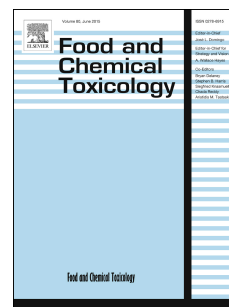


# Accepted Manuscript

Food groups for allergen risk assessment: Combining food consumption data from different countries in Europe

Sophie Birot, Charlotte B. Madsen, Astrid G. Kruizinga, Amélie Crépet, Tue Christensen, Per B. Brockhoff



PII: S0278-6915(18)30342-9

DOI: [10.1016/j.fct.2018.05.042](https://doi.org/10.1016/j.fct.2018.05.042)

Reference: FCT 9797

To appear in: *Food and Chemical Toxicology*

Received Date: 21 February 2018

Revised Date: 15 May 2018

Accepted Date: 17 May 2018

Please cite this article as: Birot, S., Madsen, C.B., Kruizinga, A.G., Crépet, Amé., Christensen, T., Brockhoff, P.B., Food groups for allergen risk assessment: Combining food consumption data from different countries in Europe, *Food and Chemical Toxicology* (2018), doi: 10.1016/j.fct.2018.05.042.

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 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.

Article

# Food groups for allergen risk assessment: combining food consumption data from different countries in Europe

Sophie BIROT <sup>1,\*</sup>, Charlotte B. Madsen <sup>2,\*\*</sup>, Astrid G. Kruizinga <sup>3</sup>, Amélie Crépet <sup>4</sup>, Tue Christensen <sup>2</sup> and Per B. Brockhoff <sup>1</sup>

<sup>1</sup> DTU Compute, Technical University of Denmark, \* present: Nordic Biosciences, Denmark;

sbi@nordicbio.com

<sup>2</sup> National Food Institute, Technical University of Denmark, Kemitorvet, DK-2800 Kgs. Lyngby, Denmark;

charm@food.dtu.dk

<sup>3</sup> The Netherlands Organization for Applied Scientific Research (TNO), Zeist, The Netherlands;

astrid.kruizinga@tno.nl

<sup>4</sup> ANSES, French Agency for Food, Environmental and Occupational Health Safety, 14 rue Pierre et Marie

Curie, 94701 Maisons-Alfort, France; Amelie.CREPET@anses.fr

\*\* Correspondence: charm@food.dtu.dk, +4535887563

## Highlights:

- A method to combine national food consumption data suitable for allergen risk assessment
- 7 and 2 days survey design comparison
- Food groups automatically created across countries
- Includes consumption data available for food allergy risk assessment

**Abstract:** To prevent allergic reactions, food producers have to be able to make a knowledge based decision on whether to label their products with precautionary labelling. As many manufactured food products are sold in different countries across Europe, the allergen risk assessment should be estimated at the European levels. As currently, there are no pan-European food data suitable for

food allergy risk assessment. The aim of this paper is to investigate if consumption data, at a meal level, from National Food Consumption Surveys, can be combined to form a common Food Consumption database. In this first attempt we developed a procedure to investigate, if national food consumption data can be combined and grouped using data from Netherlands, France and Denmark. The homogeneity of consumption patterns and the relevance of difference in risk of allergic reaction were compared, using a fixed framework of allergen concentration levels and threshold distribution. Thus, the relevance of using common consumption data across countries was verified. The food groups formed were subsequently evaluated and adjusted based on practical considerations. It resulted in designing 61 food groups that can be used for allergen risk assessment. The summary statistics and descriptive names for each food group are included.

**Keywords:** Food allergy; National Food Consumption Surveys; food groups; risk assessment

---

## 1. Introduction

Food allergy affects up to 6% of children and 3-4 % of adults and seems to be increasing (Wang and Sampson, 2011). This means that a large number of consumers must avoid consuming food products containing the harmful allergen. Checking the ingredients list is the first action to take by consumers to avoid the allergen. However, products can contain the allergen unintentionally due to contamination during production. This is potentially harmful for the allergic consumers as the unintended allergen is not included in the list of ingredients (European Directive 2003/89/EC). In order to prevent allergic reactions and inform consumers of the possible presence of allergens, precautionary labelling is used by food manufacturers when a contamination is suspected. The unintended presence of allergen could e.g. be due to shared production lines or equipment or contamination of ingredients. In EU, the use of precautionary allergen labelling (PAL) is not regulated. Therefore, PAL is used on many products that do not actually contain the unintended

allergen (DunnGalvin et al., 2015). As a consequence of the over-use and lack of accuracy in labelling, allergic consumers may exhibit dangerous behavior by ignoring precautionary labelling (DunnGalvin et al., 2015). To enable food producers to make knowledge based decisions on whether to use PAL, it is necessary to determine if the product is safe for the allergic consumer. This may be done doing a classical toxicological safety assessment, where the dose resulting from a certain contamination is compared to a dose that causes allergic symptoms. If sufficient data are available it is also possible to perform a probabilistic risk assessment allowing quantifying the risk, i.e. the number of people at risk of an allergic reaction (Madsen et al., 2014).

There are three main input parameters in a (probabilistic) food allergen risk assessment: data on the level of contamination with the allergen in the food, data from challenge studies among subjects with an allergy to the allergen, and data on how much of the food is normally consumed in one eating occasion.

Relevant food consumption data on a meal level are crucial for allergen risk assessment, since the time to reaction in case of an allergic reaction is low and allergens are metabolized within hours, making the meal the relevant input parameter.

As currently, there are no pan-European food consumption data suitable for food allergen risk assessment, a food allergen risk assessment is done on a case-by-case basis using the most appropriate food consumption data available. However, a more structured database would improve the consistency of the risk assessments. National Food Consumption Surveys create vast amounts of detailed data. Hence, selecting the correct food consumption data may be a challenge and at the same time a prerequisite for estimating the correct risk of allergic reaction in the population. In this initial step, we have developed a procedure to investigate if national food consumption data can be combined using data from the Netherlands, France and Denmark. Data are combined in food groups

with homogeneous consumption patterns. The starting point was the FoodEx2 coding system (European Food Safety Authority, 2015). This system is originally designed to group food products for nutritional and safety purposes, hence the focus is on the average consumption in one day or multiple days. This time frame is not suitable for allergens, since there the focus is on the separate eating occasions. We therefore had to make adjustments to create new food groups that are both logical for the user and based on amount consumed per meal.

In this paper, we describe a method for combining national food consumption data for food allergen risk assessment. The method developed is based on the approach developed for within country consumptions' clustering and comparison (Biro et al., 2017). A detailed description of all food groups including the accompanying consumption figures to be used in risk assessment is presented.

## **2. Materials and Methods**

### *2.1. National Food Consumption Surveys*

National Food Consumption Surveys from the Netherlands (van Rossum et al. 2011), France (Dubuisson et al., 2010) and Denmark (Pedersen et al., 2008) were merged into one database to investigate the feasibility of creating a common food consumption database suitable for allergen risk assessment. The three National Food Consumption Surveys were combined using a common coding system for food items: The FoodEx2 coding system developed by EFSA (European Food Safety Authority, 2015). Participants are aged from 18 to 75 years. The food consumption was recorded for each eating occasion, defined as breakfast, lunch, dinner and other distinct eating occasions during the days. As food allergy risk assessment is estimating the risk of allergen contamination of processed foods, data on unprocessed foods such as raw fruits and vegetables were excluded.

## *2.2. Impact of the number of days in the surveys on the maximum consumption on an eating occasion*

The relevant consumption scenario is the eating occasion. The reason is that the time between food consumption and an allergic reaction is very low. The allergen will not accumulate and therefore residual allergen will not have impact on a later allergic reaction (Madsen et al., 2014). We chose a conservative approach, and used the maximum consumption on a single eating occasion over the days, for each food item, consumed by each participant. The consumption surveys were not identical. The Netherlands performed two non-consecutive 24h recalls, whereas pre-coded 7 day food records were used in France and Denmark. However, it is expected that the maximum over 7 days is higher than the maximum over 2 days, as there are more chances, statistically, to register higher consumption for a food item over 7 days than over 2 days. It is then expected that the consumption data used for the risk assessment will be higher for France and Denmark than for Netherlands. We therefore developed a method to compare the impact of merging maximum consumption from surveys with different designs. This was done by evaluating consumption and risk of allergic reaction as described below.

Thus, the influence on number of days on the maximum consumption distribution for each food item was investigated for France and Denmark by resampling the survey days. Two non-consecutive days were randomly selected for each participant among the 7 days in the survey. Then, the eating occasion with the maximum consumption was selected per participant. In order to use the information available from all 7 days, the 2 days were sampled 50 different times. Thus, the 50 maximum consumptions for all the participants were averaged per food item, so the numbers could be directly compared to the maximum consumption over 7 days. Food item consumptions are represented by its median, as it is less sensitive to extreme values than the mean. The relative differences between the original and resampled surveys were calculated. Thus, the relative difference in median consumption per food item helps to assess the difference in consumption

between the original and the resampled data. Thus, the spread of the differences in consumption can be evaluated.

