# DATABASES AND NUTRITION

EDITED BY: Alessandra Durazzo and Massimo Lucarini
PUBLISHED IN: Frontiers in Nutrition

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-884-6
DOI 10.3389/978-2-88974-884-6

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

# DATABASES AND NUTRITION

Topic Editors:
**Alessandra Durazzo,** Council for Agricultural Research and Economics, Italy
**Massimo Lucarini,** Council for Agricultural Research and Economics, Italy

**Citation:** Durazzo, A., Lucarini, M., eds. (2022). Databases and Nutrition. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-884-6

# Table of Contents

frontiers in Nutrition

EDITORIAL
published: 18 March 2022
doi: 10.3389/fnut.2022.853600

# Editorial: Databases and Nutrition

*Alessandra Durazzo* and Massimo Lucarini**

*CREA-Research Centre for Food and Nutrition, Rome, Italy*

**Keywords: food composition databases, dedicated databases, harmonization and standardization procedures, ontology, matching process, coding procedures, data traceability, interoperability**

**Editorial on the Research Topic**

**Databases and Nutrition**

## INTRODUCTION

Studies examining the relationship between diet and health have led to a growing interest in all the biologically active components found alongside the nutrients in foods. In addition to nutritional function, these components of the diet have potentially beneficial properties: this has led to an increase in the perception of foods as functional and nutraceutical. At the same time new nutrient characteristics are emerging and the boundary between nutrients and bioactive compounds is being explored. The need to update food composition databases and dedicated databases is essential for providing information on the dietary factors contributing to chronic diseases and managing health conditions.

Food, nutrition, and health issues are strictly linked. The recent short survey of Nordhagen et al. (1) examines and highlights a set of specific causal pathways through which food safety and nutrition are interlinked across health and physiology, consumer behavior, supply chains and markets, and policy and regulation. The same authors concluded that integrating nutrition and food safety in food systems policy and programming throughout food supply chains, food environments, consumers, and food system drivers represents a new direction.

The book by Clavier and De Oliveira (2) analyzes the communication strategies of actors and the dissemination and use of information related to both food for health and health through food, considering nutrition from the point of view of public policies, educational organizations, preventive measures, consumers, and patients. For instance, the role of labeling in a sustainable food perspective is discussed as well as the social appropriation of "Diet and Health" and personalized digital tools.

The study of Siegrist et al. (3) examined the effects of psychological traits and nutrition knowledge on perceived risks related to food and nutrition: nutrition knowledge results in greater concern about the food risks that people should be most worried about. The same authors suggested that increasing people's diet-related health consciousness and nutrition knowledge may enhance people's perceptions of lifestyle risks. It is well-known that one of the major obstacles to improving health through diet and lifestyle changes is that they are difficult to implement and perhaps even more difficult to sustain in the long-term. There is now mounting evidence that one way to improve health outcomes through diet and lifestyle change is to increase health consciousness and nutrition knowledge.



***Edited and reviewed by:***
*Angela M. Zivkovic,*
*University of California, Davis,*
*United States*

****Correspondence:***
*Alessandra Durazzo*
*alessandra.durazzo@crea.gov.it*
*Massimo Lucarini*
*massimo.lucarini@crea.gov.it*

***Specialty section:***
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

***Received:*** *12 January 2022*
***Accepted:*** *31 January 2022*
***Published:*** *18 March 2022*

***Citation:***
*Durazzo A and Lucarini M (2022) Editorial: Databases and Nutrition. Front. Nutr. 9:853600. doi: 10.3389/fnut.2022.853600*

In this editorial, we provide a brief overview of the field, followed by a discussion of the research articles published in this special issue on the Research Topic "Databases and Nutrition," which highlights advances in our knowledge of the intersection between food, nutrition, health, and databases.

## LITERATURE QUANTITATIVE RESEARCH ANALYSIS: A SHOT

To provide a brief shot of this research history and the current status of the field surrounding our Research Topic, we conducted a literature search. The high level results are reported in this section. We then discuss the articles published in this Research Topic on "Databases and Nutrition" in the subsequent section.

On 20 November 2021, the Scopus database was used to carry out a search to retrieve publications that referred to databases, food, nutrition, and health relationship.

The search string: "Database*" AND "Food*" AND "Nutrition*" AND "Health*" was used to extract bibliometric data from the Scopus online database (https://www.scopus.com/home.uri, accessed on 20 November 2021) and bibliographic data, i.e., publication year, publication count, document type, countries/territories of origin, institutions, were recorded.

The functions of the Scopus web online platform named "Analyze" and "Create Citation Report" were utilized for carrying out basic analyses.

The search returned 3,213 documents covering the period 1943–2022, and mainly the subject areas of *Medicine*, *Nursing,* and *Agricultural and Biological Sciences*. The papers are distributed per typology as reported in **Figure 1** and include mainly "Articles" for the 61.9 %, followed by "Reviews" 31.6%, then "Conference papers" (3.8%), "Book chapters" (1.2%), and so on.

The oldest document was published by Mead, (4) in 1943 in the journal *Psychological Bulletin* and it is entitled "The Committee on Food Habits." The most cited document (1,770 times) is by Block et al. (5) on the development of a data-based approach to diet questionnaire design and testing.

The most recent review is a comprehensive study of history, phytochemistry, experimental pharmacology, and clinical uses of honey with special reference to Unani medicine (6).

Among the most current reviews published in 2022, it is worth mentioning the paper published by Larrick et al. (7) in the *Journal of Food Composition and Analysis* as an update on "A Partnership for Public Health: USDA Global Branded Food Products Database" with the goal of improving public health and sharing open data by expanding and enhancing the USDA National Nutrient Database (USDA FoodData Central) with nutrient composition and ingredient information on branded and private label foods to better reflect the food supply.

The most recent "article" is focused on the evaluation of micronutrient composition, antioxidant properties, and mineral safety index of selected Nigerian cooked foods (8).

A modern application is given by the USDA on the use of the Database of Flavonoid Values for USDA Food Codes 2007–2010 in assessing intake differences between the Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study and What We Eat in America (WWEIA), NHANES (9).

Within the category "Book" the following products were reported: (i) metabolomics in food and nutrition (10); a roadmap to 2050 of science and technology on public health in China (11); overview, funding issues, and trends of U.S. Nutrition and agricultural research (12).

The most cited Editorial (91 times) is entitled "Environmental impacts of food consumption and nutrition: where are we and what is next?" published by Nemecek et al. (13) in 2016 on *International Journal of Life Cycle Assessment*.

**Figure 2** reports, respectively, the most productive countries/territories: United States ($n = 960$) was the most productive country, followed by the United Kingdom ($n = 422$) and Australia ($n = 345$).

For the United States, one short survey addressing the epidemiological evidence on flavonols, flavones, flavanones, and human health is reported (14).

For the United Kingdom, the most cited review discussed the implications of nutritional antioxidants of plant polyphenols in cancer and heart disease (15), whereas the most recent review was focused on the impact of the 2008 Great Recession on dietary intake (16).

The most cited article (541 items) is a document published by Kroes et al. (17) in 2004 in the journal *Food and Chemical Toxicology* as guidance for the application of substances and their presence at low levels in the diet, with a focus on structure-based thresholds of toxicological concern.

For Australia, the most cited documents are as follows: (i) the WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement (18); definition of the Mediterranean diet (19); diet quality (20); the relationship between nutrition knowledge and dietary intake (21).

**Figure 3** reports the most productive authors with Slimani, N., publishing 31 documents, having the most. Her most cited paper (268 times) described a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study: the EPIC nutrient database project (22).

Narrowing the search toward the interface database/nutrients/bioactive components using the search string: "Database*" AND "Food*" AND "Nutrition*" AND "Health*" AND "Composition*" OR "Nutrient*" OR "bioactive compound*" OR "bioactive component*" OR "bioactive molecule*" OR "dietary supplement*" the search returned 1,469 documents.

The "full records and cited references" were exported to VOSviewer software for further bibliometric analyses and additional processing, (version 1.6.16, www.vosviewer.com) (23–25).

Out of the 9,237 keywords, 1,376 meet the selected threshold, whereas 4 of them were manually excluded, and are represented in a term map (**Figure 4**).

The most recurring terms are human/s, diet, female, nutrition, male, adult, food composition, dietary intake, food intake, caloric intake, nutritional value, nutritional assessment, food analysis, child, and nutritional status.

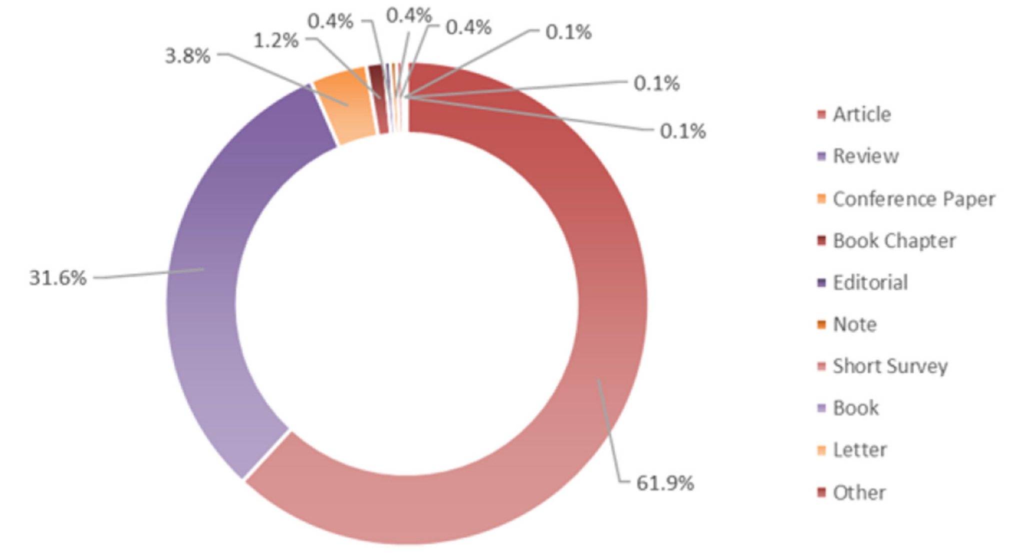


**FIGURE 1 |** Distribution of documents by type. Based on data from Scopus.

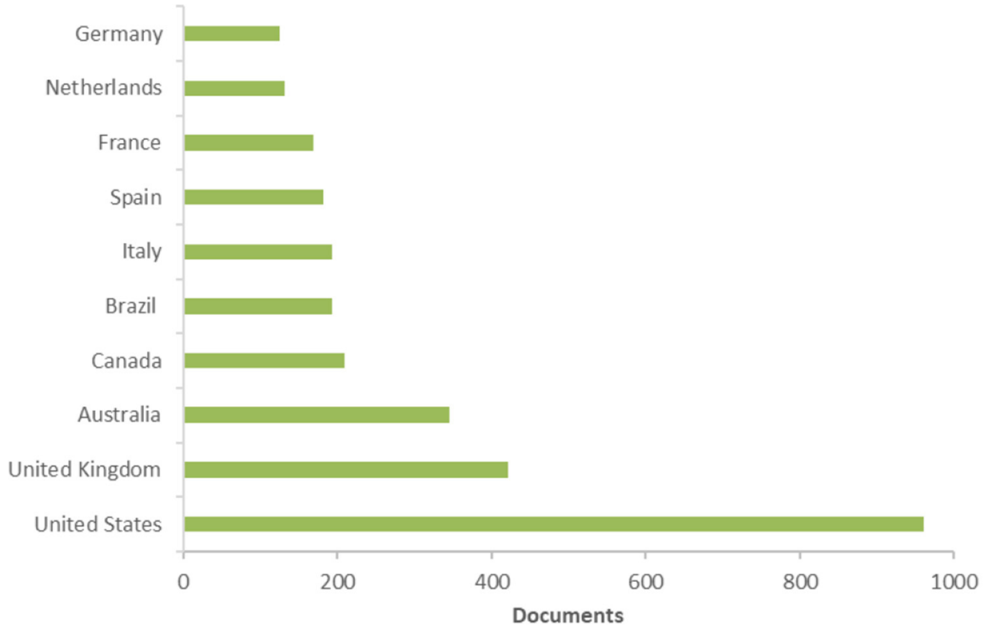


**FIGURE 2 |** Most productive countries/territories. Based on data from Scopus.

This bibliometric analysis demonstrates a growing interest in the development and refinement of databases for food, nutrients, and bioactive molecules. There is a clear and growing need to define the composition of the foods we eat to be able to increase our understanding of their effects on health outcomes. A growing interest in capturing the content of foods in accurate databases, from regional foods to branded products, was observed across the globe.

## EXPLORING THE RESEARCH TOPIC

In the collection of articles under the Research Topic "Databases and Nutrition," a total of 19 articles are published, mainly focusing on presenting new databases and/or updates for existing databases, with particular attention to their uses and applications.

It is worth mentioning the work of Samaniego-Vaesken et al. which presented an updated database and trends of declared low- and no-calorie sweeteners from foods and beverages marketed in

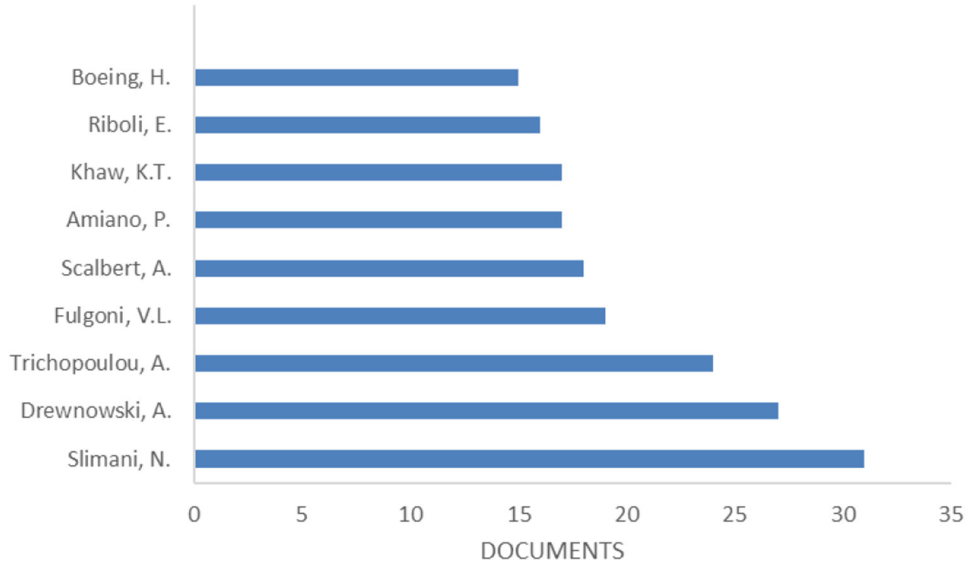


**FIGURE 3 |** Most productive Authors. Based on data from Scopus.

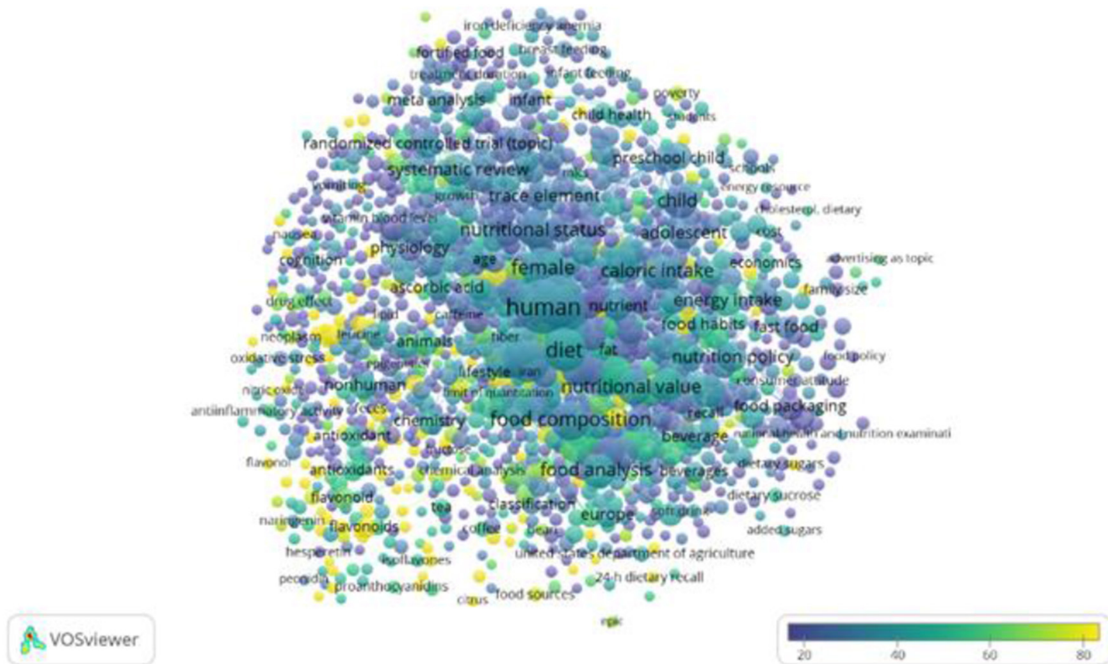


**FIGURE 4 |** Term map for search database, food, nutrition, health, composition, nutrient, bioactive compound, bioactive component, bioactive molecule, dietary supplement relationships. Bubble color represents the citations per publication (CPP). Two bubbles are closer to each other if the terms co-appeared more frequently. Based on data from Scopus and elaborated by VOSviewer software.

Spain. A total of 1,238 products were identified and the major groups were sugar and sweets (24%), non-alcoholic beverages (21%), cereals and grains (19%), and milk and dairy products (14%) accounting for >70% of total products. The authors also found that the most declared products were low- and no-calorie sweeteners were sorbitol (19.5%), sucralose (19.5%), and acesulfame K (19.2%) (Samaniego-Vaesken et al.). In a brief research report, Hafner and Pravst showed a sharp rise in the use of low- and no-calorie sweeteners in non-alcoholic beverages in Slovenia as an update based on 2020 data (Hafner and Pravst). Restrepo et al. presented a multi-dimensional dataset of open data and satellite images for the characterization of food security and nutrition. An example of database creation is given also by Palmer et al. on the development of a vitamin K database for commercially available food in Australia.

Poulain et al. presented the Malaysian Food Barometer Open Database for studying the modernization of Malaysian food patterns and their economic and health consequences.

Among discussions of the applications and uses of a database, the use of branded food composition databases is presented as a tool to support nutrition research and monitoring of the food supply, with insights from the Slovenian Composition and Labeling Information System (CLAS) (Pravst et al.) as well as for the exploitation of food fortification practices, throughout a case study on vitamin D in the Slovenian food supply (Krušič et al.). Moreover, Hafner et al. verified the use of food labeling data for compiling branded food databases throughout a case study of sugars in beverages.

In the collection, some articles were focused on the relationship between diet quality and health. Duan et al. reported the design, implementation, and major findings of the Chinese Adolescent Cohort Study. Results from the Chinese Adolescent Cohort Study baseline and the first follow-up data suggested that higher protein intake among girls and unhealthy eating habits among children might increase the risk for childhood obesity. Moreover, the same authors marked higher intakes of grain and meat and lower overall diet quality and intakes of dietary fiber and tuber might be associated with advanced pubertal development (Duan et al.). Ma et al. in a cross-sectional study of the US Population showed a saturation value association between Body Mass Index and Bone Mineral Density for people over 50 years old. The authors suggested how keeping the Body Mass Index in the slightly overweight value (around 26 kg/m$^2$) might reduce other adverse effects while obtaining optimal Bone Mineral Density in this age group (Ma et al.).

Shamah-Levy et al. analyze malnutrition, being overweight, and having obesity in children, teenagers, and adults through the National Health and Nutrition Surveys information available from public databases in Mexico from 2006 to 2020.

Some examples of dietary strategies are here presented. Neill et al. showed how vitamin D biofortification of pork may offer a food-based strategy to increase vitamin D intake in the UK Population. O'Connor et al. showed how the heterogeneity in meat food groups can meaningfully alter population-level intake estimates of red meat and poultry.

Ortenzi and Beal identified the top food sources of priority micronutrients, among minimally processed foods for the complementary feeding of children (6–23 months) in South and Southeast Asia throughout an aggregated regional food composition database for South and Southeast Asia. de Amorim et al. showed the use of databases to evaluate the prevalence of hunger among adolescents in Brazil. Meanwhile, the work of Liechti et al. proposes a multicriteria approach to select the best representative products from the market base for future reformulation by going beyond nutrition and composition information on packaging. Tseng et al. described a new, open-source ingredient list search method and applied this method to describe the presence of sensory-related industrial additives in US packaged foods. Mariscal-Arcas et al. described the evolution of nutritional habit behaviors of a Spanish population confined through social media. On the other hand, Timotijevic et al. showed the developing responsible governance for a food and nutrition e-infrastructure, throughout the case study of the determinants and intake data platform.

## CONCLUSION

This editorial has provided a general bibliographic overview of the Research Topic "Databases and Nutrition" and discusses the papers published in this special issue, including current advances in the development of databases needed for research on the relationships between food, nutrition, and health. It is important to continuously update and implement databases in a standardized and harmonized manner between various organizations and countries, in order to have resources that are more representative and that reflect changes in food environments and markets in the context of evolving consumer dietary choices and habits. New and alternative food sources, and at the same time new descriptors and markers, need to embrace different areas in the perspective of interoperability between databases. Under the One Health approach for food safety, food security, and sustainable food production (26), linking bioresources repositories (i.e., cultivar, alternative sources, etc.) to data describing the chemistry and quality of food products (i.e., food composition data, bioactive components, bioactivity, beneficial properties, etc.) represents a new frontier.

## AUTHOR CONTRIBUTIONS

AD and ML have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## ACKNOWLEDGMENTS

We would like to acknowledge the authors and the reviewers of the publications in this Research Topic for their invaluable contributions and effort. We are also grateful to the editorial board members and support staff of the journal for their kind support during the preparation of this Research Topic.

