WP3: Integrating early-life stressors

Subtask 3.1.4: Lifestyle of parents and young children

Data Harmonization Protocol

June 2018

University of Crete, Greece: *Marina Vafeiadi, Katerina Margetaki, Nikos Stratakis, Theano Roumeliotaki and Leda Chatzi*

Inserm, France: *Marie-Aline Charles, Patricia Dargent-Molina, Sandrine Lioret, Blandine de Lauzon, Barbara Heude, Sabine Plancoulaine and Maxime Cornet*

Contents

Table of Contents

Aim

The aim of subtask WP3.1.4 is to generate integrated adverse lifestyle stressors across the LifeCycle cohorts focusing on diet, physical activity, sedentary behavior and sleep.

Diet

We will generate harmonized data on early-life dietary stressors for three time-periods: pregnancy, preschool age (2-4 years), and school age (5-8 years). We focus on key food groups and nutrients. Further, we will examine exposure patterns by building 1) a well-established, easy-to-construct investigator-driven (a priori) quality index, and 2) data-driven (a posteriori) dietary patterns at the European level.

We propose to construct the Dietary Approaches to Stop Hypertension (DASH) diet score[1] for pregnant women and school-aged children. The DASH pattern emphasizes intake of fruits, vegetables, low fat dairy foods, whole-grain cereals and reduced saturated and total fat. This dietary pattern is recommended as a practical and understandable diet plan for the general public.[2] If a cohort is not able to produce the DASH index, it is not excluded from this subtask.

In a next step, we propose to derive data-driven dietary patterns based on harmonized information on food groups. We will use factor analysis (FA) methods per time period in the pooled dataset, and compare the agreement of pooled FA with cohort-specific FAs, as described in the multicenter European Prospective Investigation into Cancer and Nutrition (EPIC) study.[3]

We will also consider applying multiple imputation techniques in order to deal with missing food-groups data. Multiple imputations using chained equations (MICE) is a well-accepted and flexible method for dealing with missing data. The final decision will be made upon inspection of the percentage of missing data after the data collection is completed. Variables which are entirely missing for a cohort should not be imputed in the cohort level but could be imputed in pooled analyses.

To undertake those analyses, participating cohorts will have to upload the relevant harmonized data on their own DataSHIELD server. Finally, data transfer to UoC will possibly be requested from this subtask in order to construct the data-driven dietary patterns. In this case, we will ensure the privacy of the data by e.g. not sharing a personal ID making it impossible to link the shared data with any other personal information.

This subtask will be undertaken under the responsibility of research team led by Marina Vafeiadi from the University of Crete.

Participating cohorts:

14 cohorts have already agreed to participate in this task: ALSPAC, BiB, CHOP, DNBC, EDEN, ELFE, Gen R, INMA, MoBA, RAINE, RHEA, SWS, Piccolipiù

Proposed timeline:

- **12th of June 2018** - Presentation at the Lifecycle General Assembly (in Oulu) of the protocol and its application to the Rhea cohort
- **June 2018** - Participating cohorts receive the harmonization protocol and begin application to their data
- **July to October 2018 -** Participating cohorts finalize the harmonization process for the diet variables
- **November 2018 to May 2019** - Construction of dietary patterns

Planned papers (authorship to be decided)

- Dietary patterns in pregnant women and children and their predictors in Europe.
- The effect of diet on early child health. *Outcomes to be chosen in conjunction with WP4, 5 and 6.*
- The impact of diet on systemic metabolomic profiles. *Depending on data availability*

Physical activity sedentary behaviour and sleep

We will also generate harmonized data on other early-life stressors related to physical activity, sedentary behavior and sleep, in view of deriving additional integrated lifestyle indicators. Our major aim is to identify multi-behavioral patterns based on diet, physical activity, sedentary behavior, and sleep data collected in young children aged 2 to 4 years (before school entrance, i.e., in "preschoolers")[4]. The Inserm research team led by MA Charles will be responsible for harmonizing the data on children's physical activity, sedentary behavior and sleep across the

LifeCycle cohorts, and will then construct the combined indicators (multi-behavioral patterns) for preschoolers.

Physical activity and sleep data may include variables based on parents' reports (questionnaires, interviews) as well as constructed variables derived from accelerometry. Empirical data-driven methodologies (e.g., principal component, cluster or latent class analyses) will be used to identify multi-behavioral patterns. Different analytical methods may be tested depending on the type and distribution of variables. It is anticipated that complete harmonization will not be possible for all the behaviors/variables considered. Therefore, the multi-behavioral patterns will be derived 1/ by cohort, and 2/ on pooled data for a subset of variables/cohorts with perfectly harmonized data. If possible, analysis will be conducted separately for boys and girls.

To undertake those analyses, relevant data will need to be shared directly with INSERM, as it is expected that the current state of Datashield will not allow us to achieve our goals (particularly the multi-behavioral pattern derivation). Two main ways to proceed have been identified: Data can either be shared with INSERM with an ID allowing to link the subset with the original database, allowing us to give back directly to the original team the harmonized indicators, in order to merge them in the original database; or the data can be shared without ID, in which case we will provide the original team with R codes and protocols allowing for the reproduction of the analysis and construction of the harmonized indicators directly by the original team for their cohort.

We will include in the patterns some crucial diet behavior indicators, harmonized in each cohort under the University of Crete direction, as described in this document. As such, the INSERM team will also need to retrieve harmonized diet indicators derived in each cohort, under the same condition as the necessary data regarding sleep, sedentary behavior and physical activity. **As the INSERM team will undertake the analysis only for preschoolers, it would be useful for the participating cohorts to first harmonize diet behavior for the 2-4 years age group, defined above.**

Participating cohorts:

Twelve cohorts have already agreed to participate in this task:

ALSPAC BIB EDEN ELFE GECKO GenR INMA MOBA Piccolipiù RAINE RHEA SWS, but final status will depend of the exact age of data collection (preschool children) and the usual age at primary

school entrance in the different countries. As indicated above, participation also depends on reaching final agreement regarding data transfer/sharing with the Inserm team.

