Prospective Associations Between Socio-economic Status and Dietary Patterns in European Children

The Identification and Prevention of Dietary- and Lifestyle-induced Health Effects in Children and Infants (IDEFICS) Study

Fernandez-Alvira, Juan M.; Börnhorst, Claudia; Bammann, Karin; Gwozdz, Wencke; Krogh, Vittorio; Hebestreit, Antje; Barba, Gianvincenzo; Reisch, Lucia; Eiben, Gabriele; Iglesia, Iris; Veidebaum, Toomas; Kourides, Yiannis; Kovács, Eva; Huybrechts, Inge; Pigeot, Iris; Moreno, Luis A.

Document Version
Accepted author manuscript

Published in:
British Journal of Nutrition

DOI:
10.1017/S0007114514003663

Publication date:
2015

License
Unspecified

Citation for published version (APA):
Fernandez-Alvira, J. M., Börnhorst, C., Bammann, K., Gwozdz, W., Krogh, V., Hebestreit, A., Barba, G., Reisch, L., Eiben, G., Iglesia, I., Veidebaum, T., Kourides, Y., Kovács, E., Huybrechts, I., Pigeot, I., & Moreno, L. A. (2015). Prospective Associations Between Socio-economic Status and Dietary Patterns in European Children: The Identification and Prevention of Dietary- and Lifestyle-induced Health Effects in Children and Infants (IDEFICS) Study. *British Journal of Nutrition*, *113*(3), 517-525. https://doi.org/10.1017/S0007114514003663

Prospective Associations Between Socio-economic Status and Dietary Patterns in European Children: The Identification and Prevention of Dietary- and Lifestyle-induced Health Effects in Children and Infants (IDEFICS) Study

Juan Miguel Frenandez-Alviraa, Claudia Börnhorst, Karin Bammann, Wencke Gwozdz, Vittorio Krogh, Antje Hebestreit, Gianvincenzo Barba, Lucia Reisch, Gabriele Eiben, Iris Iglesia, Toomas Veidebaum, Yiannis Kourides, Eva Kovács, Inge Huybrechts, Iris Pigeot, and Luis A. Moreno

PROSPECTIVE ASSOCIATIONS BETWEEN SOCIO-ECONOMIC STATUS AND DIETARY PATTERNS IN EUROPEAN CHILDREN: THE IDEFICS STUDY

Juan Miguel Fernández-Alvira[1,2], Claudia Börnhorst[2], Karin Bammann[3], Wencke Gwozdz[4], Vittorio Krogh[5], Antje Hebestreit[2], Gianvincenzo Barba[6], Lucia Reisch[4], Gabriele Eiben[7], Iris Iglesia[1], Tomas Veidebaum[8], Yannis A Kourides[9], Eva Kovacs[10], Inge Huybrechts[11,12], Iris Pigeot[2,13], Luis A Moreno[1] on behalf of the IDEFICS consortium

[1]GENUD (Growth, Exercise, NUtrition and Development) Research Group, Faculty of Health Sciences, University of Zaragoza, Spain

[2]Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

[3]Institute for Public Health and Nursing Sciences (ipp), University of Bremen, Bremen, Germany

[4]Copenhagen Business School, Copenhagen, Denmark

[5]Department of Preventive and Predictive Medicine, Epidemiology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[6]Unit of Epidemiology and Population Genetics, Institute of Food Sciences, National Research Council, Avellino, Italy

[7]Department of Public Health and Community Medicine, Public Health Epidemiology Unit, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[8]Department of Chronic Diseases, National Institute for Health Development, Tallinn, Estonia

[9]Research and Education Institute of Child Health, Strovolos, Cyprus

[10]Department of Paediatrics, University of Pécs, Pécs, Hungary

[11]Department of Public Health, Ghent University, Ghent, Belgium

[12]International Agency for Research on Cancer, Dietary Exposure Assessment Group, Lyon, France

[13]Institute of Statistics, Faculty of Mathematics and Computer Science, University Bremen, Bremen, Germany

Corresponding Author: Juan Miguel Fernández Alvira

GENUD (Growth, Exercise, NUtrition and Development) Research Group. Faculty of Health Sciences, University of Zaragoza, Spain. Edificio del SAI. C/Pedro Cerbuna, s/n. 50009 Zaragoza. Spain

Phone number: 0034 876 55 40 91; Fax Number: 0034 876 55 40 9

e-mail: juanfdez@unizar.es

Running title: SES and dietary patterns in European children

Key words: cluster analysis, dietary behaviour, food frequency questionnaire, income, maternal education, paternal education

ABSTRACT

Exploring changes in children's diet over time and its relation to SES may help to understand the impact of social inequalities on dietary patterns. The current study aims to describe dietary patterns applying cluster analysis among 9,301 children participating in the baseline (2 to 9 years old) and follow-up (4 to 11 years old) surveys of the IDEFICS study and to describe the cluster memberships of children over time and their association with SES. We applied K-means clustering based on the similarities between the relative frequencies of consumption of 42 food items. Three consistent clusters were obtained at baseline and follow-up: "processed" (higher frequency of consumption of snacks and fast food), "sweet" (higher frequency of consumption of sweet foods and sweetened drinks) and "healthy" (higher frequency of consumption of fruits, vegetables and wholemeal products). The most stable pattern was the "healthy" cluster (85% of children allocated in this cluster at baseline remained in the same cluster at follow-up), followed by the "sweet" cluster. Only 46% of children allocated in the "processed" cluster at baseline remained in the same cluster at follow-up. Children with higher educated mothers and fathers and the highest income were more likely to be allocated in the "healthy" cluster at baseline and follow-up, and less likely to be allocated in the sweet cluster. Migrants were more likely to be allocated in the "processed" cluster at baseline and follow-up. Applying cluster analysis for deriving dietary patterns at two points in time allowed us to identify groups of children with lower socio-economic background presenting persistent unhealthier dietary profiles. This finding reflects the need for healthy eating interventions specifically targeting children from lower socio-economic backgrounds.