The range of differences in median was then used to elaborate risk simulation scenarios to assess the impact of the difference in consumption induced by the number of days on the risk of allergic reaction. A relative difference in median from 5% to 35% was simulated by shifting the original distribution with an incremental factor from 5% to 35%. The difference in risk was then compared to the risk calculated with the original data.

### 2.3. Creating food groups for risk assessment

The aim of the method was to design food groups across countries that could be used for allergen risk assessment. The different steps are described in **Figure 1**, and the following sections further explain it.

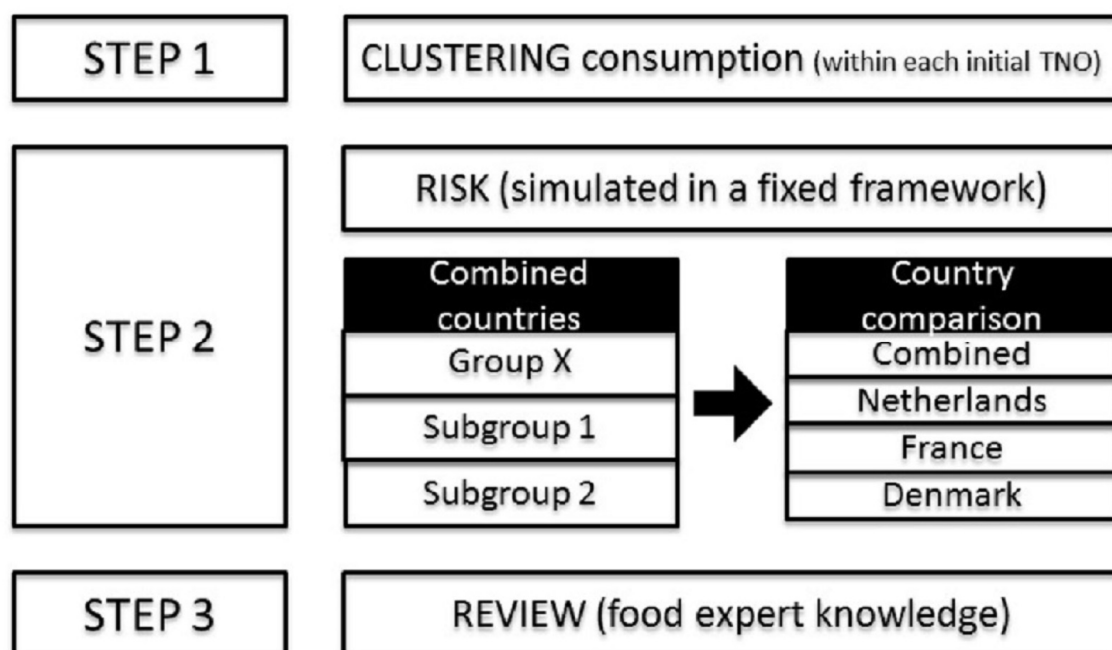


**Figure 1:** Overview of the steps of the method to design food groups across countries

A procedure to design food groups suitable for allergen risk assessment was developed and described (Biro et al., 2017). Based on this initial grouping, an adapted version was applied on the

consumption data for the three countries combined. In short in the first step, a customized clustering method was applied to the combined consumption dataset, creating subgroups with homogeneous consumption patterns (step 1 in Figure 1). To evaluate not only the consumption within the group but also the impact of difference in consumption on the allergen risk assessment outcome, the risk was estimated in a second step. This second step used a probabilistic risk assessment with the input data: distribution of food consumption, threshold distributions from peanut and soy challenges, and a fixed set of concentration of unintended allergen (1, 10, 100 and 1000 ppm allergen protein) to calculate the risk of allergic reaction patterns (step 2 in Figure 1). The framework designed this way helps to assess the consequences of differences in consumption on the allergy risk, as the concentration and the threshold remain constant in the different framework conditions (Biro et al., 2017).

In the rest of the paper, “framework” describes the set of fixed allergen concentration levels and challenge data distribution as described above.

In this paper, the framework was used for two different purposes:

- First, the consequences of the differences between the maximum consumption over the 7 days or the 2 resampled days were assessed as described in section 2.2.
- Secondly, the framework was used to assess the differences and similarities in risk, both, between the groups and subgroups designed during the clustering step, and between the countries.

The R software (R Core Team, 2015) was used to perform the simulation within the framework.

#### *2.4. Maximum relative difference and decision criterion*

The relative allergy risk was used to evaluate the difference in risk between the risk based on the food group of the combined countries and the suggested subdivision, and the risk based on consumption data from one country. Thus, the magnitude of the risk of allergic reaction has no impact on assessing the difference, and the homogeneity of consumption patterns within one group can be better assessed. Then the maximum of these relative differences is calculated, so the impact of keeping or splitting a specific food group can be assessed per food group.

Furthermore, to evaluate how grouping and subgrouping influenced the risk between the groups, a criterion to evaluate the difference in risk was developed. The selected criterion was the median of the maximum relative difference in risk calculated for the eight conditions within the framework (2 allergen x 4 concentration levels) for the combined countries, around 12,6% (step 2 on Figure 1) (Birot et al., 2017). This criterion was used in the same way for two purposes:

First, the criterion is used to decide whether an initial group was appropriate or should be divided into more groups (split). The maximum relative difference was calculated for the eight conditions of the framework in order to compare the subgroups risk to the overall group risk.

$$\text{max relative difference} = \max \left( \frac{|Risk_{Subgroup} - Risk_{Group}|}{Risk_{Group}} \right), \quad (1)$$

In a second step, the criterion was used to check if the combined countries consumption was representative enough or if a country specific consumption should be used to estimate the risk of allergic reaction for each designed group. The maximum relative difference was calculated for the eight conditions within the framework in order to compare the risk for the combined countries to the risk in Denmark, France and Netherlands.

$$\text{max relative difference} = \max \left( \frac{|Risk_{Country} - Risk_{Combined}|}{Risk_{Combined}} \right) \quad (2)$$

Finally, the criterion was used in the same way to compare the maximum relative differences: if the eight relative differences were above the criterion, then the initial group was split or it was decided to use a country specific consumption.

### *2.5. Food group revision*

New food groups based on the initial groups were formed combining the consumption data from the three countries (Denmark, France and Netherlands) where possible. The last step was to name and describe the food groups, and to verify their validity by experts (step 3 on Figure 1). The purpose was to have food groups that are logical and can easily be communicated to the end users for food allergen risk assessment.

In some cases a group formed based on risk, could not be described in a logical way to the end user and the statistical analyses were overruled by practical considerations. This was done in the following way: in order to form a logical group we decided to use the consumption in the subgroup with the highest risk to characterize the food consumption in both groups as described in detail in Appendix C.

### *2.6. Consumption's summary statistics per food groups*

To describe the consumption in each food group, summary statistics were calculated from the distribution of food items within each food group defined after the three step procedure. Thus, distribution of food groups was defined combining the distribution of food items within the food group of interest. Then, the summary statistics were calculated for consumers only, as food allergy risk assessment is done on the actual amount consumed at an eating occasion, making it necessary to exclude non consumers. The mean, the standard deviation and percentiles are calculated per food group.

## **3. Results**

### 3.1. Initial groups when combining National Food Consumption Surveys

The total number of participants was 8472 when the National Food Consumption surveys were merged (3819 in Netherlands, 2624 in France and 2029 in Denmark). The data included 703 different food items at the FoodEx2 level.

| Country     | Number of consumers | Number of food items | Number of groups based on the initial grouping method |
|-------------|---------------------|----------------------|-------------------------------------------------------|
| Netherlands | 3819                | 496                  | 42                                                    |
| France      | 2624                | 367                  | 40                                                    |
| Denmark     | 2029                | 168                  | 39                                                    |

**Table 1: Number of consumers and food items in each National Food Consumption Survey and number of groups based on the initial grouping method**

### 3.2. 2 vs 7 days comparison

To investigate the influence of having different survey designs; 2-day (The Netherlands) or 7-day (Denmark and France) the difference in maximum consumption using the original 7-day or the resampled 2-day consumption data was investigated. The analyses showed that the relative difference in median between the original and the resampled distribution is small for both France and Denmark. Most of the differences are lower than 5% (Figure 2). Therefore, the number of days in the survey does not have high impact on the reported amount of food consumed. .