## REFERENCES


1. Nordhagen S, Lambertini E, DeWaal CS, McClafferty B, Neufeld LM. Integrating nutrition and food safety in food systems policy and programming. *Global Food Security*. (2022) 32:100593. doi: 10.1016/j.gfs.2021.100593
2. Clavier V, De Oliveira JP. Food and health: actor strategies in information and communication. *Food Health*. (2019) 2019:1–363. doi: 10.1002/9781119476252
3. Siegrist M, Bearth A, Hartmann C. The impacts of diet-related health consciousness, food disgust, nutrition knowledge, and the Big Five personality traits on perceived risks in the food domain. *Food Qual Prefer*. (2022) 96:10444. doi: 10.1016/j.foodqual.2021.104441
4. Mead M. The committee on food habits. *Psychol Bull*. (1943) 40:290–3. doi: 10.1037/h0059103
5. Block G, Hartman AM, Dresser CM, Gannon J, Gardner L. A data-based approach to diet questionnaire design and testing. *Am J Epidemiol*. (1986) 124:453–69. doi: 10.1093/oxfordjournals.aje.a114416
6. Nikhat S, Fazil M. History, phytochemistry, experimental pharmacology and clinical uses of honey: A comprehensive review with special reference to Unani medicine. *J Ethnopharmacol*. (2022) 282:114614. doi: 10.1016/j.jep.2021.114614
7. Larrick B, Kretser A, McKillop K. Update on "a partnership for public health: USDA global branded food products database". *J Food Comp Anal*. (2022) 105:104250. doi: 10.1016/j.jfca.2021.104250
8. Akinbule OO, Onabanjo OO, Sanni SA, Adegunwa MO, Akinbule AS. Micronutrient composition, antioxidant properties, and mineral safety index of selected Nigerian cooked foods. *Food Chem*. (2022) 373:131386. doi: 10.1016/j.foodchem.2021.131386
9. Sebastian RS, Fanelli Kuczmarski M, Wilkinson Enns C, Zonderman AB, Evans MK. Application of the database of flavonoid values for USDA Food Codes 2007–2010 in assessing intake differences between the healthy aging in neighborhoods of diversity across the life span (HANDLS) study and what we eat in America (WWEIA), NHANES. *J Food Com Anal*. (2021) 104:104124. doi: 10.1016/j.jfca.2021.104124
10. Weimer BC, Slupsky CM. Metabolomics in food and nutrition. *Metabol Food Nutr*. (2013) 2013:1–237. doi: 10.1533/9780857098818
11. Chen K, Lin Q, Wu, J. *Science and Technology on Public Health in China: A Roadmap to 2050 Science and Technology on Public Health in China: A Roadmap to 2050*. Berlin; Heidelberg: Springer. (2010).
12. Ramos A. *U.S. Nutrition Agricultural Research: Overview, Funding Issues Trends. U.S. Nutrition and Agricultural Research: Overview, Funding Issues and Trends*. Nova Science Pub Inc. (2015).
13. Nemecek T, Jungbluth N, Canals L.M, Schenck R. Environmental impacts of food consumption and nutrition: where are we and what is next? *Int J Life Cycle Ass*. (2016) 21:607–20. doi: 10.1007/s11367-016-1071-3
14. Graf BA, Milbury PE, Blumberg JB. Flavonols, flavones, flavanones, and human health: epidemiological evidence. *J Med Food*. (2005) 8:281–90. doi: 10.1089/jmf.2005.8.281
15. Duthie GG, Duthie SJ, Kyle JAM. Plant polyphenols in cancer and heart disease: Implications as nutritional antioxidants. *Nutr Res Rev*. (2000) 13:79–106. doi: 10.1079/095442200108729016
16. Jenkins RH, Vamos EP, Taylor-Robinson D. Impacts of the 2008 Great Recession on dietary intake. *Int J Behav Nutr Phys Activ*. (2021) 18:57. doi: 10.1186/s12966-021-01125-8
17. Kroes R, Renwick AG, Cheeseman M, Vos JG, Würtzen G. Structure-based thresholds of toxicological concern (TTC): Guidance for application to substances present at low levels in the diet. *Food Chem Toxicol*. (2004) 42:65–83. doi: 10.1016/j.fct.2003.08.006
18. Langford R, Bonell CP, Jones HE, Magnus D, Campbell R. The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement. *Cochrane Database Syst Rev*. (2014) 2014:CD008958. doi: 10.1002/14651858.CD008958.pub2
19. Davis C, Bryan J, Hodgson J, Murphy K. Definition of the mediterranean diet: A literature review. *Nutrients*. (2015) 7:9139–53. doi: 10.3390/nu7115459
20. Wirt A, Collins CE. Diet quality - What is it and does it matter? *Pub Health Nutr*. (2009) 12:2473–92. doi: 10.1017/S136898000900531X
21. Spronk I, Kullen C, Burdon C, O'Connor H. Relationship between nutrition knowledge and dietary intake. *Br J Nutr*. (2014) 111:1713–26. doi: 10.1017/S0007114514000087
22. Slimani N, Deharveng G, Unwin I, Subar AF, Riboli E. The EPIC nutrient database project (ENDB): A first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. *Eur J Clin Nutr*. (2007) 61:1037–56. doi: 10.1038/sj.ejcn.1602679
23. Van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. (2010) 84:523–38. doi: 10.1007/s11192-009-0146-3
24. Van Eck NJ, Waltman L. Text mining and visualization using VOSviewer. *arXiv preprint* arXiv:1109.2058.
25. Waltman L, Van Eck NJ, Noyons EC. A unified approach to mapping and clustering of bibliometric networks. *J Inform*. (2011) 4:629–35. doi: 10.1016/j.joi.2010.07.002
26. Garcia Sara N, Osburn Bennie I, Jay-Russell Michele T. One health for food safety, food security, and sustainable food production. *Front Sust Food Sys*. (2020) 4:1. doi: 10.3389/fsufs.2020.00001


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 29 July 2021
doi: 10.3389/fnut.2021.670422

Check for updates

# Updated Database and Trends of Declared Low- and No-Calorie Sweeteners From Foods and Beverages Marketed in Spain

Mª de Lourdes Samaniego-Vaesken[1], Beatriz González-Fernández[2], Teresa Partearroyo[1], Rafael Urrialde[1,3] and Gregorio Varela-Moreiras[1,4]*

[1] Departamento de Ciencias Farmacéuticas y de la Salud, Facultad de Farmacia, Universidad San Pablo-CEU, CEU Universities, Urbanización Montepríncipe, Madrid, Spain, [2] Dietista-Nutricionista, Torres y Carrera, Madrid, Spain, [3] Departamento de Genética, Fisiología y Microbiología, Facultad de Ciencias Biológicas, Universidad Complutense de Madrid, Madrid, Spain, [4] Fundación Española de la Nutrición, c/General Álvarez de Castro 20, Madrid, Spain

OPEN ACCESS

**Edited by:**
*Massimo Lucarini, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Lorenza Mistura, Council for Agricultural and Economics Research (CREA), Italy*
*Irina Kovalskys, Favaloro University, Argentina*
*Tommaso Filippini, University of Modena and Reggio Emilia, Italy*

***Correspondence:**
*Gregorio Varela-Moreiras*
*gvarela@ceu.es*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** 21 February 2021
**Accepted:** 07 June 2021
**Published:** 29 July 2021

**Citation:**
*Samaniego-Vaesken ML, González-Fernández B, Partearroyo T, Urrialde R and Varela-Moreiras G (2021) Updated Database and Trends of Declared Low- and No-Calorie Sweeteners From Foods and Beverages Marketed in Spain. Front. Nutr. 8:670422. doi: 10.3389/fnut.2021.670422*

**Background:** The past few years have witnessed an increase in the availability of food products containing one or more low- and no-calorie sweeteners (LNCS) in the Spanish market, mostly due to the new massive reformulation plan. However, these are not included in food composition tables or databases, and, therefore, assessment of their intake among the population is complex. This study aims to update a database including commercialized foods and beverages.

**Method:** A systematic search of ingredients information from the different food and beverage categories was undertaken during 2019 by recording the availability and type of LNCS declared in the information of the product from labels and online shopping platforms of retailers from Spain to update a previous food composition database compiled in 2017.

**Results:** A total of 1,238 products were identified. The major groups were sugar and sweets (24%), non-alcoholic beverages (21%), cereals and grains (19%), and milk and dairy products (14%) accounting for >70% of total products. The mainly declared LNCS were sorbitol (19.5%), sucralose (19.5%), and acesulfame K (19.2%).

**Conclusion:** There is a wide variety of products that include LNCS as a main ingredient with higher availability than when compared with the results of database of 2017, consequently, it might be expected that LNCS are commonly consumed at present in the Spanish diet.

**Keywords: food composition database, food groups, Spain, reformulation, low- and no-calorie sweeteners**

## BACKGROUND

Low- and no-calorie sweeteners (LNCS), also known as artificial, non-nutritive, or intense sweeteners, comprise a group of food additives that provide high sweetness intensity per gram of food and beverage products (1). They are used in very small quantities and deliver no or fewer calories, replacing added sugars in a variety of food products (2).

The use of LNCS has become more common for manufacturers to develop new products and to comply with food and beverage reformulation practices to decrease energy resulting from added sugars. Furthermore, there is general consumer interest in reducing energy intake (TE), and food products containing LNCS have become more popular choice (2, 3). At present, we can find LNCS as ingredients in products labeled as "sugar-free" or "without added sugars" but also in regular products together with low amount of added sugars.

There are currently 19 compounds authorized by the European regulations for use in food products: sorbitol (E-420), mannitol (E-412), acesulfame K (E-950), aspartame (E-951), cyclamate (E-952), isomalt (E-953), saccharine and its sodium, potassium, and calcium salts (E-954), sucralose (E-955), thaumatin (E-957), neohesperidine DC (E-959), steviol glycosides (E-960), neotame (E-961), salt of aspartame-acesulfame (E-962), polyglycitol syrup (E-964), maltitols (E-965), lactitol (E-966), xylitol (E-967), erythritol (E-968), and advantame (E-969) (4). The use of sweeteners in the European Union (EU) is in accordance with Commission Regulations numbers 231/2012 (5) and 1169/2011 (6), which indicate that food labeling must specify the presence of these additives in the list of ingredients as well as next to the name of the product, "with LNCS" (EU). Therefore, the identified sweeteners have only been at the level of declaration of presence, not of quantity present in the food product.

Previous research has identified the most consumed food groups containing LNCS by a representative sample of the Spanish population aged 9–75 years from the ANIBES study ("anthropometric data, macronutrients, and micronutrients intake, as well as the practice of physical activity, socioeconomic data, and lifestyles of the population") (7) where we described that from 1,164 products analyzed, 10% contained LNCS in their composition and 5.1% declared a combined use of added sugars and LNCS declared on their labels. LNCS were mainly present in diet soft drinks (100%), chewing gum and candies (89%), soya drinks (45%), and yogurt and fermented milk (18%). For this reason, a comprehensive food composition database was developed in 2017 that surveyed all available products from the Spanish market that declared LNCS in their ingredient lists (8).

The main source of information for consumers and researchers about the presence of LNCS in packaged foods are food labels and, specifically, the ingredients list, which allows their identification. However, manufacturers are not mandated to declare added levels of LNCS, which should always remain under the quantities authorized by the European Food Safety Authority (9). At present, the lack of updated and comprehensive data on the type of LNCS in different food products challenges the assessment and monitorization of LNCS consumption, which remains unavailable in most dietary surveys (10, 11).

In 2017, the Spanish Ministry of Health, Social Services and Equality, through the Spanish Agency for Food Safety and Nutrition (AESAN) launched the "Collaboration plan to improve the composition of food and beverages and other measures 2017–2020" within the framework of the "Strategy of Nutrition, Physical Activity and Prevention of Obesity" (NAOS) of AESAN (12). This Plan included reformulation commitments from different sectors of the manufacture and distribution of food products, mainly foods and drinks that are regularly consumed by children, adolescents, and their families, and it was mainly focused on the reduction of added sugar, sodium, and saturated fats. The plan of AESAN was aimed at improving the nutritional quality of the diet, promoting a healthier food intake to prevent or reduce the incidence of overweight and obesity, and related pathologies. The agreed reformulation measures affected food and beverages belonging to 12 groups: soft drinks, pastries and cakes, breakfast cereals, creams, meat products, cookies, ice cream, fruit nectars, specially packaged bread, ready-to-eat-meals, dairy products, and sauces. In any case, it should be highlighted that food and beverage groups included in the proposed plan are responsible for providing 44.5% of the total daily energy from foods with added sugars in Spain (12). Therefore, the overall aim of the plan of AESAN was to reduce from 5 to 10%, depending on the groups, food contents in added sugars, saturated fats, and salt by the end of 2020 (12). In this regard, the use of LNCS as added sugar substitutes is highly relevant but still has not been assessed. It is, therefore, of importance to monitor potential exposures following the requirement to reduce the level of sugar intake, to ensure there is no shift in intakes, particularly for high-risk individuals, such as diabetics and children with specific dietary requirements (13).

For all the aforementioned, in the present study, we aimed to update a previously existing database designed and compiled by the research group in 2017, including all foods and beverages declaring LNCS commercialized in Spain (8).

# MATERIALS AND METHODS

## Study Design and Data Collection

This cross-sectional study comprised of a systematic search and screening of label ingredient information of food and beverage product to identify one or more of the EU authorized LNCS, available through three major online retailer shopping platforms at the national level, accounting for ≥85% of the market (14, 15). The presence and type of LNCS declared, which were included in the ingredient lists of available foods and beverages, was identified and recorded, all of these products were processed foods as they contain several ingredients and additives (with the E-number identification). Data was collected throughout 2019, and all the products which declared an LCNS within their ingredient information were included. The acquired information was compiled in an in-house database comprising Microsoft™ Excel 2013 spreadsheets, where each product was coded and classified according to label denomination into nine groups and 28 subgroups. Data including product name, denomination, brand, ingredient list, nutritional information were recorded.

## Identification of LCNS in Packaged Foods, Database Compilation, and Data Management

Low- and no calorie sweeteners were reviewed in the online platform information in accordance with the list of 19 LNCS types authorized by the EU food labeling regulations (EU).

Product names, categories, ingredients, and food additives found on the description of each food and beverage were compiled using a Microsoft$^{TM}$ Excel 2013 spreadsheet for analysis. Quality assurance control was performed by two researchers, who checked the data independently. The presence and frequency of the different LCNS were analyzed globally and within each food and beverage subgroup. In addition, the individual or combined use of each LCNS in each food group was also evaluated.

## RESULTS

The database was updated and included a total of 1,238 products declaring LNCS in their ingredient lists, which were identified from the survey performed among a representative sample of online food and drink retailers. The major groups identified included the following (**Figure 1**): sugar and sweets (24%), non-alcoholic beverages (21%), cereals and grains (19%), and milk and dairy products (14%), all of which accounted for more than 70% of total available products in the compiled database.

The different food and beverage groups and subgroups distribution, together with the number of added LNCS is shown in **Table 1.**

The main food and beverage subgroups including LNCS were bakery and pastry ($n$ = 186; 15%), yogurt and fermented milks ($n$ = 126; 10.1%), chewing gums, candies, and sweets ($n$ = 111; 9%), food supplements and substitutes ($n$ = 109; 8.8%), diet soft drinks ($n$ = 91; 7.3%), sugar soft drinks ($n$ = 90; 7.2%) and meat and meat products ($n$ = 73; 6%), which together accounted for more than 60% of total products.

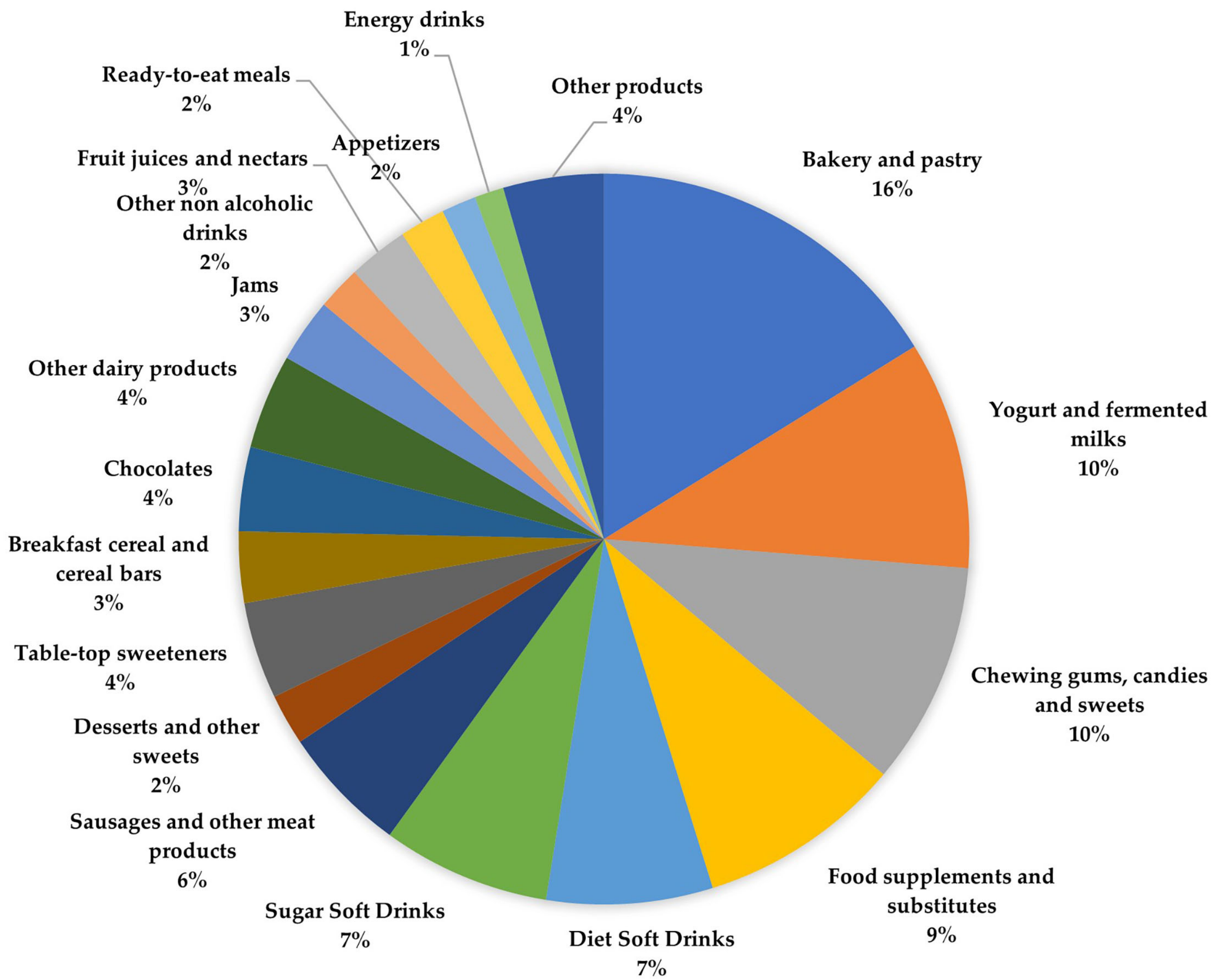


**FIGURE 1 |** Distribution of food and beverage subgroups (%) available in the Spanish market containing one or more low and no-calorie sweeteners (LCNS) as declared in their ingredient information.

**TABLE 1 |** Distribution of food and beverage groups and subgroups available in the Spanish market (*n*, %), according to the number of declared low and no-calorie sweeteners (LCNS).

| Food groups and subgroups | Number of declared LNCS | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Sugar and sweets (*n* = 299; 24.1%) | 97 (32.4%) | 88 (29.4%) | 54 (18%) | 23 (7.7%) | 12 (4%) | 17 (5.7%) | 7 (2.3%) | 1 (0.3%) |
| Chewing gum, candies, and sweets (*n* = 111) | 37 | 11 | 13 | 13 | 12 | 17 | 7 | 1 |
| Desserts and other sweets (*n* = 55) | 12 | 11 | 31 | 1 | 0 | 0 | 0 | 0 |
| Table-top sweeteners (*n* = 52) | 28 | 22 | 2 | 0 | 0 | 0 | 0 | 0 |
| Chocolates (*n* = 46) | 17 | 24 | 4 | 1 | 0 | 0 | 0 | 0 |
| Jams (*n* = 35) | 3 | 20 | 4 | 8 | 0 | 0 | 0 | 0 |
| Non-alcoholic beverages (*n* = 266; 21.5%) | 113 (42.5%) | 100 (37.6%) | 51 (19.2%) | 2 (0.75%) | 0 | 0 | 0 | 0 |
| Diet Soft Drinks (*n* = 91) | 14 | 43 | 34 | 0 | 0 | 0 | 0 | 0 |
| Sugar Soft Drinks (*n* = 90) | 61 | 20 | 7 | 2 | 0 | 0 | 0 | 0 |
| Other non-alcoholic drinks (*n* = 32) | 9 | 14 | 9 | 0 | 0 | 0 | 0 | 0 |
| Fruit juices and nectars (*n* = 30) | 15 | 14 | 1 | 0 | 0 | 0 | 0 | 0 |
| Energy drinks (*n* = 14) | 7 | 7 | 0 | 0 | 0 | 0 | 0 | 0 |
| Coffee and herbal teas (*n* = 6) | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sport drinks (*n* = 3) | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Cereals and grains (*n* = 239; 19.3%) | 196 (82%) | 35 (14.6%) | 8 (3.3%) | 0 | 0 | 0 | 0 | 0 |
| Bakery and pastry (*n* = 186) | 151 | 30 | 5 | 0 | 0 | 0 | 0 | 0 |
| Breakfast cereal and cereal bars (*n* = 48) | 41 | 4 | 3 | 0 | 0 | 0 | 0 | 0 |
| Bread (*n* = 5) | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Milk and dairy products (*n* = 169; 13.6%) | 30 (17.7%) | 122 (72.2%) | 8 (5%) | 8 (5%) | 1 (0.6%) | 0 | 0 | 0 |
| Yogurt and fermented milks (*n* = 126) | 20 | 99 | 7 | 0 | 0 | 0 | 0 | 0 |
| Other dairy products (*n* = 41) | 10 | 21 | 1 | 8 | 1 | 0 | 0 | 0 |
| Cheese (*n* = 2) | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Food supplements and substitutes (*n* = 109; 8.8%) | 79 (72.5%) | 26 (24%) | 4 (3.6%) | 0 | 0 | 0 | 0 | 0 |
| Meat and meat products (*n* = 73; 6%) | 73 (100%) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ready to eat meals (*n* = 28; 2.2%) | 28 (100%) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Appetizers (*n* = 19; 1.5%) | 19 (100%) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sauces and condiments (*n* = 11; 1%) | 9 (82%) | 2 (8%) | 0 | 0 | 0 | 0 | 0 | 0 |
| Canned fruit (*n* = 9; 0.7%) | 2 (22%) | 6 (66%) | 1 (11%) | 0 | 0 | 0 | 0 | 0 |
| Fish and shellfish (*n* = 9; 0.7%) | 9 (100%) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Alcoholic beverages (*n* = 6; 0.5%) | 0 | 5 (83.3%) | 1 (16.6%) | 0 | 0 | 0 | 0 | 0 |
| Vegetables (*n* = 1; 0.08%) | 0 | 1 (100%) | 0 | 0 | 0 | 0 | 0 | 0 |
| *n* total products = 1,238 | 655 | 385 | 127 | 33 | 13 | 17 | 7 | 1 |
| % of each combination of LNCS | 52.9 | 31.1 | 10.3 | 2.7 | 1.1 | 1.4 | 0.6 | 0.1 |

We found that 52.9% of products declared the use of only one LNCS, while 31.1% declared two LNCS and 10.3% declared three LNCS. The major groups that included just one LNCS were cereals and grains (82%), non-alcoholic beverages (42.5%), and sugars and sweets (32.4%), expressed as within-group proportion. The main groups that included two LNCS were milk and dairy products (72.2%), canned fruit (66%), and non-alcoholic beverages (37.6%). Non-alcoholic beverages (19.2%) and sugars and sweets (18%) showed the highest proportion of products including three LNCS. We recorded combinations of up to eight different LNCS, but these were minor and found in the chewing gum, candies, and sweets subgroup. The comprehensive distribution of the specific types and combinations of LNCS declared across the assessed food groups and subgroups can be retrieved from the **Supplementary Material** section, **Supplementary Table 1**.

**Table 2** shows the different types of LNCS declared among the food and beverage subgroups included in the database. Globally, we observed that sorbitol, sucralose, and acesulfame K were the most declared LNCS with 423 (19.5%), 421 (19.5%), and 415 (19.2%) label declarations, respectively. Sugars and sweets, which were the main group found in our survey, declared the use of a wide variety of LNCS across their subgroups: chocolates mainly include sorbitol and steviol glycosides; jams include sucralose, steviol glycosides, and maltitol; chewing gum, candies and sweets primarily declared sorbitol, acesulfame, K aspartame, and maltitol. Table-top sweeteners identified included steviol glycosides, saccharine, and cyclamate, among others.

**TABLE 2 |** Declared types and distribution (n) of LCNS across assessed food groups and subgroups available in the Spanish market.

| | Sorbitol (E-420) n | Mannitol (E-421) n | Acesulfame K (E-950) n | Aspartame (E-951) n | Cyclamate (E-952) n | Isomalt (E-953) n | Saccharine (E-954) n | Sucralose (E-955) n | Taumatine (E-957) n | Neohesperidine DC (E-959) n | Steviol glycosides (E-960) n | Neotame (E-961) n | Maltitol (E-965) n | Lactitol (E-966) n | Xylitol (E-967) n | Eritritol (E-968) n |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Sugars and sweets (*n* = 299)** | | | | | | | | | | | | | | | | |
| Chewing gum, candies, and sweets (*n* = 111) | 91 | 31 | 74 | 62 | 0 | 44 | 2 | 39 | 0 | 3 | 7 | 0 | 49 | 0 | 29 | 0 |
| Desserts and other sweets (*n* = 55) | 11 | 1 | 16 | 6 | 0 | 1 | 0 | 14 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 |
| Tabletop sweeteners (*n* = 52) | 0 | 0 | 4 | 7 | 15 | 0 | 16 | 6 | 2 | 0 | 22 | 0 | 0 | 0 | 3 | 5 |
| Chocolates (*n* = 46) | 0 | 0 | 7 | 4 | 0 | 3 | 0 | 3 | 0 | 1 | 17 | 0 | 39 | 2 | 0 | 2 |
| Jams (*n* = 35) | 12 | 0 | 8 | 0 | 0 | 4 | 0 | 28 | 0 | 0 | 19 | 0 | 16 | 0 | 0 | 0 |
| **Non-alcoholic beverages (*n* = 266)** | | | | | | | | | | | | | | | | |
| Diet soft drinks (*n* = 91) | 0 | 0 | 67 | 40 | 31 | 0 | 10 | 48 | 0 | 2 | 5 | 0 | 0 | 0 | 0 | 0 |
| Sugar soft drinks (*n* = 90) | 0 | 1 | 25 | 7 | 10 | 0 | 6 | 59 | 0 | 3 | 22 | 0 | 0 | 0 | 0 | 0 |
| Other non-alcoholic drinks (*n* = 32) | 0 | 0 | 12 | 1 | 4 | 0 | 3 | 13 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 |
| Fruit juices and nectars (*n* = 30) | 0 | 0 | 14 | 1 | 5 | 0 | 4 | 20 | 0 | 0 | 7 | 0 | 1 | 0 | 0 | 0 |
| Energy drinks (n = 14) | 0 | 0 | 9 | 2 | 0 | 0 | 0 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Coffee and herbal teas (*n* = 6) | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| Sport drinks (*n* = 3) | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Cereals and grains (*n* = 239)** | | | | | | | | | | | | | | | | |
| Bakery and pastry (*n* = 186) | 120 | 0 | 13 | 0 | 1 | 12 | 1 | 8 | 0 | 0 | 1 | 0 | 82 | 1 | 1 | 2 |
| Breakfast cereal and cereal bars (*n* = 48) | 24 | 0 | 3 | 0 | 0 | 1 | 0 | 4 | 0 | 0 | 1 | 0 | 14 | 0 | 2 | 0 |
| Bread (*n* = 5) | 4 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Milk and dairy products (*n* = 169)** | | | | | | | | | | | | | | | | |
| Yogurt and fermented milks (*n* = 126) | 0 | 0 | 105 | 30 | 0 | 0 | 0 | 87 | 0 | 0 | 14 | 3 | 0 | 0 | 0 | 0 |
| Other dairy products (*n* = 41) | 10 | 1 | 31 | 11 | 6 | 0 | 4 | 23 | 0 | 2 | 5 | 0 | 21 | 10 | 0 | 2 |
| Cheese (*n* = 2) | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Food supplements and substitutes (*n* = 109) | 36 | 4 | 18 | 8 | 7 | 2 | 5 | 37 | 0 | 0 | 15 | 0 | 10 | 0 | 6 | 1 |
| **Meat and meat products (*n* = 73)** | | | | | | | | | | | | | | | | |
| Sausages and other meat products (*n* = 70) | 70 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Meat (*n* = 3) | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ready-to-eat meals (*n* = 28) | 23 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Appetizers (*n* = 19) | 8 | 2 | 0 | 3 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| Sauces and condiments (*n* = 11) | 1 | 0 | 0 | 1 | 2 | 0 | 4 | 3 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Fish and shellfish (*n* = 9) | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Canned fruit (*n* = 9) | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Alcoholic beverages (*n* = 6) | 0 | 0 | 4 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Vegetables (*n* = 1) | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *n* total products = 1,238/LNCS total declarations | 423 | 41 | 415 | 188 | 83 | 71 | 56 | 421 | 2 | 13 | 143 | 3 | 239 | 13 | 41 | 12 |
| LNCS % within total declarations (*n* = 2,164) | 19.5% | 2% | 19.2% | 8.7% | 3.8% | 3.3% | 2.6% | 19.5% | 0.1% | 0.6% | 6.6% | 0.1% | 11% | 0.6% | 2% | 0.6% |

***n**, number of products included in each subgroup and types of LNCS are declared on labels of the products. One or more LNCS could be declared in products from each subgroup. For further description, refer to **Supplementary Table 1**.*

Desserts (i.e., gelatin) were comprised mainly of acesulfame K, sucralose, sorbitol, and maltitol and were the most widely used LNCS in the cereal and grain group. Within the non-alcoholic beverages group, acesulfame K and sucralose were the most frequent, especially in sugar soft drinks, while diet soft drinks declared the use of both but also the presence of aspartame and cyclamate. Meat and meat products declared only one type of LNCS, Sorbitol. Within the milk and dairy products group, where yogurt and fermented milk were the major subgroups, the mainly declared LNCS were acesulfame K, sucralose, and aspartame. Ready-to-eat meals, as well as appetizers and supplements and meal replacements, included sorbitol, and the latter also declared mainly sucralose, acesulfame K, and steviol glycosides. The minor group of LNCS was taumatine and (0.1%) and neotame (0.1%) that were present in table-top sweeteners and yogurt and fermented milk, respectively.