Proposed timeline:

- Transfer of data from each cohort to Inserm: Feb-August 2018
- Data description in preparation for harmonisation
    - On Elfe and Eden data: 15 March 2018
    - On other cohorts when data sets are retrieved: Oct 2018
- Inventory and comparison of available statistical methods for pattern analysis: May 2018
- Selection of method and pattern analysis: June 2018-June 2019

Planned papers:

- Energy balance-related multibehavioural patterns in European pre-schoolers: Description and associations with sociodemographic parental characteristics
- Energy balance-related multibehavioural patterns in European pre-schoolers: associations with parental overweight status and child adiposity development.

These papers will mobilize the 'core' variables harmonized by WP1. Analysis may be done centrally or remotely through DataShield if feasible.

Variables

Food groups

In **Annexes 2 and 3**, you can find examples of food products that belong to main food and sub-food groups, respectively.

Main Food Groups (servings/day)

1) Vegetables without potatoes
2) Fruits
3) Milk and milk products
4) Fish and fish products
5) Meat and meat products
6) Legumes, nuts and their products
7) Sugar, sugar products, chocolate products and confectionery
8) Egg and egg products
9) Grains and grain products

Sub-Groups (servings/day)

1) Low fat dairy
2) Fatty Fish
3) Non Fatty Fish
4) Red Meat
5) Processed meat
6) Whole grain cereals
7) Sugar-sweetened beverages
8) Potatoes
9) Savory biscuits and crisps

Nutrients

1) Total Daily Kcal intake (kcal/day): Calculated using the conversion factors: protein 4 kcal/g, fat 9 kcal/g, carbohydrate (available, expressed as monosaccharides) 3.75 kcal/g and alcohol 7 kcal/g.
2) Total Fat (gr/day)
3) Percentage of Total Fat (% of energy intake)
4) Saturated Fats (% of total food)
5) Polyunsaturated fats (% of total food)
6) Trans Fats (% of total food)
7) Total Protein (gr/day)
8) Percentage of Total Protein (% of energy intake)

9) Total carbohydrate (gr/day)
10) Percentage of Total carbohydrate (% energy intake)
11) Sodium (mg/day)

Diet quality index

DASH diet index: pregnant women and school-aged children (≥5 years old)

This index consists of eight food groups. Table 1 presents the food groups and a summary of food items per each group.

Table 1. Summary of included and excluded items for each of the DASH score components

Component	Summary of included items	Summary of excluded items
Vegetables	✓ Fresh vegetables, frozen or canned vegetables *(See Annex 1, for examples of this food group)*	× Potatoes
Fruits	✓ Fresh, frozen or canned fruits, fresh fruit juice *(See Annex 1, for examples of this food group)*	× Carbonated fruit drink with added sugars × Fruit juice with added sugars
Whole-grains	✓ See Annex 2, for examples of this sub-group	
Low-fat dairy products	✓ See Annex 2, for examples of this sub-group	
Nuts, legumes	✓ See Annex 1, for examples of this food group	
Red and processed meat	✓ See Annex 2, for examples of these sub-groups	
Sweetened beverages	✓Carbonated and noncarbonated sweetened beverages (soft drinks; eg, cola, fanta, sprite, seven up, dr pepper, orangina)	
Sodium		

Table 2 presents the scoring criteria of the DASH score. For each of the components, we need to classify individuals into quintiles of consumption frequencies. Component score for fruits, vegetables, nuts and legumes, low-fat dairy products, and whole grains is the participants' quintile ranking. For example, quintile 1 is assigned 1 point and quintile 5, 5 points. For sodium, red and processed meats, and sweetened beverages, low intake is desired. Therefore, the lowest quintile is given a score of 5 points and the highest quintile, 1 point. We then sum up the component scores to obtain an overall DASH score ranging from 8 to 40 (Table 1).

In Annex 3 you can find an example of STATA code for the construction of the proposed index.

Table 2. Scoring criteria for the DASH diet score

DASH diet score	Scoring Criteria
Dietary components for which greater intakes receive higher scores	
	Q1=1 point Q2=2 points Q3=3 points Q4=4 points Q5=5 points
Total fruit (servings/day)	
Vegetables without potatoes (servings/day)	
Whole grains (servings/day)	
Low-fat dairy products (servings/day)	
Nuts, legumes (servings/day)	
Dietary components for which lower intakes receive higher scores (Reverse scoring)	
	Q5=1 point Q4=2 points Q3=3 points Q2=4 points Q1=5 points
Red and processed meat (servings/day)	
Sugar-sweetened beverages (servings/day)	
Sodium (mg/day)	
Total score (points)	**8–40**

Dietary behaviour

1) Skipping Breakfast: *times/week*
2) Family dinner or dinner with at least one adult: *times/week*
3) Eating with the TV on: *times/week*
4) Main meals frequency: Average number of main meals per day. *Count breakfast, lunch and dinner.*
5) Snacking frequency: Average number of snacks per day. *Exclude breakfast, lunch and dinner.*
6) Visiting fast food restaurant: *times/week*

Supplements

1) Any supplements: *Yes/No*

Physical activity, sedentary behaviour and sleep

We will consider all available data on frequency/duration (e.g., times/min. per week), intensity (e.g., time spent in MVPA), and type/domains (e.g., outside play, organized sports in clubs or associations) of preschoolers' physical activity. If a specific and validated questionnaire has been

used, the reference should be provided. If the data were collected by accelerometry, the device and process used for data collection as well as the procedures used for data reduction should be described (or reference should be made to a published method or protocol article containing this information).

For sedentary behaviour, we will consider data on screen time (different types of electronic devices), non-screen sedentary behaviour (e.g., reading) as well as contextual data such as whether the child has a TV in his/her bedroom.

Regarding sleep, we will consider sleep duration, indicators of sleep quality as well as data on sleep routines.

It is anticipated that we will be able to at least derive the following harmonized indicators for the large majority of the participating cohorts, for the pre-schoolers age group:

1) Physical Activity:
   - Time spent playing outside (quantitative or semi qualitative value)
2) Sleep:
   - Time spent sleeping per day (night + day time; quantitative or semi qualitative value)
3) Sedentary behaviour:
   - Total screen time per day (Time spent watching TV / playing video games / using a tablet; a smartphone. Depending on the year of the study, and the available variables).

Key References

1. Fung TT, Chiuve SE, McCullough ML, Rexrode KM, Logroscino G, Hu FB. Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women. Arch Intern Med 2008;168:713-20.
2. Scientific report of the 2015 Dietary Guidelines Advisory Committee. Washington, DC: Department of Agriculture; 2015.
3. Moskal A, Pisa PT, Ferrari P, et al. Nutrient Patterns and Their Food Sources in an International Study Setting: Report from the EPIC Study. PloS one 2014;9.