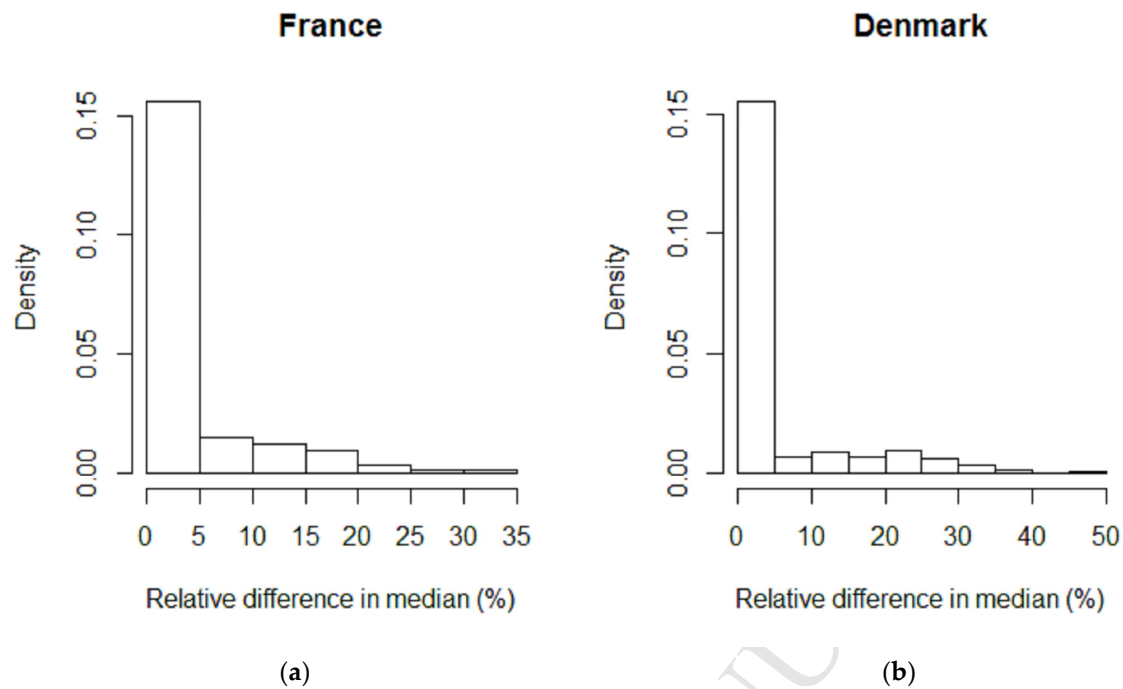


**Figure 2:** Relative difference in median between the original and the resampled food items consumption distributions (a) in France; (b) in Denmark

In addition the consequence for allergy risk was investigated. This was only done for the food items with a relative difference in median consumption higher than 25% ( 8 food items in France, 13 in Denmark). The average difference in allergy risk is presented in Table 1 for the conditions within the fixed framework (2 allergens and 4 concentration levels). As expected, the risk for the resampled data is small, on average lower from 0 to 0.9 points in France and from 0.1 to 1.4 points in Denmark. The data behind are presented in appendix A (France) and the appendix B (Denmark).

| Allergen | Concentration levels |         |        |         |         |         |          |         |
|----------|----------------------|---------|--------|---------|---------|---------|----------|---------|
|          | 1 ppm                |         | 10 ppm |         | 100 ppm |         | 1000 ppm |         |
|          | France               | Denmark | France | Denmark | France  | Denmark | France   | Denmark |
| Peanut   | -0.1                 | -0.2    | -0.3   | -0.4    | -0.5    | -0.9    | -0.9     | -1.4    |
| Soy      | 0.0                  | -0.1    | -0.1   | -0.1    | -0.2    | -0.3    | -0.4     | -0.6    |

**Table 2:** Average difference in allergy risk between the original 7-days and the resampled 2-days consumption's distribution. Only food items with relative difference in median higher than 25% are included.

Table 2 shows that the risk of having an allergic reaction calculated from the resampled consumption is slightly lower because the maximum consumption over the resampled 2 days is lower than the one over the original 7 days. However, this difference is too low to impact the risk management decisions following the results of the risk assessment. Thus, having data from two different survey designs available within this study, 2 days in Netherlands vs. 7 days in France and Denmark, will have a limited impact on the group design.

### 3.3. Group subdivision based on clustering and risk assessment analysis

Each initial food group was suggested to be divided into one or more subdivisions by a clustering method described in detail in Birot et al 2017 (step 1 in Figure 1).

| Number of subdivisions suggested | Based on clustering | Based on risk assessment analysis |
|----------------------------------|---------------------|-----------------------------------|
| 0                                | 5                   | 25                                |
| 1                                | 29                  | 13                                |
| 2 or more                        | 8                   | 4                                 |

**Table 3:** number of subdivisions based on clustering and risk assessment analysis for the combined countries within each of the 42 initial groups

There were five initial groups for which no further split was suggested because they had a low number of food item (i.e. one or two), 29 with one subdivision and eight with two and more subdivisions (Table 3).

In order to establish the appropriateness of the initial groups for food allergen risk assessment the allergy risk for the subgroups within the group was calculated and compared with the allergy risk of the overall group (step 2 in Figure 1). This was done as follows: The median of the maximum relative difference in risk was calculated across the conditions within the framework (2 allergens x 4 concentration levels) (Biro et al., 2017). After the risk analysis, the number of initial groups that did not need subdivisions increased to 25. Simultaneously, the number of groups that needed one subdivision fell to 13 and the number of groups that needing two subdivisions fell to four (Table 3).

### 3.4. Initial groups with subdivisions

Based on the criteria described above it was necessary to subdivide 40% of the 42 initial groups. This resulted in 17 food groups to be divided (Table 4). The number of subdivisions for each group that were found to be divided and their name is also indicated in Table 4.

| Group name                                       | Number of subgroups |
|--------------------------------------------------|---------------------|
| Chestnut paste and coconut milk                  | 2                   |
| Milk powder and Cocoa powder                     | 2                   |
| Milk(products), yoghurt(products), desserts      | 3                   |
| Potato chips                                     | 2                   |
| Meal replacements, meat imitates and supplements | 2                   |
| Small sweets                                     | 2                   |
| Sugar                                            | 2                   |
| Sweet confectionary (jam, marmalade)             | 3                   |
| Sauces                                           | 2                   |
| Fish products                                    | 3                   |
| Meat products                                    | 2                   |
| Herbs and spices                                 | 2                   |
| Alcoholic drinks (excl. beer)                    | 3                   |
| Cakes                                            | 2                   |
| Breakfast products                               | 2                   |
| Fruit and vegetables, processed                  | 2                   |
| Eggs and egg based dishes                        | 2                   |

**Table 4:** Groups with required a subdivision

The group division and the ones which were not divided resulted in 63 different food groups. The groups were named to best describe the foods included in the newly designed groups.

Below we have used the breakfast group to illustrate the subdividing of a group. A large difference in risk (according to the criterion) can be seen between the two subgroups (Table 5 and Figure 3). It can be noted that the first subgroup has a higher risk than the second subgroup, so if the overall consumption will be used, it will result in an underestimation of the risk for the first subgroup and an overestimation for the second one. Moreover, the maximum relative differences between the risk of the breakfast products group and its subdivisions for each contamination level and each allergen are all above the criteria set for subdivision of 12.6%.

| Allergen | Concentration levels of contamination |        |         |          |
|----------|---------------------------------------|--------|---------|----------|
|          | 1 ppm                                 | 10 ppm | 100 ppm | 1000 ppm |
| Peanut   | 41.7%                                 | 40.4%  | 35.9%   | 27.5%    |
| Soy      | 43.5%                                 | 43.1%  | 42.1%   | 39.6%    |

**Table 5:** Maximum relative difference between the risk of the breakfast product group and its subgroups

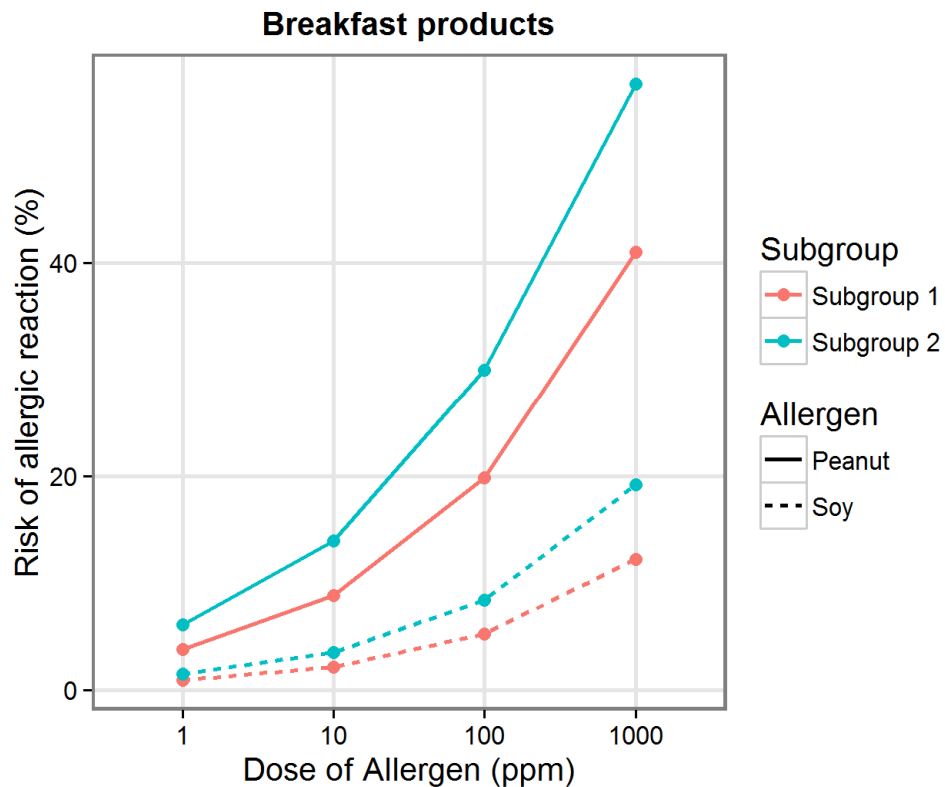


**Figure 3:** Risk of allergic reaction for two subgroups in the breakfast product group

### 3.5. Country comparison

The analyses resulted in 61 food groups. From these groups, only eight groups were country specific e.g. the consumption data from the three countries could not be merged based on the criterion chosen (Table 5 and 12).

