient Databases across the 10 European Countries Participating in the EPIC Study. *Eur. J. Clin. Nutr.* **2007**, *61*, 1037–1056. [CrossRef] [PubMed]
52. Longvah, T.; Anantan, I.; Bhaskarachary, K.; Venkaiah, K.; Longvah, T. *Indian Food Composition Tables*; National Institute of Nutrition, Indian Council of Medical Research: Hyderabad, Indian, 2017; ISBN 93-5267-677-7.

53. Park, S.-H.; Kim, S.-N.; Lee, S.H.; Choe, J.-S.; Choi, Y. Development of 9th Revision Korean Food Composition Table and Its Major Changes. *Korean J. Community Nutr.* **2018**, *23*, 352–365. [CrossRef]
54. Coelho, K.C.; Giuntini, B.G.; Grande, F.; Dias, J.S.; Purgatto, E.; Franco, B.D.G.M.; Lajolo, F.M.; Menezes, E.W. Brazilian Food Composition Table (TBCA): Development and functionalities of the online version. *J. Food Comp. Anal.* **2019**, *84*, 103287. [CrossRef]
55. Naveja, J.J.; Rico-Hidalgo, M.P.; Medina-Franco, J.L. Analysis of a Large Food Chemical Database: Chemical Space, Diversity, and Complexity. *F1000Research* **2018**, *7*, 993. [CrossRef] [PubMed]
56. Pérez-Burillo, S.; Hinojosa-Nogueira, D.; Pastoriza, S.; Rufián-Henares, J.A. Nutrición personalizada inteligente. *Alimentaria* **2019**, *500*, 25–29.
57. Charrondiere, U.R.; Rittenschober, D.; Nowak, V.; Nicodemi, C.; Bruggeling, P.; Petracchi, C. FAO/INFOODS e-Learning Course on Food Composition Data. *Food Chem.* **2016**, *193*, 6–11. [CrossRef]
58. Castanheira, I.; André, C.; Oseredczuk, M.; Ireland, J.; Owen, L.; Robb, P.; Earnshaw, A.; Calhau, M.A. Improving Data Quality in Food Composition Databanks: A EuroFIR Contribution. *Accredit. Qual. Assur.* **2007**, *12*, 117–125. [CrossRef]
59. Dahdouh, S.; Grande, F.; Espinosa, S.N.; Vincent, A.; Gibson, R.; Bailey, K.; King, J.; Rittenschober, D.; Charrondière, U.R. Development of the FAO/INFOODS/IZINCG Global Food Composition Database for Phytate. *J. Food Compos. Anal.* **2019**, *78*, 42–48. [CrossRef]
60. Klensin, J.C.; Feskanich, D.; Lin, V.; Truswell, A.S.; Southgate, D.A. *Identification of Food Components for INFOODS Data Interchange*; United Nations University: Tokyo, Japan, 1989; ISBN 92-808-0734-X.
61. Knaze, V.; Rothwell, J.A.; Zamora-Ros, R.; Moskal, A.; Kyrø, C.; Jakszyn, P.; Skeie, G.; Weiderpass, E.; Santucci de Magistris, M.; Agnoli, C.; et al. A New Food-Composition Database for 437 Polyphenols in 19,899 Raw and Prepared Foods Used to Estimate Polyphenol Intakes in Adults from 10 European Countries. *Am. J. Clin. Nutr.* **2018**, *108*, 517–524. [CrossRef] [PubMed]
62. Maclean, W.; Harnly, J.; Chen, J.; Chevassus-Agnes, S.; Gilani, G.; Livesey, G.; Warwick, P. Food energy—Methods of analysis and conversion factors. In *Food and Agriculture Organization of the United Nations Technical Workshop Report*; Agricultural Research Service, U.S. Department of Agriculture: Beltsville, MD, USA, 2003; Volume 77, pp. 8–9.
63. Martinez-Victoria, E.; de Victoria, I.M.; Martinez-Burgos, M.A. Intake of Energy and Nutrients; Harmonization of Food Composition Databases. *Nutr. Hosp.* **2015**, *31*, 168–176.