4. Leech RM, McNaughton SA, Timperio A. The clustering of diet, physical activity and sedentary behavior in children and adolescents: a review. Int J Behav Nutr Phys Act 2014;11:4.

Annex 1: Food items included in each requested food group

1) **Vegetables** include:
   - **a.** *Leaf vegetables (Endive, Lettuce, Lamb's lettuce, Swiss chard, Spinach, Garden orache, Cress seedling, Mustard seedling, Land cress, Watercress, Vine leaf, Dandelion leaf, Nettle, Sorrel, Purslane,Parsley)*
   - ***b.*** Fruit vegetables *(Tomato, Aubergine, Sweet pepper, Chilli pepper, Cucumber, Courgette, Cucurbita squash, Other gourds, Ackee, Breadfruit, Matoki, Plantain, Avocado, Olive)*
   - ***c.*** Brassicas *(Broccoli, Broccoli tops, Cauliflower, Cabbage, Red cabbage, Chinese cabbage, Cabbage penca, Brussels sprouts, Brussels tops, Turnip tops, Kohlrabi, Curly kale)*
   - **d.** Pod and seed vegetables *(Sweet corn, Okra)*
   - **e.** Stalk vegetables *(Celery, Fennel, Sea kale, Rhubarb)*
   - ***f.*** Edible fungi *(Cultivated mushroom, Field mushroom, Honey mushroom, Boletus, Truffle, Morel, Cantharelle, Orange agaric, Oyster mushroom, Shiitake mushroom, Straw mushroom)*
   - ***g.*** Shoot vegetables *(Asparagus, Chicory, Globe artichoke, Bamboo shoot, Palm heart)*
   - ***h.*** Seaweeds *(Irish moss, Kombu, Laver, Wakame)*
   - ***i.*** Onion-family vegetables *(Onion, Spring onion, Shallot, Leek, Garlic, Chives)*
   - ***j.*** Vegetable mixtures *(Vegetable mixes, Mustard and cress, Pot-herb)*
   - ***k.*** Vegetable products *(Mushy peas, Garlic purée, Tomato purée, Vegetable purée, Pickled gherkins, Pickled onion, Pickled red cabbage, Sauerkraut)*
   - **l.** Root vegetables *(Carrot, Salsify, Celeriac, Parsnip, Turnip, Swede, Radish, Beetroot, Parsley root)*

2) **Fruits** include**:**
   - ***a.*** Malaceous fruit *(Dessert apple, Cooking apple, Pear, Nashi pear, Quince, Medlar, Loquat)*
   - ***b.*** Prunus species fruit *(Apricot, Peach, Nectarine, Plum, Damson, Mirabelle, Greengage, Sweet cherry, Sour cherry, Chickasaw plum, Susina, Sloe)*
   - ***c.*** Other stone fruit *(Date, Lychee, Persimmon plum, Barbados cherries)*
   - ***d.*** Berries *(White grapes, Black grapes, Strawberries, Raspberries, Loganberries, Blackberries, Dewberries, Cloudberries, Gooseberries, Black currants, Red currants, White currants, Cranberries, Bilberries, Cowberry, Blueberries, Elderberries, Rowanberries, Physalis, Mulberries, Bearberries, Sea buckthorn)*
   - ***e.*** Citrus fruit *(Lemon, Orange, Tangerine, Grapefruit, Pomelo, Lime, Kumquat)*
   - ***f.*** Miscellaneous fruit *(Banana, Pineapple, Kiwi fruit, Melon, Water melon, Fig, Mango, Pomegranate, Passionfruit, Cashew fruit, Guava, Papaya, Custard apple, Prickly pear, Rose hip, Sapodilla, Carambola, Durian, Jack fruit, Chayote, Rambutan, Tamarillo)*
   - ***g.*** Fruit mixtures *(Fruit cocktail, Fruit salad)*
   - ***h.*** Fruit products *(Dried mixed fruit, Mixed peel, Glacé cherry, Crystallised pineapple, Apple sauce, Cranberry sauce)*

i. Fresh fruit juice *(orange juice, grapefruit juice, grape juice, mixed fruit juice etc.)*

3) Milk and milk products include:

a. *Liquid milks*
b. *Processed milks (e.g. Chocolate-flavoured milk, Fruit-flavoured milk, Evaporated milk, Condensed milk, Dried milk, Filled milk, Buttermilk, Acidophilus milk, Whey),*
c. *Cream, Yogurt, Other fermented milk products (Alcoholic fermented milk products, Lactic fermented milk products)*
d. *Fresh cheese*
e. *Soft cheese*
f. *Hard cheese*
g. *Semi-hard cheese*
h. *Blue cheese*
i. *Smoked cheese*
j. *Processed cheese*
k. *Whey cheese*
l. *Imitation milk and cream (Soya milk, Non-dairy coffee creamer, Imitation cream, Soya yogurt, Soya cheese),*
m. *Milk beverage powders (Milk shake powder, Malt beverage powder, Drinking chocolate powder), Ices (Dairy ice cream, Non-dairy ice cream, Water ice, Granita, Sorbet)*

4) Fish and fish products include**:**

a. Clupeiformes *(Herring, Sprat, Sardine and pilchard, Anchovy, Shad, Salmon and trout, Char, Smelt, Whitefish)*
b. Perciformes *(Perch, Bass, Surgeon-fish, Mackeral, Tuna, Sea catfish and wolf-fish, Grey mullet)*
c. Gadiformes *(Cod and whiting, Hake)*
d. Pleuronectiformes *(Flounder, Halibut, Plaice, Sole)*
e. Cypriniformes and related sub-orders *(Roach, Carp, Babel, Bream)*
f. Other fish *(Eels, Zeomorphi, Lophiiformes, Selachoidei, Rays, Acipenseriformes)*
g. Crustaceans *(Crab, Lobster, Norway lobster, Prawns, Shrimps, Crawfish, Crayfish)*
h. Molluscs *(Squid, Octopus, Cuttlefish, Abalone, Clam, Cockle, Mussel, Oyster, Queen scallop, Scallop, Razor clam, Other bivalves, Limpet, Whelk, Winkle, Snail, Barnacle)*
i. Miscellaneous marine and aquatic foods *(Sea slug)*
j. Fish offal *(Herring roe, Salmon roe, Cod roe, Mullet roe, Lumpfish roe, Caviar, Hard roe from other fish, Herring milt, Milt from other fish)*
k. Dried and salted fish *(Dried cod, Bombay duck, Shark's fin, Jellyfish seaweed)*
l. Smoked fish *(Smoked herring, Smoked sprat, Smoked haddock, Smoked salmon and trout, Smoked mackeral, Smoked halibut, Smoked eel, Smoked sturgeon)*
m. Canned fish *(Canned herring, Canned sardine, Canned pilchard, Canned anchovy, Canned salmon, Canned mackeral, Canned tuna, Canned crab, Canned abalone, Canned mussels)*