| Initial group name                   | Group name after subdivision                          | Subgroup       |
|--------------------------------------|-------------------------------------------------------|----------------|
| Fried/warm snacks                    | No change                                             | No subdivision |
| Pancakes and waffles                 | No change                                             | No subdivision |
| Chewing gum                          | No change                                             | No subdivision |
| Sugar                                | No change                                             | Subgroup 1     |
| Sweet confectionary (jam, marmalade) | No change                                             | Subgroup 1     |
| Meat products                        | Meat products - mean 65 g                             | Subgroup 1     |
| Herbs/spices                         | Herbs and spices mixes, bouillon cubes, yeast extract | Subgroup 1     |
| Breakfast products                   | Breakfast products, porridge                          | Subgroup 2     |

**Table 5:** Groups with country specific consumption data

To illustrate the decision of having a country specific consumption, the chewing-gum group is taken as an example. In Table 6, it can be seen that the maximum relative difference between risk, for each contamination level and allergens are all above the threshold of 12.6% (from 27.0% to 33.3%). It was therefore decided to create groups with country specific consumption. In Figure 4, it can be noticed that the difference in risk between the three countries is large (according to the criterion).

| Allergen | Concentration levels |        |         |          |
|----------|----------------------|--------|---------|----------|
|          | 1 ppm                | 10 ppm | 100 ppm | 1000 ppm |
| Peanut   | 29.9%                | 30.1%  | 28.8%   | 27.0%    |
| Soy      | 33.3%                | 30.5%  | 31.1%   | 30.2%    |

**Table 6:** Maximum relative difference in risk between combined and individual countries for the chewing-gum group

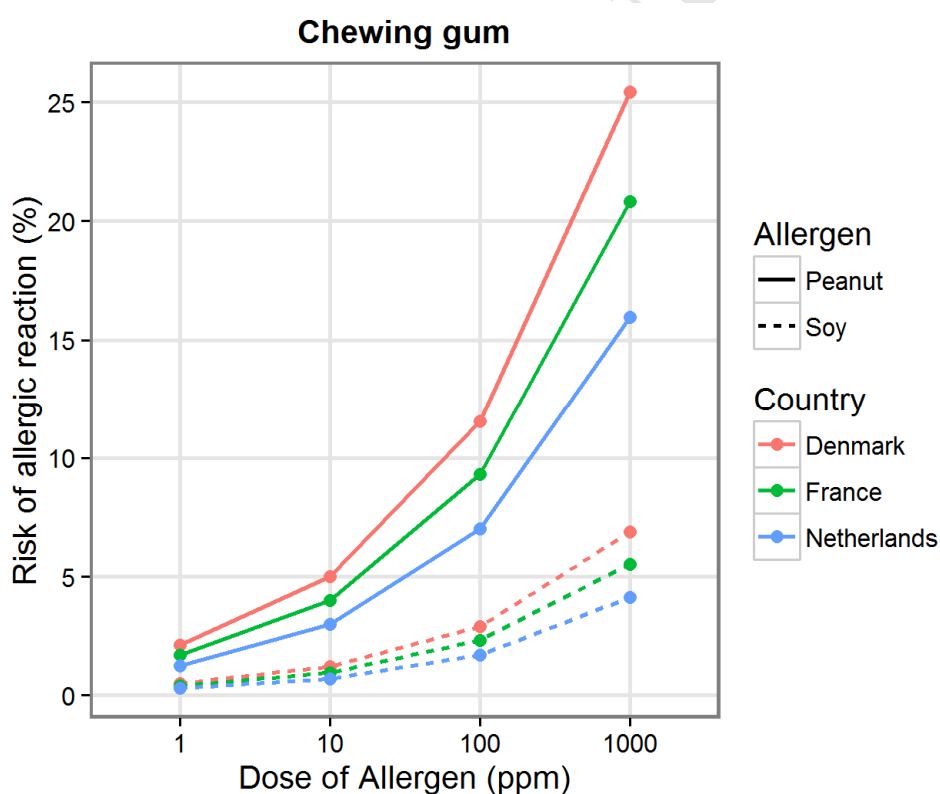


**Figure 4:** Risk of allergic reaction for the chewing-gum group in each country

### Group for combined countries and without need of subdivision

Finally, 21 of the 42 initial groups were found to not require a subdivision and country combined consumption can be used. These food groups are shown in the Table 7.

| Group name                             |
|----------------------------------------|
| Cheese                                 |
| Cream and coffee milk                  |
| Ice cream                              |
| Peanuts, nuts and dried fruits         |
| Soups                                  |
| Chocolate and chocolate products       |
| Cereal bars                            |
| Mashed potato powder                   |
| Potato product (excl. powder)          |
| Vegetable oils and animal fat          |
| Butter/halvarine/margarine             |
| Crackers, crisp bread, rusk and toast  |
| Bread, bread rolls and bread doughs    |
| Beer                                   |
| Syrups                                 |
| Drinks without alcohol (excl. syrup)   |
| Cookies (biscuits)                     |
|                                        |
| Pasta, rice, couscous and other grains |
| Legumes                                |
| Sandwich and pizza                     |
| Composite dishes                       |

**Table 7:** Groups for combined countries and without need of subdivision

The cookies group illustrates a group where the allergy risk between the overall group and the subgroups suggested by clustering is low, (Figure 5 and Tables 8 and 9). In Table 8, it can be seen that all maximum differences between the group risk and the subgroups is lower than the criterion of 12.6%. Furthermore, Table 9 shows that the maximum relative difference between the combined countries risk and each individual country risk is lower than the criterion. It was therefore decided

that for the cookies group, the combined countries consumption is representative for all three countries.

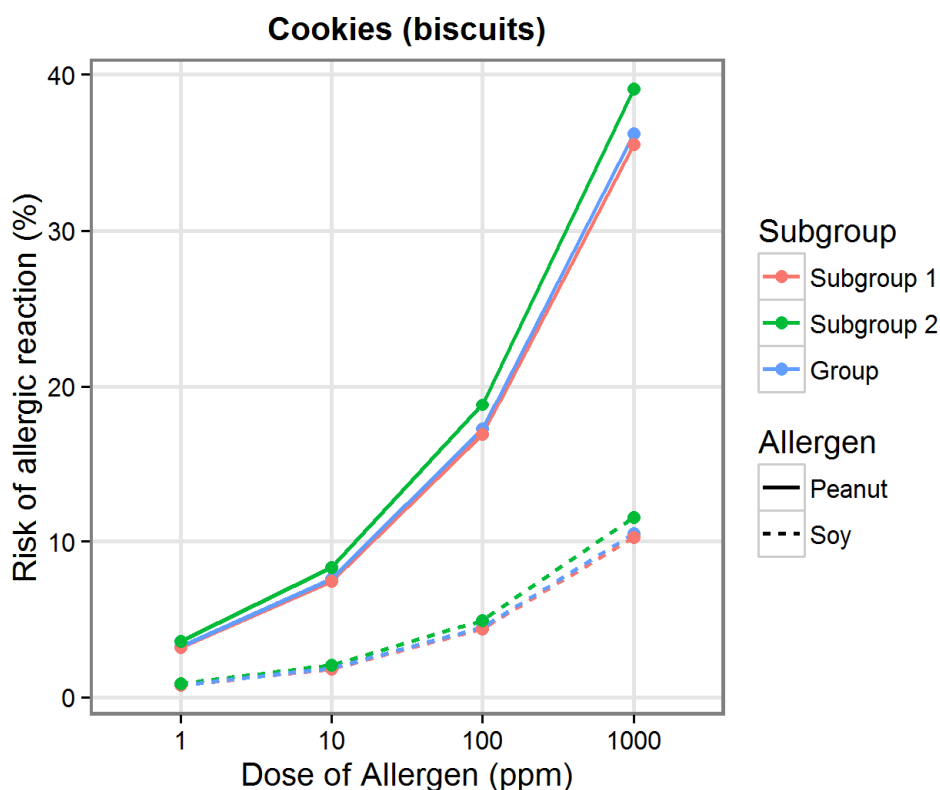


**Figure 5:** Risk of allergic reaction for the cookies group and the two subgroups

| Allergen | Concentration levels |        |         |          |
|----------|----------------------|--------|---------|----------|
|          | 1 ppm                | 10 ppm | 100 ppm | 1000 ppm |
| Peanut   | 9.9%                 | 10.0%  | 9.0%    | 8.0%     |
| Soy      | 10.4%                | 10.3%  | 10.0%   | 9.8%     |