INTRODUCTION

The influence of socio-economic status (SES) on health has been observed for all age groups. Due to differences in health-related behaviours, health knowledge, housing conditions, psychosocial stressors, access to health care, etc, people living under lower socio-economic conditions have a heavier burden of disease compared to their better-off counterparts[1; 2]. More specifically, diet quality and food consumption have been shown to be related to several indicators of SES (e.g. income, educational attainment) and to factors leading to social vulnerability (e.g. migration), which in turn can affect overall health and increase the pre-disposition to develop certain disorders like overweight and obesity[3; 4].

Previous studies focusing on the association between SES indicators and food intake in children and adolescents reported lower intake of fruits and vegetables and higher intake of energy-dense foods in lower SES groups[5; 6]. Other studies focused on dietary patterns instead and their association with SES indicators[7; 8]. Indeed, considering diet as a whole is of great relevance for describing groups at higher risk of developing overweight and obesity, as the overall diet seems to be a more important determinant of weight gain compared to single dietary components [9; 10].

Dietary pattern analysis has been increasingly applied in the recent years in order to assess the relationship between the overall diet and the risk of chronic diseases[10]. The One commonly applied method to derive dietary patterns is cluster analysis, which clusters individuals into non-overlapping groups that reflect relatively homogeneous dietary patterns within groups and distinct dietary patterns between groups. Several studies applied this method for deriving dietary patterns in children and adolescents, and explored their associations with SES indicators[11; 12; 13]. Moreover, exploring changes in children's diet over time and its relation to SES may help to identify changes in dietary patterns and / or children changing their dietary patterns, and thus to better understand the impact of social inequalities on diet. Changes in diet over time have been previously explored using principal components analysis (PCA), but to the best of our knowledge there is only one report examining children's dietary patterns over time using cluster analysis yet[14]. While PCA provides linear combinations of foods instead of refering to identifiable groups of individuals, cluster analysis identifies relatively homogeneous groups of children based on their food consumption. Applying cluster analysis for describing longitudinal changes in dietary patterns can provide further insight about dietary changes within children, identifying groups with persistent unhealthier diets.

Therefore, the present study aims firstly to describe dietary patterns applying cluster analysis among children participating in the IDEFICS ('Identification and prevention of Dietary- and lifestyle induced health EFfects In Children and infantS') study baseline and follow-up surveys. The second aim is to describe the cluster memberships of children over time and their association with SES.

SUBJECTS AND METHODS

The IDEFICS study is a multi-centre population-based intervention study of children aged 2 to 9 years upon recruitment in selected regions of eight European countries (Belgium, Cyprus, Estonia, Germany, Hungary, Italy, Spain and Sweden)[15; 16]. Each participating country included one intervention region, where the community intervention program took place, and an equivalent control region. Two main surveys (baseline (T0) and follow-up after the intervention (T1)) were conducted in pre-schools and primary school classes (1st and 2nd grades at baseline). The baseline survey (September 2007 - May 2008) included 16,228 children aged 2 to 9 years (median=6.3; range=7.7). The follow-up survey (September 2009 - May 2010) reached an overall response proportion of 68% and included 11,038 children aged 4 to 11 years. The general design of the IDEFICS study has been described elsewhere [15; 16]. The present study includes only children with less than 50% of missing values in the food frequency data at baseline and follow-up and also having socio-economic and anthropometric information (n=9,301 children, 50.3 % boys) (see figure 1). Each participating centre obtained ethical approval by its respective responsible authority. All children provided oral and their parents written informed consent for all examinations and the collection of samples, analysis and storage of personal data and collected samples.

Measurements

Dietary data were obtained in both T0 and T1 by the food frequency section of the so-called Children's Eating Habits Questionnaire-Food Frequency Questionnaire (CEHQ-FFQ)[17], a validated screening tool in which the frequency of the child's consumption of selected food items during the preceding four weeks was reported by their parents. In order to assess meals under parental control, the questionnaire referred to meals outside the school canteen or childcare meal provision settings only [17; 18]. The CEHQ-FFQ consisted of 43 food items clustered into 14 food groups. It was applied as a screening instrument to investigate the consumption of foods shown to be related, either positively or negatively, to overweight and obesity in children. The CEHQ-FFQ was not designed to provide an estimate of total energy intake or total food intake [18]. Response options displayed from left to right were as follows: 'Never/less than once a week', '1-3 times a week', '4-6 times a week', '1 time per day', '2 times per day', '3 times per day', '4 or more times per day' and 'I have no idea'. For the dietary patterns analysis, a conversion factor was used to transform the questionnaire answers into weekly consumption frequencies, represented by a number ranging from 0 to 30. Only children with less than 50% of missing values and with valid data on anthropometric measures and socio-economic variables were included the analyses. Multiple imputation was applied using gender, age, BMI and country as predictors for the remaining missing values (The median number of available items was 43, SD=2.55) [19].

During the baseline and follow-up surveys parents completed a self-administered questionnaire on parental attitudes, children's behaviour and social environment. Parental education and income were self-reported. Parental education level was categorised according to the International Standard Classification of Education (ISCED97) [20]. Household income was assessed with nine country-specific categories based on the median equivalent income. The gained amount was equalised to the number of household members using the OECD (Organization for Economic Co-operation and Development) square root scale[21]. Additionally, migrant background was assessed. A migrant background was assumed if one or both of the parents were born in another country.

Trained staff carried out anthropometric measurements at baseline and follow-up following a standardised procedure. Body height (cm) was measured without shoes and all braids undone using a portable stadiometer (model: telescopic height measuring instrument SECA 225). Weight (kg) was measured by means of a child-adapted version of the electronic scale TANITA BC 420 SMA with the children in fasting status (more than 8 hours since last meal) and wearing only underwear[22]. Body mass index (BMI) and age-and gender specific BMI z-scores were calculated and categorised according to the criteria proposed by the International Obesity Task Force[23].