## DISCUSSION

In a global level, when comparing the food product database of 2019 with 2017, we currently observed that a higher number of food products included LNCS since there were more food products included in the webs of food retailers and, potentially, an increased number of reformulated products that have total or partially substituted added sugar contents.

Among the first studies that described LNCS consumption is the ANIBES Study (7), back in 2103, where food and beverage products consumed by participants were extracted from food records obtained from a three-day dietary record using a tablet device. Label data from 1,164 products of different brands were collected and reviewed for the content of LNCS. The results showed the diversity of food groups including these ingredients that the population consumed but failed to describe the market availability.

Then, in 2017, we started the compilation of a food composition database that included all products declaring LNCS in their ingredients lists (8). The present study follows the same methodology and, therefore, comparisons might be made with great precision. Concerning major food groups available, soft drinks, fruit juices and nectars, yogurts and fermented milk, and chewing gum and candies were coincident in their distribution. In turn, the new update shows a major proportion of bakery and pastry products (16% *vs*. 8%), supplement and meal replacers (9% *vs*. 4%), breakfast cereals and bars (3% *vs*. 1%), and ready-to-eat meals (2% *vs*. 1%). The food and beverage groups and subgroups containing LNCS are available more in number in the Spanish market, new food groups found in the present survey included meat and meat products ($n = 73$), fish and shellfish ($n = 9$), alcoholic beverages ($n = 6$), coffee and herbal teas ($n = 6$), and bread products ($n = 5$) and cheese ($n = 2$).

When assessing the types and combinations of LNCS added, the previous database showed that the most frequent combination of LNCS found was acesulfame K with sucralose (22%; $n = 81$), followed by individual addition of sucralose (13%; $n = 49$) and sorbitol (9%; $n = 34$), whereas mannitol, xylitol, neohesperidine DC, and lactitol were minor (≤1%). Similarly, in the present update, we found that the most frequent combination was acesulfame K with sucralose. In addition, sorbitol (19.5%; $n = 423$), sucralose (19.5%; $n = 421$) and acesulfame K (19.2%; $n = 415$) were the most declared LNCS. In addition, mannitol and xylitol increased their frequency of addition. We found that 52.9% of products from the current database declared the use of one LNCS, while 31.1% declared two LNCS and 10.3% declared three LNCS. The higher number of combinations determines that because a variety of LNCS is used, lower quantities of each one might be needed to be added, and this leads to the possibility of not exceeding the acceptable daily intake levels (ADI) (regarded as a safety threshold). In contrast, the larger availability of different types of foods containing LNCS reinforces the importance of continuous monitoring of the food consumption by the Spanish Population to follow up the dietary model weighing benefits and risks in the new post-reformulation era.

The contribution of these food and beverage groups containing LNCS to the daily total TE of the Spanish diet might be studied by reviewing the findings from the ANIBES study (16) that assessed the dietary intake of free and intrinsic sugars and their major food sources in the Spanish Population aged 9–75 years. The median total daily sugar consumption in the Spanish population was 71.5 g, contributing 17% of the TE intake. Of these, free sugars (i.e., added) accounted for 28.8 g (0.0–189.8 g; min–max) and 7.3% of TE. Therefore, the potential substitution of added sugars with LNCS, as shown within the major food and beverage groups, found in the present study could represent a decrease of up to 7.3% of TE, assuming that all sugar contents from these products were replaced by LNCS. The ANIBES findings indicate that the major sources contributing to free sugar intakes were sugar soft drinks (25.5%), sugar (17.8%), bakery and pastry items (15.2%), chocolates (11.4%), yogurt and fermented milk (6.44%), other dairy products (5.99%), jams (3.58%), juices and nectars (2.91%), and breakfast cereals and cereal bars (2.78%). Within these food and beverage products, we found in the present study that soft drinks, bakery and pastry, chocolates, and yogurt and fermented milk included LNCS in a higher percentage. Interestingly, a report from the Spanish Ministry of Agriculture, Fisheries, and Food in 2019 (17) showed that consumption of LNCS increased 1.6% from 2018 but did not give any further information on types of LNCS or the food groups to which they are added.

It is well-known that the packaged food and beverage segments of the food supply are dynamic and are characterized by continuous change and turnover as new products are introduced and less-favored products are removed from the retail market (18). These frequent changes require food composition databases to be continuously updated, however, so far LNCS are not included in most food composition databases worldwide, Spain is no exception. In addition, it is noteworthy that there is a fewer number of European studies assessing the presence of LNCS in food products.

In an Italian study (19) that collected 326 products containing LNCS from a label survey, non-alcoholic beverages, table-top sweeteners, and food supplements were the major contributors for almost all sweeteners. The most consumed sweeteners were acesulfame K, aspartame, and cyclamate. A study led in Belgium that was also conducted by a label survey between December 2009 and February 2010 (20), included 270 products, found that aspartame was the most frequently used LNCS (34%) as a single sweetener or primarily in combination with acesulfame K, followed by saccharine (24%), and cyclamate (22%). In addition, it was observed that the most important group were beverages, including non-alcoholic drinks (44%) and beers (12%), which, together, represented more than half of the total supply of sweetened foods. Sweets accounted for 12%, chewing gums for 4% and a noteworthy result was that only one cereal product contained LNCS as opposed to this study results, where cereals and grains were 19% of total products surveyed (20).

In countries like Chile, where heavy regulations to label high sugar content products are in place, the use of LNCS has increased as a part of food reformulation by manufacturers (21). Researchers highlight that there might be overuse of these additives and that it might represent a risk for some key populations, such as children, but a recent study (22) showed that this group does not exceed the ADI of any of the six LNCS authorized in Chile. The major food groups containing LNCS were non-alcoholic beverages (38.2%), 28.8% dairy products, 15.6% sweets and other desserts, 14.5% cereal products, and 2.9% processed fruits. Regarding the number of LNCSs present in these products, 42% contained only one LNCS, 48.1% had two LNCS, 8.6% had three LNCS, and only 1.4% had four LNCS (21). As for the types of LNCS from the Chilean study, sucralose and steviol glycosides were the most widely used LNCS, these sweeteners are present, either alone or mixed with other, in 73.5 and 39.7% of the LNCS-containing products, respectively, while the use of saccharin and cyclamate was low (21).

It seems evident from these data that LNCS distribution across food and beverage groups and the type of LNCS used is highly variable depending on the country studied. In addition, as reformulation practices are increasing to comply with plans promoting the reduction of added sugars, it could be relevant that EU regulations mandate the quantities of LNCS to be included in the ingredient composition to evaluate the potential level of ADI achieved by target populations, such as children.

At present, there is a controversy regarding the potential health benefits and risks associated with the use of LNCS. In a recent review by our group, we found a limited number of representative studies on the consumption of LNCS and its effect on health (3). However, these mostly indicate that the consumption of LNCS can be a useful tool along with other nutritional strategies in the treatment of overweight, obesity (23), diabetes (24), and the prevention of caries (25), when used appropriately in the context of a balanced diet and physical activity. However, it is necessary to be cautious with the consumption of certain sweeteners since the effects of LNCS on the intestinal microbiota (26) or its effect on premature deliveries (27), among others, have not been fully elucidated. Further studies should be undertaken to clarify the safety and value of sweeteners as food additives in the medium/long-term, in a model of increasing consumption as a consequence of the reformulation of many foods.

## Strengths and Limitations

Among the main strengths of this study, we may highlight the representativity of food and beverage products assessed, as the main Spanish retailer online platforms were scanned for LNCS containing products. In addition, the detailed description per food group and subgroup enables a further assessment of the intakes of the population. A possible limitation of the present work is the potential lack of updated information on the label of the product available from these retailers: availability of new products, reformulation by manufacturers, and elimination of others might hinder the comprehensive collection of data. Furthermore, to date, data regarding LNCS concentration among different food groups remain unavailable, given that manufacturers are not mandated to declare this information. Therefore, the actual content of LNCS (and not only presence) across food and beverage products remains unknown in our country. One major limitation of this study was the inability to evaluate the actual content of LNCS to be able to estimate dietary intakes. In this regard, biomarker analysis could be combined with dietary assessment as suggested by Logue et al. (28), who used a novel urinary biomarker approach for the determination of up to five LNCS to characterize consumption in adults.

# CONCLUSION

This study showed that the variety of LNCS in the present Spanish diet increased when compared with the previous cross-survey in 2017, where mainly soft drinks, fruit juices and nectars, yogurts and fermented milks, and chewing gum and candies contained LNCS; as they are now included in further subgroups such as meat and meat products, fish and shellfish, alcoholic beverages, coffee and herbal teas, bread products and cheese. The number and type of LNCS combinations have also improved, and this is remarkable because manufacturers can diversify the use of LNCS and decrease the amounts used in each of them. In addition, data concerning LNCS contents and presence in food and beverages are still not compiled in the food composition databases, and they should be periodically updated to enable their assessment and monitoring in nutritional surveys. As a result, further studies are needed in the future to continuously evaluate the evolution of the presence of LNCS in the Spanish diet. In addition, studies assessing the content of LNCS in foods are required, using the present investigation as a guide on which compounds should be investigated in food groups.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

BG-F and RU created the database. GV-M, RU, MS-V, and TP performed analyses and wrote the first draft of the manuscript. All authors contributed to read and approved the final manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.670422/full#supplementary-material

## REFERENCES


1. Sylvetsky AC, Rother KI. Trends in the consumption of low-calorie sweeteners. *Physiol Behav.* (2016) 164:446–50. doi: 10.1016/j.physbeh.2016.03.030
2. Dunford E, Taillie L, Miles D, Eyles H, Tolentino-Mayo L, Ng S. Non-nutritive sweeteners in the packaged food supply—an assessment across 4 countries. *Nutrients.* (2018) 10:257. doi: 10.3390/nu10020257
3. Samaniego-Vaesken MdL, Partearroyo T, Varela-Moreiras G. Low and no calorie sweeteners, diet and health: an updated overview. *Nutr Hospitalaria.* (2020) 37:24–7. doi: 10.20960/nh.03352
4. *EU Commission Regulation (EU) No 1129/2011 of 11 November 2011 Amending Annex II to Regulation (EC) No 1333/2008 of the European Parliament and of the Council by Establishing a Union List of Food Additives*. Official Journal of the European Union L295. Available online at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32011R1129andfrom=EN (accessed November 24, 2020).
5. EU Commission Regulation (EU) No 231/2012 of 9 March 2012 laying down specifications for food additives listed in Annexes II and III to Regulation (EC) No 1333/2008 of the European Parliament and of the Council Text with EEA relevance. *OJ L.* (2012) 83:1–295. Available online at: http://data.europa.eu/eli/reg/2012/231/oj
6. EU European Union. Regulation (EU) 1169/2011 of the European parliament and of the Council of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No 1924/2006 and(EC) No 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004 (Text with EEA relevance). Official Journal of the European Union, 2011; L304, 18-63. Available online at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32011R1169 (accessed November 24, 2020).
7. Samaniego-Vaesken MdL, Ruiz E, Partearroyo T, Aranceta-Bartrina J, Gil Á, González-Gross M, et al. Added sugars and low- and no-calorie sweeteners in a representative sample of food products consumed by the Spanish ANIBES study population. *Nutrients.* (2018) 10:1265. doi: 10.3390/nu10091265
8. Samaniego-Vaesken MdL, Partearroyo T, Cano A, Urrialde R, Varela-Moreiras G. Novel database of declared low- and no-calorie sweeteners from foods and beverages available in Spain. *J Food Compos Anal.* (2019) 82:103234. doi: 10.1016/j.jfca.2019.103234
9. European Food Safety Authority (EFSA). *Web Page for Sweeteners*. Available online at: https://www.efsa.europa.eu/en/topics/topic/sweeteners (accessed January 11, 2021).
10. Serra-Majem L, Raposo A, Aranceta-Bartrina J, Varela-Moreiras G, Logue C, Laviada H, et al. Ibero–American consensus on low- and no-calorie sweeteners: safety, nutritional aspects and benefits in food and beverages. *Nutrients.* (2018) 10:818. doi: 10.3390/nu10070818
11. Padilla A, Nestares T, Urrialde R. Actualización sobre aspectos científico-técnicos y regulatorios de los edulcorantes bajos en o sin calorías. *Rev Esp Nutr Comunitaria.* (2020) 26:80–92. doi: 10.14642/RENC.2020.26.2.5305
12. AESAN. Ministry of Health, Social Services and Equality. *Spanish Agency for Consumer Affairs, Food Safety and Nutrition. 2018. Collaboration Plan for the Improvement of the Composition of Food and Beverages and Other Measures 2020*. (2018). Available online at: http://www.aecosan.msssi.gob.es/AECOSAN/docs/documentos/nutricion/Plan_Colaboracion_INGLES.pdf (accessed November 22, 2020).
13. Martyn D, Darch M, Roberts A, Lee H, Yaqiong Tian T, Kaburagi N, et al. Low-/no-calorie sweeteners: a review of global intakes. *Nutrients.* (2018) 10:357. doi: 10.3390/nu10030357
14. Kantar. Kantar World Panel. *Balance de la Distribución y el Gran Consumo 2017*. (2017). Available online at: https://www.kantarworldpanel.com/es/Noticias/el-gran-consumo-crece-un-13-en-2017 (accessed December 1, 2019).
15. Kantar. *Kantar World Panel. Brand Footprint España 2018*. Las marcas de Gran Consumo más elegidas (2018). Available online at: https://www.kantarworldpanel.com/es/Noticias/Brand-Footprint-2018-Las-marcas-de-gran-consumo-mas-elegidas (accessed December 1, 2019).
16. Ruiz E, Rodriguez P, Valero T, Ávila J, Aranceta-Bartrina J. Gil Á, et al. Dietary intake of individual (free and intrinsic) sugars and food sources in the Spanish population: findings from the ANIBES study. *Nutrients.* (2017) 9:275. doi: 10.3390/nu9030275
17. MAPA. *Ministerio de Agricultura, Pesca y Alimentación. Informe del Consumo Alimentario en España 2019*. Madrid: Centro de Publicaciones del MAPA. Gobierno de España (2020).
18. Poti JM, Yoon E, Hollingsworth B, Ostrowski J, Wandell J, Miles DR, et al. Development of a food composition database to monitor changes in packaged foods and beverages. *J Food Composit Anal.* (2017) 64:18–26. doi: 10.1016/j.jfca.2017.07.024
19. Le Donne C, Mistura L, Goscinny S, Janvier S, Cuypers K, D'Addezio L, et al. Assessment of dietary intake of 10 intense sweeteners by the Italian population. *Food Chem Toxicol.* (2017) 102:186–97. doi: 10.1016/j.fct.2017.02.014
20. Huvaere K, Vandevijvere S, Hasni M, Vinkx C, Van Loco J. Dietary intake of artificial sweeteners by the Belgian population. *Food Addit Contam A.* (2012) 29:54–65. doi: 10.1080/19440049.2011.627572
21. Sambra V, López-Arana S, Cáceres P, Abrigo K, Collinao J, Espinoza A, et al. Overuse of non-caloric sweeteners in foods and beverages in Chile: a threat to consumers' free choice? *Front Nutr.* (2020) 7:68). doi: 10.3389/fnut.2020.00068
22. Martínez X, Zapata Y, Pinto V, Cornejo C, Elbers M, Graaf Mvd, et al. Intake of non-nutritive sweeteners in chilean children after enforcement of a new food labeling law that regulates added sugar content in processed foods. *Nutrients.* (2020) 12:1594. doi: 10.3390/nu12061594
23. Higgins KA, Mattes RD. A randomized controlled trial contrasting the effects of 4 low-calorie sweeteners and sucrose on body weight in adults with overweight or obesity. *Am J Clin Nutr.* (2019) 109:1288–301. doi: 10.1093/ajcn/nqy381
24. Lohner S, Kuellenberg de Gaudry D, Toews I, Ferenci T, Meerpohl JJ. Non-nutritive sweeteners for diabetes mellitus. *Cochrane Database Syst Rev.* (2020) 5:CD012885. doi: 10.1002/14651858.CD012885.pub2
25. Burt BA. The use of sorbitol-and xylitol-sweetened chewing gum in caries control. *J Am Dent Assoc.* (2006) 137:190–6. doi: 10.14219/jada.archive.2006.0144

26. Ruiz-Ojeda FJ, Plaza-Díaz J, Sáez-Lara MJ, Gil A. Effects of sweeteners on the gut microbiota: a review of experimental studies and clinical trials. *Adv Nutr.* (2019) 10:S31–48. doi: 10.1093/advances/nmy037
27. Bailón-Uriza R, Ayala Méndez JA, Cavagnari BM, Celis González C, Chapa Tellez R, Chávez Brambila J, et al. Edulcorantes no calóricos en la mujer en edad reproductiva: documento de consenso. *Nutr Hosp.* (2020) 37:211–22. doi: 10.20960/nh.2870
28. Logue C, Dowey LRC, Verhagen H, Strain J, O'Mahony M, Kapsokefalou M, et al. A novel urinary biomarker approach reveals widespread exposure to multiple low-calorie sweeteners in adults. *J Nutr.* (2020) 150:2435–41. doi: 10.1093/jn/nxaa184


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 15 October 2021
doi: 10.3389/fnut.2021.763677

# The Saturation Effect of Body Mass Index on Bone Mineral Density for People Over 50 Years Old: A Cross-Sectional Study of the US Population

*Ming Ma[1,2,3†], Zhiwei Feng[1,2†], Xiaolong Liu[1,2], Gengxin Jia[1,2], Bin Geng[2,3*] and Yayi Xia[2,3*]*

*[1] The Second School of Clinical Medical, Lanzhou University, Lanzhou, China, [2] Orthopaedics Key Laboratory of Gansu Province, Lanzhou, China, [3] Department of Orthopaedics, Lanzhou University Second Hospital, Lanzhou, China*

OPEN ACCESS

**Edited by:**
*Massimo Lucarini, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Geng-dong Chen, Foshan Women and Children Hospital, China*
*Leonor Gallardo, University of Castilla La Mancha, Spain*
*Haolong Zhou, Huazhong University of Science and Technology, China*

***Correspondence:**
*Bin Geng*
*cxxxf@qq.com*
*Yayi Xia*
*xiayy@lzu.edu.cn*

*[†] These authors have contributed equally to this work*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** *24 August 2021*
**Accepted:** *20 September 2021*
**Published:** *15 October 2021*

**Citation:**
*Ma M, Feng Z, Liu X, Jia G, Geng B and Xia Y (2021) The Saturation Effect of Body Mass Index on Bone Mineral Density for People Over 50 Years Old: A Cross-Sectional Study of the US Population. Front. Nutr. 8:763677. doi: 10.3389/fnut.2021.763677*

**Background:** Previous studies had revealed that Body Mass Index (BMI) positively affected Bone Mineral Density (BMD). However, an excessively high BMI was detrimental to health, especially for the elderly. Moreover, it was elusive how much BMI was most beneficial for BMD in older adults to maintain.

**Objective:** To investigate whether there was a BMI saturation effect value that existed to maintain optimal BMD.

**Methods:** A cross-sectional study was conducted using the datasets of the National Health and Nutrition Examination Survey (NHANES) 2005–2006, 2007–2008, 2009–2010, 2013–2014, and 2017–2018. After adjusting for covariates, an analysis of the association between BMI and BMD in different femoral regions (Total femur, Femoral neck, Trochanter, Intertrochanter, and Ward's triangle) and lumbar spine regions (Total spine, L1, L2, L3, and L4) in the whole population was performed using the multivariate linear regression models, smoothing curve fitting, and saturation effects analysis models. Then, subgroup analyses were performed according to gender, age, and race.

**Results:** A total of 10,910 participants (5,654 males and 5,256 females) over 50 years were enrolled in this population-based study. Multivariate linear regression analyses in the population older than 50 years showed that BMI was positively associated with femoral BMD and lumbar spine BMD ($P < 0.001$, respectively). Smoothing curve fitting showed that the relationship between BMI and BMD was not simply linear and that a saturation value existed. The saturation effect analysis showed that the BMI saturation value was 26.13 (kg/m$^2$) in the total femur, 26.82 (kg/m$^2$) in the total spine, and showed site-specificity in L1 (31.90 kg/m$^2$) and L2 (30.89 kg/m$^2$). The saturation values were consistent with the whole participants in males, while there was high variability in the females. BMI saturation values remained present in subgroup analyses by age and race, showing specificity in some age (60–70 years old) groups and in some races.

**Conclusions:** Our study showed a saturation value association between BMI and BMD for people over 50 years old. Keeping the BMI in the slightly overweight value (around 26 kg/m$^2$) might reduce other adverse effects while obtaining optimal BMD.

**Keywords: Body Mass Index, Bone Mineral Density, femur, lumbar spine, National Health and Nutrition Examination Survey (NHANES)**

## INTRODUCTION

Osteoporosis was a common disease of the skeletal system that imposed a substantial economic burden on society (1, 2) and posed a risk of fracture in the elderly (3, 4). As the most common clinical examination index, BMD was the gold standard for assessing osteoporosis (5). Obesity was one of the major public health issues to face in today's society, especially in the United States, with the highest adult obesity prevalence globally (6). Similarly, BMI was used in clinical situations to assess overweight and obesity (7).

Most of the available studies showed a positive association between BMI and BMD. A study in a 70-year-old population showed that patients with high BMI had a higher BMD and a lower risk of osteoporosis than those with a normal BMI (8). In older men and postmenopausal women, an increase in BMI was accompanied by BMD (9–11). However, an excessive BMI caused various other systemic diseases and complications such as hypertension (12), coronary heart disease (13), and type 2 diabetes (14). Possible mechanisms involved in the above diseases were the activation of the sympathetic nervous system (15), damage to the vascular endothelium (16), or insulin resistance (17).

Therefore, it was essential to strike a balance between BMI and BMD. However, existing researches were unclear exactly how much BMI was most beneficial for BMD while reducing the occurrence of other obesity-related complications. Therefore, we investigated the association between BMI and BMD in this work using a cross-sectional population survey sample from the NHANES (18) database for participants aged above 50 years, representing all regions and major ethnicities in the United States. We hypothesize that there was a saturation value of BMI and that keeping BMI at this value would result in an optimal balance between BMI and BMD.

## MATERIALS AND METHODS

### Data Sources

The present study used the data sets from the NHANES database to carry out a cross-sectional study. NHANES is a national health information source to collect the health and nutrition of adults and children in the USA. The NHANES project had approved by the National Center of Health Statistics (NCHS) Research Ethics Review Board (ERB), and every willing participant signed a consent document before starting. (NCHS IRB/ERB Protocol Number: Protocol #2005-06, Continuation of Protocol #2005-06, Continuation of Protocol #2011-17, Protocol #2018-01, https://www.cdc.gov/nchs/nhanes/irba98.htm).