**Table 8:** Maximum relative difference in risk between the cookies group and the subgroups

| Allergen | Concentration levels |        |         |          |
|----------|----------------------|--------|---------|----------|
|          | 1 ppm                | 10 ppm | 100 ppm | 1000 ppm |
| Peanut   | 10.9%                | 10.6%  | 9.2%    | 7.6%     |
| Soy      | 11.2%                | 11.0%  | 10.6%   | 9.9%     |

**Table 9:** Maximum relative difference in risk between combined and individual countries for the cookies group

### 3.6. Adjustments based on practical considerations and groups' overestimation

The overall aim of the project was to make food groups comprising similar foods with similar consumption. In some instances the grouping resulted in similar foods in different groups, making it impossible to create logical groups. To correct for this it was necessary to adjust 12 of the 61 groups. This adjustment was made by AK and CBM. The adjustments are summarized in Table 10. The reasons and criteria behind the adjustments are detailed in Appendix C. As an example the sweet confectionary group was split into two groups with e.g. jam and 'other fruit spreads' in one subgroup and e.g. marmalade, rose hip jam and 'fruit spreads' in the other subgroup. In order to form a logical group, the consumption in the subgroup with the highest risk (jam) is used to characterize the food consumption in both groups. Furthermore, the group was described as such: Sweet confectionary (jam, marmalade).

A similar procedure was applied when the country with the highest risk was used to describe the food consumption of a food group. As an example, the differences in risk between the three countries for the porridge group (breakfast product group, subgroup 2) are high according to the criterion. Thus, it would be necessary to use different consumption distributions for each country. However, in order not to increase the complexity of the food groups, the country with the highest risk is used to calculate the risk of allergic reaction in the two other countries. For the porridge group, the consumption in Denmark is used to estimate the risk of allergic reaction in the Netherlands and France.

Finally, in the "Drinks without alcohol" group, 73 different food items from the FoodEx2 coding system were used. Due to the high number of food items in this food group, it was not possible to cluster food items in a meaningful way, for examples apple juice and orange juice ended up in two different subgroups. The food items were aggregated on the next level of the FoodEx2 coding system, so suggested subgroups could be coherent.

| Food group                                                   | Adjustments                   |                              |                                    |
|--------------------------------------------------------------|-------------------------------|------------------------------|------------------------------------|
|                                                              | Use subgroup with higher risk | Use country with higher risk | Use less detailed food item coding |
| Milk and milk products consumed with a spoon                 | ✓                             |                              |                                    |
| Potato and other starch based chips (including salty sticks) | ✓                             |                              |                                    |
| Sugar                                                        |                               | ✓                            |                                    |
| Sweet confectionary (jam, marmalade)                         | ✓                             |                              |                                    |
| Meat products – mean consumption 65 g                        |                               | ✓                            |                                    |
| Herbs and spices mixes, bouillon cubes, yeast extract        |                               | ✓                            |                                    |
| Alcoholic drinks, alcohol $\leq 15\%$                        | ✓                             |                              |                                    |
| Drinks without alcohol (excl. syrup)                         |                               |                              | ✓                                  |
| Cakes (including pastry)                                     | ✓                             |                              |                                    |
| Breakfast products, porridge                                 |                               | ✓                            |                                    |
| Fruit and vegetables, processed                              | ✓                             |                              |                                    |
| Egg based dishes                                             | ✓                             |                              |                                    |

**Table 10:** Summary of the adjustment made for practical consideration per food group

The relative difference in risk between a food group and the one used for the allergen risk assessment was average across the conditions within the framework (Table 11). For example, the consequences of using the French data for the “Meat 65g” consumption is that the Danish and Dutch consumption is overestimated with an average relative difference of 20% and 23%, respectively. It can be seen that using different food consumptions have different range of overestimation, from 6% to 74%. And on average, after adjustment for practical consideration, a food group will have a risk 27% higher than the one that would have been calculated with the original consumption data.

| Group overestimated     | Group name                                                   | Average relative difference between risks |
|-------------------------|--------------------------------------------------------------|-------------------------------------------|
| Subgroup 3, Combined    | Milk and milk products consumed with a spoon                 | 47%                                       |
| Subgroup 2, Combined    | Potato and other starch based chips (including salty sticks) | 74%                                       |
| Subgroup 2, Combined    | Sugar                                                        | 24%                                       |
| Subgroup 1, Netherlands | Sugar                                                        | 28%                                       |
| Subgroup 2, Combined    | Sweet confectionary (jam, marmalade)                         | 6%                                        |
| Subgroup 3, Combined    | Sweet confectionary (jam, marmalade)                         | 19%                                       |
| Subgroup 1, Denmark     | Meat products – mean consumption 65 g                        | 20%                                       |
| Subgroup 1, Netherlands | Meat products – mean consumption 65 g                        | 23%                                       |
| Subgroup 1, France      | Herbs and spices mixes, bouillon cubes, yeast extract        | 44%                                       |
| Subgroup 1, Netherlands | Herbs and spices mixes, bouillon cubes, yeast extract        | 7%                                        |
| Subgroup 1, Combined    | Alcoholic drinks, alcohol $\leq 15\%$                        | 8%                                        |
| Subgroup 1, Combined    | Cakes (including pastry)                                     | 22%                                       |
| Subgroup 2, France      | Breakfast products, porridge                                 | 29%                                       |
| Subgroup 2, Netherlands | Breakfast products, porridge                                 | 14%                                       |
| Subgroup 1, Combined    | Fruit and vegetables, processed                              | 31%                                       |
| Subgroup 1, Combined    | Egg based dishes                                             | 34%                                       |

**Table 11:** The relative difference in risk (overestimation) between a food group and the one used for the allergen risk assessment (see also appendix C for details of the groups)

A complete list of the 61 final groups created with this procedure, including consumption figures, can be found in Table 12.

#### 4. Conclusion/discussion

In this paper, a procedure to create a combined consumption database using data from three European countries, suitable for use in allergen risk assessment, was presented. The study was a part of the risk assessment work package in the EU project “Integrated Approaches to Food Allergen and Allergy Risk Management” (iFAAM). The study showed that it is possible to merge consumption data and create common food groups relevant for food allergen risk assessment. The summary data presented in Table 12 may be used by food industry and authorities, for food allergy risk assessment (Madsen et al., 2009).

The study had two main challenges. The first challenge was that the study survey in The Netherlands was different than in Denmark and France leaving us with the question, if data from different survey types could be merged. We therefore had to develop a method allowing us to compare the maximum intake in a 2-days survey with the maximum intake in a 7-days survey. Luckily we found that the lower maximum in a 2-days survey was small enough to allow us using data from both survey types. To our knowledge this is the first time consumption data obtained by the two survey methods have been compared using statistical methods.

The second challenge was the large amount of data including differences in detail i.e. the FoodEx2 level used for recording food consumption data. The consumption survey in The Netherlands used a much higher level of detail (496 foods) compared to Denmark (168) with France in the middle (367). As an example in Denmark 97% of cheese consumption was registered on FoodEx2 code level 2 (cheese), whereas the France only 23% of the cheese data were recorded on this level. The majority of French cheese data were recorded at the FoodEx2 level 5 e.g. camembert, emmental etc.

Because of the large amount of data the method developed to create the food groups needed to be based on an automated criterion. Depending on the homogeneity of the initial group, the difference in allergy risk between the overall group and its suggested subdivisions can be low or high. In order to evaluate the magnitude of the risk differences across food groups, a criterion was selected. The criterion choice was made, so no subjective information had influence, as there was no previous knowledge on such relative risk differences. Based on some investigation (not detailed in this paper), the criterion was decided to be the median of the maximum relative difference in risk across food groups or 12.6%. In order to evaluate differences in risk across contamination levels and allergens magnitude various summary statistics (mean, median and different quantiles) were calculated. These investigations helped to evaluate what is a high difference in risk between a group and its suggested subdivisions. As the median was the only summary statistics stable across the

framework conditions (allergen x concentration level), the criterion was decided to be the overall median of the relative differences in risks. This criterion best fit the purpose of significant difference in risk between a group and its subgroup and the homogeneity of the consumption patterns. Thus, comparing the criterion within food groups makes sure the suggested subgroups with the clustering method actually detects a real difference in consumption distribution and not a difference in risk induced by the contamination level or allergen, since the framework only changes the food group consumptions, the relevant difference is measured. It is also worth noting that the criterion is based on the probabilistic risk assessment using the consumption per meal from the survey participants of the combined distribution of food items within the food group and its food groups. Furthermore, the criterion also helps to assess the order of magnitude of risk differences across food groups and assure that differences between the group and its subgroup risk is high (according to the criterion). Thus, this criterion was found to be the best trade-off between homogeneity in consumption patterns within food groups and significant difference in risk across food groups.