Statistical methods

To identify clusters of children with similar dietary patterns, k-means cluster analysis was performed [24]. First, all the variables in the food frequency questionnaire were checked for their suitability in the cluster analysis in terms of relevance. The item "Meat replacement products" was not included in the set of variables as more than 95% of the subjects reported "never / less than once per week" as frequency of consumption. Second, correlations between single food items were checked to assess the problem of multicollinearity. No redundant variables were found by assessing their correlations. Therefore, all remaining (42) food items were taken into account. The relative frequency of consumption was calculated for each food item by dividing the frequency of the consumption of a specific food item by the sum of the consumption frequencies of all food items reported for each single subject. Z-scores of the relative consumption frequency were calculated to standardise the data set before clustering, as otherwise differences in variances of the variables may affect the resulting clusters [25]. The K-means algorithm was applied with a pre-defined maximum of 100 iterations to generate separate cluster solutions for 2 to 6 clusters. In order to find a stable clustering, several solutions were obtained with different starting seeds. Iterations were generated until no change in clusters centroids was observed. Stability of the final solution was examined by randomly splitting the database in halves and repeating the same clustering procedure, where satisfactory results were observed (a maximum of 327 children in baseline clustering and 495 children in follow-up were allocated into different clusters, representing 3.5% and 5.3% of the total sample). This procedure was applied for both baseline and follow-up data sets.

The stability of the cluster solutions and the interpretability of the clusters were considered as criteria to choose the final number of clusters to retain. Based on the food items z-scores, labelling of the clusters was conducted.

Distribution of children in the different clusters was calculated stratified by gender, age, BMI categories and country, both at baseline and follow-up. For assessing the changes in dietary patterns over time, children's cluster memberships at baseline and follow-up were cross-tabulated, showing the proportion of children being allocated to the same or to different clusters. Based on logistic regression models, odds ratios for being allocated in the same cluster in T0 and T1 (i.e. both times "healthy", "sweet" or "processed; three models) or for changing the cluster ("processed/sweet" to "healthy" or vice versa; two models) were calculated, where the reference category consisted of all remaining combinations of cluster memberships in each model. Gender, age group, BMI status, migrant status, maternal and paternal education level, household income, country and a dummy variable indicating intervention vs. control region were assessed at both points in time and included as covariates in all models. Statistical significant level was set at $p \leq 0.05$. The analyses were performed using the Statistical Package for the Social Sciences (SPSS) (Version 20.0, SPSS Inc., Chicago, IL).

RESULTS

Based on the 42-food items and their relative frequency of consumption, the three cluster solutions were considered the most interpretable and stable for both baseline and follow-up data, and therefore were retained. The following labels were assigned to the three clusters: "processed" (n=4,427 in T0, n=2,554 in T1), "sweet" (n=1,910 in T0, n=1,939 in T1) and "healthy" (n=2,964 in T0, n=4,808 in T1). Tables 1 and 2 present the mean z-scores and standard deviations of all food items in the three clusters at baseline and follow-up. Dietary data in both surveys were more likely to be available for children with lower educated parents and lower income and for children with lower BMI compared to the complete IDEFICS sample (data not shown). The obtained cluster solutions were similar in terms of interpretability at both points in time. The mean values of the majority of the food items differed markedly between the three clusters (tables 1 and 2).

The “processed” cluster presented at both points in time higher relative frequencies of consumption of take away and high-fat foods, such as “savoury pastries, fritters”, “pizza as main dish”, “fried potatoes”, “hamburger, hot dog, kebab and wraps” and “crisps, corn crisps and popcorn” compared to the other clusters. Products as “whole meal bread”, “cooked vegetables”, “raw vegetables” and “fresh fruits without added sugar” scored lowest. The “sweet” cluster had at both points in time higher values of sugar-rich products, like “chocolate or nut based spread”, “sweetened drinks”, “fruit juices”, “diet drinks”, “candies, loose candies, marshmallows” and “biscuits, packaged cakes, pastries, puddings” and had the lowest scores of “water”, “porridge, oat meal, gruel, cereals, muesli, unsweetened”, “raw vegetables” and “plain unsweetened milk” and “plain unsweetened yoghurt, kefir”. The third cluster labelled as “healthy” had at both points in time higher values of low-fat foods, foods rich in vitamins and whole grain foods, e.g. “raw vegetables”, “fresh fruits without added sugar”, “porridge, oat meal, gruel, cereals, muesli, unsweetened” and “plain unsweetened milk”, and lower values of high-fat, high-sugar products, such as “fried potatoes”, “sweetened drinks”, “sweetened milk”, “mayonnaise and mayonnaise based products”, “chocolate or nut based spread”, “crisps, corn crisps, popcorn” and “biscuits, packaged cakes, pastries, puddings”.

Table 3 shows the distributions of age, gender, BMI category and country in the three clusters at baseline and follow-up. The proportion of girls in the “healthy” cluster was slightly higher compared to the other two clusters, while a higher percentage of boys were allocated to the “processed” and “sweet” clusters. Older children represented a higher percentage in the “processed” and “sweet” clusters compared to younger children. The “processed” cluster included a lower proportion of normal weight children and a higher proportion of obese children compared to the other two clusters. The biggest differences were observed between countries, i.e. certain countries represented up to 46% of one cluster. This way, the “sweet” cluster was mainly represented by Belgian and German children, the “processed” cluster by Italian, Cypriot, Estonian and Spanish children, while the “healthy” cluster included a high proportion of Swedish children.

Table 4 shows the proportions of children being allocated to the same cluster at T0 and T1, and those being allocated to different clusters, respectively (see Appendix, table 1 for the same proportions taking into account only subjects with complete information). With 85% of children being allocated in the “healthy” cluster at both T0 and T1, this cluster was the one with the greatest stability. Only 46% of the children in the “processed” cluster at baseline remained in this cluster at T1, while 43% switched to the “healthy” cluster at T1. Also 382 children (20%) being allocated to the “sweet” cluster at T0 changed to the “healthy” cluster at T1. No differences in the proportion of children allocated in the same or different clusters at T0 and T1 were found between intervention and control regions (data not shown).