### Participants Selected

Data sets were used from NHANES 2005–2006, NHANES 2007–2008, NHANES 2009–2010, NHANES 2013–2014, and NHANES 2017–2018 because femur BMD data were only collected the above period. Before the study began, the following people were excluded from BMD testing: (1) Participants refused to measure BMD. (2) Pregnancy. (3) History of radiographic contrast agents in the past 7 days. (4) Test weight over 450 pounds (Beyond the measurement range of DXA equipment). (5) History of bilateral hip fractures, bilateral hip replacements, and pins or steel in both hips. (6) Participants with degenerative spine disease include severe scoliosis, stiffening of the spine, previous spinal fusion, laminectomy, and spine fracture. (7) Insufficient scans of the participant's vertebrae and hips to complete the whole scan. (8) Overlapping of body parts during BMD measurement, e.g., overlapping of hands and legs. (9) Other causes were affecting BMD measurements, such as cardiac stenting, morbid obesity producing excessive x-ray noise. A total of 50,463 participants were included in this study initially. After excluded 36,297 participants under 50 years old, 869 participants without BMI data, and 2,387 participants without BMD data, leaving 10,910 participants included in this study finally (**Figure 1**).

### Definition of Lead Exposure BMI

BMI was calculated as weight (kg) divided by height (m) squared. Participants would not be excluded during the body measurements protocol for medical, safety, or other reasons. Values above the 99th percentile or below the 1st percentile were reviewed for data reasonableness. Data were reviewed for reasonableness based on a combination of height, weight, age, and gender of the participants. Data that were determined to be unrealistic after review will be artificially removed. All data during the BMI measurement would be the original data without any modifications.

### Definition of Outcomes Femoral BMD and Lumbar Spine BMD

The primary outcomes of this study were femoral BMD and lumbar spine BMD. Dual-energy X-ray absorptiometry (DXA) was used to measure the BMD of the participants (19). BMD measurement regions include femoral regions (Total femur, Femoral neck, Trochanter, Intertrochanter, and Ward's triangle) and lumbar spine regions (Total spine, L1, L2, L3, and L4). All measurement data would be subject to data review.

### Definition of Covariates

Demographic variables included Gender (Male, Female), Age, Race (Mexican American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black, and Other Race), Education level (<9th

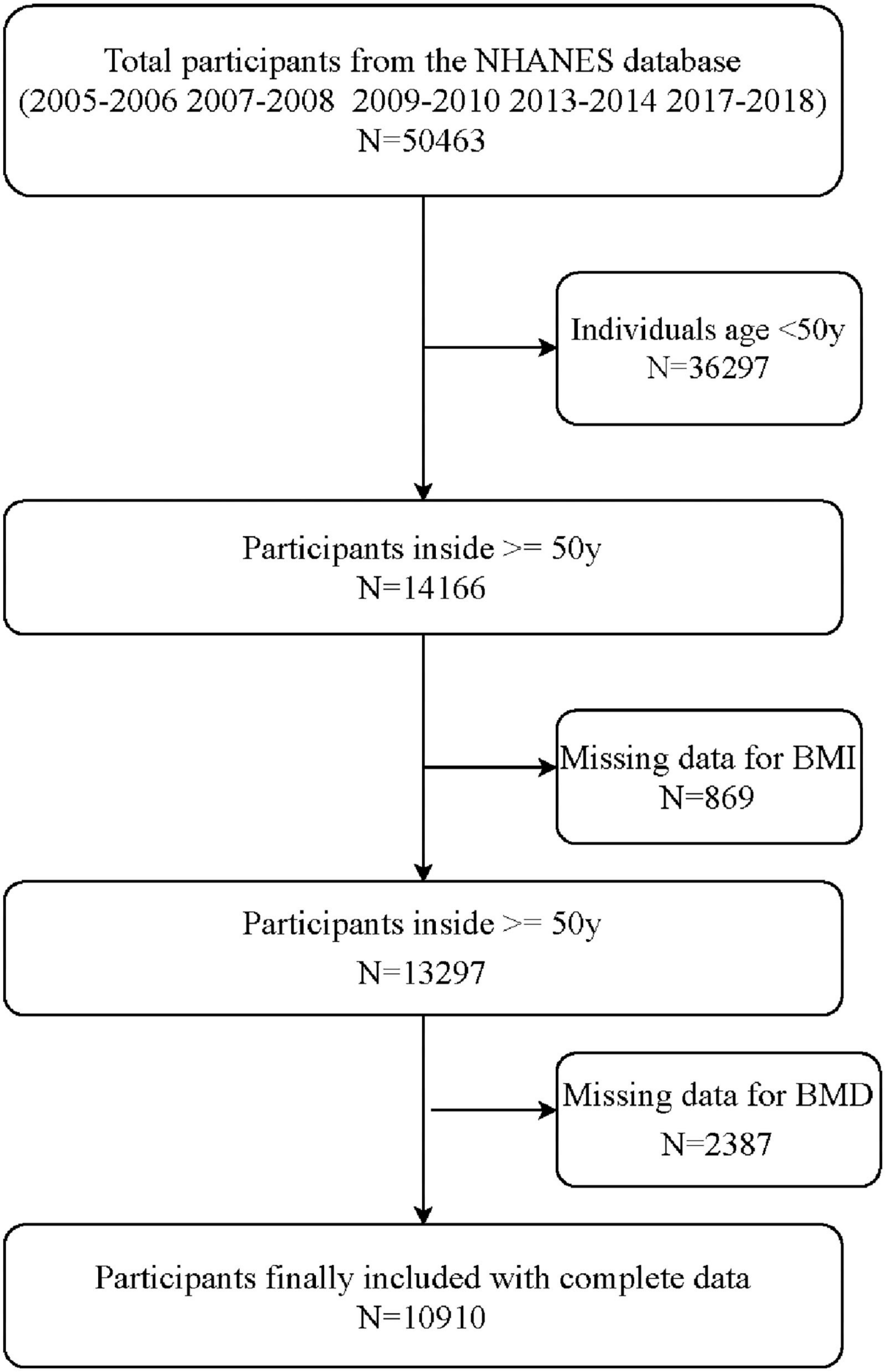


**FIGURE 1 |** Flow chart of participants selected from the NHANES database.

grade, 9–11th grade, High school graduate, Some college or AA degree, and College graduate or above), and Ratio of family income to poverty (0–5). The body covariates included Weight, Standing Height, Arm Circumference, and Waist Circumference. Furthermore, other covariates included Biochemistry (Albumin refrigerated serum, Globulin, Glucose refrigerated serum, Cholesterol refrigerated serum, and Triglycerides refrigerated serum), Moderate work activity, and Smoking-Cigarette Use (Smoked at least 100 cigarettes in life).

## Equipment Information

The Hologic Discovery A: Participants' BMD was examined with the *Hologic Discovery A*. *The Discovery A* used a low level of x-rays, and under standard operating conditions, the entrance dose to the examinee for a whole-body scan is less than one mR1 (a standard x-ray is ∼35 mR). Digital weight scale: The participants' weight was weighed using a digital weight scale placed on a flat surface. Portable weight scales: If the participant weighed beyond the scope of the digital weight scale, he or she was weighed on a portable scale. Stadiometer: The participants' height was measured used a stadiometer. Height adjustment ruler: A ruler ∼15 cm long was used to correct height if the participant's hairstyle interfered with the measurement or if they were unwilling to remove their shoes.

## Statistical Analysis

Statistical software R (version 3.6.1) and EmpowerStats Software were used to carry out the statistical analysis. Characteristics of the study population were conducted according to BMI subgroup (1) (Underweight, <18.5 kg/$m^2$; Normal, 18.5–24.9 kg/$m^2$; Overweight, 25–29.9 kg/$m^2$; and Obese, ≥30 kg/$m^2$). Mean ± standard deviation (SD), and Linear regression model for continuous variables (Age, Ratio of family income to poverty, Albumin refrigerated serum, Globulin, Glucose refrigerated serum, Cholesterol refrigerated serum, Triglycerides refrigerated serum, Weight, Standing Height, Waist Circumference, Arm Circumference, Total femur BMD, Femoral neck BMD, Trochanter BMD, Intertrochanter BMD, Wards triangle BMD, Total spine BMD, L1 BMD, L2 BMD, L3 BMD, L4 BMD. Frequencies (%) and a chi-squared test for categorical variables (Gender, Race, Education level, Moderate work activity, and Smoked at least 100 cigarettes in life). Multivariate linear regression analyses were conducted between the BMI and BMD to calculate the $\beta$ and *95% confidence interval* (*CI*). Three models were used to construct the multivariate test: Model 1, no covariates were adjusted; Model 2, Gender, Age, and Race were adjusted; Model 3, all covariates were adjusted. All covariates included Gender, Race, Age, Education level, Ratio of family income to poverty, Smoked at least 100 cigarettes in life, Moderate work activity, Albumin refrigerated serum, Globulin, Glucose refrigerated serum, Cholesterol refrigerated serum, Cholesterol Triglycerides, Standing Height, Arm Circumference, Waist Circumference. Whether covariates were adjusted in Model 3, we added them to the basic model or removed them from the full model, changing the BMI $\beta$ at least 10% or the covariate $P < 0.1$ in the univariate model (**Supplementary Table 1**). Smoothed curve fittings were performed simultaneously by adjusting the covariates. The association between BMI and BMD was analyzed with a saturated effects analysis model, and the results were expressed as the BMI Turning point (K), effect-$\beta$ (*95% CI*), and log-likelihood ratio test (*LRT*-test). In the saturated effect analysis model, the covariates were adjusted according to criteria 2 in **Supplementary Table 1**. Finally, the same multivariate linear regression, smoothed curve fitting, and saturation effects analyses were performed in the gender, age, and race subgroups. $P < 0.05$ was considered statistically significant. Sample weight: The 2-year sample weights were used for all NHANES analyses.

# RESULTS

## Characteristics of the Study Participants

The characteristics of study participants were shown in **Table 1**. In total, 5,654 male and 5,256 female adults above 50 years old were included. According to the BMI, 163 participants were underweight, 2,781 were normal, 4,185 were overweight, and 3,781 were obese. In each BMI group, the mean for age was 63.236 ± 10.085 years, 63.498 ± 9.887 years, 63.336 ± 9.410 years, and 61.805 ± 8.548 years, respectively. With increased BMI, the femur BMD and spine BMD significantly increased ($P < 0.00001$, respectively). Age, Gender, Race, Weight, Standing Height, Ratio of family income to poverty, Albumin refrigerated serum, Globulin, Glucose refrigerated serum, Cholesterol refrigerated serum, Cholesterol Triglycerides, Waist Circumference, Arm Circumference, Education level, Moderate work activity, and Smoke behavior were also presented in **Table 1**.

## The Association Between BMI and BMD

Multivariate linear regression analyses showed BMI was positively associated with Total femur BMD ($\beta = 0.010$, *95% CI* = 0.010, 0.011), Femoral neck BMD ($\beta = 0.014$, *95% CI* = 0.012, 0.01), Trochanter BMD ($\beta = 0.013$, *95% CI* = 0.012, 0.014), Intertrochanter BMD ($\beta = 0.016$, *95% CI* = 0.008, 0.009), Wards triangle BMD ($\beta = 0.013$, *95% CI* = 0.011, 0.015), Total spine BMD ($\beta = 0.008$, *95% CI* = 0.012, 0.01), L1 BMD ($\beta = 0.011$, *95% CI* = 0.009, 0.013), L2 BMD ($\beta = 0.013$, *95% CI* = 0.011, 0.015), L3 BMD ($\beta = 0.013$, *95% CI* = 0.011, 0.016), and L4 BMD ($\beta = 0.014$, *95% CI* = 0.012, 0.017) in the adjusted model (**Table 2**). When the BMI was categorized for analysis, the higher categorical had a higher BMD in the Total femur in adjusted model (BMI 18.5, 25: $\beta = 0.116$, *95% CI* = 0.095, 0.138; BMI 25, 30: $\beta = 0.186$, *95% CI* = 0.165, 0.207; BMI ≥ 30: $\beta = 0.247$, *95% CI* = 0.226, 0.269). Similarly, Femoral neck, Trochanter, Intertrochanter, Wards triangle, Total spine, L1, L2, L3, and L4 also had a higher BMD in BMI categorical analysis ($\beta$-Value, *95% CI*, and *P*-value were showed in **Table 2**).

When a smoothing curve fitting was conducted in the adjusted model (**Figure 2**), we found a saturation effect value between BMI and BMD. We further used the saturation effect analysis model to investigate the BMI turning point and found that the saturation effect value was 26.13 (kg/$m^2$) in the Total femur BMD, 26.59 (kg/$m^2$) in the Femoral neck BMD, 26.44 (kg/$m^2$) in the Trochanter BMD, 26.06 (kg/$m^2$) in the Intertrochanter BMD, 29.60 (kg/$m^2$) in the Wards triangle BMD, 26.82 (kg/$m^2$) in the

**TABLE 1 |** Characteristics of the study participants.

| | BMI (kg/m$^2$) categorical | | | | *P*-value |
|---|---|---|---|---|---|
| | **Underweight, ≤18.5 (*N* = 163)** | **Normal, >18.5, ≤25 (*N* = 2,781)** | **Overweight, >25, ≤30 (*N* = 4,185)** | **Obese, >30 (*N* = 3,781)** | |
| Age (years) | 63.236 ± 10.085 | 63.498 ± 9.887 | 63.336 ± 9.410 | 61.805 ± 8.548 | <0.00001 |
| Gender (%) | | | | | <0.00001 |
| Male | 38.683 | 39.426 | 55.319 | 48.435 | |
| Female | 61.317 | 60.574 | 44.681 | 51.565 | |
| Race (%) | | | | | <0.00001 |
| Mexican American | 1.847 | 3.120 | 5.605 | 5.604 | |
| Other Hispanic | 0.324 | 3.089 | 4.219 | 3.687 | |
| Non-Hispanic White | 72.688 | 75.851 | 76.225 | 76.063 | |
| Non-Hispanic Black | 13.944 | 7.384 | 7.885 | 11.474 | |
| Other Race | 11.197 | 10.557 | 6.066 | 3.173 | |
| Weight (kg) | 48.342 ± 5.896 | 62.685 ± 9.117 | 78.265 ± 10.377 | 97.658 ± 15.646 | <0.00001 |
| Standing height (cm) | 165.485 ± 8.853 | 166.370 ± 9.710 | 168.483 ± 10.309 | 167.711 ± 10.045 | <0.00001 |
| Ratio of family income to poverty | 2.387 ± 1.624 | 3.278 ± 1.598 | 3.301 ± 1.588 | 3.270 ± 1.570 | <0.00001 |
| Albumin, refrigerated serum (g/L) | 41.639 ± 4.164 | 42.084 ± 3.177 | 41.958 ± 3.003 | 40.936 ± 3.032 | <0.00001 |
| Globulin (g/L) | 29.952 ± 6.323 | 29.118 ± 5.203 | 29.176 ± 4.657 | 29.954 ± 4.718 | <0.00001 |
| Glucose, refrigerated serum (mmol/L) | 5.371 ± 1.953 | 5.734 ± 2.248 | 6.015 ± 2.229 | 6.423 ± 2.582 | <0.00001 |
| Cholesterol, refrigerated serum (mmol/L) | 5.135 ± 0.980 | 5.257 ± 1.058 | 5.209 ± 1.141 | 5.072 ± 1.122 | <0.00001 |
| Triglycerides, refrigerated serum (mmol/L) | 1.092 ± 0.626 | 1.409 ± 0.873 | 1.839 ± 1.298 | 2.089 ± 1.639 | <0.00001 |
| Waist Circumference (cm) | 72.738 ± 4.940 | 85.639 ± 7.571 | 99.079 ± 7.143 | 114.007 ± 10.537 | <0.00001 |
| Arm Circumference (cm) | 35.402 ± 2.428 | 36.510 ± 2.655 | 37.637 ± 2.767 | 38.283 ± 2.728 | <0.00001 |
| Total femur BMD (g/cm$^2$) | 0.719 ± 0.130 | 0.834 ± 0.136 | 0.929 ± 0.149 | 0.990 ± 0.150 | <0.00001 |
| Femoral neck BMD (g/cm$^2$) | 0.623 ± 0.116 | 0.694 ± 0.118 | 0.763 ± 0.129 | 0.815 ± 0.140 | <0.00001 |
| Trochanter BMD (g/cm$^2$) | 0.542 ± 0.112 | 0.632 ± 0.115 | 0.707 ± 0.129 | 0.750 ± 0.130 | <0.00001 |
| Intertrochanter BMD (g/cm$^2$) | 0.852 ± 0.159 | 0.992 ± 0.166 | 1.104 ± 0.177 | 1.176 ± 0.179 | <0.00001 |
| Wards triangle BMD (g/cm$^2$) | 0.449 ± 0.115 | 0.510 ± 0.135 | 0.563 ± 0.149 | 0.611 ± 0.166 | <0.00001 |
| Total spine BMD (g/cm$^2$) | 0.829 ± 0.150 | 0.932 ± 0.148 | 1.009 ± 0.158 | 1.059 ± 0.155 | <0.00001 |
| L1 BMD (g/cm$^2$) | 0.758 ± 0.153 | 0.856 ± 0.159 | 0.936 ± 0.165 | 0.999 ± 0.165 | <0.00001 |
| L2 BMD (g/cm$^2$) | 0.837 ± 0.161 | 0.934 ± 0.163 | 1.012 ± 0.173 | 1.066 ± 0.169 | <0.00001 |
| L3 BMD (g/cm$^2$) | 0.862 ± 0.152 | 0.975 ± 0.166 | 1.041 ± 0.172 | 1.092 ± 0.174 | <0.00001 |
| L4 BMD (g/cm$^2$) | 0.877 ± 0.159 | 0.978 ± 0.164 | 1.048 ± 0.175 | 1.097 ± 0.176 | <0.00001 |
| Moderate work activity (%) | | | | | 0.14931 |
| Yes | 33.601 | 38.005 | 40.236 | 38.435 | |
| No | 66.399 | 61.882 | 59.734 | 61.565 | |
| Education level (%) | | | | | <0.00001 |
| <9th grade | 9.289 | 6.346 | 6.970 | 5.453 | |
| 9–11th grade | 15.359 | 9.484 | 10.521 | 10.316 | |
| High school graduate | 31.251 | 23.436 | 25.371 | 27.820 | |
| Some college or AA degree | 17.784 | 26.408 | 27.462 | 31.744 | |
| College graduate or above | 26.317 | 34.325 | 29.676 | 24.667 | |
| Smoked at least 100 cigarettes in life (%) | | | | | 0.00007 |
| Yes | 68.455 | 48.844 | 50.008 | 48.972 | |
| No | 31.545 | 51.156 | 49.992 | 51.028 | |

*Mean ± SD for: Age, Ratio of family income to poverty, Albumin refrigerated serum (g/L), Globulin (g/L), Glucose refrigerated serum (mmol/L), Cholesterol refrigerated serum (mmol/L), Triglycerides refrig serum (mmol/L), Weight (kg), Standing Height (cm), Waist Circumference (cm), Arm Circumference (cm), Total femur BMD (g/cm$^2$), Femoral neck BMD (g/cm$^2$), Trochanter BMD (g/cm$^2$), Intertrochanter BMD (g/cm$^2$), Wards triangle BMD (g/cm$^2$), Total spine BMD (g/cm$^2$), L1 BMD (g/cm$^2$), L2 BMD (g/cm$^2$), L3 BMD (g/cm$^2$), L4 BMD (g/cm$^2$). P-value was calculated by weighted linear regression model.*

*Frequencies (%) for Gender, Race, Education level, Moderate work activity, Smoked at least 100 cigarettes in life. The weighted chi-square test calculated the P-value.*

*Weighted by: Full sample 2 year interview weight.*

**TABLE 2 |** Multiple linear regression analysis of the associations between BMI (kg/m$^2$) and BMD (g/cm$^2$) in different models.

| | Model | BMI (kg/m$^2$) | BMI (kg/m$^2$) categorical | | | | *P* for trend |
|---|---|---|---|---|---|---|---|
| | | | ≤18.5 | 18.5, 25 | 25, 30 | ≥30 | |
| Total femur BMD | Model 1, *β* (*95% CI*) | 0.012 (0.011, 0.012) | 0 | 0.115 (0.090, 0.139) | 0.210 (0.185, 0.234) | 0.271 (0.247, 0.296) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.011 (0.010, 0.011) | 0 | 0.118 (0.097, 0.139) | 0.191 (0.170, 0.211) | 0.254 (0.233, 0.275) | |
| | Model 3, *β* (*95% CI*) | 0.010 (0.010, 0.011) | 0 | 0.116 (0.095, 0.138) | 0.186 (0.165, 0.207) | 0.247 (0.226, 0.269) | |
| Femoral neck BMD | Model 1, *β* (*95% CI*) | 0.009 (0.009, 0.010) | 0 | 0.071 (0.049, 0.092) | 0.140 (0.118, 0.161) | 0.192 (0.170, 0.214) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.008 (0.008, 0.009) | 0 | 0.076 (0.056, 0.095) | 0.131 (0.112, 0.151) | 0.180 (0.161, 0.200) | |
| | Model 3, *β* (*95% CI*) | 0.014 (0.012, 0.010) | 0 | 0.048 (0.029, 0.060) | 0.079 (0.058, 0.100) | 0.086 (0.062, 0.109) | |
| Trochanter BMD | Model 1, *β* (*95% CI*) | 0.009 (0.008, 0.009) | 0 | 0.090 (0.069, 0.111) | 0.165 (0.144, 0.186) | 0.208 (0.187, 0.229) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.008 (0.008, 0.008) | 0 | 0.092 (0.074, 0.111) | 0.149 (0.130, 0.167) | 0.195 (0.176, 0.213) | |
| | Model 3, *β* (*95% CI*) | 0.013 (0.012, 0.014) | 0 | 0.065 (0.046, 0.084) | 0.097 (0.077, 0.117) | 0.103 (0.081, 0.125) | |
| Intertrochanter BMD | Model 1, *β* (*95% CI*) | 0.014 (0.013, 0.014) | 0 | 0.140 (0.111, 0.169) | 0.252 (0.223, 0.281) | 0.324 (0.295, 0.353) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.013 (0.012, 0.013) | 0 | 0.143 (0.118, 0.169) | 0.230 (0.204, 0.255) | 0.304 (0.279, 0.329) | |
| | Model 3, *β* (*95% CI*) | 0.016 (0.014, 0.018) | 0 | 0.100 (0.074, 0.126) | 0.145 (0.118, 0.173) | 0.155 (0.125, 0.186) | |
| Wards triangle BMD | Model 1, *β* (*95% CI*) | 0.008 (0.007, 0.008) | 0 | 0.061 (0.036, 0.086) | 0.114 (0.089, 0.140) | 0.163 (0.137, 0.188 | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.007 (0.006, 0.007) | 0 | 0.066 (0.043, 0.090) | 0.114 (0.091, 0.137) | 0.153 (0.130, 0.177) | |
| | Model 3, *β* (*95% CI*) | 0.013 (0.011, 0.015) | 0 | 0.047 (0.022, 0.071) | 0.078 (0.052, 0.103) | 0.085 (0.057, 0.114) | |
| Total spine BMD | Model 1, *β* (*95% CI*) | 0.010 (0.009, 0.010) | 0 | 0.103 (0.073, 0.133) | 0.180 (0.150, 0.210) | 0.230 (0.200, 0.260) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.009 (0.008, 0.009) | 0 | 0.107 (0.078, 0.135) | 0.169 (0.141, 0.197) | 0.220 (0.192, 0.248) | |
| | Model 3, *β* (*95% CI*) | 0.008 (0.008, 0.009) | 0 | 0.112 (0.083, 0.141) | 0.171 (0.142, 0.200) | 0.218 (0.189, 0.247) | |
| L1BMD | Model 1, *β* (*95% CI*) | 0.010 (0.010, 0.011) | 0 | 0.099 (0.070, 0.127) | 0.178 (0.150, 0.206) | 0.241 (0.213, 0.269) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.010 (0.009, 0.010) | 0 | 0.100 (0.074, 0.126) | 0.159 (0.134, 0.185) | 0.227 (0.201, 0.252) | |
| | Model 3, *β* (*95% CI*) | 0.011 (0.009, 0.013) | 0 | 0.073 (0.047, 0.100) | 0.107 (0.079, 0.136) | 0.135 (0.103, 0.167) | |
| L2 BMD | Model 1, *β* (*95% CI*) | 0.009 (0.009, 0.010) | 0 | 0.098 (0.068, 0.127) | 0.176 (0.146, 0.205) | 0.229 (0.200, 0.259) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.009 (0.008, 0.009) | 0 | 0.101 (0.073, 0.128) | 0.160 (0.133, 0.187) | 0.215 (0.187, 0.242) | |
| | Model 3, *β* (*95% CI*) | 0.013 (0.011, 0.015) | 0 | 0.081 (0.053, 0.110) | 0.121 (0.090, 0.151) | 0.143 (0.108, 0.177) | |
| L3 BMD | Model 1, *β* (*95% CI*) | 0.009 (0.008, 0.009) | 0 | 0.113 (0.082, 0.143) | 0.178 (0.148, 0.209) | 0.229 (0.199, 0.260) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.008 (0.008, 0.009) | 0 | 0.117 (0.087, 0.146) | 0.168 (0.139, 0.197) | 0.222 (0.193, 0.251) | |
| | Model 3, *β* (*95% CI*) | 0.013 (0.011, 0.016) | 0 | 0.101 (0.070, 0.131) | 0.136 (0.104, 0.169) | 0.161 (0.125, 0.198) | |
| L4 BMD | Model 1, *β* (*95% CI*) | 0.009 (0.008, 0.009) | 0 | 0.101 (0.069, 0.133) | 0.171 (0.140, 0.203) | 0.220 (0.188, 0.252) | *P* < 0.001 |
| | Model 2, *β* (*95% CI*) | 0.008 (0.008, 0.009) | 0 | 0.107 (0.077, 0.138) | 0.164 (0.133, 0.194) | 0.217 (0.187, 0.248) | |
| | Model 3, *β* (*95% CI*) | 0.014 (0.012, 0.017) | 0 | 0.089 (0.058, 0.121) | 0.127 (0.094, 0.161) | 0.146 (0.108, 0.184) | |

*Results: β (95% CI).*
*Outcome: Total femur BMD, Femoral neck BMD, Trochanter BMD, Intertrochanter BMD, Wards triangle BMD, Total spine BMD, L1 BMD, L2 BMD, L3 BMD, and L4 BMD.*
*Exposure: BMI (kg/m$^2$).*
*Model 1 adjusts for None.*
*Model 2 adjusted for Gender, Age, Race.*
*Model 3 adjusted for: Gender; Race; Gender, Race, Age, Education level, Ratio of family income to poverty, Smoked at least 100 cigarettes in life, Moderate work activity, Albumin refrigerated serum, Globulin, Glucose refrigerated serum, Cholesterol refrigerated serum, Cholesterol Triglycerides, Standing Height, Arm Circumference, Waist Circumference.*
*Weighted by: Full sample 2 year interview weight.*

Total spine BMD, 31.90 (kg/m$^2$) in the L1 BMD, 30.89 (kg/m$^2$) in the L2 BMD, 25.60 (kg/m$^2$) in the L3 BMD, and 25.60 (kg/m$^2$) in the L4 BMD (**Table 3**). When BMI was <26.13 (kg/m$^2$), the total femur BMD increased by 0.023 (*95% CI* = 0.022, 0.025) g/cm$^2$ for each unit increase in BMI. However, when BMI exceeded 26.13 (kg/m$^2$), the total femur BMD increased by only 0.007 (*95% CI* = 0.007, 0.008) g/cm$^2$ for each unit increase in BMI. Similarly, when BMI was less than the turning point, BMD of the Femoral neck, Trochanter, Intertrochanter, Wards triangle, Total spine, L1, L2, L3, and L4 increased by 0.016 (*95% CI* = 0.015, 0.017), 0.018 (*95% CI* = 0.017, 0.020), 0.028 (*95% CI* = 0.026, 0.030), 0.013 (*95% CI* = 0.011, 0.014), 0.018 (*95% CI* = 0.016, 0.020), 0.015 (*95% CI* = 0.014, 0.016), 0.015 (*95% CI* = 0.014, 0.016), 0.018 (*95% CI* = 0.016, 0.020), and 0.018 (*95% CI* = 0.016, 0.020) g/cm$^2$, respectively, for each unit increase in BMI. However, when BMI exceeds the turning point, BMD at the sites mentioned above increases very slowly (**Table 3**).