To cope with the large dataset an automated approach was necessary, but it also resulted in groups that were not logical. A detailed assessment of the foods in the groups resulted in a manual adjustment of some of the groups in a way where the consumption in the combined new group was always based on the subgroup with the highest consumption. The disadvantage is that the risk of some products having a lower consumption will be overestimated. However, this principle ensures that consumption of a certain food will be overestimated rather than underestimated. The most extreme example is salty stickers (pretzels) where the risk will be 74% higher based on using the group data compared to the individual food. In average the manual approach for the twelve groups result in an overestimation of risk of 27%.

In hindsight adding salty stickers to the potato chip group, may not have been the wisest, but on the other hand creating a separate group for salty stickers also seems to be overkill. From appendix C it can be seen that the three food groups with an over estimation of more than 31% are;

three products added to the “milk with a spoon” groups, one product added to “egg based dishes” and the heterogeneous “Herbs and spices mixes, bouillon cubes, yeast extract” where the French consumption will be overestimated. For the other groups the overestimation of risk will be from 6 to 31%. We find that the excess risk is an acceptable price to pay for getting a relatively simple food grouping that is logical and can be communicated.

In this paper, we present a method that may be used to build European consumption databases suitable for allergen risk assessment. We have created the first database based on food consumption data from Denmark, France and The Netherlands. We have developed a system where food consumption survey data from different countries on a meal level coded with FoodEx codes can be merged and logical food groups created. The summary statistics are presented and can be used for risk assessment. To extend the database to more countries, it will be possible to organize similar coded food consumption data from other countries into the defined food groups, and analyze similarities and differences. Having consumption data from across Europe and using the proposed procedures, it will be possible to investigate, if one pan-European food consumption data set for use in food allergen risk assessment can be created or consumption databases based on regional eating habits are necessary. Hopefully this will allow performing food allergen risk assessment covering Europe as a whole, benefiting the consumers.

**Acknowledgments:** This work was part of the iFAAM project (Integrated Approaches to Food Allergen and Allergy Risk Management, Grant Agreement No. 322147), the national food consumption data for France and Denmark were kindly provided by the work package partners: ANSES and DTU. The national food consumption data of the Netherlands was kindly provided by the National Institute for Public Health and the Environment in the Netherlands. The challenge data for peanut and soy was kindly provided by TNO and FARRP (University of Nebraska – Food Allergy Research and Resource Program).

**Author Contributions:** Sophie BIROT, Charlotte B. Madsen, Astrid G Kruizinga, Amélie Crépet, Tue Christensen and Per B. Brockhoff conceived and designed the experiments; Sophie Birot developed the methods and analyzed the data with input from Charlotte B. Madsen, Astrid G Kruizinga, Amélie Crépet, Tue

Christensen and Per B. Brockhoff; Astrid G Kruizinga, Amélie Crépet, Tue Christensen and Per B. Brockhoff contributed materials and analysis tools; Sophie Birot wrote the paper

ACCEPTED MANUSCRIPT

**Table 12:** Food consumption summary statistics per food group (in g) \*

| Name                                                                   | Country               | Mean  | SD    | P75   | P90   |
|------------------------------------------------------------------------|-----------------------|-------|-------|-------|-------|
| Chestnut paste and coconut milk                                        | Combined              | 63.7  | 55.3  | 100.0 | 116.5 |
| Peanut butter                                                          | Combined              | 27.4  | 17.5  | 35.0  | 45.0  |
| Cheese                                                                 | Combined              | 39.8  | 31.3  | 48.0  | 80.0  |
| Milk powder and Cocoa powder                                           | Combined              | 18.5  | 14.1  | 26.4  | 33.6  |
| Coffee creamer                                                         | Combined              | 4.3   | 4.2   | 6.0   | 8.0   |
| Cream and coffee milk                                                  | Combined              | 22.5  | 26.1  | 30.0  | 40.0  |
| Ice cream                                                              | Combined              | 88.2  | 47.3  | 100.0 | 150.0 |
| Milk and milk products for drinking                                    | Combined              | 264.5 | 163.7 | 317.5 | 432.0 |
| Milk and milk products consumed with a spoon                           | Combined              | 156.9 | 76.4  | 200.0 | 250.0 |
| Peanuts, nuts and dried fruits                                         | Combined              | 33.3  | 29.5  | 40.0  | 60.0  |
| Potato and other starch based chips (including salty sticks/pretzels)) | Combined              | 43.4  | 38.2  | 59.0  | 79.0  |
| Fried/warm snacks                                                      | DK                    | 162.8 | 103.1 | 180.0 | 270.0 |
| Fried/warm snacks                                                      | FR                    | 109.0 | 89.2  | 140.0 | 210.0 |
| Fried/warm snacks                                                      | NL                    | 77.4  | 50.5  | 85.5  | 140.0 |
| Meal replacements and meat imitates                                    | Combined              | 105.1 | 111.6 | 113.0 | 250.0 |
| Supplements                                                            | Combined              | 1.7   | 2.6   | 2.0   | 3.0   |
| Pancakes and waffles                                                   | DK                    | 151.5 | 104.6 | 200.0 | 300.0 |
| Pancakes and waffles                                                   | FR                    | 152.7 | 102.3 | 200.0 | 300.0 |
| Pancakes and waffles                                                   | NL                    | 87.1  | 100.1 | 100.0 | 210.0 |
| Soups                                                                  | Combined              | 318.9 | 161.0 | 400.0 | 500.0 |
| Small sweets - sweet confectionary unspecified/Combined                | Combined              | 47.9  | 42.7  | 60.0  | 100.0 |
| Small sweets - sweet confectionary specified                           | Combined              | 25.5  | 31.0  | 28.0  | 60.0  |
| Sugar                                                                  | Combined <sup>1</sup> | 21.4  | 12.9  | 24.0  | 33.0  |
| Chocolate and chocolate products                                       | Combined              | 32.1  | 33.4  | 40.0  | 60.0  |
| Sweet confectionary (jam, marmalade)                                   | Combined              | 33.4  | 25.0  | 35.0  | 60.0  |
| Cereal bars                                                            | Combined              | 31.7  | 27.1  | 32.1  | 50.0  |
| Chewing gum                                                            | DK                    | 10.6  | 7.8   | 10.0  | 20.0  |
| Chewing gum                                                            | FR                    | 5.9   | 7.4   | 6.0   | 10.0  |
| Chewing gum                                                            | NL                    | 2.9   | 2.5   | 4.0   | 5.0   |
| Mashed potato powder                                                   | Combined              | 177.3 | 84.0  | 200.0 | 300.0 |
| Potato product (excl. powder)                                          | Combined              | 172.2 | 108.2 | 225.0 | 300.0 |
| Vegetable oils and animal fat                                          | Combined              | 14.8  | 11.9  | 20.0  | 30.0  |
| Butter/halvarine/margarine                                             | Combined              | 14.3  | 10.6  | 20.0  | 25.0  |
| Sauces used as condiments and dessert sauces                           | Combined              | 22.3  | 20.7  | 30.0  | 46.5  |
| Sauces , savory, chutneys and pickles                                  | Combined              | 57.1  | 47.7  | 75.0  | 105.0 |
| Fish products - mean 35 g such as fish fingers, fish paté              | Combined              | 34.2  | 29.7  | 40.0  | 62.6  |
| Fish products - mean 75 g such as smoked salmon, canned fish in oil    | Combined              | 74.2  | 49.5  | 100.0 | 136.2 |