Table 5 shows odds ratios (OR) and 95% confidence intervals (95% CI) for the associations between cluster membership over time and socio-economic characteristics. Girls (OR: 0.88; 95% CI: 0.79-0.98) and children with higher educated fathers (OR: 0.73; 95% CI: 0.59-0.91) were less likely to be included in the “processed” cluster at baseline and follow-up, while the odds ratios were higher for older children (OR: 1.23; 95% CI: 1.10-1.38) and migrants (OR: 1.24; 95% CI: 1.05-1.46) compared to younger children and non-migrants. Girls (OR: 0.78; 95% CI: 0.66-0.92), migrants (OR: 0.40; 95% CI: 0.31-0.52), children with the highest educated mothers (OR: 0.65; 95% CI: 0.47-0.89) and fathers (OR: 0.73; 95% CI: 0.54-0.99) and highest income (OR: 0.77; 95% CI: 0.61-0.97) were less likely to be allocated to the “sweet” cluster at baseline and follow-up. Obese children (OR: 1.37; 95% CI: 1.08-1.74) and children with higher educated mothers (OR: 1.61; 95% CI: 1.28-2.04) and fathers (OR: 1.51; 95% CI: 1.20-1.90) were more likely to be allocated to the “healthy” cluster at both points in time. Girls (OR: 1.16; 95% CI: 1.04-1.31) and children with the highest household income

(OR: 1.31; 95% CI: 1.12-1.53) were also more likely to be allocated to the "healthy" cluster at baseline and follow-up. Older children (OR: 0.65; 95% CI: 0.58-0.73) were less likely to be allocated to the "healthy" cluster. Girls (OR: 1.18; 95% CI: 1.07-1.31), obese children (OR: 1.41; 95% CI: 1.12-1.78) and children with higher educated fathers (OR: 1.24; 95% CI: 1.02-1.50) were more likely to change from the "processed"/"sweet" cluster at T0 to the "healthy" cluster at T1. Finally, obese children (OR: 0.54; 95% CI: 0.35-0.85) were less likely to change from the "healthy" cluster at T0 to the "processed"/"sweet" cluster at T1.

DISCUSSION

This paper derived dietary patterns based on a cluster analysis performed at two different points in time in 2 to 9 year old children participating in the IDEFICS study. Three consistent dietary patterns were identified at baseline and at follow-up: a "processed" cluster, showing higher frequencies of consumption of snacks, fast food and lower frequencies of vegetables and whole meal products; a "sweet" cluster with higher frequencies of consumption of biscuits and sweet products, candies and sweetened drinks, and a "healthy" cluster, showing higher frequencies of consumption of fruits, vegetables and wholemeal products and lower frequency of consumption of processed products. These three patterns presented similar profiles of relative frequencies of food consumption at each point in time, allowing us to assess which children remained in the same patterns and who changed their dietary pattern between baseline and follow-up. Cluster membership was additionally found to be associated with a number of socio-economic indicators, namely paternal and maternal education levels, household income and migrant status.

Although dietary patterns are dependent of the population considered and therefore not completely comparable between studies, previous reports extracting children's dietary patterns using cluster analysis found similar results. A British study in children also identified three clusters that were labelled as "healthy diet", "convenience diet" and "traditional diet"[(11)]. Another recent British study in 7-years-old children described "processed", "plant-based" and "traditional British" clusters[(13)]. A study among Chinese children also found three clusters, a “healthy” pattern, a “transitive” pattern and a “western” pattern[26]. But also different numbers of dietary patterns are described, ranging from two up to seven clusters[12; 14; 27; 28; 29; 30; 31]. The heterogeneity of the reference populations from different countries and continents, the different dietary assessment methods (FFQ vs. dietary records), the different number and types of food items included and the use of different clustering algorithms (e.g. k-means, Ward’s method) are likely explanations for the different results. Nevertheless, similar variations of certain patterns have been repeatedly reported across different populations. This is especially true for the patterns labelled as “healthy” or “health-conscious”[24].

A previous study derived four dietary patterns from the IDEFICS baseline data applying PCA[32]. The first pattern was labelled as “snacking”, with highest loadings for hamburgers, hotdogs, butter, savoury pastries and white bread. The “sweet and fat” pattern showed the highest loadings for sweet products like chocolate or nut-based spread, cakes, pudding and cookies. The third pattern was labelled as “vegetables and wholemeal”, with highest loadings for vegetables, fruits and wholemeal bread. Finally, the “protein and water” pattern presented highest loadings for fish, water, eggs and meat. Our cluster solution presents groupings that are pretty similar to the PCA solution. Nevertheless, it also reflects different aspects and detects a different number of factors / clusters. There are studies comparing dietary patterns obtained by applying PCA and cluster analysis to the same samples[13; 33; 34]. The results showed a general coincidence between the methods, although these two methods describe diet in a different way.

Although it was not the focus of the study, we found a higher percentage of overweight/obese children allocated in the healthy cluster compared to the sweet pattern. The results also show that obese children were more likely to be allocated in the healthy cluster at both points in time. A plausible explanation is that our dietary instrument, like most instruments assessing children’s diet, reflects the information provided by proxy reporters (parents), and therefore only includes meals under parental control. Therefore, this questionnaire might have been unable to adequately capture the consumption of certain high-fat high-sugar foods, potentially out of parental control[32].

The present study found that children's membership in a specific cluster was associated with parental education. Specifically, children with higher educated mothers and fathers were more likely to remain in the "healthy" cluster at two points in time or to change from the "processed"/"sweet" cluster to the "healthy" one. It is noteworthy that the association was found to be stronger for paternal education. Previous results of IDEFICS reports also pointed out the association of parental education and children's food consumption[35]. A recent publication describing four clusters at three different points in time in a sample of British children (i.e. "processed", "healthy", "traditional" and "packed lunch" clusters) also found associations between children's cluster membership over time and maternal education level[14]. In particular, children with lower educated mothers were more likely to be allocated in the "processed" cluster at all points in time, while children with higher educated mothers were more likely to remain in the "healthy" cluster. In the present study this association was also found in the case of paternal education. Although similar associations have been reported previously using dietary patterns derived from PCA [36], the use of cluster analysis for describing dietary patterns over time allows to track which children remain in a specific cluster, and therefore provides more insight about specific subgroups showing consistently unhealthier dietary patterns.

The present study is subject to a number of limitations. First, the IDEFICS study was not designed to be nationally representative. The participation in the study was voluntary, and some groups of the populations may have been less keen to take part in the study. Having no systematic information about non-participants, and being the direction of the bias usually pointing in opposite directions among subjects with lower and higher socioeconomic characteristics, the direction of a possible bias cannot be predicted. A further limitation is the fact that 43% of the initial baseline cohort did not participate at follow-up and / or did not provide complete data and therefore were not taken into account for the present study. Excluded participants showed a higher prevalence of overweight/obesity and included a higher proportion of lower educated parents (see Appendix table 2). Therefore, a selection bias cannot be ruled out. Additionally, participants without valid information on maternal education were more likely to be allocated to the "processed" cluster at two points in time, and therefore a selection bias cannot be ruled out. The CEHQ-FFQ was not designed to reflect total food intake, but to capture information on parent-supervised meals. The number of meals under parental control did vary between countries (e.g. higher number of meals and higher percentage of children eating at school in Sweden). This might partially explain the observed dietary patterns differences between countries. Nonetheless, we were able to describe socioeconomic differences in dietary patterns, as socioeconomic characteristics of the family are mainly influencing meals under parental control, rather than meals at school.