n. Pickled fish *(Matjes herring, Rollmop herring)*
o. Restructured fish and fish analogues *(Crabsticks)*
p. Fish products *(Fishballs, Fishcakes, Fish fingers, Fish paste, Fish pâté, Taramasalata)*
q. Marine mammals
r. Amphibians
s. Reptiles

5) Meat and meat products include**:**

a. Beef, carcass meat *(Tenderloin, Striploin, Fore-rib, Topside, Silverside, Shoulder clod, Chuck tender)*
b. Veal, carcass meat *(Veal, cuts)*
c. Pork / piglet, carcass meat *(Loin, Tenderloin, Neck, Belly, Chump, Leg)*
d. Mutton / lamb, carcass meat *(Mutton / lamb, cuts)*
e. Mammals, other *(Horse, Goat / kid, Rabbit, Hare, Wild pig, boar, Venison, Elk, Reindeer, Chamois, Kangaroo)*
f. Chicken *(Chicken breast, Chicken leg, Chicken wing)*
g. Turkey *(Turkey breast, Turkey leg, Turkey wing)*
h. Birds, other *(Duck, Goose, Pigeon, Guinea fowl, Pheasant, Partridge, Quail, Snipe, Grouse, Ptarmigan, Ostrich)*
i. Liver *(Beef liver, Veal liver, Pork liver, Mutton / lamb liver, Chicken liver, Turkey liver, Duck liver, Goose liver)*
j. Kidney *(Beef kidney, Veal kidney, Pork kidney, Mutton / lamb kidney)*
k. Other offal *(Tongue, Heart, Brain, Lungs, Stomach, Intestines, Pancreas, Spleen, Thymus, Marrowbone, Tail, Totters and feet, Giblets)*
l. Preserved meats *(Ham, Bacon, Preserved beef, Tongue (preserved), Preserved poultry)*
m. Restructured meat and meat analogues
n. Meat products *(Pastes, pâtés and terrines, Minced meat products, Dry, smoked sausages (Rohwurst), Fresh and lightly cooked sausages (Bratwurst), Cooked sausages (Kochwurst), Other meat products, Blood and offal products)*
o. Meat dishes *(Meat burger, Meat balls, Meat pasty, Meat pie)*

6) Legumes, nuts and products include**:**

a. *Legumes (Dried pea, Chick pea, Dried broad bean, Lentil, Common bean, Dried lima bean, Mung bean, Urd bean, Black eye bean, Soya beans, Tofu, Carob fruit, Lupin(e)s, Unfermented soya paste, Fermented soya paste)*
b. *Nuts (Walnut, Hazelnut, Filbert, Coconut, Brazil nut, Hickory nut, Cashew nut, Almond, sweet, Almond, bitter, Pistachio nut, Sweet chestnut, peanut)*
c. *Nut products (Peanute butter, Chestnut puré)*

7) Sugar, sugar products, chocolate products and confectionery include***:***

a. *Sugar (sucrose) (White sugar, Brown sugar, Sugar crystals)*
b. *Other sugars (Glucose, Fructose, Malt sugar, Milk sugar)*

- **c.** *Sugar substitutes (Non-nutritive sweeteners, Nutritive sweeteners)*
- **d.** *Honey*
- **e.** *Syrups (Molasses, Black treacle, Golden syrup, Cane syrup, Maple syrup, Fruit syrup, Glucose syrup, Sugar syrup)*
- **f.** *Jams, marmalades and spreads (Fruit jam, Fruit jelly preserve, Marmalade)*
- **g.** *Jelly*
- **h.** *Non-chocolate dessert topping*
- **i.** *Chocolate and chocolate products (Cocoa powder, Slab cooking chocolate, Milk chocolate bar, Plain chocolate bar, White chocolate bar, Chocolate flake, Chocolate button, Chocolate egg, Filled chocolate, Other chocolate goods)*
- **j.** *Chocolate-coated confectionery bars*
- **k.** *Non-chocolate confectionery (Boiled sweet, Chew sweet, Gum sweet, Liquorice sweet, Mint sweet, Sherbet sweet, Fudge, Toffee, Marshmallow, Nougat, Turkish delight, Cereal chewy bar, Cereal crunchy bar, Chewing gum)*
- **l.** *Sugar products (Marzipan, Marzipan fruit, Candied fruit, Candied angelica, Preserved ginger)*

8) *Egg and egg products* include*:*

- **a.** *Chicken eggs (Egg yolk, Egg white)*
- **b.** *Turkey eggs*
- **c.** *Duck eggs*
- **d.** *Goose eggs*
- **e.** *Quail eggs*
- **f.** *Ostrich eggs*
- **g.** *Seagull eggs*
- **h.** *Egg products (Scotch egg)*
- **i.** *Egg dishes (Soufflé, Meringue, Egg nog)*

3) *Grains and grain products* include:

- **a.** *Wheat basic products (Whole grain wheat, Wheat flour)*
- **b.** *Rye basic products (Rye flour)*
- **c.** *Oats basic products (Groats, Rolled oats, Oatmeal, Oatflour)*
- **d.** *Barley basic products (Whole grain Barley, Pearl Barley, Barley flakes, Barley meal, Barley flour)*
- **e.** *Maize basic products (Hominy, Maize rice, Cornmeal, Maize flour, Cornflour, Custard powder)*
- **f.** *Rice basic products (Rice flour, Rice flakes, Brown rice, Basmati rice, Carolina rice, Glutinous rice, Parboiled rice, Polished rice, Red rice, Wild rice)*
- **g.** *Basic products of other cereals (Buckwheat, Buckwheat flour, Millet, Millet flour, Sorghum, Spelt)*
- **h.** *Substitute flours and starches (Soya flour, Potato flour, Carob flour, Lotus root flour, Arrowroot, Sago, Tapioca)*

i. *Pasta and noodles (Dried main-dish pasta, Dried minature pasta, Fresh main-dish pasta, Fresh minature pasta, Egg noodles, Plain noodles, Rice noodles, Transparent noodles),*
j. *Leavened breads (Wheat bread, Naan bread, Soda bread, Rye bread, Potato bread, Other flour Bread, Mixed flour Bread)*
k. *Unleavened breads and crispbreads (Bannock, Pitta bread, Matzo, Tortilla, Crispbread, rye, Crispbread, wheat)*
l. *Bread products (Breadcrumbs, Rusks, Bread stuffing, Bread pudding),*
m. *Fine bakery wares (Savoury biscuits, Sweet biscuits and cookies, Croissants, Currant bun, Dough cakes, Scone, Doughnut, Danish pastry, Greek pastry, Tart, Mince pie, Baked cake, Cream cake, Sponge cake)*
n. *Savoury cereal dishes (Dumpling, Savoury pancake, Pizza, Savoury pie),*
o. *Sweet puddings (Custard, Trifle, Fruit crumble, Fruit pie, Milk pudding, Rice pudding, Sponge pudding, Suet pudding)*
p. *Breakfast cereals (Cereals, wheat based, Cereals, rye based, Cereals, maize based, Cereals, oats based, Cereals, rice based, Cereals, mixed grain, Muesli)*

Annex 2: Food items included in each requested food subgroups

1) ***Low fat dairy*** include:

   a. Liquid milks *(Semi-skimmed Milk (1 - 2.9% fat), Skimmed Milk (< 1% fat))*
   b. Yogurt *(Yogurt 1 - 3% fat, Yogurt < 1% fat)*
   c. *Fat-free cheese*
   d. *Low fat cheese (<5% fat, eg cottage cheese)*
   e. *Fat-free cream or cream cheeses*
   f. *Low fat cream or cream cheeses (<5% fat, eg philadephia low-fat soft cheese)*
   g. Imitation milk and *cream (eg, Soya milk, Soya yogurt, Soya cheese)*

2) ***Fatty Fish*** include:

   a. Clupeiformes *(Herring, Sprat, Sardine and pilchard, Anchovy, Shad, Salmon and trout, lake Whitefish)*
   b. Perciformes *(Mackerel, Tuna)*
   c. Cypriniformes and related sub-orders *(Carp)*
   d. Other fish *(Eels)*
   e. Fish offal *(Herring roe, Salmon roe, Herring milt)*
   f. Smoked fish *(Smoked herring, Smoked sprat, Smoked salmon and trout, Smoked mackerel, Smoked eel)*
   g. Canned fish *(Canned herring, Canned sardine, Canned anchovy, Canned salmon, Canned mackerel, Canned tuna (albacore), Canned pilchard)*
   h. Pickled fish *(Matjes herring, Rollmop herring)*

3) ***Non Fatty Fish*** include:

   a. Clupeiformes *(Smelt)*
   b. Perciformes *(Perch, Bass, Surgeon-fish, Sea catfish and wolf-fish, Grey mullet)*
   c. Gadiformes *(Cod and whiting, Hake)*
   d. Pleuronectiformes *(Flounder, Halibut, Plaice, Sole)*
   e. Cypriniformes and related sub-orders *(Roach, Babel, Bream)*
   f. Other fish *(Zeomorphi, Lophiiformes, Selachoidei, Rays, Acipenseriformes)*
   g. Fish offal *(Cod roe, Mullet roe, Lumpfish roe, Caviar, Hard roe from other fish, Milt from other fish)*
   h. *Canned light tuna (skipjack)*
   i. Dried and salted fish *(Dried cod, Bombay duck, Shark's fin, Jellyfish seaweed)*
   j. Smoked fish *(Smoked haddock, Smoked, Smoked halibut, Smoked eel, Smoked sturgeon)*
   k. Fish products *(Fishballs, Fishcakes, Fish fingers, Fish paste, Fish pâté, Taramasalata)*

4) ***Red Meat*** include:

   a. *Beef, carcass meat (Tenderloin, Striploin, Fore-rib, Topside, Silverside, Shoulder clod, Chuck tender)*
   b. *Veal, carcass meat (Veal, cuts)*
   c. *Pork / piglet, carcass meat (Loin, Tenderloin, Neck, Belly, Chump, Leg)*
   d. *Mutton / lamb, carcass meat (Mutton / lamb, cuts)*

e. *Mammals, other (Horse, Goat / kid, Hare, Wild pig, boar, Venison, Elk, Reindeer, Chamois, Kangaroo)*
f. *Liver (Beef liver, Veal liver, Pork liver, Mutton / lamb liver, Chicken liver, Turkey liver, Duck liver, Goose liver)*
g. *Kidney (Beef kidney, Veal kidney, Pork kidney, Mutton / lamb kidney)*
h. *Other offal (Tongue, Heart, Brain, Lungs, Stomach, Intestines, Pancreas, Spleen, Thymus, Marrowbone, Tail, Totters and feet, Giblets)*
i. *Meat dishes (Meat burger, Meat balls, Meat pasty, Meat pie)*

5) ***Processed meat*** include:

a. *Preserved meats (Ham, Bacon, Preserved beef, Tongue (preserved), Preserved poultry)*
b. *Restructured meat and meat analogues*
c. *Meat products (Pastes, pâtés and terrines, Minced meat products, Dry, smoked sausages (Rohwurst), Fresh and lightly cooked sausages (Bratwurst), Cooked sausages (Kochwurst), Other processed meat products, Blood and offal products)*

6) ***Whole grain cereals*** include:

a. *Wheat products (Whole grain wheat, Bulgar; wholemeal, brown)*
b. *Rye basic products (Rye flour, whole)*
c. *Oats basic products (Groats, Rolled oats, Oatmeal, Oatflour)*
d. *Barley basic products (Whole grain, barley; Pearl, barley; Barley flakes, Barley meal, Barley flour)*
e. *Brown rice, wild rice*
f. *Basic products of other cereals (Buckwheat, Buckwheat flour, Millet, Millet flour, Sorghum, Spelt)*
g. *Leavened breads (Wheat bread wholemeal; wheat bread brown; rye bread dark)*
h. *Breakfast cereals (Cereals, whole wheat based; Cereals, rye based; Cereals, oats based; Cereals, mixed grain, Muesli)*