## Subgroup Analysis

The subgroup analysis conducted by gender was shown in **Supplementary Table 2** (Male) and **Supplementary Table 3** (Female). After adjusting the covariates, the smoothing curve fitting showed a saturation effect value between BMI and BMD in male participants (**Supplementary Figure 1**). After adjusting for covariates and using a saturated effects model, we found that the BMI saturation value was 26.25 (kg/m$^2$)

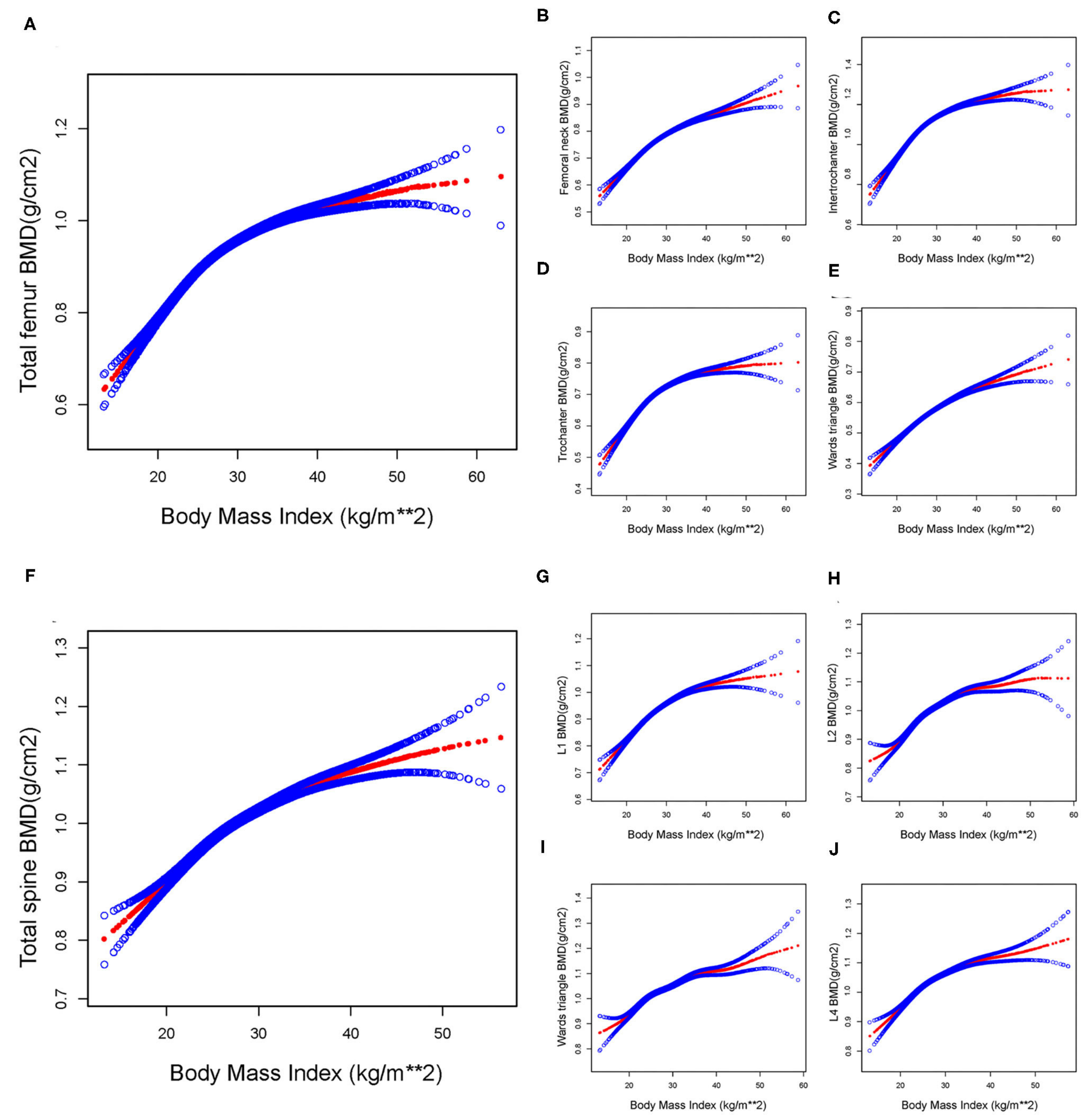


**FIGURE 2 |** The association between BMI and BMD in total participants. By saturation effect analysis, saturation values were found in all subgroups (Log-likelihood ratio test, $P < 0.001$, respectively). **(A)** Total femur group, BMI saturation value = 26.13 (kg/m$^2$). **(B)** Femoral neck group, BMI saturation value = 26.59 (kg/m$^2$). **(C)** Trochanter group, BMI saturation value = 26.44 (kg/m$^2$). **(D)** Intertrochanter group, BMI saturation value = 26.06 (kg/m$^2$). **(E)** Ward triangle group, BMI saturation value = 29.60 (kg/m$^2$). **(F)** Total spine group, BMI saturation value = 26.82 (kg/m$^2$). **(G)** L1 group, BMI saturation value = 31.90 (kg/m$^2$). **(H)** L2 group, BMI saturation value = 30.89 (kg/m$^2$). **(I)** L3 group, BMI saturation value = 25.60 (kg/m$^2$). **(J)** L4 group, BMI saturation value = 25.60 (kg/m$^2$). The red line represents the $\beta$-value, and the blue line represents the *95% CI*. Adjust for: Gender, Race, Age, Education level, Ratio of family income to poverty, Smoked at least 100 cigarettes in life, Moderate work activity, Albumin refrigerated serum, Globulin, Glucose refrigerated serum, Cholesterol refrigerated serum, Cholesterol Triglycerides, Standing Height, Arm Circumference, Waist Circumference.

in total femur BMD and 26.84 (kg/m$^2$) in the total spine in males participants. This was consistent with the results obtained for the whole participants. However, saturation effects were not evident in the female participants, and saturation values for each femur and lumbar spine region varied widely (**Supplementary Figure 2**).

**TABLE 3 |** Saturation effect analysis of BMI (kg/m$^2$) on BMD (g/cm$^2$) in whole participants.

| Outcome | Model: Saturation effect analysis | | | | *LRT*-test |
|---|---|---|---|---|---|
| | BMI turning point (K), kg/m$^2$ | <K, effect 1 | >K, effect 2 | Effect 2 – 1 | |
| Total femur BMD | 26.13 | 0.023 (0.022, 0.025) | 0.007 (0.007, 0.008) | −0.016 (−0.018, −0.014) | <0.001 |
| Femoral neck BMD | 26.59 | 0.016 (0.015, 0.017) | 0.007 (0.006, 0.007) | −0.009 (−0.011, −0.008) | <0.001 |
| Trochanter BMD | 26.44 | 0.018 (0.017, 0.020) | 0.005 (0.004, 0.006) | −0.013 (−0.015, −0.012) | <0.001 |
| Intertrochanter BMD | 26.06 | 0.028 (0.026, 0.030) | 0.008 (0.007, 0.009) | −0.020 (−0.022, −0.017) | <0.001 |
| Wards triangle BMD | 29.60 | 0.013 (0.011, 0.014) | 0.006 (0.005, 0.007) | −0.006 (−0.008, −0.005) | <0.001 |
| Total spine BMD | 26.82 | 0.018 (0.016, 0.020) | 0.006 (0.005, 0.007) | −0.012 (−0.015, −0.010) | <0.001 |
| L1 BMD | 31.90 | 0.015 (0.014, 0.016) | 0.004 (0.003, 0.005) | −0.012 (−0.014, −0.010) | <0.001 |
| L2 BMD | 30.89 | 0.015 (0.014, 0.016) | 0.002 (0.001, 0.003) | −0.013 (−0.015, −0.011) | <0.001 |
| L3 BMD | 25.60 | 0.018 (0.016, 0.020) | 0.006 (0.005, 0.007) | −0.012 (−0.014, −0.009) | <0.001 |
| L4 BMD | 25.60 | 0.018 (0.016, 0.020) | 0.006 (0.005, 0.006) | −0.012 (−0.015, −0.010) | <0.001 |

*Results in the table: β (95% CI).*
*Outcome: Total femur BMD, Femoral neck BMD, Femoral neck BMD, Trochanter BMD, Intertrochanter BMD, Wards triangle BMD, Total spine BMD, L1 BMD, L2 BMD, L3 BMD, L4 BMD.*
*Exposure: BMI (kg/m$^2$).*
*Adjust for: Gender, Race, Age, Education level, Ratio of family income to poverty, Smoked at least 100 cigarettes in life, Moderate work activity, Albumin refrigerated serum, Globulin, Glucose refrigerated serum, Cholesterol refrigerated serum, Cholesterol Triglycerides, Standing Height, Arm Circumference, Waist Circumference.*
*LRT-test: Log-likelihood ratio test.*

The subgroup analysis conducted by age was shown in **Supplementary Table 4**. After adjusting the covariates, the smoothing curve fitting also showed a saturation effect value. The saturation effect model showed that in the participants aged 60–70 years, the saturation values of BMI in the Femoral neck, Ward triangle, Total spine, and L1–L4 reached 32 (kg/m$^2$), and the results in other age groups were similar in the whole participants (**Supplementary Figure 3**).

The subgroup analysis conducted by race was shown in **Supplementary Table 5**. The results showed that the non-Hispanic population had a higher BMI saturation value at the lumbar spine region than at around 30 (kg/m$^2$), while others had a BMI saturation value at around 24 (kg/m$^2$). Mexican Americans and other Hispanics had higher BMI saturation values than other races at several sites in the femoral region, reaching around 35 (kg/m$^2$) in some areas (**Supplementary Figure 4**).

## DISCUSSION

In the present study, a BMI saturation value (around 26 kg/m$^2$) was founded in the femur BMD and spine BMD in all participants over 50 years old. A positive association was found between BMI and BMD at BMI level <26 (kg/m$^2$) and a minimal increase in BMI level above 26 (kg/m$^2$), this being important for maintaining optimal BMD.

A positive association was found between BMI and BMD. When BMI was <26 (kg/m$^2$), for each unit increased in BMI, the BMD of the femur and spine increased by 0.023 (g/cm$^2$) and 0.018 (g/cm$^2$), which was consistent with the previous studies (20, 21). A cross-sectional study in a Polish population also showed a positive association between BMI and BMD (22). Another study by Morin et al. indicated that low BMI predicted the future occurrence of osteoporosis and increased fracture risk (23). In the middle-aged and elderly female population, maintaining obesity could prevent the onset of postmenopausal bone loss (24). Similarly, a 10.5-year prospective cohort study suggested that obesity may be associated with the delayed bone loss (25). There were several possible mechanisms for a positive association between BMI and BMD as follows: (1) The increased static mechanical compliance produced by excessive fat accumulation was one of the possible mechanisms. Excessive fat accumulation and high body weight could impose greater static mechanical loads on the bones, and bone tissue produced a series of changes when it felt the mechanical forces exerted by the body (26, 27). (2) More body fat in patients with high BMI was accompanied by an increase in various hormones, such as estrogen (28), insulin (29), and leptin (30), which had a beneficial effect on BMD by inhibiting bone resorption (31) and bone remodeling (32, 33). (3) Androgens in adipose tissue were converted to estrogen, which would increase bone mass (34). (4) Studies also reported the effect of some genes (35) on BMD, such as the mutation of the Pro10 allele in tumor necrosis factor-β1 (TGF-β1), which is more frequent in obese patients (36).

Nevertheless, when BMI exceeded a specific value of 26 (kg/m$^2$), the BMD of the femur and spine increased with each unit of BMI by only 0.007 (g/cm$^2$) and 0.006 (g/cm$^2$). The mechanisms of keeping the BMI around 26 to have the most optimal BMD had not been fully explained. A review of the literature allowed us to draw some possible clues. (1) Genetic determinism. Bone growth trajectories (37) and peak bone mass (38) were determined early in life, which might be one of the possible explanations that BMD no longer increased after a limited increase in value in adults. Pocock et al. (39) explored genetic effects on BMD in 38 identical and 27 dizygotic twins and found that genetics determines adult bone mass. Genetic influences explained 75% of BMD variance, regardless of whether the twin was male (40) or female (41). Meanwhile, other acquired environmental factors, such as increased BMI (21), calcium intake (42), estrogen intake (43), and physical exercise (44), have a limit to the increase in BMD in adults. (2) Multi-factors co-leading. Researches had shown a specific bone-adipose axis (45) between adipose and bone tissues within the body, connected by a variety of bioactive molecules and maintained bone homeostasis, which might be another possible mechanism for the presence of BMI saturation effects. Available studies clarified that bone and adipocytes originated from a common stem cell precursor and were competitive, with excess fat gain leading to bone loss (46, 47). Several experiments in animal models induced by high-fat diets had confirmed that BMD in obese animals decreased with increasing obesity (48–50). The PPAR-γ (Peroxisome Proliferator-Activated Receptor-γ) pathway (46) was involved in adipose and bone differentiation in obesity animal models *in vivo*. Activation of PPAR-γ (51) stimulates adipogenesis and bone loss, while inhibition of PPAR-γ prevents (52) bone loss. Other studies had shown that adipose was an endocrine organ that secreted inflammatory cells such as interleukin-1 (IL-1) and tumor necrosis factor-α (TNF-α) (53). The above cytokines inhibit BMD via the OPG/RANKL/RANK pathway (54). We hypothesized that the multiple factors mentioned above combined to cause an increase in BMI to a certain range without a significant increase in BMD, but direct experimental evidence was insufficient, so more advanced studies were still needed.

Our results showed that although the BMI saturation values were similar in the total femur (BMI saturation value, 26.13 $kg/m^2$) and total spine (BMI saturation value, 26.82 $kg/m^2$), they were significantly higher in the L1 (BMI saturation value, 31.09 $kg/m^2$) and L2 (BMI saturation value, 30.89 $kg/m^2$) than in the other sites, showed site-specificity. This might be related to the increased spondylarthrosis in the elderly, which led to reactive changes and increasing BMD in the lumbar (55). Unfortunately, we could not extract osteoarthritis-related data for a related study because of limitations in the database itself. Subgroup analysis by gender showed that the BMI saturation values for male participants' femur and spine were concentrated around 26 $kg/m^2$. However, the BMI saturation values for female participants varied greatly by each site, which might be linked to the different levels of sex hormones in males and females (56). Age-subgrouped analysis showed that people aged 60–70 years had higher BMI saturation values for the spine than other age groups. A 10-year survey of bone loss rates in the elderly population from Japan suggested that this was related to the prevalence of spine osteophytes in this age group (57). In addition, race subgroup analysis revealed significant differences in BMI saturation values across races, suggesting that different genetic backgrounds and ancestry might be associated with this phenomenon (58).

Indeed, excessive BMI was detrimental to the elderly population. For one, high BMI brought a range of bone-related diseases, such as increased bone fragility (59) and increased fracture risk (60). On the other hand, obesity could lead to a variety of chronic diseases and complications, such as cardiovascular disease (61), type 2 diabetes (62), gallbladder disease (63), and fatty liver (64). More seriously, excessive obesity might be associated with increased cancer risk (65) and cancer-related mortality (66). Therefore, we believed that keeping BMI at a reasonable value (around 26 $kg/m^2$) would maintain optimal BMD and reduce the risk of other obesity-related diseases and complications.

There were several limitations in this study. Firstly, this was a cross-sectional study, and we could only conclude whether there was an association between BMI and BMD, not a direct cause-and-effect correlation. Secondly, due to limitations in the database itself, we could not identify the medication-taking participants, menstrual and menopausal (female participants), and osteoarthritis participants, so our conclusions need to be interpreted with caution. Thirdly, we performed covariates detection to control confounders, but there might still be unpredictable covariates, such as medicine use, menstrual status (female), body fat percentage, etc. Last but not least, to the best of our knowledge, the present study was conducted in a US population, and the findings should be cautiously extended to other populations.

## CONCLUSION

In the present study, we used the multivariate linear regression models, smoothing curve fitting, and saturation effects analysis models to investigate the association between BMI and BMD of people over 50 years old in the US. We found not only a simple linear positive association between BMI and BMD, but a saturation value existed, and this saturation value persisted during site-specific analysis, sex, age, and race subgroup analysis. This work indicated that keeping BMI at a reasonable value (around 26 $kg/m^2$) will provide the most significant benefit to older adults in maintaining optimal BMD and reducing other obesity-related diseases.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by NCHS IRB/ERB Protocol Number: Protocol #2005-06, Continuation of Protocol #2005-06, Continuation of Protocol #2011-17, Protocol #2018-01. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

MM and ZF conceived the study, data curation, data analysis, and draft writing. XL and GJ completed images and tables preparation. BG conceived the study and writing Instruction. YX conceived the study, funding acquisition, and writing-review editing. All the authors participated in critical revision of the manuscript, contributed to the article, and approved the submitted version.


## FUNDING

This work was supported by the National Natural Science Foundation of China (81874017, 81960403, and 82060405); Natural Science Foundation of Gansu Province of China (20JR5RA320); and Cuiying Scientific and Technological Innovation Program of Lanzhou University Second Hospital (CY2017-ZD02).


## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.763677/full#supplementary-material

## REFERENCES


1. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. *J Bone Miner Res.* (2007) 22:465–75. doi: 10.1359/jbmr.061113
2. Lane JM, Russell L, Khan SN. Osteoporosis. *Clin Orthop Relat Res.* (2000) 372:139–50. doi: 10.1097/00003086-200003000-00016
3. Black DM, Rosen CJ. Clinical practice. Postmenopausal osteoporosis. *N Engl J Med.* (2016) 374:254–62. doi: 10.1056/NEJMcp1513724
4. Seeman E, Delmas PD. Bone quality–the material and structural basis of bone strength and fragility. *N Engl J Med.* (2006) 354:2250–61. doi: 10.1056/NEJMra053077
5. Hauk L. Treatment of low BMD osteoporosis to prevent fractures: updated guideline from the ACP. *Am Fam Phys.* (2018) 97:352–3.
6. Mujica OJ, Victora CG. Obesity inequality among adults in Latin America and the Caribbean. *Lancet Glob Health.* (2019) 7:e1589–90. doi: 10.1016/S2214-109X(19)30460-7
7. Daniels SR. The use of BMI in the clinical setting. *Pediatrics.* (2009) 124(Suppl. 1):S35–41. doi: 10.1542/peds.2008-3586F
8. Barrera G, Bunout D, Gattás V, de la Maza MP, Leiva L, Hirsch S. A high body mass index protects against femoral neck osteoporosis in healthy elderly subjects. *Nutrition.* (2004) 20:769–71. doi: 10.1016/j.nut.2004.05.014
9. Głogowska-Szelag J. [Assessment of the relationship between BMD and body mass index BMI in women with postmenopausal osteoporosis]. *Wiad Lek.* (2018) 71:1714–8.
10. Edelstein SL, Barrett-Connor E. Relation between body size and bone mineral density in elderly men and women. *Am J Epidemiol.* (1993) 138:160–9. doi: 10.1093/oxfordjournals.aje.a116842
11. Ravn P, Cizza G, Bjarnason NH, Thompson D, Daley M, Wasnich RD, et al. Low body mass index is an important risk factor for low bone mass and increased bone loss in early postmenopausal women. Early Postmenopausal Intervention Cohort (EPIC) study group. *J Bone Miner Res.* (1999) 14:1622–7. doi: 10.1359/jbmr.1999.14.9.1622
12. Seravalle G, Grassi G. Obesity and hypertension. *Pharmacol Res.* (2017) 122:1–7. doi: 10.1016/j.phrs.2017.05.013
13. Garrison RJ, Higgins MW, Kannel WB. Obesity and coronary heart disease. *Curr Opin Lipidol.* (1996) 7:199–202. doi: 10.1097/00041433-199608000-00005
14. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature.* (2006) 444:840–6. doi: 10.1038/nature05482
15. DeLalio LJ, Sved AF, Stocker SD. Sympathetic nervous system contributions to hypertension: updates and therapeutic relevance. *Can J Cardiol.* (2020) 36:712–20. doi: 10.1016/j.cjca.2020.03.003
16. Li M, Qian M, Kyler K, Xu J. Adipose tissue-endothelial cell interactions in obesity-induced endothelial dysfunction. *Front Cardiovasc Med.* (2021) 8:681581. doi: 10.3389/fcvm.2021.681581
17. Greenfield JR, Campbell LV. Insulin resistance and obesity. *Clin Dermatol.* (2004) 22:289–95. doi: 10.1016/j.clindermatol.2004.01.011
18. Patel CJ, Pho N, McDuffie M, Easton-Marks J, Kothari C, Kohane IS, et al. A database of human exposomes and phenomes from the US National Health and Nutrition Examination Survey. *Sci Data.* (2016) 3:160096. doi: 10.1038/sdata.2016.96
19. Baran DT, Faulkner KG, Genant HK, Miller PD, Pacifici R. Diagnosis and management of osteoporosis: guidelines for the utilization of bone densitometry. *Calcif Tissue Int.* (1997) 61:433–40. doi: 10.1007/s002239900362
20. Song J, Zhang R, Lv L, Liang J, Wang W, Liu R, et al. The relationship between Body Mass Index and bone mineral density: a Mendelian randomization study. *Calcif Tissue Int.* (2020) 107:440–5. doi: 10.1007/s00223-020-00736-w
21. Dogan A, Nakipoglu-Yüzer GF, Yildizgören MT, Ozgirgin N. Is age or the body mass index (BMI) more determinant of the bone mineral density (BMD) in geriatric women and men? *Arch Gerontol Geriatr.* (2010) 51:338–41. doi: 10.1016/j.archger.2010.01.015
22. Wiacek M, Skrzek A, Ignasiak Z, Zubrzycki IZ. The changes of bone mineral density in relation to body mass index and aging among Polish and different ethnic women in the United States: cross-sectional studies. *J Clin Densitom.* (2010) 13:307–14. doi: 10.1016/j.jocd.2010.03.002
23. Morin S, Tsang JF, Leslie WD. Weight and body mass index predict bone mineral density and fractures in women aged 40 to 59 years. *Osteoporos Int.* (2009) 20:363–70. doi: 10.1007/s00198-008-0688-x
24. Ribot C, Tremollieres F, Pouilles JM, Bonneu M, Germain F, Louvet JP. Obesity and postmenopausal bone loss: the influence of obesity on vertebral density and bone turnover in postmenopausal women. *Bone.* (1987) 8:327–31. doi: 10.1016/8756-3282(87)90062-7
25. Saarelainen J, Kiviniemi V, Kröger H, Tuppurainen M, Niskanen L, Jurvelin J, et al. Body mass index and bone loss among postmenopausal women: the 10-year follow-up of the OSTPRE cohort. *J Bone Miner Metab.* (2012) 30:208–16. doi: 10.1007/s00774-011-0305-5
26. Lanyon LE. Control of bone architecture by functional load bearing. *J Bone Miner Res.* (1992) 7(Suppl. 2):S369–75. doi: 10.1002/jbmr.5650071403
27. Hla MM, Davis JW, Ross PD, Wasnich RD, Yates AJ, Ravn P, et al. A multicenter study of the influence of fat and lean mass on bone mineral content: evidence for differences in their relative influence at major fracture sites. Early Postmenopausal Intervention Cohort (EPIC) study group. *Am J Clin Nutr.* (1996) 64:354–60. doi: 10.1093/ajcn/64.3.345
28. Movérare-Skrtic S, Wu J, Henning P, Gustafsson KL, Sjögren K, Windahl SH, et al. The bone-sparing effects of estrogen and WNT16 are independent of each other. *Proc Natl Acad Sci USA.* (2015) 112:14972–7. doi: 10.1073/pnas.1520408112
29. Costantini S, Conte C. Bone health in diabetes and prediabetes. *World J Diabetes.* (2019) 10:421–45. doi: 10.4239/wjd.v10.i8.421
30. Isaia GC, D'Amelio P, Di Bella S, Tamone C. Is leptin the link between fat and bone mass? *J Endocrinol Invest.* (2005) 28:61–5.
31. Guo L, Chen K, Yuan J, Huang P, Xu X, Li C, et al. Estrogen inhibits osteoclasts formation and bone resorption via microRNA-27a targeting PPARγ and APC. *J Cell Physiol.* (2018) 234:581–94. doi: 10.1002/jcp.26788
32. Krishnan A, Muthusami S. Hormonal alterations in PCOS and its influence on bone metabolism. *J Endocrinol.* (2017) 232:R99–113. doi: 10.1530/JOE-16-0405
33. Martin, de Vittoris R, David V, Moraes R, Bégeot M, Lafage-Proust MH, et al. Leptin modulates both resorption and formation while preventing disuse-induced bone loss in tail-suspended female rats. *Endocrinology.* (2005) 146:3652–9. doi: 10.1210/en.2004-1509
34. Bélanger C, Luu-The V, Dupont P, Tchernof A. Adipose tissue intracrinology: potential importance of local androgen/estrogen metabolism in the regulation of adiposity. *Horm Metab Res.* (2002) 34:737–45. doi: 10.1055/s-2002-38265