64. Rand, W.M.; Pennington, J.A.; Murphy, S.P.; Klensin, J.C. *Compiling Data for Food Composition Data Bases*; United Nations University Press: Tokyo, Japan, 1991; ISBN 92-808-0772-2.
65. Roe, M.A.; Bell, S.; Oseredczuk, M.; Christensen, T.; Westenbrink, S.; Pakkala, H.; Presser, K.; Finglas, P.M. *Updated Food Composition Database for Nutrient Intake*; EFSA: Palma, Italy, 2013; Volume 10. [CrossRef]
66. Slimani, N.; Ruth Charrondière, U.; van Staveren, W.; Riboli, E. Standardization of Food Composition Databases for the European Prospective Investigation into Cancer and Nutrition (EPIC): General Theoretical Concept. *J. Food Compos. Anal.* **2000**, *13*, 567–584. [CrossRef]
67. Uusitalo, U.; Kronberg-Kippilä, C.; Aronsson, C.A.; Schakel, S.; Schoen, S.; Mattisson, I.; Reinivuo, H.; Silvis, K.; Sichert-Hellert, W.; Stevens, M.; et al. Food Composition Database Harmonization for Between-Country Comparisons of Nutrient Data in the TEDDY Study. *J. Food Compos. Anal.* **2011**, *24*, 494–505. [CrossRef]
68. Aouachria, N.B.; Bakker, M.; König, J.; Leblanc, J.-C.; Lindtner, O.; Tlustos, C.; Arcella, D.; Tard, A.; Tasiopoulou, S. European Food Safety Authority Use of the EFSA Comprehensive European Food Consumption Database in Exposure Assessment. *EFSA J.* **2011**, *9*, 2097. [CrossRef]
69. Gestión de Salud y Nutrición, S.L. I-Diet Food Composition Database, Updated from Original Version of G. Martín Peña FCD. Asturias, Spain, 2019.
70. Hartmann, B.M.; Bell, S.; Vásquez-Caicedo, A.L.; Götz, A.; Erhardt, J.; Brombach, C. *Der Bundeslebensmittelschlüssel*; German Nutrient Database; Federal Research Centre for Nutrition and Food (BfEL): Karlsruhe, Germany, 2005.
71. Trichopoulou, A.; Georga, K. *Composition Tables of Foods and Greek Dishes*, 3rd ed.; Parisianou Publications: Athens, Greece, 2004.
72. Gnagnarella, P.; Salvini, S.; Parpinel, M. Food Composition Database for Epidemiological Studies in Italy. Available online: http://www.bda-ieo.it/ (accessed on 31 January 2020).
73. Neveu, V.; Perez-Jimenez, J.; Vos, F.; Crespy, V.; du Chaffaut, L.; Mennen, L.; Knox, C.; Eisner, R.; Cruz, J.; Wishart, D.; et al. Phenol-Explorer: An Online Comprehensive Database on Polyphenol Contents in Foods. *Database* **2010**, *2010*, bap024. [CrossRef]
74. Rothwell, J.A.; Urpi-Sarda, M.; Boto-Ordonez, M.; Knox, C.; Llorach, R.; Eisner, R.; Cruz, J.; Neveu, V.; Wishart, D.; Manach, C.; et al. Phenol-Explorer 2.0: A Major Update of the Phenol-Explorer Database Integrating Data on Polyphenol Metabolism and Pharmacokinetics in Humans and Experimental Animals. *Database* **2012**, *2012*, bas031. [CrossRef]
75. Rothwell, J.A.; Perez-Jimenez, J.; Neveu, V.; Medina-Remon, A.; M'Hiri, N.; Garcia-Lobato, P.; Manach, C.; Knox, C.; Eisner, R.; Wishart, D.S.; et al. Phenol-Explorer 3.0: A Major Update of the Phenol-Explorer Database to Incorporate Data on the Effects of Food Processing on Polyphenol Content. *Database* **2013**, *2013*, bat070. [CrossRef] [PubMed]
76. Reinivuo, H.; Bell, S.; Ovaskainen, M.-L. Harmonisation of Recipe Calculation Procedures in European Food Composition Databases. *J. Food Compos. Anal.* **2009**, *22*, 410–413. [CrossRef]
77. Vásquez-Caicedo, A.; Bell, S.; Hartmann, B. *Report on Collection of Rules on Use of Recipe Calculation Procedures Including the Use of Yield and Retention Factors for Imputing Nutrient Values for Composite Foods*; European Food Information Resource Network: Brussels, Belgium, 2008.