|                                                                         |                       |       |       |       |       |
|-------------------------------------------------------------------------|-----------------------|-------|-------|-------|-------|
| Fish products - mean 115 g such as fish cake, fish balls                | Combined              | 115.6 | 75.0  | 150.0 | 190.0 |
| Meat products - mean 65 g such as bacon, salami, pâté                   | Combined <sup>2</sup> | 64.5  | 46.2  | 75.0  | 125.0 |
| Meat products -mean 105 g such as meat loaf, sausages                   | Combined              | 107.4 | 63.1  | 126.3 | 178.0 |
| Crackers, crisp bread, rusk and toast                                   | Combined              | 22.9  | 19.3  | 28.0  | 45.0  |
| Bread, bread rolls and bread doughs                                     | Combined              | 90.9  | 51.3  | 120.0 | 150.0 |
| Herbs and spices mixes, bouillon cubes, yeast extract                   | Combined <sup>1</sup> | 18.1  | 32.4  | 20.0  | 20.0  |
| Spices and salt                                                         | Combined              | 2.9   | 2.5   | 3.0   | 4.0   |
| Alcoholic drinks, alcohol ≤15%                                          | Combined              | 222.1 | 144.5 | 282.5 | 420.0 |
| Alcoholic drinks, alcohol above 15%                                     | Combined              | 68.9  | 69.6  | 83.8  | 120.0 |
| Beer                                                                    | Combined              | 567.1 | 520.3 | 660.0 | 990.0 |
| Syrups                                                                  | Combined              | 26.6  | 30.8  | 33.8  | 60.4  |
| Drinks without alcohol (excl. syrup)                                    | Combined              | 362.3 | 252.1 | 483.3 | 600.0 |
| Cookies (biscuits)                                                      | Combined              | 32.8  | 27.8  | 42.0  | 60.0  |
| Cakes (including pastry)                                                | Combined              | 144.4 | 78.0  | 180.0 | 250.0 |
|                                                                         |                       |       |       |       |       |
| Breakfast products eaten unprocessed (e.g. müsli, oat and maize flakes) | Combined              | 46.9  | 28.1  | 60.0  | 83.2  |
| Breakfast products, porridge                                            | Combined <sup>1</sup> | 168.0 | 163.2 | 202.0 | 257.0 |
| Pasta, rice, couscous and other grains                                  | Combined              | 155.4 | 91.2  | 200.0 | 270.0 |
| Legumes                                                                 | Combined              | 132.2 | 67.3  | 175.0 | 215.0 |
| Fruit and vegetables, processed                                         | Combined              | 139.1 | 86.7  | 190.0 | 238.0 |
| Eggs                                                                    | Combined              | 40.9  | 29.0  | 55.0  | 80.0  |
| Egg based dishes such as omelet                                         | Combined              | 123.8 | 69.0  | 180.0 | 200.0 |
| Sandwich and pizza                                                      | Combined              | 270.4 | 209.9 | 335.0 | 500.0 |
| Composite dishes such as lasagna, quiche, vegetable casserole           | Combined              | 238.2 | 155.5 | 320.0 | 450.0 |

<sup>1</sup> The group is combined but the consumption data used are the Danish (see also appendix C for explanation)

<sup>2</sup> The group is combined but the consumption data used are the French (see also appendix C for explanation)

\*For a description of the foods in the food groups reference is being made to Birot et al. 2017.

**Appendix A:**

| Allergen | Item  | Name                  | Concentration levels |           |          |           |          |           |          |           |
|----------|-------|-----------------------|----------------------|-----------|----------|-----------|----------|-----------|----------|-----------|
|          |       |                       | 1 ppm                |           | 10 ppm   |           | 100 ppm  |           | 1000 ppm |           |
|          |       |                       | Original             | Resampled | Original | Resampled | Original | Resampled | Original | Resampled |
| Peanut   | A004Q | Wheat germ            | 3.1%                 | 2.9%      | 7.3%     | 6.9%      | 16.5%    | 15.7%     | 34.8%    | 33.3%     |
| Peanut   | A033J | Honey                 | 2.8%                 | 2.7%      | 6.5%     | 6.3%      | 14.9%    | 14.5%     | 31.9%    | 31.1%     |
| Peanut   | A03KA | Coffee beverages      | 6.4%                 | 6.2%      | 14.7%    | 14.3%     | 31.5%    | 30.8%     | 59.0%    | 58.0%     |
| Peanut   | A03TJ | Soya drink            | 6.5%                 | 6.3%      | 14.8%    | 14.3%     | 31.6%    | 30.8%     | 58.9%    | 58.0%     |
| Peanut   | A03YA | Veggie pot pie        | 5.9%                 | 5.8%      | 13.5%    | 13.3%     | 29.4%    | 28.9%     | 56.4%    | 55.6%     |
| Peanut   | A043V | Savoury sauces        | 3.1%                 | 3.0%      | 7.2%     | 7.1%      | 16.4%    | 16.1%     | 34.7%    | 34.2%     |
| Peanut   | A044V | Pesto                 | 2.9%                 | 2.9%      | 6.9%     | 6.8%      | 15.7%    | 15.6%     | 33.2%    | 32.9%     |
| Peanut   | A04NH | Fresh smoked sausages | 5.2%                 | 5.0%      | 12.0%    | 11.6%     | 26.3%    | 25.7%     | 51.8%    | 50.7%     |
| Soy      | A004Q | Wheat germ            | 0.7%                 | 0.7%      | 1.8%     | 1.7%      | 4.3%     | 4.0%      | 10.1%    | 9.5%      |
| Soy      | A033J | Honey                 | 0.7%                 | 0.6%      | 1.6%     | 1.5%      | 3.8%     | 3.7%      | 9.0%     | 8.7%      |
| Soy      | A03KA | Coffee beverages      | 1.6%                 | 1.5%      | 3.8%     | 3.6%      | 8.9%     | 8.6%      | 20.2%    | 19.7%     |
| Soy      | A03TJ | Soya drink            | 1.6%                 | 1.5%      | 3.8%     | 3.7%      | 8.9%     | 8.7%      | 20.3%    | 19.7%     |
| Soy      | A03YA | Veggie pot pie        | 1.4%                 | 1.4%      | 3.4%     | 3.3%      | 8.1%     | 8.0%      | 18.6%    | 18.3%     |
| Soy      | A043V | Savoury sauces        | 0.7%                 | 0.7%      | 1.8%     | 1.7%      | 4.2%     | 4.1%      | 10.0%    | 9.8%      |
| Soy      | A044V | Pesto                 | 0.7%                 | 0.7%      | 1.7%     | 1.7%      | 4.0%     | 4.0%      | 9.5%     | 9.5%      |
| Soy      | A04NH | Fresh smoked sausages | 1.2%                 | 1.2%      | 3.0%     | 2.9%      | 7.1%     | 6.9%      | 16.5%    | 16.1%     |

**Table A1:** Difference in risk between the original and the resampled consumption's distribution for the food items that were found to have a high difference in mean consumption, France