35. Nguyen TV, Howard GM, Kelly PJ, Eisman JA. Bone mass, lean mass, and fat mass: same genes or same environments? *Am J Epidemiol.* (1998) 147:3–16. doi: 10.1093/oxfordjournals.aje.a009362
36. Rosmond R, Chagnon M, Bouchard C, Björntorp P. Increased abdominal obesity, insulin and glucose levels in nondiabetic subjects with a T29C polymorphism of the transforming growth factor-beta1 gene. *Horm Res.* (2003) 59:191–4. doi: 10.1159/000069323
37. Jones G, Dwyer T. Birth weight, birth length, and bone density in prepubertal children: evidence for an association that may be mediated by genetic factors. *Calcif Tissue Int.* (2000) 67:304–8. doi: 10.1007/s002230001148
38. Weiler HA, Yuen CK, Seshia MM. Growth and bone mineralization of young adults weighing less than 1500 g at birth. *Early Hum Dev.* (2002) 67:101–12. doi: 10.1016/S0378-3782(02)00003-8
39. Pocock NA, Eisman JA, Hopper JL, Yeates MG, Sambrook PN, Eberl S. Genetic determinants of bone mass in adults. A twin study. *J Clin Invest.* (1987) 80:706–10. doi: 10.1172/JCI113125
40. Videman T, Levälahti E, Battié MC, Simonen R, Vanninen E, Kaprio J. Heritability of BMD of femoral neck and lumbar spine: a multivariate twin study of Finnish men. *J Bone Miner Res.* (2007) 22:1455–62. doi: 10.1359/jbmr.070606
41. Flicker L, Hopper JL, Rodgers L, Kaymakci B, Green RM, Wark JD. Bone density determinants in elderly women: a twin study. *J Bone Miner Res.* (1995) 10:1607–13. doi: 10.1002/jbmr.5650101102
42. Calvin Hirsch. ACP Journal Club. Review: dietary or supplemental calcium increase BMD by ≤1.8% in persons >50 years of age. *Ann Intern Med.* (2016) 164:JC5. doi: 10.7326/ACPJC-2016-164-2-005
43. Watts NB, Nolan JC, Brennan JJ, Yang HM. Esterified estrogen therapy in postmenopausal women. Relationships of bone marker changes and plasma estradiol to BMD changes: a two-year study. *Menopause.* (2000) 7:375–82. doi: 10.1097/00042192-200011000-00002
44. Tønnesen R, Schwarz P, Hovind PH, Jensen LT. Physical exercise associated with improved BMD independently of sex and vitamin D levels in young adults. *Eur J Appl Physiol.* (2016) 116:1297–304. doi: 10.1007/s00421-016-3383-1
45. Gómez-Ambrosi J, Rodríguez A, Catalán V, Frühbeck G. The bone-adipose axis in obesity and weight loss. *Obes Surg.* (2008) 18:1134–43. doi: 10.1007/s11695-008-9548-1
46. Akune T, Ohba S, Kamekura S, Yamaguchi M, Chung UI, Kubota N, et al. PPARgamma insufficiency enhances osteogenesis through osteoblast formation from bone marrow progenitors. *J Clin Invest.* (2004) 113:846–55. doi: 10.1172/JCI200419900
47. Gregoire FM, Smas CM, Sul HS. Understanding adipocyte differentiation. *Physiol Rev.* (1998) 78:783–809. doi: 10.1152/physrev.1998.78.3.783
48. Halade GV, Rahman MM, Williams PJ, Fernandes G. High fat diet-induced animal model of age-associated obesity and osteoporosis. *J Nutr Biochem.* (2010) 21:1162–9. doi: 10.1016/j.jnutbio.2009.10.002
49. Iwaniec UT, Dube MG, Boghossian S, Song H, Helferich WG, Turner RT, et al. Body mass influences cortical bone mass independent of leptin signaling. *Bone.* (2009) 44:404–12. doi: 10.1016/j.bone.2008.10.058
50. Halade GV, El Jamali A, Williams PJ, Fajardo RJ, Fernandes G. Obesity-mediated inflammatory microenvironment stimulates osteoclastogenesis and bone loss in mice. *Exp Gerontol.* (2011) 46:43–52. doi: 10.1016/j.exger.2010.09.014
51. David V, Martin A, Lafage-Proust MH, Malaval L, Peyroche S, Jones DB, et al. Mechanical loading down-regulates peroxisome proliferator-activated receptor gamma in bone marrow stromal cells and favors osteoblastogenesis at the expense of adipogenesis. *Endocrinology.* (2007) 148:2553–62. doi: 10.1210/en.2006-1704
52. Li X, Ning L, Ma J, Xie Z, Zhao X, Wang G, et al. The PPAR-γ antagonist T007 inhibits RANKL-induced osteoclastogenesis and counteracts OVX-induced bone loss in mice. *Cell Commun Signal.* (2019) 17:136. doi: 10.1186/s12964-019-0442-3
53. Abrahamsen B, Bonnevie-Nielsen V, Ebbesen EN, Gram J, Beck-Nielsen H. Cytokines and bone loss in a 5-year longitudinal study–hormone replacement therapy suppresses serum soluble interleukin-6 receptor and increases interleukin-1-receptor antagonist: the Danish Osteoporosis Prevention Study. *J Bone Miner Res.* (2000) 15:1545–54. doi: 10.1359/jbmr.2000.15.8.1545
54. Khosla S. Minireview: the OPG/RANKL/RANK system. *Endocrinology.* (2001) 142:5050–5. doi: 10.1210/endo.142.12.8536
55. Salminen H, Sääf M, Ringertz H, Strender LE. Bone mineral density measurement in the calcaneus with DXL: comparison with hip and spine measurements in a cross-sectional study of an elderly female population. *Osteoporos Int.* (2005) 16:541–51. doi: 10.1007/s00198-004-1719-x
56. Nguyen HT, von Schoultz B, Nguyen TV, Thang TX, Chau TT, Duc PT, et al. Sex hormone levels as determinants of bone mineral density and osteoporosis in Vietnamese women and men. *J Bone Miner Metab.* (2015) 33:658–65. doi: 10.1007/s00774-014-0629-z
57. Yoshimura N, Kinoshita H, Danjoh S, Takijiri T, Morioka S, Kasamatsu T, et al. Bone loss at the lumbar spine and the proximal femur in a rural Japanese community, 1990-2000: the Miyama study. *Osteoporos Int.* (2002) 13:803–8. doi: 10.1007/s001980200111
58. Nam HS, Kweon SS, Choi JS, Zmuda JM, Leung PC, Lui LY, et al. Racial/ethnic differences in bone mineral density among older women. *J Bone Miner Metab.* (2013) 31:190–8. doi: 10.1007/s00774-012-0402-0
59. Dolan E, Swinton PA, Sale C, Healy A, O'Reilly J. Influence of adipose tissue mass on bone mass in an overweight or obese population: systematic review and meta-analysis. *Nutr Rev.* (2017) 75:858–70. doi: 10.1093/nutrit/nux046
60. Palermo A, Tuccinardi D, Defeudis G, Watanabe M, D'Onofrio L, Lauria Pantano A, et al. BMI and BMD: the potential interplay between obesity and bone fragility. *Int J Environ Res Public Health.* (2016) 13:544. doi: 10.3390/ijerph13060544
61. Kim MS, Kim WJ, Khera AV, Kim JY, Yon DK, Lee SW, et al. Association between adiposity and cardiovascular outcomes: an umbrella review and meta-analysis of observational and Mendelian randomization studies. *Eur Heart J.* (2021) 42:3388–403. doi: 10.1093/eurheartj/ehab454
62. Scheen AJ, Van Gaal FL. Combating the dual burden: therapeutic targeting of common pathways in obesity and type 2 diabetes. *Lancet Diabetes Endocrinol.* (2014) 2:911–22. doi: 10.1016/S2213-8587(14)70004-X
63. Cortés VA, Barrera F, Nervi F. Pathophysiological connections between gallstone disease, insulin resistance, and obesity. *Obes Rev.* (2020) 21:e12983. doi: 10.1111/obr.12983
64. Fabbrini E, Sullivan S, Klein S. Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, clinical implications. *Hepatology.* (2010) 51:679–89. doi: 10.1002/hep.23280
65. Iyengar NM, Gucalp A, Dannenberg AJ, Hudis CA. Obesity and cancer mechanisms: tumor microenvironment and inflammation. *J Clin Oncol.* (2016) 34:4270–6. doi: 10.1200/JCO.2016.67.4283
66. Majumder K, Gupta A, Arora N, Singh PP, Singh S. Premorbid obesity and mortality in patients with pancreatic cancer: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol.* (2016) 14:355–68.e; quiz: e32. doi: 10.1016/j.cgh.2015.09.036


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 05 November 2021
doi: 10.3389/fnut.2021.747088

# The Chinese Adolescent Cohort Study: Design, Implementation, and Major Findings

Ruonan Duan[1], Xiaoyu Wang[1], Shufang Shan[1], Li Zhao[2], Jingyuan Xiong[2], Lars Libuda[3] and Guo Cheng[1]*

[1] Laboratory of Molecular Translational Medicine, Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University), Department of Pediatrics, Center for Translational Medicine, Ministry of Education, West China Second University Hospital, Sichuan University, Chengdu, China, [2] Healthy Food Evaluation Research Center, Sichuan University, Chengdu, China, [3] Faculty of Natural Sciences, Institute of Nutrition, Consumption and Health, Nutrition Sciences, Paderborn University, Paderborn, Germany

The importance of diet quality on children's growth is being increasingly recognized. The Chinese Adolescent Cohort (CAC) is a longitudinal cohort study to comprehensively investigate the health impacts of nutritional factors on child growth. From 2013 to 2018, 6,967 children aged 6–8 years have been recruited from 23 primary schools in Sichuan, Guizhou, and Chongqing, which have been planned to be followed up annually until their age of 15 years. Regular assessments included the measurement of height, weight, waist circumference, and skinfold thicknesses; pubertal development was examined by trained investigators according to Tanner stages; dietary intake was obtained by three 24-h recalls and food frequency questionnaire; validated questionnaires were used to estimate socio-demographic characteristics, physical activity, and sedentary behaviors. Findings from the CAC baseline and the first follow-up data suggested that higher protein intake among girls and unhealthy eating habits among children might increase the risk for childhood obesity. Also, higher intakes of grain and meat and lower overall diet quality and intakes of dietary fiber and tuber might be associated with advanced pubertal development. Those results indicated that the CAC study could contribute to the development of strategies for optimizing Chinese children's health.

**Keywords: cohort study, nutrition, growth, obesity, puberty, children, China**

OPEN ACCESS

***Edited by:***
*Massimo Lucarini,*
*Council for Agricultural Research and Economics, Italy*

***Reviewed by:***
*Mario Cicero Falcao,*
*University of São Paulo, Brazil*
*Jenny Vilchis-Gil,*
*Hospital Infantil de México Federico Gómez, Mexico*

***Correspondence:***
*Guo Cheng*
*ehw_cheng@126.com*

***Specialty section:***
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

***Received:*** 25 July 2021
***Accepted:*** 15 October 2021
***Published:*** 05 November 2021

***Citation:***
*Duan R, Wang X, Shan S, Zhao L, Xiong J, Libuda L and Cheng G (2021) The Chinese Adolescent Cohort Study: Design, Implementation, and Major Findings. Front. Nutr. 8:747088. doi: 10.3389/fnut.2021.747088*

## INTRODUCTION

The impacts of diet quality on children's growth have been investigated for decades. Evidence from some recent studies has suggested that dietary intake during childhood may have long-term impacts on health, such as obesity in adulthood (1). Such findings have consistently emphasized the importance of diet quality in early life. In recent years, the prevalence of childhood obesity has risen alarmingly in China (2, 3). Meanwhile, a secular trend of earlier puberty onset has been observed in both Chinese boys (4) and girls (5). Both an earlier onset of puberty and childhood obesity have been considered to be risk factors for chronic diseases in later life (6–8). Although previous studies have revealed the relationship between diet quality and growth during childhood, little is known about the mechanisms underlying the impacts of diet quality on the development of obesity or early pubertal development among Chinese children. Since dietary habits in China are in a transition period (9), knowledge on dietary determinants that might pronounce the trends in obesity and puberty onset is of major public health relevance. Also, whether childhood obesity may contribute to early onset of puberty remains controversial (10). Thus, high-quality longitudinal cohort studies with prospectively collected information on childhood exposures are greatly needed in China.

To date, although an array of child cohorts have been established worldwide to address the effects of childhood exposures on subsequent health outcomes (11–18), only a few of them have focused on childhood nutrition and pubertal development. For example, the DOrtmund Nutritional and Anthropometric Longitudinally Designed (DONALD) Study (11), a German cohort, has collected detailed data on diet, anthropometry, and endocrinology of children from infancy to adolescence or early adulthood. However, the sample size was relatively small and non-representative. The 1993 Pelotas (Brazil) Birth Cohort has covered information on lifestyles (e.g., dietary patterns, physical activity) and child's growth (e.g., anthropometry, age at menarche) during the follow-up visits at 4–11 years (18, 19). However, the indicator of puberty timing was limited to menarche, which only represents a relatively late stage of pubertal development in girls (18). The China-Anhui Birth Cohort study (17) is another example of birth cohorts, which has focused on environmental exposures for children's development (including puberty timing), with limited information collected on childhood nutrition. To our knowledge, few large-scale child cohort studies addressing nutrition on children's growth existed in Southwest China. To fill this gap, the Chinese Adolescent Cohort (CAC) study aims to provide comprehensive and reliable data that can be used to investigate nutritional and lifestyle determinants of children's growth, with emphasis on overweight/obesity and pubertal development.

## MATERIALS AND METHODS

### Design and Study Population

The CAC study has been designed in Sichuan Province, Guizhou Province, and Chongqing Municipality, which are located in Southwest China, a large area of 2.5 million square kilometers with a population of 199 million (including about 2.5 million children aged 6–8 years) (**Figure 1**). The CAC is a longitudinal cohort collecting detailed data on diet, lifestyle behaviors, and children's growth across the entire pubertal period. Subjects were recruited at their age of 6–8, and would be followed up until their age of 15. Regular assessments include the measurement of anthropometry and puberty status every year, and the recording of diet, lifestyle behaviors, and parental characteristics every 2 years.

The recruitment of study participants started in January 2013. The starting study sample included children recruited from a baseline cross-sectional study (i.e., Diet Quality during Childhood), which aimed to collect information on the diet, anthropometry, and lifestyle of children in Southwest China (20). A sampling design stratified by urban and rural locations was used to obtain the study sample. Twenty-three study sites (10 primary schools in urban areas and 13 primary schools in rural areas) were included as of December 2018. The eligible participants were children aged 6–8 years, cooperative and voluntary at the time of recruitment, in the selected primary schools. Those who were multiple births, preterm (<37 weeks of gestation), or late term (>42 weeks of gestation), as well as those with birth weight lower than 2,500 g or major organ diseases (e.g., heart, liver, or kidney disease), were excluded. This study was approved by the Ethics Committee of Sichuan University. All methods were performed in accordance with the relevant guidelines and regulations. All parents gave written informed consent and all children gave assent (21).

### Follow-Up Assessments

Children in the CAC study have been planned to be followed up from 6–8 years until the age of 15. Thus, the entire pubertal period would be captured (i.e., before, during, and after the pubertal period). In the first examination, information on socio-demographic issues, dietary intake and eating behaviors, physical activity and sedentary behaviors, and anthropometry and pubertal development were obtained. Follow-up assessments of anthropometry and puberty status were conducted every year, while data of dietary intake and physical activity were collected every other year. The study protocol was summarized in an assessment timeline (**Figure 2**). By the time of conducting this study, seven follow-up assessments have been completed for participants recruited in 2013, with the eighth follow-up survey ongoing in 2021. Follow-ups of children recruited in 2014 to 2018 have been carried out for six times, five times, four times, three times and twice, respectively (**Figure 3**).

### Data Collection

The CAC study assessment procedures were based on several modules: dietary intake and eating behaviors, physical activity and sedentary behaviors, anthropometric measurements, pubertal development examination, and some additional information asked in face-to-face interviews by trained investigators using a parent questionnaire (**Table 1**).

### Dietary Intake and Eating Behaviors

Participants were interviewed with respect to their diet (i.e., one food frequency questionnaire covering the consumption over the past 12 months, and three 24-h dietary recalls). The dietary assessment in young children was challenging, as they may not be able to provide relevant information by themselves. In the CAC study, children were asked to recall their consumption of all foods and beverages only when they were aged 9 years or older. When they were younger than nine at the time of interviewing, children provided only the dietary intake information from school by themselves, with the information on food consumption by parents at home.

Three 24-h recalls were used to collect dietary intake data by trained investigators. Details on recipes and brands of all food items reported were inquired. Food models for foods, standard tableware including bowls, plates, and glasses were provided to improve the accuracy of the portion-size estimates. In addition, children have been given a photo book, which contained photos

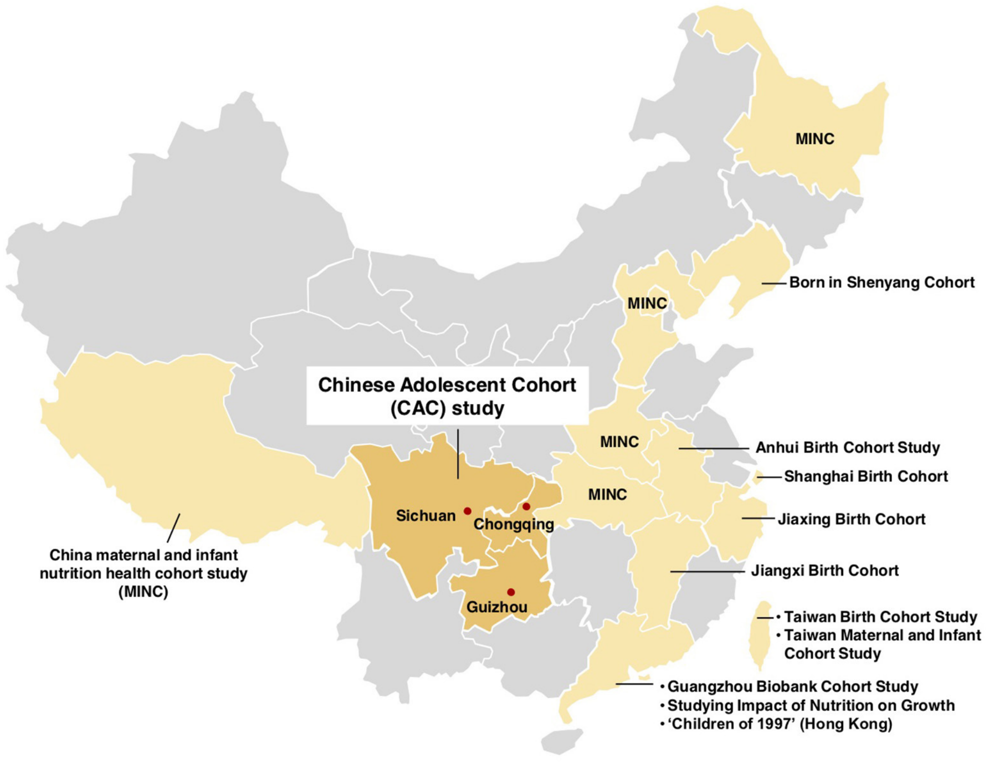


**FIGURE 1 |** Location of the Chinese Adolescent Cohort study and other child cohorts in China.

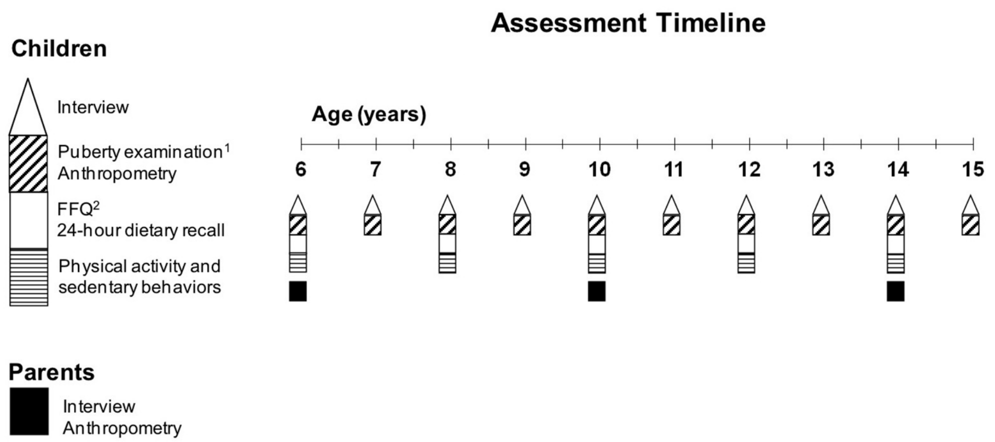


**FIGURE 2 |** Assessment timeline of the Chinese Adolescent Cohort study. [1]Puberty examination includes breast development and pubic hair development examinations in girls and testicular development and pubic hair development examinations in boys. In addition, children are asked whether menarche (girls), spermarche or voice break (boys) has already occurred and the respective month and year is recorded. [2] Food frequency questionnaires.

of snacks and beverages and pictures of the commonly used commercial packaging (e.g., one carton) to improve the diet recall. The dietary intake data from 24-h recalls were converted into energy and nutrient intake data, using the continuously updated in-house nutrient database that reflected the China Food Composition (22) in NCCW software (version 11.0, 2014).