To the best of our knowledge, this is the first multi-centre European study assessing dietary patterns over time using cluster analysis. The large sample size, the wide variety of dietary habits and cultural backgrounds across eight European regions and the use of a validated dietary instrument, shown to provide reproducible estimates of the consumption frequencies, are the main strengths of the study. The use of cluster analyses for deriving dietary patterns at two points in time allowed us to identify groups of children with persistent unhealthier dietary profiles and to characterise them according to socio-economic indicators. Healthy eating interventions may benefit of the results of the study, and may take the results into consideration to specifically address groups presenting unhealthier dietary patterns in a persistent manner.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the children and parents who participated in the IDEFICS project. This study is part of the IDEFICS project and is published on behalf of the IDEFICS European Consortium (http://www.idefics.eu). This work was supported by the European Community's Sixth RTD Framework Programme Contract N. 016181 (FOOD).

Conflict of interest: none

JMFA carried out the statistical analyses with the help of CB and IP and drafted the manuscript. KB supervised the quality control study protocol. KB, WG, VK, LR and IP developed the measurement instruments. AH, GB, GE, II, TV, YK, EK, IH and LAM supervised the national data collection procedures. All authors read and critically reviewed the manuscript.

REFERENCES

1. Marmot M (2005) Social determinants of health inequalities. *Lancet* **365**, 1099-1104.

2. Mackenbach JP, Stirbu I, Roskam AJ *et al.* (2008) Socioeconomic inequalities in health in 22 European countries. *The New England journal of medicine* **358**, 2468-2481.

3. Nicklas TA, Baranowski T, Cullen KW *et al.* (2001) Eating patterns, dietary quality and obesity. *J Am Coll Nutr* **20**, 599-608.

4. Darmon N, Drewnowski A (2008) Does social class predict diet quality? *The American journal of clinical nutrition* **87**, 1107-1117.

5. van der Horst K, Oenema A, Ferreira I *et al.* (2007) A systematic review of environmental correlates of obesity-related dietary behaviors in youth. *Health Educ Res* **22**, 203-226.

6. Sausenthaler S, Kompauer I, Mielck A *et al.* (2007) Impact of parental education and income inequality on children's food intake. *Public Health Nutr* **10**, 24-33.

7. Richter A, Heidemann C, Schulze MB *et al.* (2012) Dietary patterns of adolescents in Germany--associations with nutrient intake and other health related lifestyle characteristics. *BMC Pediatr* **12**, 35.

8. Craig LC, McNeill G, Macdiarmid JI *et al.* (2010) Dietary patterns of school-age children in Scotland: association with socio-economic indicators, physical activity and obesity. *The British journal of nutrition* **103**, 319-334.

9. Moreira P, Santos S, Padrao P *et al.* (2010) Food patterns according to sociodemographics, physical activity, sleeping and obesity in Portuguese children. *Int J Environ Res Public Health* **7**, 1121-1138.

10. Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* **13**, 3-9.

11. Pryer JA, Rogers S (2009) Dietary patterns among a national sample of British children aged 1 1/2-4 1/2 years. *Public Health Nutr* **12**, 957-966.

12. Okubo H, Miyake Y, Sasaki S *et al.* (2012) Dietary patterns in infancy and their associations with maternal socio-economic and lifestyle factors among 758 Japanese mother-child pairs: the Osaka Maternal and Child Health Study. *Matern Child Nutr*.

13. Smith AD, Emmett PM, Newby PK *et al.* (2011) A comparison of dietary patterns derived by cluster and principal components analysis in a UK cohort of children. *Eur J Clin Nutr* **65**, 1102-1109.

14. Northstone K, Smith AD, Newby PK *et al.* (2013) Longitudinal comparisons of dietary patterns derived by cluster analysis in 7- to 13-year-old children. *The British journal of nutrition* **109**, 2050-2058.

15. Ahrens W, Bammann K, Siani A *et al.* (2011) The IDEFICS cohort: design, characteristics and participation in the baseline survey. *Int J Obes (Lond)* **35 Suppl 1**, S3-15.

16. Bammann K, Peplies J, Sjöström M *et al.* (2006) Assessment of diet. physical activity and biological, social and environmental factors in a multi-centre European project on diet- and lifestyle-related disorders in children (IDEFICS). *J Public Health* **14**, 279-289.

17. Huybrechts I, Bornhorst C, Pala V *et al.* (2011) Evaluation of the Children's Eating Habits Questionnaire used in the IDEFICS study by relating urinary calcium and potassium to milk consumption frequencies among European children. *Int J Obes* **35 Suppl 1**, S69-78.