7) ***Sugar sweetened beverages*** include*:*

a. *Carbonated soft drinks: Carbonated lemonade (Seven Up, Sprite)Carbonated fruit drink (Fanta, Orangina, Tango, Cherryade), Ginger ale, Root beer, Cola, Dr Pepper, Lucozade*
b. *Fruit nectars, packed fruit juice*

8) ***Potatoes*** include:

a. ***Tumbers****: New potato, Main-crop potato, Jerusalem artichoke, Sweet potato, Yam, Cassava, Taro*

9) ***Savory biscuits and crisps*** include:

a. *Savoury biscuits*
b. *Potato-based snacks, Maize-based snacks, Wheat-based snacks, Rice-based snacks*

Annex 3: Example of Stata code for the DASH index during pregnancy

```
* Calculate red and processed meat as the sum of the two food subgroups (red meat and processed meat)

gen rpmeat_pgn= redmeat_pgn+ procmeat_pgn

*Create quintiles for each component

xtile fruitpgnQ= fruit_pgn, nq(5)
xtile vegpgnQ=veg_pgn, nq(5)
xtile wgrainspgnQ=whgrains_pgn, nq(5)
xtile lfdairypgnQ=lfdairy_pgn, nq(5)
xtile pulsespgnQ=pulses_pgn, nq(5)
xtile rpmeatpgnQ=rpmeat_pgn, nq(5)
xtile ssbevpgnQ=swebev_pgn, nq(5)
xtile napgnQ=na_pgn, nq(5)


*Reverse scoring for the 3 detrimental items
gen rpmeatpgnR=1 if rpmeatpgnQ ==5
replace rpmeatpgnR =2 if rpmeatpgnQ ==4
replace rpmeatpgnR =3 if rpmeatpgnQ ==3
replace rpmeatpgnR =4 if rpmeatpgnQ ==2
replace rpmeatpgnR =5 if rpmeatpgnQ ==1

gen ssbevpgnR=1 if ssbevpgnQ ==5
replace ssbevpgnR =2 if ssbevpgnQ ==4
replace ssbevpgnR =3 if ssbevpgnQ ==3
replace ssbevpgnR =4 if ssbevpgnQ ==2
replace ssbevpgnR =5 if ssbevpgnQ ==1

gen napgnR=1 if  napgnQ ==5
replace napgnR=2 if  napgnQ ==4
replace napgnR =3 if napgnQ ==3
replace napgnR =4 if  napgnQ ==2
replace napgnR =5 if  napgnQ ==1

*Sum all the scores to generate the DASH index
gen dash_pgn= fruitpgnQ +vegpgnQ + wgrainspgnQ + lfdairypgnQ + pulsespgnQ + rpmeatpgnR +
ssbevpgnR + napgnR

*All intermediate variables can be discarded
drop fruitpgnQ  vegpgnQ  wgrainspgnQ  lfdairypgnQ  pulsespgnQ  rpmeatpgnQ  ssbevpgnQ  napgnQ
rpmeatpgnR  ssbevpgnR  napgnR
```

Instructions for Quality control

of Harmonized WP3.1.4 Dietary and Lifestyle Variables

IMPORTANT NOTE: There were two different approaches for data harmonisation and transfer within WP3.1.4. For dietary variables, participating cohorts were asked to harmonise their data locally, and then to upload the relevant harmonized data on their own Opal server. Regarding lifestyle variables, the planned analysis required the necessary behavioral data to be transferred to INSERM and the participating cohorts have agreed to share their data under various conditions, in accordance with their regulation. After deriving the multi-behavioral patterns centrally by the INSERM team, participating cohorts received the relevant files for reconstructing and/or uploading the relevant variables on their own Opal server.

Step1: Verify list of variables and formats

Please, verify that your cohort-specific harmonized dietary and lifestyle variables completely match with the information provided in the WP3.1.4 Variable Lists (attached at the end of this document).

The *variable name* , *data type* and *unit* of each of the variables must correspond exactly to the WP3.1.4 Variable Lists. Also, please check the requirements for *type of harmonization (complete, partial, na)*. Variables considered "*completely harmonised*" must match the information provided in the file: *LifeCycle_Subtask_3.1.4_Harmonisation_Protocol.pdf*

Additionally, the INSERM team has shared with the cohorts two Excel files with the cohort-specific information for the derived variables (*LifeCycle_COHORT_SourceVariables.xlsx* & *LifeCycle_COHORT_Harmonizations.xlsx*). Please refer to **those** files for the following quality control checks on lifestyle variables.

If any mismatch is observed, please correct the errors accordingly.

Example 1:

1. Check the variable name (i.e. "*dietass_pgn*")
2. Check that values are *Categorical*
3. Check the coding is in accordance with the specified *values*

Example 2:

1. Check the variable name (i.e. "*veg_pgn*")
2. Check that values are *Decimal*
3. Check that values are expressed in *servings/day*

variable	label	datatype	values	unit
dietass_pgn	Method of dietary assessment in pregnancy	Categorical	1=FFQ 2=Recall diary 3=Other	
veg_pgn	Vegetables without potatoes during pregnancy	Decimal		servings/day

Note: all variables are harmonised for prenatal period (i.e. pregnancy) and postnatally (preschool and shool aged children). Hence, examples presented in this document for pregnancy, also apply to corresponding postnatal variables.

Step 2: Check distribution of variables

Before running any of the following checks, please ensure that *child_id* variable has no duplicates.

Please, generate distributions for all variables, and check for outliers and improbable values. Also, for variables that have been reported in papers and/or publications of your cohort, verify that distributions or summary statistics of the harmonized LifeCycle variable match those of the reported/published variables.

For categorical variables, check that there are no improbable values, i.e. values not corresponding to the categories defined in the Variable List. Please, correct errors where relevant.