In addition, children's consumption of foods and food groups over the past year was collected with a validated food frequency questionnaire (23), which included 53 foods or food groups that were most representative of local consumption and ultimately grouped into 17 categories: whole grains, refined grains, tubers, vegetables, fruits, nuts, meat, fish and shrimp, animal viscera,

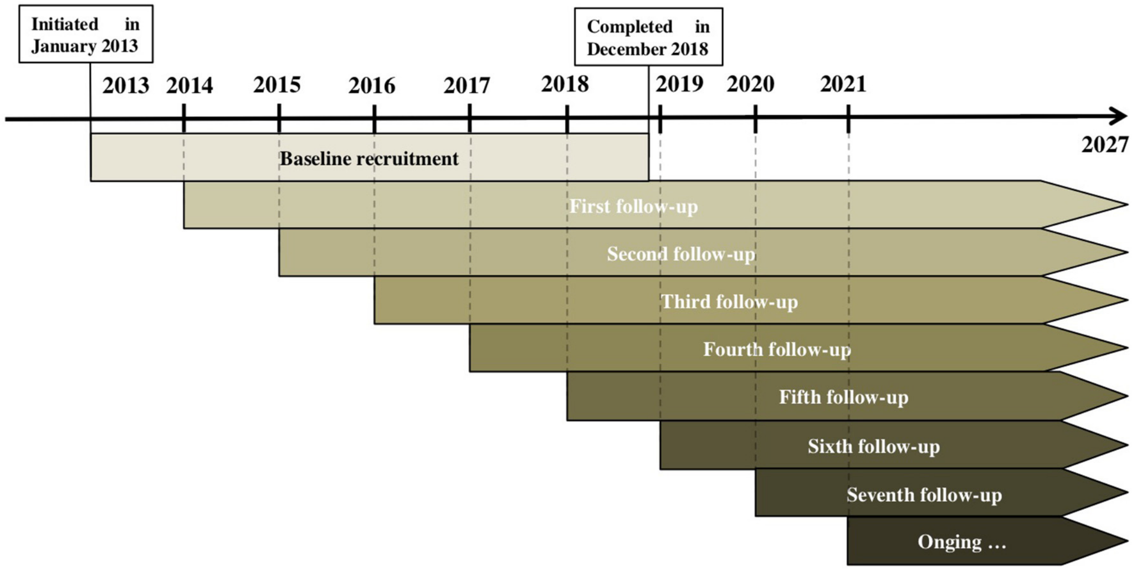


**FIGURE 3 |** Follow-up periods of the Chinese Adolescent Cohort study. Each participant will be followed up for 8–10 times (from age of 6–8 years to 15 years). Up to now, seven follow-up assessments have been completed for participants recruited in 2013, and the eighth follow-up survey is ongoing in 2021. Follow-ups of children recruited from 2014 to 2018 have been carried out six times, five times, four times, three times and twice, respectively.

**TABLE 1 |** Variables and instruments to be used in the Chinese Adolescent Cohort (CAC) study.

| Type of Exposure Data | Instruments | Variables for Exposure/Outcome |
|---|---|---|
| Dietary data | Three 24-h recalls and food frequency questionnaire | Energy intake, macro- and micro-nutrients, food and food groups |
| Physical activity | Physical activity questionnaire | Met[a]-hours, time spent on sedentary behaviors per day |
| Anthropometry | Ultrasonic Weight and Height Instrument, non-elastic tape, Holtain caliper | Height, weight, waist circumference, skinfold thickness |
| Pubertal development | Prader orchidometer and questionnaire | Pubertal maturation for breast (girls) and pubic hair (girls and boys) stages, testicular volume, menarche (girls), spermarche or voice break (boys) |
| Children's birth characteristics | Questionnaire | Gestational age, birth weight, birth length, breastfeeding and complementary feeding practices |
| Socio-demographics | Questionnaire | Residence, household income, family size, passive smoking, parental age, parental occupation and education levels, parental height and weight, mother's menarche age |

[a] *MET, metabolic equivalent.*

eggs, dairy and dairy products, soybeans and its products, fried foods, sugary snacks, sugar-sweetened beverages, fruit juice, and dietary supplements. Participants were asked to report the frequency and amounts of their consumption for each food item.

For eating behaviors relevant to diet quality of children, participants were asked about whether or not they often ate breakfast (at least 5 days per week) or had dinner with parents/grandparents regularly (at least 5 days per week), with the corresponding time and places of breakfast, lunch, dinner and snacks recorded.

## Physical Activity and Sedentary Behaviors

A validated physical activity questionnaire (PAQ) (24) was used to collect information on the frequency, duration, and type of physical activities inside and outside school settings and leisure-time activities. The PAQ included a checklist of 38 items, e.g., walking, running, climbing stairs, ball games, dancing, and so on. In addition, participants were asked to report the usual time spent on sedentary behaviors (i.e., watching television, using computers, using smart phones, and doing homework) on weekdays and weekends.

## Anthropometric Measurements

Children in the CAC study have been undergoing the anthropometric measurement every year, which were performed by investigators who have been trained according to standard procedures, with the subjects dressed lightly and barefoot. An Ultrasonic Weight and Height Instrument (DHM-30, Dingheng Ltd, Zhengzhou, China) was used to assess the standing height to the nearest 0.1 cm and weight to the nearest 0.1 kg. Waist circumference at the umbilicus was measured to the nearest 0.1 cm by a calibrated measuring tape. In addition, skinfold thickness of biceps, triceps, anterior superior iliac spines, and subscapular angle sites were measured on the right side of the body to the nearest 0.1 mm using a Holtain caliper (Holtain Ltd, Crymych, UK). All anthropometric measurements were performed twice with the average calculated for each participant.

## Pubertal Development Examinations

During the annual physical examination for children, pubertal maturation for breast (girls) and pubic hair (girls and boys) stages was assessed by trained investigators according to the standardized criteria of Tanner stages (25). Testicular volume was assessed by comparative palpation with the Prader orchidometer. If the testicular volumes of the two testes were not equal, the volume of the larger one was recorded. Testicular volume <1 mL was recorded as 1 mL. In addition, children were asked at each examination about whether or not menarche (girls), spermarche, or voice break (boys) had already occurred, with the corresponding month and year information recorded.

## Additional Information

Information on children's birth characteristics and feeding practices (i.e., gestational age, birth weight, birth length, exclusive breastfeeding duration, and timing of complementary feeding) and socio-demographic data (i.e., place of residence, household income, family size, parental age, and parental occupation and education levels) were provided by parents. In addition, parents' height and weight were self-reported, and information on mother's menarche age was also collected.

## Quality Control

To ensure the accurate and standard measurement throughout the CAC study, several quality assurance measures have been undertaken. Prior to study implementation, a 3-day training was conducted in study offices at each site. A training manual describing all of the standardized procedures in details was developed, and all instruments for data collection (i.e., Ultrasonic Weight and Height Instrument, non-elastic tape, and Holtain caliper) were calibrated. During the study, data checks were conducted for completeness and accuracy immediately after questionnaires were completed. For data entry and management, double entries were carried out and corrections were made as needed. In addition, a pencil box or a notebook was delivered to each participant to encourage him/her to complete the interview.

# RESULTS

Between January 2013 and December 2018, 7,119 children from 23 primary schools in Southwest China have been invited for the baseline data collection. Among them, 152 children refused to enroll due to parental disapproval. They did not differ in their basic characteristics (e.g., age and gender) from those who have participated. However, non-respondents were more likely to have lower parental education level and to live in rural areas (data not shown). Thanks to the efforts of the participating primary schools in assisting with data collection, 6,967 children completed the baseline questionnaire (2,320, 912, 926, 921, 986, and 902 in 2013–2018, respectively). Their baseline characteristics were shown in **Table 2**. Of the 6,967 participants, 52.6% were girls, with an average age of 7.8 years and moderate socioeconomic status. Those who agreed to participate were representative of the general children population in terms of age and parental education level and socioeconomic status (26).

**Figure 3** showed the numbers of participants recruited and followed up until April 2021. A total of 1,262 children (18.1%) were lost to first follow-up due to several reasons (e.g., refusal to complete time-consuming questionnaires, moving away from the study area, parental disapproval). Children who were lost to follow-up were generally younger, with lower parental education and family income levels, spending more time on sedentary behaviors than their counterparts who retained in the study (**Table 3**).

As the follow-up assessments of CAC study are still ongoing, it is still at an early stage with data analysis in progress. To date, more than 10 peer-reviewed publications based on the CAC baseline and the first follow-up data have been published, covering different aspects of exposures and health outcomes. A summary of those findings was presented in **Table 4**, with the key findings representing multidisciplinary nature of CAC data collection presented below.

**TABLE 2** | Baseline characteristics of 6,967 children in the Chinese Adolescent Cohort (CAC) study.

| Characteristics | Value |
|---|---|
| Age, y | 7.8 (0.6) |
| Female, % | 52.6 |
| Urban residence, % | 52.1 |
| Birth weight, kg | 3.3 (0.5) |
| Breastfeeding for 6 months, % | 69.2 |
| Complementary feeding from 6th month, % | 31.7 |
| Family size, n | 4 (1) |
| High family income level[a], % | 23.9 |
| High maternal education level[b], % | 21.2 |
| High paternal education level[b], % | 26.3 |
| Weight, kg | 24.4 (4.5) |
| Height, cm | 121.3 (6.0) |
| BMI-SDS[c] | 0.2 (1.0) |
| Percent body fat, % | 16.2 (12.9, 23.1) |
| Overweight (including obesity), % | 19.9 |
| Energy intake, MJ/d | 7.5 (6.4, 8.9) |
| Protein, % of energy | 16.5 (14.3, 17.8) |
| Fat, % of energy | 26.3 (21.3, 31.9) |
| Carbohydrate, % of energy | 57.2 (52.8, 65.2) |
| MVPA energy expenditure[d], MJ/day | 0.5 (0.2) |
| Sedentary behaviors, h/d | 2.1 (0.6) |

[a]*Average family income more than 35,000 Yuan every year.*
[b]*At least 12 years of school education.*
[c]*BMI-SDS calculated according to the Chinese reference curves.*
[d]*MVPA energy expenditure, energy expended on moderate-to-vigorous physical activities (MJ/day).*

**TABLE 3 |** Comparison of basic characteristics between children who remained in the cohort for at least once follow-up and those lost to follow-ups.

| Characteristics | Children who were followed up | Children lost to follow-up | *P* value |
|---|---|---|---|
| N, % | 5,705 (81.9) | 1,262 (18.1) | |
| Age, y | 7.9 (0.5) | 7.5 (0.7) | 0.04 |
| Female, % | 52.9 | 51.6 | 0.7 |
| Urban, % | 52.2 | 51.8 | 0.6 |
| Birth weight, kg | 3.3 (0.5) | 3.2 (0.6) | 0.8 |
| Breastfeeding for 6 months, % | 69.2 | 65.1 | 0.6 |
| Complementary feeding from 6th month, % | 32.1 | 29.9 | 0.08 |
| Family size, n | 4 (1) | 5 (1) | 0.07 |
| High family income level[a], % | 26.3 | 12.9 | 0.04 |
| High maternal education level[b], % | 22.2 | 16.5 | 0.03 |
| High paternal education level[b], % | 28.1 | 18.3 | 0.03 |
| Weight, kg | 24.2 (4.3) | 24.5 (4.6) | 0.1 |
| Height, cm | 121.6 (6.2) | 120.7 (5.9) | 0.2 |
| BMI-SDS[c] | 0.5 (0.8) | 0.5 (1.1) | 0.9 |
| Percent body fat, % | 16.3 (13.1, 23.2) | 15.8 (12.7, 22.9) | 0.2 |
| Energy intake, MJ/d | 7.4 (6.3, 8.8) | 7.6 (6.5, 9.2) | 0.3 |
| Protein, % of energy | 16.9 (14.6, 18.1) | 15.8 (14.1, 17.3) | 0.06 |
| Fat, % of energy | 26.3 (21.1, 32.2) | 26.4 (21.4, 31.8) | 0.08 |
| Carbohydrate, % of energy | 56.8 (53.3, 66.8) | 57.8 (52.8, 65.2) | 0.07 |
| MVPA energy expenditure[d], MJ/day | 0.7 (0.3) | 0.4 (0.5) | 0.051 |
| Sedentary behaviors, h/d | 1.7 (0.5) | 2.3 (0.7) | 0.04 |

[a] *Average family income more than 35,000 Yuan every year.*
[b] *At least 12 years of school education.*
[c] *BMI-SDS calculated according to the Chinese reference curves.*
[d] *MVPA energy expenditure, energy expended on moderate-to-vigorous physical activities (MJ/day).*

## Diet Quality

According to the Chinese Dietary Guidelines and Dietary Reference Intakes, the Chinese Children Dietary Index (CCDI) was developed to assess the overall diet quality among the Chinese children. It showed the need to improve the diet quality among children in Southwest China, especially in boys, because their consumption of grains and meat were excessive with inadequate consumption of soybeans, fish, shrimp, eggs, and dietary fiber. In addition, socioeconomic factors were identified to be correlated with the CCDI score, which revealed an impact of paternal, rather than maternal, education level on children's diet quality. A positive association between the CCDI score and family size was further observed (21).

## Obesity

We found that the prevalence of overweight/obesity among children in Southwest China was relatively high: 10.7 and 6.5% of CAC children were observed as overweight and obesity, respectively, compared to the overall situation that the prevalence of childhood overweight and obesity has risen from 4.5 and 2.3% in 2002 to just 9.6 and 6.4% in 2012, respectively (27).

As a modifiable factor, dietary intake has played an important role in the development of childhood overweight/obesity. In the CAC study, girls, but not boys, with higher dietary protein intake may have higher body fat mass, which, further modified by paternal occupation, was observed only in girls with part-time working fathers (28). Also, children with healthy dietary patterns (i.e., beans, potatoes, whole grains, dairy) were at lower risk for overweight/obesity (OR = 0.59, 95% CI [0.39, 0.90]) (30). Unhealthy eating habits, characterized by sugar-sweetened beverages, not having breakfast, and not having dinner with parents/grandparents, have increased the risk for childhood obesity (31).

## Pubertal Development

Compared with children in western countries and in eastern and central cities of China, genital/breast development among children in Southwest China was later (33). Low dietary fiber (especially fruit fiber) (34) and tuber intake (35), and high grain (35) and meat intake (36) may be associated with advanced pubertal development, which were more pronounced in girls. Additionally, based on data from 3,983 CAC study participants (1,752 girls and 2,231 boys) with at least 2 follow-up assessments by the end of 2019, the prospective relevance of diet quality on puberty timing were examined. We found that both girls and boys with higher diet quality were more likely to enter their puberty later than their counterparts with lower diet quality, which were independent of pre-pubertal body fat (37).

## Sedentary Behaviors

In the CAC study, children with high paternal or maternal education level tended to spend less time on watching TV, while

**TABLE 4 |** Key published findings from the baseline and the first follow-up data of Chinese Adolescent Cohort (CAC) study.

| Key variables | Main findings |
|---|---|
| **Diet quality** | |
| Development of dietary index | The Chinese Children Dietary Index was successfully developed and used to assess overall diet quality among Chinese children (21). |
| Diet quality and socio-demographic factors | Diet quality of children was associated with age negatively and family size positively (21). |
| **Obesity** | |
| Prevalence of overweight/obesity | Prevalence of overweight and obesity was relatively high among children in Southwest China (27). |
| Obesity and dietary energy density | Dietary energy density was not associated with body composition of children in Southwest China (20). |
| Obesity and protein intake | Girls, but not boys, with higher dietary protein intake might have higher body fat mass (28). |
| Obesity and dietary calcium/dairy | Boys with higher dietary calcium intake had the lowest prevalence of overweight (29). |
| Obesity and dietary pattern | Dietary pattern consisting of beans, potatoes, whole grains and dairy was associated with a lower risk of overweight and obesity (30). |
| Obesity and eating behaviors | Poor eating behaviors, characterized by sugar-sweetened beverages, not eating breakfast regularly and not having dinner with parents, are associated with child obesity (31). |
| Obesity and socioeconomic status | Urban girls living in relatively affluent circumstances were at greatest risk for being overweight (32). |
| **Pubertal development** | |
| Status of pubertal development | Genital/breast development among children in Southwest China starts later, while spermarche/menarche is earlier (33). |
| Pubertal development and dietary fiber | Dietary fiber intake, especially fruit fiber, is lower in children with earlier pubertal development (34). |
| Pubertal development and grain/tubers | Grain intake was positively associated with pubertal development among boys. Both grain and tubers intake were negatively associated with pubertal development among girls (35). |
| Pubertal development and meat and dairy | Level of pubertal development was positively associated with meat intake among boys, and negatively associated with dairy intake among girls (36). |
| Pubertal development and overall diet quality | Girls and boys with higher diet quality were more likely to enter their puberty later than their counterparts with lower diet quality, which were independent of pre-pubertal body fat (37). |
| Pubertal development and eating behaviors | Level of pubertal development was positively associated the frequency of snacks and sugar-sweetened beverages consumption among both boys and girls (38). |
| Pubertal development and socio-demographic factors | Pubertal development is related to parental educational level and occupation (33). |
| **Sedentary behaviors** | |
| Sedentary behaviors and socio-demographic factors | Time spent on watching TV was negatively associated with paternal or maternal education level. Time spent on using computers was positively associated with family income (39). |
| Sedentary behaviors and childhood body composition | Sedentary behaviors were positively and independently related to fat mass among Chinese children, and were more pronounced in girls (40). |

those living in high-income families seemed to spend more time on computer use (39). Furthermore, sedentary behaviors were positively and independently related to fat mass among children in Southwest China, which were also more pronounced in girls (40).

## DISCUSSION

This study described the design and implementation procedures of the CAC cohort, and summarized the baseline characteristics of the participants and major findings. The CAC was the first child cohort study in Southwest China that has used an annual multimodal assessment approach to prospectively address research questions on the associations between nutritional factors and later growth in children. The subjects would be followed up from the age of 6–8 till the age of 15, with follow-up assessments of anthropometry and puberty status every year and of diet and physical activity every 2 years.

To date, more than 10 child cohort studies have been launched in China. Each of those existing cohorts focused on different exposures or outcomes, starting their recruitment at different childhood periods and following up their participants in different time frames. For example, some child cohorts have examined the impacts of environmental factors [e.g., lead exposure and micronutrient deficiency (41, 42), phthalate exposure (43), early life feeding practice (44), and eating behaviors (45)] on children's cognitive and behavior outcomes [e.g., China Jintan Child Cohort Study (41, 42), Taiwan Maternal and Infant Cohort Study (43)], adiposity [e.g., Jiaxing Birth Cohort (44), Studying Impact of Nutrition on Growth (13), Childhood Obesity Study in China Mega-cities (45)], and physical, psychological, behavioral and sexual development [e.g., China-Anhui Birth Cohort study (17)]. Among them, only the China-Anhui Birth Cohort study has considered puberty timing as one of their outcomes. Although nutrition was also considered an exposure in that study, it did not specifically focus on nutritional factors as their main exposures. Moreover, it did not provide repeated and detailed exposure assessments across the entire pubertal period. Therefore, the capacity of estimating the longitudinal associations between dietary intake, body composition, and timing of puberty in that cohort was not comparable with that in the CAC study. Compared to the aforementioned child cohorts, a major advantage of the CAC study was the frequent follow-ups and detailed assessments of socio-demographic characteristics, nutrition, lifestyles, anthropometry, and pubertal development across the entire pubertal period (i.e., before, during, and after the pubertal period). In addition, information on characteristics of the participants' parents was also frequently collected. Furthermore, the CAC dataset was of high quality since all the interviews and examinations have been conducted by trained investigators. Due to the abundance of data and the long duration of follow-up, the temporal changes in children's body composition and processes of pubertal development could be tracked. Furthermore, the prospective associations between nutritional factors and the gradual changes in growth during childhood would be able to be determined. Despite the lack of comparable data from Southwest China, only minor differences were found in socio-demographic characteristics (characterized by household income and maternal/paternal education level) between participants in this study and those in the existing studies conducted in Sichuan (46) or Chongqing (47, 48), which might ensure generalizability of our findings. The results of baseline and the first follow-up data indicated that the CAC study would be a strong evidence-based platform for public health research.

Some limitations should be mentioned. First, attrition is a concern. Although we have made efforts to maintain the participants in the cohort, in common with most child cohort studies, the response rate has decreased over time. Until April 2021, 18.1% of the CAC participants were lost to the first follow-up, especially those with lower parental education and family income levels, which may result in the selection of a more "health conscious" study sample by the end of the follow-ups. It required more administrative efforts to partner with schools to enhance children's participation. Second, objective measures of physical activity such as pedometer are not used in the CAC study. However, the physical activity questionnaire adopted in the study has been validated and widely used in the literature (24, 49). Furthermore, serum or plasma biomarkers for vitamin C, calcium, iron, zinc, or cholesterol would have been preferable to verify the nutritional status of children. However, the inclusion of blood sampling is difficult, especially in most large observational studies of children. Finally, due to a focus on the period of middle childhood in the CAC study, we could not provide detailed data on early childhood. However, some important early-life exposures (i.e., children's birth weight, exclusive breast-feeding duration, and timing of complementary feeding) are collected in our study.

In future work, we hope to broaden both breadth and depth of measures in our study. We are also planning to include new domains in the CAC study, such as genetic factors, environmental exposures, and some biomarkers in at least one sub-cohort, although impossible for the entire cohort due to limited financial support. With these comprehensive and detailed data, we would be able to examine interactions between genetic and nutritional factors and their influences on children's growth and the underlying mechanisms. Furthermore, if funding comes through, we also hope to follow up the participants in the sub-cohort(s) until their adulthood.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Sichuan University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

GC conceived and supervised the project. RD performed the statistical analysis and wrote the manuscript. RD, XW, and SS took part in the data collection. LZ, JX, and LL coordinated the study centers and revised the manuscript. All authors contributed substantially to the work and approved the final version of the manuscript.

## FUNDING

This research was funded by a research grant from the New Century Excellent Talents in University Program (NCET-12-0377), Sichuan Outstanding Young Scholars Foundation (2014JQ0005), and the National Natural Science Foundation of China (81673158).

## ACKNOWLEDGMENTS

The participation of all children and their families is gratefully acknowledged. The authors also thank the staff of the Healthy Food Evaluation Research Center for carrying out the interview, anthropometric measurements and pubertal development examinations. We also thank Dr. Peng Jia and Dr. Liang Wang for their support and advice on the writing of this manuscript.