## Appendix B

| Allergen | Item  | Name                                         | Concentration levels |           |          |           |          |           |          |           |
|----------|-------|----------------------------------------------|----------------------|-----------|----------|-----------|----------|-----------|----------|-----------|
|          |       |                                              | 1 ppm                |           | 10 ppm   |           | 100 ppm  |           | 1000 ppm |           |
|          |       |                                              | Original             | Resampled | Original | Resampled | Original | Resampled | Original | Resampled |
| Peanut   | A004Y | Wheat bread and rolls, white (refined flour) | 4.8%                 | 4.5%      | 11.2%    | 10.4%     | 24.7%    | 23.1%     | 49.3%    | 46.7%     |
| Peanut   | A023G | Cooked cured meat                            | 3.3%                 | 3.1%      | 7.6%     | 7.2%      | 17.3%    | 16.4%     | 36.5%    | 34.8%     |
| Peanut   | A025J | Cooked sausages                              | 3.3%                 | 3.2%      | 7.7%     | 7.4%      | 17.6%    | 16.8%     | 36.9%    | 35.6%     |
| Peanut   | A02MA | Cow milk, skimmed (low fat)                  | 7.9%                 | 7.5%      | 17.8%    | 17.1%     | 37.4%    | 36.1%     | 67.1%    | 65.6%     |
| Peanut   | A032J | White sugar                                  | 2.3%                 | 2.2%      | 5.5%     | 5.2%      | 12.7%    | 12.1%     | 27.7%    | 26.5%     |
| Peanut   | A034V | Sweet confectionery                          | 3.8%                 | 3.6%      | 8.9%     | 8.4%      | 19.9%    | 19.0%     | 40.8%    | 39.3%     |
| Peanut   | A034X | Candies                                      | 2.8%                 | 2.6%      | 6.5%     | 6.1%      | 14.7%    | 13.9%     | 31.5%    | 29.9%     |
| Peanut   | A03KC | Coffee (average strenght) beverage           | 9.2%                 | 8.7%      | 20.7%    | 19.7%     | 42.6%    | 40.8%     | 73.2%    | 71.3%     |
| Peanut   | A03ND | Cider                                        | 7.6%                 | 7.4%      | 17.2%    | 16.9%     | 36.3%    | 35.9%     | 65.9%    | 65.4%     |
| Peanut   | A03VN | Hummus                                       | 3.0%                 | 3.0%      | 7.1%     | 7.0%      | 16.3%    | 15.9%     | 34.6%    | 34.0%     |
| Peanut   | A042D | Mixed vegetable salad                        | 4.6%                 | 4.3%      | 10.6%    | 10.1%     | 23.3%    | 22.4%     | 46.1%    | 44.6%     |
| Peanut   | A044F | Table-top condiments                         | 3.1%                 | 3.0%      | 7.3%     | 7.1%      | 16.6%    | 16.2%     | 35.3%    | 34.3%     |
| Peanut   | A044X | Mayonnaise                                   | 2.4%                 | 2.3%      | 5.6%     | 5.4%      | 13.0%    | 12.5%     | 28.4%    | 27.3%     |
| Soy      | A004Y | Wheat bread and rolls, white (refined flour) | 1.2%                 | 1.1%      | 2.8%     | 2.6%      | 6.7%     | 6.1%      | 15.5%    | 14.3%     |
| Soy      | A023G | Cooked cured meat                            | 0.8%                 | 0.7%      | 1.9%     | 1.7%      | 4.5%     | 4.2%      | 10.6%    | 10.0%     |
| Soy      | A025J | Cooked sausages                              | 0.8%                 | 0.7%      | 1.9%     | 1.8%      | 4.5%     | 4.3%      | 10.7%    | 10.2%     |
| Soy      | A02MA | Cow milk, skimmed (low fat)                  | 1.9%                 | 1.8%      | 4.6%     | 4.4%      | 10.9%    | 10.4%     | 24.4%    | 23.4%     |
| Soy      | A032J | White sugar                                  | 0.5%                 | 0.5%      | 1.3%     | 1.3%      | 3.2%     | 3.0%      | 7.6%     | 7.2%      |
| Soy      | A034V | Sweet confectionery                          | 0.9%                 | 0.9%      | 2.2%     | 2.1%      | 5.2%     | 5.0%      | 12.3%    | 11.7%     |
| Soy      | A034X | Candies                                      | 0.6%                 | 0.6%      | 1.6%     | 1.5%      | 3.8%     | 3.5%      | 9.0%     | 8.4%      |
| Soy      | A03KC | Coffee (average strenght) beverage           | 2.3%                 | 2.1%      | 5.5%     | 5.1%      | 12.8%    | 12.1%     | 28.3%    | 26.9%     |
| Soy      | A03ND | Cider                                        | 1.8%                 | 1.8%      | 4.4%     | 4.4%      | 10.5%    | 10.3%     | 23.6%    | 23.3%     |
| Soy      | A03VN | Hummus                                       | 0.7%                 | 0.7%      | 1.7%     | 1.7%      | 4.2%     | 4.1%      | 9.9%     | 9.7%      |
| Soy      | A042D | Mixed vegetable salad                        | 1.1%                 | 1.0%      | 2.6%     | 2.5%      | 6.3%     | 6.0%      | 14.6%    | 13.9%     |
| Soy      | A044F | Table-top condiments                         | 0.7%                 | 0.7%      | 1.8%     | 1.7%      | 4.3%     | 4.2%      | 10.1%    | 9.8%      |
| Soy      | A044X | Mayonnaise                                   | 0.6%                 | 0.5%      | 1.4%     | 1.3%      | 3.3%     | 3.1%      | 7.8%     | 7.5%      |

**Table A2:** Difference in risk between the original and the resampled consumption's distribution for the food items that were found to have a high difference in mean consumption, Denmark

## **Appendix C: Detailed adjustments made for practical considerations**

### **Milk and milk products consumed with a spoon**

Three foods, that had their own subgroup, but fit the description of this group and had a lower risk, were moved into this group without including them in the combined consumption of the group.

### **Potato and other starch based chips (including salty sticks)**

Salty sticks and Tapioca starch-based snacks with a lower consumption were moved into the group without contributing to the group consumption. This means that the risk connected to these two foods will be overestimated.

### **Sugar**

The risk analyses divided the sugar group in two groups both including white sugar. To solve this one group was formed using the Danish sugar consumption data to characterise the group. This means that the FR and NL sugar consumption will be overestimated.

### **Sweet confectionary (jam, marmalade)**

This group was split in two groups with e.g. jam and 'other fruit spreads' in one subgroup and e.g. marmalade, rose hip jam and 'fruit spreads' in the other subgroup. In order to form a logical group the two subgroups were combined into one group: Sweet confectionary (jam, marmalade). The consumption in the subgroup with the highest risk was used to characterise the food consumption in the combined group.

### **Meat products – mean consumption 65 g**

To be able to form two logical groups of the meat consumption data the FR data was used to characterise this group. This means that the DK and NL risk for this food group will be overestimated. In the other meat group the combined data from the three countries was used.

**Herbs and spices mixes, bouillon cubes, yeast extract**

This is a small and heterogeneous group. To make a combined group the Danish consumption data was used: This means that the FR and NL risk for this group will be overestimated.

**Alcoholic drinks, alcohol  $\leq 15\%$** 

The risk analyses resulted in two groups with wine in one group and red wine in the other (different levels of reporting consumption). The two groups were merged, but the consumption used is the one representing the subgroup with the highest risk.

**Drinks without alcohol (excl. syrup)**

A less detailed level of coding was used due to a very detailed food items coding in this group. The very detailed coding made it impossible to form meaningful group. Instead of the original 73 food items, the food items were grouped based on a higher level of coding. The higher level of coding organised the 73 food items in seven levels:

- Coffee, cocoa, tea and herbal drinks
- Coffee, cocoa, tea and herbal ingredient
- Drinking water
- Fruit juices and nectars
- Other fruit and vegetable juices or nectars
- Vegetable juices

This group did not need to be subdivided.

**Cakes (including pastry)**

The risk analyses resulted in two groups with many cakes but no logical division e.g. 'croissant' in one group and 'croissant filled with cream' in the other. The two groups were merged, but the consumption used is the one representing the subgroup with the highest risk.

**Breakfast products, porridge**

This is a small group. To make a combined group the Danish consumption data was used: This means that the FR and NL risk for this group will be overestimated.

**Fruit and vegetables, processed**

This is a large and heterogeneous group. The risk analyses resulted in two groups with no logical division: e.g. 'fruit compote' in one group and 'fruit compote, apple' in the other. The two groups were merged, but the consumption data used for the group is the one representing the subgroup with the highest risk.

**Egg based dishes**

The risk analyses resulted in two groups. One food item was moved from the group with the lower risk to the group with the higher risk without changing the consumption data in order to make two groups that could be described logical.

**References**

- Biro, S., Madsen, C. B., Kruizinga, A. G., Christensen, T., Crépet, A., & Brockhoff, P. B. (2017). A procedure for grouping food consumption data for use in food allergen risk assessment. *Journal of Food Composition and Analysis*, 59, 111-123. <http://doi.org/10.1016/j.jfca.2017.01.008>
- Dubuisson, C., Lioret, S., Touvier, M., Dufour, A., Calamassi-Tran, G., Volatier, J.-L., & Lafay, L. (2010). Trends in food and nutritional intakes of French adults from 1999 to 2007: results from the INCA surveys. *The British Journal of Nutrition*, 103, 1035-48. <http://doi.org/10.1017/S0007114509992625>
- DunnGalvin, A., Chan, C. H., Crevel, R., Grimshaw, K., Poms, R., Schnadt, S., ... Roberts, G. (2015). Precautionary allergen labelling: Perspectives from key stakeholder groups. *Allergy*, 70, 1039-1051. <http://doi.org/10.1111/all.12614>
- European Food Safety Authority. (2015). The food classification and description system FoodEx2

(revision 2). *EFSA Supporting Publications*, 8(12), n/a--n/a.

<http://doi.org/10.2903/sp.efsa.2011.EN-215>

Madsen, C. B., Hattersley, S., Buck, J., Gendel, S. M., Houben, G. F., Hourihane, J. O. B., ... Crevel, R.

W. R. (2009). Approaches to risk assessment in food allergy: Report from a workshop

developing a framework for assessing the risk from allergenic foods". *Food and Chemical*

*Toxicology*, 47, 480–489. <http://doi.org/10.1016/j.fct.2008.12.001>

Madsen, C.B., Houben, G.F., Hattersley, S, Crevel, R.W.R., Remington, B.C., Baumert, J.L., From

hazard to risk – Assessing the risk: in Madsen C.B., Crevel, R.W.R., Mills, E.N.C., Taylor S.L. Editors.

*Risk management for food allergy*, Elsevier, 2014, ISBN 978-0-12-381988-8

Pedersen, A. N., Fagt, S., Groth, M. V., Christensen, T., Biloft-Jensen, A., Matthiessen, J., ... Trolle, E.

(2008). *Danskernes kostvaner 2003-2008. Danskernes kostvaner 2003-2008*.

R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Retrieved from <https://www.r-project.org/>

van Rossum, C. T. M., Fransen, H., Verkaik-Kloosterman, J., Buurma-Rethans, E., & Ocke, M. (2011).

*RIVM 2011. Dutch National Food Consumption Survey 2007-2010 : Diet of children and adults aged 7*

*to 69 years. RIVM-Rapport 350050006*.

Wang, J. and H. A. Sampson. (2011). Food allergy. *The Journal of Clinical Investigation*, 121, 827–835.

<http://doi.org/10.1172/JCI45434>