Example1: Method of dietary assessment in pregnancy

Check that data is coded into a maximum of 3 categories, and falls into the range 1-3:

```
. tab dietass_pgn
```

Method of dietary assesment in pregnancy	Freq.	Percent	Cum.
1	1,104	100.00	100.00
Total	1,104	100.00	

Example2: Supplements intake during pregnancy

Check that data is binary, and takes values 0-1:

```
. tab supp_pgn
```

Supplements during pregnancy	Freq.	Percent	Cum.
0	87	6.83	6.83
1	1,187	93.17	100.00
Total	1,274	100.00	

For continuous variables, check that there are no outliers, i.e. values out of the expected range for cohort and/or national daily consumption distributions for each food group. For example, no negative values should exist in daily consumption of harmonised food items or food groups, sleep time or time spent watching TV; minimum expected value is 0. Additionally, use your scientific knowledge and practical sense when making this quality check so as not to

drop interesting outliers. Probable errors may be caused by an error in your harmonization script, please check carefully and correct where required.

Example 1: Vegetables without potatoes during pregnancy

Check whether values of veg_pgn are in the expected range (example from Rhea cohort data).

```
. sum veg_pgn
```

Variable	Obs	Mean	Std. Dev.	Min	Max
veg_pgn	1069	3.953734	1.944816	.066	19.414

Example 2: Time spent playing outdoor and time spent watching tv

Check whether values of outdoorp_psc and tv_psc are in the expected range.

Notice that outdoorp_psc, that has been returned by the INSERM team, is a z-score with a mean close to 0 and standard deviation close to 1, and tv_spc is expressed as h/day (example from Rhea cohort data).

```
. sum  tv_psc
```

Variable	Obs	Mean	Std. Dev.	Min	Max
tv_psc	857	1.31573	.939187	0	6

Step 3: Check internal validation

Internal validation is an important part of the local quality control. Thus, within reason cross-check the variables against other variables to check for consistency.

Example: Fish consumption

There are three variables related to daily fish consumption. "*Fish and fish products during pregnancy*" cannot be less than its subgroups "*Fatty*" and "*Non Fatty Fish*" during pregnancy.

fish_pgn	Fish and fish products during pregnancy
ffish_pgn	Fatty Fish during pregnancy
nffish_pgn	Non Fatty Fish during pregnancy

```
. sum fish_pgn ffish_pgn nffish_pgn
```

Variable	Obs	Mean	Std. Dev.	Min	Max
fish_pgn	1079	.1876562	.3380482	0	10
ffish_pgn	1079	.087823	.13555	0	3
nffish_pgn	1079	.0622308	.1156006	0	3

Step 4: Complete the Online Catalogue

As part of the Quality Control, each cohort must also ensure that the Online Catalogue is aligned accordingly. This means that in the Online Catalogue description of harmonization is complete and information in all three tabs (*description*, *variables used* and *script/syntax*) are completed in full.

For dietary variables each cohort please complete and upload to the Catalogue the harmonized dietary variables and the corresponding source variables (Excel templates to be used have been shared via email for your convenience). Please follow the same procedure as for variables from other WPs. For the lifestlyle variables, please use the files sent by INSERM (described in Step 1 of this QC manual).

LifeCycle WP3.1.4 Diet Variable list

variable	label	datatype	values	unit
child_id	Child identifier	Integer		
dietass_pgn	Method of dietary assesment in pregnancy	Categorical	1=ffq, 2=recall diary, 3=other	
veg_pgn	Vegetables without potatoes during pregnancy	Decimal		servings/day
fruit_pgn	Fruits during pregnancy	Decimal		servings/day
dairy_pgn	Milk and milk products during pregnancy	Decimal		servings/day
fish_pgn	Fish and fish products during pregnancy	Decimal		servings/day
meat_pgn	Meat and meat products during pregnancy	Decimal		servings/day
pulses_pgn	Legumes, nuts and their products during pregnancy	Decimal		servings/day
sugar_pgn	Sugar, sugar products, chocolate products and confectionery during pregnancy	Decimal		servings/day
egg_pgn	Egg and egg products during pregnancy	Decimal		servings/day
grain_pgn	Grains and grain products during pregnancy	Decimal		servings/day
lfdairy_pgn	Low fat dairy during pregnancy	Decimal		servings/day
ffish_pgn	Fatty Fish during pregnancy	Decimal		servings/day
nffish_pgn	Non Fatty Fish during pregnancy	Decimal		servings/day
redmeat_pgn	Red meat during pregnancy	Decimal		servings/day
procmeat_pgn	Processed meat during pregnancy	Decimal		servings/day
whgrains_pgn	Whole grain cereals during pregnancy	Decimal		servings/day
swebev_pgn	Sugar-sweetened beverages during pregnancy	Decimal		servings/day
potat_pgn	Potatoes during pregnancy	Decimal		servings/day
sav_pgn	Savory biscuits and crisps during pregnancy	Decimal		servings/day
dietga_pgn	Gestational week at dietary assesment	Decimal		weeks
veg_psc	Vegetables without potatoes in preschool age	Decimal		servings/day

variable	label	datatype	values	unit
fruit_psc	Fruits in preschool age	Decimal		servings/day
dairy_psc	Milk and milk products in preschool age	Decimal		servings/day
fish_psc	Fish and fish products in preschool age	Decimal		servings/day
meat_psc	Meat and meat products in preschool age	Decimal		servings/day
pulses_psc	Legumes, nuts and their products in preschool age	Decimal		servings/day
sugar_psc	Sugar, sugar products, chocolate products and confectionery in preschool age	Decimal		servings/day
egg_psc	Egg and egg products in preschool age	Decimal		servings/day
grain_psc	Grains and grain products in preschool age	Decimal		servings/day
lfdairy_psc	Low fat dairy in preschool age	Decimal		servings/day
ffish_psc	Fatty Fish in preschool age	Decimal		servings/day
nffish_psc	Non Fatty Fish in preschool age	Decimal		servings/day
redmeat_psc	Red meat during in preschool age	Decimal		servings/day
procmeat_psc	Processed meat in preschool age	Decimal		servings/day
whgrains_psc	Whole grain cereals in preschool age	Decimal		servings/day
swebev_psc	Sugar-sweetened beverages in preschool age	Decimal		servings/day
potat_psc	Potatoes in preschool age	Decimal		servings/day
sav_psc	Savory biscuits and crisps in preschool age	Decimal		servings/day
dietass_psc	Method of dietary assesment in preschool age	Categorical	1=ffq, 2=recall diary, 3=other	
dietage_psc	Exact age at dietary assesment in preschool age	Decimal		years
veg_sch	Vegetables without potatoes in school-age children	Decimal		servings/day
fruit_sch	Fruits in school-age children	Decimal		servings/day
dairy_sch	Milk and milk products in school-age children	Decimal		servings/day
fish_sch	Fish and fish products in school-age children	Decimal		servings/day
meat_sch	Meat and meat products in school-age children	Decimal		servings/day