18. Lanfer A, Hebestreit A, Ahrens W *et al.* (2011) Reproducibility of food consumption frequencies derived from the Children's Eating Habits Questionnaire used in the IDEFICS study. *Int J Obes* **35 Suppl 1**, S61-68.
19. Graham JW (2009) Missing data analysis: making it work in the real world. *Annu Rev Psychol* **60**, 549-576.
20. UNESCO (2006) ISCED 1997. International Standard Classification of Education [UNE SaCOIfS, editor]. Montreal.
21. OECD (2008) *Growing unequal? Income distribution and poverty in OECD countries.* Paris.
22. Stomfai S, Ahrens W, Bammann K *et al.* (2011) Intra- and inter-observer reliability in anthropometric measurements in children. *Int J Obes* **35 Suppl 1**, S45-51.
23. Cole TJ, Lobstein T (2012) Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatr Obes* **7**, 284-294.
24. Newby PK, Tucker KL (2004) Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* **62**, 177-203.
25. Everitt B, Landau S, Leese M *et al.* (2011) *Cluster Analysis*. 5th ed. London: Wiley.
26. Shang X, Li Y, Liu A *et al.* (2012) Dietary pattern and its association with the prevalence of obesity and related cardiometabolic risk factors among Chinese children. *PLoS ONE* **7**, e43183.
27. Knol LL, Haughton B, Fitzhugh EC (2005) Dietary patterns of young, low-income US children. *J Am Diet Assoc* **105**, 1765-1773.
28. Rodriguez-Ramirez S, Mundo-Rosas V, Garcia-Guerra A *et al.* (2011) Dietary patterns are associated with overweight and obesity in Mexican school-age children. *Arch Latinoam Nutr* **61**, 270-278.
29. Rasanen M, Lehtinen JC, Niinikoski H *et al.* (2002) Dietary patterns and nutrient intakes of 7-year-old children taking part in an atherosclerosis prevention project in Finland. *J Am Diet Assoc* **102**, 518-524.
30. Campain AC, Morgan MV, Evans RW *et al.* (2003) Sugar-starch combinations in food and the relationship to dental caries in low-risk adolescents. *Eur J Oral Sci* **111**, 316-325.
31. Kim JA, Kim SM, Lee JS *et al.* (2007) Dietary patterns and the metabolic syndrome in Korean adolescents: 2001 Korean National Health and Nutrition Survey. *Diabetes Care* **30**, 1904-1905.
32. Pala V, Lissner L, Hebestreit A *et al.* (2013) Dietary patterns and longitudinal change in body mass in European children: a follow-up study on the IDEFICS multicenter cohort. *Eur J Clin Nutr* **67**, 1042-1049.
33. Reedy J, Wirfalt E, Flood A *et al.* (2010) Comparing 3 dietary pattern methods--cluster analysis, factor analysis, and index analysis--With colorectal cancer risk: The NIH-AARP Diet and Health Study. *Am J Epidemiol* **171**, 479-487.
34. Crozier SR, Robinson SM, Borland SE *et al.* (2006) Dietary patterns in the Southampton Women's Survey. *Eur J Clin Nutr* **60**, 1391-1399.
35. Fernandez-Alvira JM, Mouratidou T, Bammann K *et al.* (2013) Parental education and frequency of food consumption in European children: the IDEFICS study. *Public Health Nutr* **16**, 487-498.
36. Northstone K, Emmett P (2005) Multivariate analysis of diet in children at four and seven years of age and associations with socio-demographic characteristics. *Eur J Clin Nutr* **59**, 751-760.

Appendix 1. Cross-tabulation between cluster memberships at baseline (T0) and follow-up (T1) for children with complete data (n, %)

	Cluster membership at baseline (T0)						
	Processed		Sweet		Healthy		Total
	n	%	n	%	n	%	n
Cluster membership at follow-up (T1)							
Processed	1524	44	184	11	188	8	1896
Sweet	367	11	1106	69	144	6	1617
Healthy	1542	45	321	20	2176	86	4039
Total	3433		1611		2508		7552

Appendix table 2. Description of the included and excluded subjects participating in the IDEFICS baseline study (T0).

	Included		Excluded		
	n	%	n	%	p-value
Total	9301		6927		
Gender					
Boys	4691	51	3571	52	0.170
Girls	4605	49	3356	48	
Age					
<6 years	4250	46	2730	46	0.482
≥6 years	5046	54	3166	54	
BMI categories					
Underweight	1006	11	652	11	<0.001
Normal weight	6587	71	4003	68	
Overweight	1111	12	803	14	
Obese	592	6	438	7	
Parental Education					
Low	571	7	522	9	<0.001
Medium	4101	48	2831	51	
High	3814	45	2221	40	

Table 1. Mean (standard deviation) z-scores of relative consumption frequencies in the three clusters at baseline (T0)

Food item	Processed n 4427		Sweet n 1910		Healthy n 2964	
	Mean	SD	Mean	SD	Mean	SD
Cooked vegetables, potatoes, beans	-0,36	0,74	0,42	1,06	0,27	1,10
Fried potatoes, potato croquettes	0,21	1,13	0,08	0,95	-0,37	0,67
Raw vegetables	-0,32	0,72	-0,35	0,73	0,70	1,15
Fresh fruits without added sugar	-0,24	0,88	-0,24	0,81	0,51	1,09
Fresh fruits with added sugar	0,23	1,19	-0,17	0,73	-0,24	0,73
Water	0,21	0,98	-0,49	0,99	0,00	0,92
Fruit juices	0,08	1,02	0,21	1,15	-0,26	0,79
Sweetened drinks	-0,10	0,76	0,58	1,64	-0,22	0,50
Diet drinks	-0,14	0,45	0,59	1,95	-0,17	0,33
Breakfast cereals, muesli, sweetened	0,15	1,09	0,10	1,01	-0,29	0,76
Porridge, oat meal, gruel, cereals, muesli, unsweetened	-0,21	0,72	-0,40	0,49	0,57	1,31
Plain unsweetened milk	-0,22	0,90	-0,23	0,83	0,48	1,07
Sweetened milk	0,29	1,13	-0,02	0,94	-0,42	0,60
Plain unsweetened yoghurt or kefir	-0,08	0,85	-0,29	0,66	0,31	1,27
Sweet yoghurt, fermented milk beverages	-0,01	1,02	0,16	1,02	-0,09	0,94
Fresh or frozen fish, not fried	0,02	1,00	-0,33	0,80	0,19	1,07
Fried fish, fish fingers	0,03	1,02	-0,06	0,94	-0,01	1,00
Cold cuts, preserved, ready to cook meat products	-0,04	0,87	0,60	1,19	-0,32	0,86
Fresh meat, not fried	0,21	1,01	-0,30	0,96	-0,12	0,95
Fried meat	-0,15	0,96	0,35	1,09	0,00	0,95
Fried or scrambled eggs	0,25	1,09	-0,26	0,82	-0,21	0,85
Boiled or poached eggs	0,07	1,27	-0,14	0,72	-0,02	0,62
Mayonnaise, mayonnaise based products	-0,04	0,86	0,52	1,53	-0,28	0,52
Cheese	0,07	1,05	-0,17	0,90	0,00	0,98
Jam, honey	-0,08	0,90	0,29	1,24	-0,07	0,94
Chocolate or nut based spread	-0,11	0,72	0,95	1,44	-0,44	0,48
Butter, margarine on bread	-0,08	0,88	0,03	1,09	0,10	1,09
Reduced-fat products on bread	-0,28	0,58	0,13	1,12	0,34	1,26
Ketchup	-0,07	1,00	0,11	1,04	0,03	0,97
White bread, white roll, white crispbread	0,31	1,06	-0,08	0,96	-0,42	0,74
Whole meal bread, dark roll, dark crispbread	-0,36	0,73	0,27	1,17	0,37	1,04
Pasta, noodles, rice	-0,03	1,04	-0,27	0,72	0,22	1,04
Dish of milled cereals	0,03	1,06	-0,22	0,55	0,10	1,10
Pizza as main dish	0,23	1,22	-0,12	0,81	-0,27	0,58
Hamburger, hot dog, kebab, wrap, falafel	0,32	1,16	-0,48	0,52	-0,17	0,79
Nuts, seeds, dried fruits	0,01	0,95	-0,27	0,69	0,16	1,19
Crisps, corn crisps, popcorn	0,20	1,13	-0,04	0,92	-0,26	0,74
Savoury pastries, fritters	0,37	1,20	-0,38	0,55	-0,31	0,62
Chocolate, candy bars	0,19	1,10	0,13	1,10	-0,37	0,60
Candies, loose candies, marshmallows	-0,17	0,78	0,72	1,48	-0,20	0,63
Biscuits, packaged cakes, pastries, puddings	-0,15	0,86	0,70	1,30	-0,24	0,73
Ice cream, milk or fruit based bars	0,12	1,14	-0,16	0,92	-0,07	0,78