78. Bell, S.; Becker, W.; Vásquez-Caicedo, A.; Hartmann, B.; Møller, A.; Butriss, J. *Report on Nutrient Losses and Gains Factors Used in European Food Composition Databases*; European Food Information Resource Network: Brussels, Belgium, 2006.
79. Bognár, A. *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, 2002; ISBN 0933-5463.
80. Charrondiere, U.R.; Sivieri, A.; Burlingame, B. Differences in nutrient values of recipes due to different calculation methods and sets of nutrient retention factors. In *Quality Food Composition Data: Key for Health and Trade, Proceedings of the 8th International Food Data Conference, Bangkok, Thailand, 1–3 October 2009*; Judprasong, K., Puwastien, P., Jittinandana, S., Eds.; Bangkok, Thailand, 2009.
81. Bawajeeh, A.; Kalendar, S.; Scarpa, G.; Hancock, N.; Beer, S.; Gibson, L.; Williams, G.; Dashti, B.; Albar, S.; Ensaff, H.; et al. Development of an Arabic Food Composition Database for Use in an Arabic Online Dietary Assessment Tool (Myfood24). *J. Food Compos. Anal.* **2021**, *102*, 104047. [CrossRef]
82. Carter, M.; Hancock, N.; Albar, S.; Brown, H.; Greenwood, D.; Hardie, L.; Frost, G.; Wark, P.; Cade, J. Development of a New Branded UK Food Composition Database for an Online Dietary Assessment Tool. *Nutrients* **2016**, *8*, 480. [CrossRef]
83. Nowak, V.; Persijn, D.; Rittenschober, D.; Charrondiere, U.R. Review of Food Composition Data for Edible Insects. *Food Chem.* **2016**, *193*, 39–46. [CrossRef]
84. Yoo, Y.M.; Atkin, R.; Pachón, H. Development of a Food Composition Table to Analyze Senegalese Food Expenditure Data. *Afr. J. Food Agric. Nutr. Dev.* **2019**, *19*, 1–21.
85. Durazzo, A.; Camilli, E.; Marconi, S.; Lisciani, S.; Gabrielli, P.; Gambelli, L.; Aguzzi, A.; Lucarini, M.; Kiefer, J.; Marletta, L. Nutritional Composition and Dietary Intake of Composite Dishes Traditionally Consumed in Italy. *J. Food Compos. Anal.* **2019**, *77*, 115–124. [CrossRef]
86. Ireland, J.D.; Møller, A. Review of International Food Classification and Description. *J. Food Compos. Anal.* **2000**, *13*, 529–538. [CrossRef]
87. Chin, E.L.; Simmons, G.; Bouzid, Y.Y.; Kan, A.; Burnett, D.J.; Tagkopoulos, I.; Lemay, D.G. Nutrient Estimation from 24-Hour Food Recalls Using Machine Learning and Database Mapping: A Case Study with Lactose. *Nutrients* **2019**, *11*, 3045. [CrossRef]
88. Ispirova, G.; Eftimov, T.; Korošec, P.; Koroušić Seljak, B. MIGHT: Statistical Methodology for Missing-Data Imputation in Food Composition Databases. *Appl. Sci.* **2019**, *9*, 4111. [CrossRef]
89. Presser, K.; Weber, D.; Norrie, M. FoodCASE: A System to Manage Food Composition, Consumption and TDS Data. *Food Chem.* **2018**, *238*, 166–172. [CrossRef] [PubMed]
90. Møller, A.; Christensen, T. *Danish Food Information: EuroFIR Web Services-EuroFIR Food Data Transport Package, Version 1.3*; DTU: Lyngby, Denmark, 2008.