variable	label	datatype	values	unit
pulses_sch	Legumes, nuts and their products in school-age children	Decimal		servings/day
sugar_sch	Sugar, sugar products, chocolate products and confectionery in school-age children	Decimal		servings/day
egg_sch	Egg and egg products in school-age children	Decimal		servings/day
grain_sch	Grains and grain products in school-age children	Decimal		servings/day
lfdairy_sch	Low fat dairy in school-age children	Decimal		servings/day
ffish_sch	Fatty Fish in school-age children	Decimal		servings/day
nffish_sch	Non Fatty Fish in school-age children	Decimal		servings/day
redmeat_sch	Red meat during in school-age children	Decimal		servings/day
procmeat_sch	Processed meat in school-age children	Decimal		servings/day
whgrains_sch	Whole grain cereals inschool-age children	Decimal		servings/day
swebev_sch	Sugar-sweetened beverages in school-age children	Decimal		servings/day
potat_sch	Potatoes in school-age children	Decimal		servings/day
sav_sch	Savory biscuits and crisps in school-age children	Decimal		servings/day
dietass_sch	Method of dietary assesment in school-age children	Categorical	1=ffq, 2=recall diary, 3=other	
dietage_sch	Exact age at dietary assesment of school-age children	Decimal		years
kcal_pgn	Total Daily Kcal intake during pregnancy	Decimal		Kcal/day
totfat_pgn	Total Fat intake during pregnancy	Decimal		gr/day
percfat_pgn	Percentage of Total Fat intake during pregnancy	Decimal		% energy intake
satfat_pgn	Saturated Fats intake during pregnancy	Decimal		% total food
pufas_pgn	Polyunsaturated fats intake during pregnancy	Decimal		% total food
transfat_pgn	Trans Fats intake during pregnancy	Decimal		% total food
totprot_pgn	Total Protein intake during pregnancy	Decimal		gr/day
percprot_pgn	Percentage of Total Protein intake during pregnancy	Decimal		% energy intake
totcarb_pgn	Total carbohydrate intake during pregnancy	Decimal		gr/day

variable	label	datatype	values	unit
perccarb_pgn	Percentage of Total carbohydrate intake during pregnancy	Decimal		% energy intake
na_pgn	Sodium intake during pregnancy	Decimal		mg/day
kcal_psc	Daily Kcal intake in preschool age	Decimal		Kcal/day
totfat_psc	Total Fat intake in preschool age	Decimal		gr/day
percfat_psc	Total Fat intake in preschool age	Decimal		% energy intake
satfat_psc	Saturated Fats intake in preschool age	Decimal		% total food
pufas_psc	Polyunsaturated fats intakein preschool age	Decimal		% total food
transfat_psc	Trans Fats intake in preschool age	Decimal		% total food
totprot_psc	Total Protein intake in preschool age	Decimal		gr/day
percprot_psc	Total Protein intake in preschool age	Decimal		% energy intake
totcarb_psc	Total carbohydrate intake in preschool age	Decimal		gr/day
perccarb_psc	Total carbohydrate intake in preschool age	Decimal		% energy intake
na_psc	Sodium intake in preschool age	Decimal		mg/day
kcal_sch	Daily Kcal intake in school-age children	Decimal		Kcal/day
totfat_sch	Total Fat intake in school-age children	Decimal		gr/day
percfat_sch	Total Fat intake in school-age children	Decimal		% energy intake
satfat_sch	Saturated Fats intake during pregnancy	Decimal		% total food
pufas_sch	Polyunsaturated fats intake in school-age children	Decimal		% total food
transfat_sch	Trans Fats intake in school-age children	Decimal		% total food
totprot_sch	Total Protein intake in school-age children	Decimal		gr/day
percprot_sch	Total Protein intake in school-age children	Decimal		% energy intake
totcarb_sch	Total carbohydrate intake in school-age children	Decimal		gr/day
perccarb_sch	Total carbohydrate intake in school-age children	Decimal		% energy intake
na_sch	Sodium intake in school-age children	Decimal		mg/day
dash_pgn	Fung's DASH diet index in pregnancy	Integer		
dash_sch	Fung's DASH diet index in school-aged children	Integer		

variable	label	datatype	values	unit
skipbreakf_psc	Skipping Breakfast	Decimal		times/week
famdinner_psc	Family dinner or dinner with at least one adult	Decimal		times/week
tveat_psc	Eating with the TV on	Decimal		times/week
mainmeal_psc	Main meals frequency per day	Decimal		times/day
snacks_psc	Snacking frequency per day	Decimal		times/day
fastfood_psc	Visiting fast food restaurant	Decimal		times/week
skipbreakf_sch	Skipping Breakfast	Decimal		times/week
famdinner_sch	Family dinner or dinner with at least one adult	Decimal		times/week
tveat_sch	Eating with the TV on	Decimal		times/week
mainmeal_sch	Main meals frequency per day	Decimal		times/day
snacks_sch	Snacking frequency per day	Decimal		times/day
fastfood_sch	Visiting fast food restaurant	Decimal		times/week
supp_pgn	Supplements during pregnancy	Binary	0=No, 1=Yes	
supp_psc	Supplements in preschool age	Binary	0=No, 1=Yes	
supp_sch	Supplements in school-age children	Binary	0=No, 1=Yes	

LifeCycle WP3.1.4 Lifestyle Variable list

variable	label	datatype	values	unit
sleept_psc	sleep (day + night), h/day, preschool age	Decimal		h/day
sleeptage_psc	Age at sleept_psc collection	Decimal		years
outdoorp_psc	Season adjusted time spent playing outdoors Z score preschool age	Decimal		
outdoorpage_psc	Age at outdoop_psc collection	Decimal		years
tv_psc	time spent watching TV, h/day, preschool age	Decimal		h/day
screenoth_psc	time spent watching screens (except TV) h/day, preschool age	Decimal		h/day

screenage_psc	Age at tv_psc and screenoth_psc collection	Decimal		years
patternA_snackscreen_psc	child's score on the snack screen pattern (relative, derived from PCA loadings), preschool age	Decimal		
patternB_psc	child's score on the second multibehavioral pattern (relative, derived from PCA loadings), preschool age	Decimal		