Table 2. Mean (standard deviation) z-scores of relative consumption frequencies in the three clusters at follow-up (T1)

Food items	Processed n 2554		Sweet n 1939		Healthy n 4808	
	Mean	SD	Mean	SD	Mean	SD
Cooked vegetables, potatoes, beans	-0,45	0,64	0,31	1,12	0,11	1,02
Fried potatoes, potato croquettes	0,42	1,22	0,16	1,01	-0,29	0,74
Raw vegetables	-0,40	0,71	-0,41	0,66	0,37	1,10
Fresh fruits without added sugar	-0,46	0,74	-0,25	0,81	0,35	1,06
Fresh fruits with added sugar	0,27	1,13	-0,13	0,79	-0,09	0,98
Water	0,06	0,99	-0,46	0,96	0,15	0,97
Fruit juices	0,15	1,05	0,11	1,15	-0,12	0,89
Sweetened drinks	-0,05	0,73	0,68	1,71	-0,25	0,47
Diet drinks	-0,11	0,51	0,55	1,92	-0,16	0,40
Breakfast cereals, muesli, sweetened	0,31	1,21	-0,02	0,87	-0,16	0,88
Porridge, oat meal, gruel, cereals, muesli, unsweetened	-0,19	0,77	-0,40	0,54	0,26	1,16
Plain unsweetened milk	-0,22	0,85	-0,26	0,86	0,22	1,07
Sweetened milk	0,28	1,12	-0,05	0,95	-0,13	0,92
Plain unsweetened yoghurt or kefir	0,04	0,94	-0,35	0,61	0,12	1,12
Sweet yoghurt, fermented milk beverages	-0,13	0,91	0,14	1,04	0,01	1,02
Fresh or frozen fish, not fried	0,00	0,96	-0,39	0,82	0,16	1,04
Fried fish, fish fingers	0,08	1,04	-0,04	0,89	-0,02	1,02
Cold cuts, preserved, ready to cook meat products	-0,16	0,81	0,48	1,18	-0,11	0,95
Fresh meat, not fried	0,17	0,94	-0,23	1,08	0,00	0,98
Fried meat	-0,28	0,74	0,59	1,23	-0,09	0,92
Fried or scrambled eggs	0,30	1,50	-0,14	0,69	-0,10	0,70
Boiled or poached eggs	0,10	1,13	-0,23	0,82	0,04	0,98
Mayonnaise, mayonnaise based products	0,09	1,07	0,50	1,42	-0,25	0,59
Cheese	0,01	0,95	-0,23	0,96	0,09	1,03
Jam, honey	-0,09	0,86	0,13	1,10	-0,01	1,02
Chocolate or nut based spread	0,10	0,89	0,84	1,38	-0,39	0,55
Butter, margarine on bread	-0,16	0,76	-0,05	1,03	0,11	1,08
Reduced-fat products on bread	-0,23	0,59	0,05	1,04	0,10	1,13
Ketchup	0,37	1,44	0,04	0,85	-0,21	0,65
White bread, white roll, white crispbread	0,06	1,00	0,04	1,01	-0,05	0,99
Whole meal bread, dark roll, dark crispbread	-0,35	0,68	0,15	1,13	0,12	1,04
Pasta, noodles, rice	0,00	1,05	-0,21	0,81	0,09	1,03
Dish of milled cereals	0,17	1,29	-0,23	0,55	0,01	0,95
Pizza as main dish	0,63	1,49	-0,16	0,68	-0,27	0,53
Hamburger, hot dog, kebab, wrap, falafel	0,31	1,15	-0,36	0,70	-0,02	0,97
Nuts, seeds, dried fruits	0,16	1,13	-0,28	0,65	0,03	1,02
Crisps, corn crisps, popcorn	0,42	1,23	0,17	1,06	-0,29	0,70
Savoury pastries, fritters	0,78	1,41	-0,31	0,54	-0,29	0,56
Chocolate, candy bars	0,24	1,12	0,36	1,25	-0,27	0,69
Candies, loose candies, marshmallows	-0,13	0,78	0,78	1,48	-0,25	0,64
Biscuits, packaged cakes, pastries, puddings	-0,07	0,79	0,83	1,44	-0,30	0,63
Ice cream, milk or fruit based bars	0,30	1,26	-0,08	0,95	-0,12	0,81

Table 3. Description of the study population by cluster membership at baseline (T0) and follow-up (T1)