91. Durazzo, A.; Camilli, E.; D'Addezio, L.; Le Donne, C.; Ferrari, M.; Marconi, S.; Marletta, L.; Mistura, L.; Piccinelli, R.; Scalvedi, M.L.; et al. Food Groups and Individual Foods: Nutritional Attributes and Dietary Importance. In *Reference Module in Food Science*; Elsevier: Amsterdam, The Netherlands, 2018; p. B9780081005965214000. ISBN 978-0-08-100596-5.
92. Jimoh, F.; Lund, E.K.; Harvey, L.J.; Frost, C.; Lay, W.J.; Roe, M.A.; Berry, R.; Finglas, P.M. Comparing Diet and Exercise Monitoring Using Smartphone App and Paper Diary: A Two-Phase Intervention Study. *JMIR mHealth uHealth* **2018**, *6*, e17. [CrossRef] [PubMed]
93. Kwon, D.Y. Personalized Diet Oriented by Artificial Intelligence and Ethnic Foods. *J. Ethn. Foods* **2020**, *7*, 10. [CrossRef]
94. Cespedes, E.M.; Hu, F.B. Dietary Patterns: From Nutritional Epidemiologic Analysis to National Guidelines. *Am. J. Clin. Nutr.* **2015**, *101*, 899–900. [CrossRef] [PubMed]
95. Ocké, M.; de Boer, E.; Brants, H.; van der Laan, J.; Niekerk, M.; van Rossum, C.; Temme, L.; Freisling, H.; Nicolas, G.; Casagrande, C.; et al. *PANCAKE—Pilot Study for the Assessment of Nutrient Intake and Food Consumption Among Kids in Europe*; EFSA: Palma, Italy, 2012; Volume 9. [CrossRef]
96. Wark, P.A.; Hardie, L.J.; Frost, G.S.; Alwan, N.A.; Carter, M.; Elliott, P.; Ford, H.E.; Hancock, N.; Morris, M.A.; Mulla, U.Z.; et al. Validity of an Online 24-h Recall Tool (Myfood24) for Dietary Assessment in Population Studies: Comparison with Biomarkers and Standard Interviews. *BMC Med.* **2018**, *16*, 136. [CrossRef] [PubMed]
97. European Food Safety Authority (EFSA). *Dietary Reference Values for Nutrients Summary Report*; EFSA: Palma, Italy, 2017.
98. Harrington, R.A.; Adhikari, V.; Rayner, M.; Scarborough, P. Nutrient Composition Databases in the Age of Big Data: FoodDB, a Comprehensive, Real-Time Database Infrastructure. *BMJ Open* **2019**, *9*, e026652. [CrossRef] [PubMed]
99. Blasco, T.; Pérez-Burillo, S.; Balzerani, F.; Hinojosa-Nogueira, D.; Lerma-Aguilera, A.; Pastoriza, S.; Cendoya, X.; Rubio, Á.; Gosalbes, M.J.; Jiménez-Hernández, N.; et al. An Extended Reconstruction of Human Gut Microbiota Metabolism of Dietary Compounds. *Nat. Commun.* **2021**, *12*, 4728. [CrossRef] [PubMed]
100. Kovarich, S.; Ceriani, L.; Ciacci, A.; Baldin, R.; Perez Miguel, M.; Gibin, D.; Carnesecchi, E.; Roncaglioni, A.; Mostrag, A.; Tarkhov, A.; et al. *OpenFoodTox: EFSA's Chemical Hazards Database*; EFSA: Palma, Italy, 2020.
101. Rufian-Henares, J.A.; De la Cueva, S.P. Assessment of Hydroxymethylfurfural Intake in the Spanish Diet. *Food Addit. Contam.* **2008**, *25*, 1306–1312. [CrossRef]
102. Rufian-Henares, J.A.; Garcia-Villanova, B.; Guerra-Hernández, E. Occurrence of furosine and hydroxymethylfurfural as markers of thermal damage in dehydrated vegetables. *Eur. Food Res. Technol.* **2008**, *228*, 249–256. [CrossRef]