	Processed cluster				Sweet cluster				Healthy cluster					
	T0		T1		T0		T1		T0		T1		Total	
	n	*%*	n	*%*	n	*%*	n	*%*	n	%	n	*%*	n	*%*
Total	4427		2554		1910		1939		2964		4808		9301	
Gender														
Boys	2240	*51*	1342	*53*	999	*52*	1037	*54*	1444	*49*	2304	*48*	4683	51
Girls	2187	*49*	1212	*47*	911	*48*	902	*46*	1520	*51*	2504	*52*	4618	49
Age														
<6 years	1858	*42*	1042	*41*	895	*47*	903	*47*	1497	*51*	2305	*48*	4250	46
≥6 years	2569	*58*	1512	*59*	1015	*53*	1036	*53*	1467	*49*	2503	*52*	5051	54
BMI categories														
Underweight	435	*10*	269	*10*	251	*13*	261	*14*	317	*10*	473	*9*	1003	11
Normal weight	2997	*68*	1751	*69*	1417	*74*	1443	*74*	2180	*74*	3400	*71*	6594	71
Overweight	626	*14*	338	*13*	166	*9*	170	*9*	320	*11*	604	*13*	1112	12
Obese	369	*8*	196	*8*	166	*4*	65	*3*	320	*5*	331	*7*	592	6
Country														
Italy	1032	*23*	579	*23*	181	*10*	221	*11*	261	*9*	674	*14*	1474	16
Estonia	749	*17*	393	*15*	100	*5*	110	*6*	397	*13*	743	*16*	1246	13
Cyprus	795	*18*	680	*27*	6	*1*	8	*1*	235	*8*	348	*7*	1036	11
Belgium	72	*2*	43	*2*	877	*46*	867	*45*	141	*5*	180	*4*	1090	12
Sweden	64	*1*	35	*1*	34	*2*	53	*3*	1257	*42*	1267	*26*	1355	15
Germany	161	*4*	98	*4*	558	*29*	464	*24*	259	*9*	416	*9*	978	10
Hungary	680	*15*	328	*13*	99	*5*	148	*8*	207	*7*	510	*11*	986	11
Spain	874	*20*	398	*16*	55	*3*	68	*4*	207	*7*	670	*14*	1136	12

Table 4. Cross-tabulation between cluster memberships at baseline (T0) and follow-up (T1) (n, %)

	Cluster membership at baseline (T0)						
	Processed		Sweet		Healthy		Total
	n	%	n	%	n	%	n
Cluster membership at follow-up (T1)							
Processed	2046	46	228	12	280	9	2554
Sweet	474	11	1300	68	165	6	1939
Healthy	1907	43	382	20	2519	85	4808
Total	4427		1910		2964		9301

Table 5. Odds ratios (OR) and 95% confidence intervals (95% CI) for associations between cluster membership over time (each group compared to all other combinations of cluster memberships) and socio-economic characteristics. All models adjusted for country and study region (intervention vs control) and all other factors in the table

	Processed cluster at two time points (n 2046)		Sweet cluster at two time points (n 1300)		Healthy cluster at two time points (n 2519)		Processed / sweet cluster at T0, healthy cluster at T1 (n 2289)		Healthy cluster at T0, processed / sweet cluster at T1 (n 445)	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Gender										
Boys (n 4683)	1.00		1.00		1.00		1.00		1.00	
Girls (n 4618)	0.88	0.79, 0.98	0.78	0.66, 0.92	1.16	1.04, 1.31	1.18	1.07, 1.31	1.01	0.83, 1.22
Age group										
<6 years (n 4250)	1.00		1.00		1.00		1.00		1.00	
>6 years (n 5051)	1.23	1.10, 1.38	1.14	0.97, 1.35	0.65	0.58, 0.73	1.00	0.90, 1.11	1.07	0.88, 1.30
BMI status										
Normal weight (n 6462)	1.00		1.00		1.00		1.00		1.00	
Overweight (n 1098)	0.96	0.80, 1.15	0.79	0.62, 1.02	1.02	0.85, 1.23	1.25	1.04, 1.51	0.96	0.71, 1.28
Obese (n 598)	0.81	0.65, 1.03	0.81	0.57, 1.16	1.37	1.08, 1.74	1.41	1.12, 1.78	0.54	0.35, 0.85
Missing (n 1045)	0.73	0.56, 0.95	0.49	0.30, 0.81	1.52	1.15, 2.03	1.64	1.27, 2.13	0.60	0.35, 1.02
Migrant status										
Non migrant (n 7951)	1.00		1.00		1.00		1.00		1.00	
Migrant (n 1252)	1.24	1.05, 1.46	0.40	0.31, 0.52	0.98	0.82, 1.18	1.06	0.90, 1.23	0.88	0.66, 1.18
Maternal ISCED level										
Low (n 1406)	1.00		1.00		1.00		1.00		1.00	
Medium (n 4610)	1.15	0.96, 1.38	0.80	0.63, 1.01	1.31	1.07, 1.60	0.98	0.84, 1.15	1.00	0.71, 1.40
High (n 2848)	1.07	0.86, 1.33	0.65	0.47, 0.89	1.61	1.28, 2.04	0.92	0.76, 1.12	1.03	0.69, 1.52
Missing (n 339)	1.82	1.23, 2,68	0.83	0.56, 1.24	0.83	0.56, 1.22	0.95	0.68, 1.32	1.07	0.59, 1.97
Paternal ISCED level										
Low (n 1247)	1.00		1.00		1.00		1.00		1.00	
Medium (n 4814)	0.84	0.71, 0.99	0.81	0.63, 1.04	1.22	1.00, 1.48	1.27	1.09, 1.49	1.08	0.77, 1.40
High (n 2472)	0.73	0.59, 0.91	0.73	0.54, 0.99	1.51	1.20, 1.90	1.24	1.02, 1.50	1.15	0.78, 1.70
Missing (670)	0.96	0.72, 1.29	0.65	0.46, 0.92	1.15	0.85, 1.56	1.17	0.91. 1.52	1.07	0.65, 1.77
Household income										
Low (n 2993)	1.00		1.00		1.00		1.00		1.00	
Medium (n 2297)	0.98	0.84, 1.13	0.86	0.68, 1.08	1.07	0.91, 1.26	1.06	0.92, 1.21	0.79	0.61, 1.03
High (n 2934)	0.90	0.78, 1.04	0.77	0.61, 0.97	1.31	1.12, 1.53	1.11	0.97, 1.27	0.75	0.57, 0.97
Missing (n 979)	0.85	0.70, 1.03	0.71	0.52, 0.98	1.24	0.99, 1.54	1.19	0.99, 1.43	0.82	0.59, 1.14