## REFERENCES


1. Bass R, Eneli I. Severe childhood obesity: an under-recognised and growing health problem. *Postgrad Med J.* (2015) 91:639–45. doi: 10.1136/postgradmedj-2014-133033
2. Wang H, Xue H, Du S, Zhang J, Wang Y, Zhang B. Time trends and factors in body mass index and obesity among children in China: 1997-2011. *Int J Obes (Lond).* (2017) 41:964–70. doi: 10.1038/ijo.2017.53
3. Jia P, Xue H, Zhang J, Wang Y. Time trend and demographic and geographic disparities in childhood obesity prevalence in china-evidence from twenty years of longitudinal data. *Int J Environ Res Public Health*. (2017) 14:369. doi: 10.3390/ijerph14040369
4. Ma HM, Chen SK, Chen RM, Zhu C, Xiong F, Li T, et al. Pubertal development timing in urban Chinese boys. *Int J Androl.* (2011) 34:e435–45. doi: 10.1111/j.1365-2605.2011.01173.x
5. Ma HM, Du ML, Luo XP, Chen SK, Liu L, Chen RM, et al. Onset of breast and pubic hair development and menses in urban chinese girls. *Pediatrics.* (2009) 124:e269–77. doi: 10.1542/peds.2008-2638
6. Brisbois TD, Farmer AP, McCargar LJ. Early markers of adult obesity: a review. *Obes Rev.* (2012) 13:347–67. doi: 10.1111/j.1467-789X.2011.00965.x
7. Umer A, Kelley GA, Cottrell LE, Giacobbi P. Jr., Innes KE, Lilly CL. Childhood obesity and adult cardiovascular disease risk factors: a systematic review with meta-analysis. *BMC public health.* (2017) 17:683. doi: 10.1186/s12889-017-4691-z
8. Bodicoat DH, Schoemaker MJ, Jones ME, McFadden E, Griffin J, Ashworth A, et al. Timing of pubertal stages and breast cancer risk: the Breakthrough Generations Study. *Breast Cancer Res.* (2014) 16:R18. doi: 10.1186/bcr3613
9. Popkin BM. Synthesis and implications: China's nutrition transition in the context of changes across other low- and middle-income countries. *Obes Rev.* (2014)15:60–7. doi: 10.1111/obr.12120
10. Li W, Liu Q, Deng X, Chen Y, Liu S, Story M. Association between obesity and puberty timing: a systematic review and meta-analysis. *Int J Environ Res Public Health.* (2017) 14:1266. doi: 10.3390/ijerph14101266
11. Kroke A, Manz F, Kersting M, Remer T, Sichert-Hellert W, Alexy U, et al. The DONALD study. History, current status and future perspectives. *Eur J Nutr.* (2004) 43:45–54. doi: 10.1007/s00394-004-0445-7
12. Potischman N, Fang L, Hao L, Bailey RR, Berrigan D, Berry RJ, et al. The Chinese children and families cohort study: the nutrition, physical activity, and ultraviolet radiation data collection. *Nutr Today.* (2018) 53:104–14. doi: 10.1097/NT.0000000000000275
13. Lee A, Cheung CK, Lo K, Keung VM, Mui LW, Tam WWS. Studying Impact of Nutrition on Growth (SING): a prospective cohort for comparing the health outcomes of young children with the dietary quality score. *BMJ Open.* (2017) 7:e018380. doi: 10.1136/bmjopen-2017-018380
14. Ahrens W, Bammann K, de Henauw S, Halford J, Palou A, Pigeot I, et al. Understanding and preventing childhood obesity and related disorders–IDEFICS: a European multilevel epidemiological approach. *Nutr Metab Cardiovasc Dis.* (2006) 16:302–8. doi: 10.1016/j.numecd.2006.01.011
15. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van Ijzendoorn MH, et al. The Generation R Study: design and cohort update 2017. *Eur J Epidemiol.* (2016) 31:1243–64. doi: 10.1007/s10654-016-0224-9
16. Moreno LA, De Henauw S, González-Gross M, Kersting M, Molnár D, Gottrand F, et al. Design and implementation of the Healthy Lifestyle in Europe by Nutrition in Adolescence Cross-Sectional Study. *Int J Obes (Lond).* (2008) 32:S4–S11. doi: 10.1038/ijo.2008.177
17. Tao FB, Hao JH, Huang K, Su PY, Cheng DJ, Xing XY, et al. Cohort profile: the china-anhui birth cohort study. *Int J Epidemiol.* (2013) 42:709–21. doi: 10.1093/ije/dys085
18. Victora CG, Hallal PC, Araújo CL, Menezes AM, Wells JC, Barros FC. Cohort profile: the 1993 Pelotas (Brazil) birth cohort study. *Int J Epidemiol.* (2008) 37:704–9. doi: 10.1093/ije/dym177
19. Gonçalves H, Assunção MC, Wehrmeister FC, Oliveira IO, Barros FC, Victora CG, et al. Cohort profile update: The 1993 Pelotas (Brazil) birth cohort follow-up visits in adolescence. *Int J Epidemiol.* (2014) 43:1082–8. doi: 10.1093/ije/dyu077
20. Zhou X, Xue H, Duan R, Liu Y, Zhang L, Harvey L, et al. The cross-sectional association of energy intake and dietary energy density with body composition of children in Southwest China. *Nutrients.* (2015) 7:5396–412. doi: 10.3390/nu7075228
21. Cheng G, Duan RN, Kranz S, Liu Y, Xue HM. Development of a dietary index to assess overall diet quality for chinese children and adolescents: The Chinese children dietary index. *J Acad Nutr Diet.* (2014) 116:608–17. doi: 10.1016/j.jand.2015.11.010
22. Yang YX, He M, Pan XC. *China Food Composition*. 2nd ed. Peking University Medical Press. (2004).
23. Duan R, Liu Y, Xue H, Yang M, Cheng G. [Cross-sectional association between overall diet quality and overweight/obesity among children and adolescents in Chengdu]. *Zhonghua Liu Xing Bing Xue Za Zhi.* (2014) 35:994–8.
24. Huang YJ, Wong SH, Salmon J. Reliability and validity of the modified Chinese version of the Children's Leisure Activities Study Survey (CLASS) questionnaire in assessing physical activity among Hong Kong children. *Pediatr Exerc Sci.* (2009) 21:339–53. doi: 10.1123/pes.21.3.339
25. Tanner JM. Normal growth and techniques of growth assessment. *Clin Endocrinol Metab.* (1986) 15:411–51. doi: 10.1016/S0300-595X(86)80005-6
26. National Bureau of Statistics of China. (2021). Available online at: http://data.stats.gov.cn/easyquery.htm?cn=C01. *National Data*
27. Liu Y, Duan R, Xue H, Yang M, Zhou X, Cheng G. [Overweight and obesity status among children and adolescents aged 7-15 years old in Chengdu]. *Chinese Journal of School Health.* (2014) 35:1191–3.
28. Yang MZ, Xue HM, Pan J, Libuda L, Muckelbauer R, Yang M, et al. High protein intake along with paternal part-time employment is associated with higher body fat mass among girls from South China. *Eur J Nutr.* (2018) 57:1845–54. doi: 10.1007/s00394-017-1468-1
29. Chen Y, Liu Y, Xue H, Bao Y, Luo J, Tian G, et al. [Association between dietary calcium/dairy intakes and overweight/obesity]. *Wei Sheng Yan Jiu.* (2016) 45:402–8.
30. Li DT, Chen MX, Xue HM Li Y, Tian G, Gong YH, et al. [Association between dietary patterns and overweight/obesity among school-aged children in Chengdu]. *Modern Preventive Medicine.* (2018) 45:4070–3.
31. Duan R, Xue H, Liu Y, Yang M, Cheng G. [Cross-sectional association between eating behaviors and overweight/obesity among children and adolescents in Chengdu city]. *Wei Sheng Yan Jiu.* (2015) 44:576–80.
32. Zhai J, Xue H, Luo J, Zhang L, Cheng G. Associations between socioeconomic status and overweight among urban children aged 7-12 years in Chengdu, southwest China. *Asia Pac J Clin Nutr.* (2018) 27:617–23. doi: 10.6133/apjcn.022017.12
33. Yang M, Wang D, Hu JX, Liu Y, Duan R, Xue H, et al. [Pubertal development among children in Chengdu and associated family factors]. *Chin J Sch Health.* (2015) 36:226–30, 35.
34. Tian G, Liu Y, Xue HM, Luo J, Chen YR, Bao YX, et al. [Dietary fiber and pubertal development among children and adolescents–a cross-sectional study in Chengdu, Sichuan]. *Sichuan da xue xue bao Yi xue ban = Journal of Sichuan University Medical science edition.* (2016) 47:244–7.
35. Liu Y, Duan R, Xue H, Yang M, Cheng G. [Cross-sectional association between consumption of grain/tubers and pubertal development among children and adolescents in Chengdu City]. *Wei Sheng Yan Jiu.* (2015) 44:49–52.
36. Luo J, Yang MZ, Duan RN, Tian G, Bao YX, Chen YR, et al. [Consumptions of Meat and dairy products, zinc intake and pubertal development in adolescents in Chengdu]. *Sichuan da xue xue bao Yi xue ban = Journal of Sichuan University Medical science edition.* (2015) 46:726–31.
37. Duan R, Qiao T, Chen Y, Chen M, Xue H, Zhou X, et al. The overall diet quality in childhood is prospectively associated with the timing of puberty. *Eur JNutr.* (2020). doi: 10.1007/s00394-020-02425-8
38. Liu Y, Yang M, Duan R, Xue H, Cheng G. [Dietary behaviors among children and adolescents in Chengdu and its association with pubertal development]. *Chin J Sch Health.* (2015) 36:1126−9.
39. Xue H, Yang M, Duan R, Liu Y, Cheng G, Zhang X. [Relationship between family socio-economic factors and sedentary behaviors among children and adolescents in Chengdu city]. *Chin J Sch Health.* (2015) 36:47-50, 3.
40. Xue H, Tian G, Duan R, Quan L, Zhao L, Yang M, et al. Sedentary behavior is independently related to fat mass among children and adolescents in South China. *Nutrients*. (2016) 8. doi: 10.3390/nu8110667
41. Liu J, McCauley LA, Zhao Y, Zhang H, Pinto-Martin J. Cohort Profile: The China Jintan child cohort study. *Int J Epidemiol.* (2010) 39:668–74. doi: 10.1093/ije/dyp205

42. Liu J, Cao S, Chen Z, Raine A, Hanlon A, Ai Y, et al. Cohort Profile Update: The China Jintan Child Cohort Study. *Int J Epidemiol*. (2015) 44:1548. doi: 10.1093/ije/dyv119
43. Wu CF, Chen HM, Sun CW, Chen ML, Hsieh CJ, Wang SL, et al. Cohort Profile: The Taiwan Maternal and Infant Cohort Study (TMICS) of phthalate exposure and health risk assessment. *Int J Epidemiol*. (2018) 47:1047. doi: 10.1093/ije/dyy067
44. Zheng JS, Liu H, Zhao YM, Li J, Chen Y, Zhu S, et al. Complementary feeding and childhood adiposity in preschool-aged children in a large Chinese cohort. *J Pediatr*. (2015) 166:326–31.e2. doi: 10.1016/j.jpeds.2014.11.010
45. Jia P, Li M, Xue H, Lu L, Xu F, Wang Y. School environment and policies, child eating behavior and overweight/obesity in urban China: the childhood obesity study in China megacities. *Int J Obes (Lond)*. (2017) 41:813–9. doi: 10.1038/ijo.2017.2
46. Zhang J, Mou Y, Liao J, Xiong H, Duan Z, Huang Y, et al. Uptake of maternal care and childhood immunization among ethnic minority and Han populations in Sichuan province: a study based on the 2003, 2008 and 2013 health service surveys. *BMC Pregnancy Childbirth*. (2019) 19:250. doi: 10.1186/s12884-019-2371-y
47. Li W, Liu Q, Chen Y, Yang B, Huang X, Li Y, et al. Effects of indoor environment and lifestyle on respiratory health of children in Chongqing, China. *J Thorac Dis*. (2020) 12:6327–41. doi: 10.21037/jtd.2020.03.102
48. Yang B, Ostbye T, Huang X, Li Y, Fang B, Wang H, et al. Maternal age at menarche and pubertal timing in boys and girls: a cohort study from Chongqing, China. *J Adolesc Health*. (2021) 68:508–16. doi: 10.1016/j.jadohealth.2020.06.036
49. Matthews CE, Shu XO, Yang G, Jin F, Ainsworth BE, Liu D, et al. Reproducibility and validity of the Shanghai Women's Health Study physical activity questionnaire. *Am J Epidemiol*. (2003) 158:1114–22. doi: 10.1093/aje/kwg255

**Author Disclaimer:** Frontiers Media SA remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

BRIEF RESEARCH REPORT
published: 19 November 2021
doi: 10.3389/fnut.2021.778178

# The Sharp Rise in the Use of Low- and No-Calorie Sweeteners in Non-Alcoholic Beverages in Slovenia: An Update Based on 2020 Data

*Edvina Hafner*[1,2] *and Igor Pravst*[1,2,3*]

[1] Nutrition Institute, Nutrition and Public Health Research Group, Ljubljana, Slovenia, [2] Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia, [3] VIST–Higher School of Applied Sciences, Ljubljana, Slovenia

OPEN ACCESS

**Edited by:**
Alessandra Durazzo,
Council for Agricultural Research and Economics, Italy

**Reviewed by:**
Malgorzata Grembecka,
Medical University of Gdańsk, Poland
Hasnah Binti Haron,
National University of Malaysia, Malaysia
Ewa Czarniecka-Skubina,
Warsaw University of Life Sciences, Poland

***Correspondence:**
Igor Pravst
igor.pravst@nutris.org

**Specialty section:**
This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition

**Received:** 16 September 2021
**Accepted:** 25 October 2021
**Published:** 19 November 2021

**Citation:**
Hafner E and Pravst I (2021) The Sharp Rise in the Use of Low- and No-Calorie Sweeteners in Non-Alcoholic Beverages in Slovenia: An Update Based on 2020 Data. Front. Nutr. 8:778178. doi: 10.3389/fnut.2021.778178

Reducing added sugars in non-alcoholic beverages is an important public health goal, which can result in increased use of low- and no-calorie sweeteners (LNCS). The aim of this study was to investigate recent changes in the use of LNCS in non-alcoholic beverages in the Slovenian food supply. The national branded foods dataset was updated with beverages available in 2020, and compared with previous datasets. The data were extracted from food labels. In 2020, $N = 1,650$ unique beverages were found in shops from five different retailers, covering the majority of the national market. The use of LNCS increased from 13.2% in 2017 and 15.5% in 2019 to 20.2% in 2020, with a major growth in soft drinks (16.8, 19.6, and 26.7%, respectively). We observed a significant growth of beverages containing both LNCS and added sugar. Results were also consistent with sales data, which showed that increased offer of beverages with LNCS also resulted in similarly increased sales of such beverages. The average energy and total sugar content in non-alcoholic beverages decreased, which reflects both the higher percentage of beverages with LNCS, and also the reduction of the sugar content in beverages with only added sugar. Analyses of product-specific reformulation practices highlighted reduced sugar content in 16.8% of products, and in 3.6% with the use of LNCS. The most commonly used LNCS are acesulfame K, sucralose, and aspartame. Typically, combinations are used, however steviol glycosides, sucralose and saccharin are also used alone, in most cases combined with added sugar. The results indicated rapid changes in the use of LNCS in non-alcoholic beverages in the Slovenian food supply, making further monitoring of this area highly relevant.

**Keywords: low-and no-calorie sweeteners, beverages, reformulation, added sugar, food composition, database**

## INTRODUCTION

Low- and no-calorie sweeteners (LNCS) are a group of food additives that provide a sweet taste with no or fewer calories per gram of food, compared to sugar (1). They include both intensive sweeteners and polyols, which can partially or entirely replace sugar in various foods and drinks (2). Since excessive sugar intake is a major public health issue in the modern diet, its reduction is a key step in fighting against the worldwide obesity epidemic (3). The World Health Organization (WHO) recommends keeping free sugar consumption below 10% of the daily energy intake, and preferably even below 5% (4, 5).

Sugar-sweetened beverages are one of the major contributors to sugar intake (6–8), with sugar often being the only source of energy (9). Reduction of the intake of sugar-sweetened beverages is considered an important public health goal (10). In addition to raising public awareness, national authorities also encourage food manufacturers to reformulate their products to reduce added sugars. Some governments have implemented a sugar/soda tax (e.g., UK, Ireland, France, Mexico), and already results show a rapid increase in reformulation and decreased sales of sugary drinks (11–13). Since sugar is an important ingredient that provides a desired taste, its reduction could decrease consumers' interest in the product. Therefore, LNCS are often a convenient alternative, allowing formulation of products with reduced sugar and energy content, while achieving the same level of sweetness (3, 14). Consumer awareness and reformulation activities can result not only in increased availability, but also in increased intake of products with added LNCS (2, 15, 16). Although the use of LNCS is carefully regulated (17, 18), some studies have suggested possible risks related to the excessive consumption of LNCS, i.e., impacts on the microbiome (19), and increased risk of metabolic syndrome (20) and diabetes (21). However, it should be noted that because different LNCS have very different chemical structures, generalizations of health risks are not appropriate (22, 23).

With consideration of abovementioned issues, careful monitoring of changes in the food supply is crucial. Out of all foods and drinks, non-alcoholic beverages have been shown to be one of the major contributors to LNCS availability (15) and intake (24, 25). In Slovenia, monitoring of the use of LNCS in non-alcoholic beverages was undertaken for products available in 2017, following by further data collection in 2019 (26). While no significant differences were observed in a comparative study (26), signs of increased use of LNCS were found, making this area very interesting for further research. In Slovenia, there are currently no legislative restrictions or taxations for the content of sugars in sugary drinks. However, the government is promoting a reduction of added sugar through mass media and voluntary pledges (27). Although food reformulation in general is encouraged, the National Program on Nutrition and Health Enhancing Physical Activity 2015–2025 (28) specifically mentions that the use of LNCS should be also reduced. Given that a major part of the Slovenian food market consists of branded foods imported from other countries, the monitoring of the food supply will provide important insights into the European food supply. Furthermore, such insights are very important nationally for efficient, evidence-based policy decisions in the future.

Monitoring of the composition of the foods in the food supply is very challenging, because thousands of different products are available on the market, their compositions can change, and notable differences can be observed in different regions (29). This complicates data collection, and available datasets are commonly focused on nutritional food components, which are part of mandatory nutrition declarations, and less commonly on other food constituents, such as additives. It should also be mentioned that nationally representative branded food datasets are only available in some countries, and that different data collection approaches are used. A standard approach is cross-sectional food monitoring studies in food stores, where data is extracted from food labels. This methodological approach was harmonized within the Global Food Monitoring Group (30) and INFORMAS initiative (31, 32). In Slovenia, food monitoring studies are conducted within the government funded national research programme "Nutrition and Public Health," and are supported by the HORIZON2020 "Food Nutrition Security Cloud" project (FNS-Cloud; www.fns-cloud.eu), which is funded by the European Commission.

The present study aimed to investigate recent changes in the use of LNCS in non-alcoholic beverages in the Slovenian food supply. The national branded foods dataset was updated with beverages available in 2020. The dataset, containing both nutrition declaration and ingredients data, was compared with previous datasets compiled in 2017 and 2019. The use of repeated cross-sectional studies also enabled investigation of product-specific reformulation practices.

# MATERIALS AND METHODS

## Data Collection and Processing

This repeated cross-sectional study used new data collected in Slovenia in 2020, and previously reported data collected in 2017 and 2019 (26). In all 3 years, the data were collected using the Composition and Labeling Information System (CLAS, Nutrition Institute, Ljubljana, Slovenia) (33). In Slovenia, CLAS collect information on prepacked foods and drinks available at major retailers that represent most of the market share. The details regarding data collection are described elsewhere (34). In brief, pictures of all pre-packed foods and drinks with a unique European Article Number (EAN) barcode available at the time of sampling were collected. All the information on the nutritional composition and ingredients needed for the study were extracted from photographs using the online CLAS tool. The 2020 data were collected from all major retailers with nationwide networks of shops. The following shops in Ljubljana (Slovenia) were included in the data collection: two mega markets (Mercator Center, Interspar), two supermarkets (Tuš, Spar), and three discount markets (Hofer, Lidl, Eurospin). With the exception of Eurospin, the same shops were also included in the data collection in 2017 and 2019. Products were classified based on previously developed global categorisations by Dunford et al. (35), with minor adaptations considering the specifics of the European market. This study examined five categories of non-alcoholic beverages: juices, nectars, soft drinks, energy drinks, and sports drinks. Details regarding product categorization have been previously described (26).

The LNCS and added sugar were identified from the ingredient list on the food packaging. We examined the use of all 19 LNCS that are currently authorized for use in food products in European regulations (35). LNCS were identified by their name and/or their E number. Similarly, we identified added sugar, which was defined as all mono- and disaccharides added to foods, excluding fruit juices and purees. Based on this, beverages were segmented into four groups: (1) with added LNCS; (2) added sugar; (3) with added LNCS and added sugar; (4) without added LNCS and without added sugar. To further examine these

**TABLE 1 |** Sample description and availability of non-alcoholic beverages with added low- and/or no-calorie sweeteners (LNCS) in the Slovenian food supply in 2017, 2019, and 2020.

| | 2017 | | 2019 | | 2020 | |
|---|---|---|---|---|---|---|
| | **Total** *N* **(%)** | **Added LNCS** *N* **(%)** | **Total** *N* **(%)** | **Added LNCS** *N* **(%)** | **Total** *N* **(%)** | **Added LNCS** *N* **(%)** |
| Total | 1,043 (100) | 138 (13.2) | 1,221 (100) | 189 (15.5) | 1,650 (100) | 333 (20.2)[a,b] |
| Soft drinks | 555 (53.2) | 93 (16.8) | 601 (49.2) | 118 (19.6) | 898 (54.5) | 240 (26.7)[a,b] |
| Juices | 267 (25.6) | / | 330 (27.0) | / | 350 (21.2) | / |
| Nectars | 135 (12.9) | 16 (11.9) | 158 (12.9) | 11 (7.0) | 229 (13.8) | 12 (5.3)[b] |
| Energy drinks | 65 (6.2) | 11 (16.9) | 110 (9.0) | 46 (41.8) | 136 (8.2) | 56 (41.2)[b] |
| Sports drinks | 21 (2.0) | 18 (85.7) | 22 (1.8) | 14 (63.6) | 37 (2.2) | 25 (67.6) |

*N—number of products; Data for 2017 and 2019 from Hafner et al. (26).*
*[a]significant change between 2019 and 2020.*
*[b]significant change between 2017 and 2020.*

four groups with consideration of market-share differences, we compared the availability of such products in the food supply with sales data. We were able to obtain nation-wide 12-month sales data for 2017 and 2020 from retailers, representing over 50% of the national food supply. Sales data was provided in universal form including EAN barcode number, number of products sold per year, and package quantity (L). We matched products in the food supply dataset with those in sales data dataset. EAN barcode numbers were used as unique product identifiers for the matching process. There were 705 (68%) matches found in the 2017 dataset, and 1,007 (61%) in the 2020 dataset.

To provide insights into reformulation practices, we also matched products (using EAN barcode numbers) in the 2020 dataset with those in the 2017/2019 datasets. A similar approach was used by Bernstein et al. (36). All products available in the 2020 dataset were searched for matches in the 2017/2019 datasets. For those matches, we calculated the difference in total sugar content (TSC) between 2017 and 2020. If a product was not available in the 2017 dataset, we used the sugar content from the 2019 dataset.

## Statistical Analysis

Data were collected and processed using the Composition and Labeling Information System (CLAS) (Nutrition Institute, Ljubljana, Slovenia) and Microsoft$^{TM}$ Excel 2019. Statistical analysis was performed using IBM SPSS v.26. Descriptive statistics were used for reporting the prevalence of beverages with added LNCS, added sugar, both, or neither. We also assessed the prevalence of individual LNCS and their combinations. A two-tailed z-test was used to compare changes in the LNCS use between time periods. A t-test was used to compare the mean energy value (EV) and TSC within each category between time periods. For sale-weighting, we calculated the total amount (L) of beverages sold per year (separately for years 2017, and 2020) using package quantity (L) and the number of sold products. We presented the sale-weighted proportions and compared them with the food supply offer (products available at the time of sampling).

# RESULTS

## Presence of LNCS and Their Effect on Energy Value and Total Sugar Content

The 2020 dataset consisted of $N = 1{,}650$ non-alcoholic beverages. The between category distribution was similar to that observed in 2017 and 2019 (26). Soft drinks had the largest share in the sample ($N = 898$; 54.5%), followed by juices ($N = 350$; 21.2%), nectars ($N = 229$; 13.8%), energy drinks ($N = 136$; 8.2%), and sports drinks ($N = 37$; 2.2%) (**Table 1**). At least one LNCS was present in 333 products. The proportion of products with LNCS therefore increased from 15.5% in 2019 to 20.2% in 2020 ($p < 0.01$). The biggest difference between 2019 and 2020 was observed in soft drinks; the LNCS beverages increased from 19.6% to 26.7% ($p < 0.01$). A comparison with the 2017 data also provided interesting insights. In energy drinks, the proportion of products with LNCS increased between 2017 and 2019 (from 16.9% to 41.8%) (26), and stayed at a comparable level in 2020 (41.2%). Interestingly, the overall number of nectars increased considerably (from 135 in 2017 to 229 in 2020), but the number of such products with LNCS stayed almost the same ($n < 20$), which resulted in a decreased proportion of nectars with LNCS ($p < 0.05$). Beverages with LNCS are most often sports (67.6%) and energy drinks (41.2%), but due to the small number of products, they do not contribute much to the overall supply of non-alcoholic beverages with LNCS. No use of LNCS was observed in juices.

The results of the segmentation of beverages based on added sugar and/or LNCS are presented in **Table 2**. The largest contribution to the increase in the use of LNCS was made by beverages that use a combination of LNCS and added sugar. In 2017 and 2019, the proportions of such products were 7.6 and 7.5, respectively, increasing to 11.0% in 2020 ($p < 0.01$, comparison with 2017 data). Notable changes were also observed in the use of LNCS alone (5.7, 7.9, and 9.2, respectively).Similar trends were also observed, when 12-months sales values were considered, to account for market-share differences between different beverages in the food supply (**Supplementary Figure 1**). Increased availability of beverages with LNCS was also reflected in the increased sales of such beverages. From 2017 to 2020 the

**TABLE 2 |** Comparison of the energy value and total sugar content in non-alcoholic beverages, based on the presence of added sugar and low and/or no-calorie sweeteners (LNCS).

| | Added sugar | Added LNCS | 2017 | 2019 | 2020 | Energy value (kJ/100 mL) | | | Total sugar content (g/100 mL) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | 2017 | 2019 | 2020 | 2017 | 2019 | 2020 |
| | | | *N* (%) | *N* (%) | *N* (%) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Total | | | 1043 (100%) | 1221 (100%) | 1650 (100%) | 152.2 (66.5) | 147.9 (70.9) | 140.1 (72.2)*** | 8 (3.6) | 7.6 (3.8) | 7.3 (3.9)*** |
| | – | – | 300 (28.7%) | 398 (32.6%) | 435 (26.4%) | 174.8 (69.1) | 174.9 (67.6) | 179.9 (91.7) | 8.3 (3.8) | 8.2 (3.8) | 8.3 (3.9) |
| | ✓ | – | 605 (58.0%) | 634 (52.9%) | 882 (53.5%) | 161.8 (50.8) | 159 (52.7) | 154 (54)** | 8.9 (2.7) | 8.7 (2.8) | 8.5 (3)** |
| | – | ✓ | 59 (5.7%) | 97 (7.9%) | 152 (9.2%) | 22.3 (37.9) | 14.3 (27.9) | 13 (20.3) | 1 (2.1) | 0.4 (1.4) | 0.3 (1)* |
| | ✓ | ✓ | 79 (7.6%) | 92 (7.5%) | 181 (11.0%)** | 92.8 (28.1) | 96.2 (40) | 98.4 (41.1) | 4.8 (1.5) | 5.1 (2.3) | 5.3 (2.3) |
| Soft drinks | | | 555 | 601 | 898 | 126.8 (61.8) | 119.8 (63.3) | 109*** (61.2) | 7 (3.5) | 6.5 (3.5) | 6 (3.5)*** |
| | – | – | 26 | 43 | 45 | 60.2 (66.8) | 70 (55.4) | 40.8 (51.5) | 2.5 (3.5) | 2.4 (2.5) | 1.5 (2.3) |
| | ✓ | – | 436 | 440 | 613 | 146.3 (48.4) | 142.5 (50.9) | 134.8** (47.4) | 8.2 (2.7) | 7.9 (2.7) | 7.6 (2.7)** |
| | – | ✓ | 39 | 50 | 105 | 4.4 (4.8) | 5.8 (8) | 9.4** (14.6) | 0.1 (0.2) | 0.2 (0.8) | 0.2 (0.6)** |
| | ✓ | ✓ | 54 | 68 | 135 | 90.4 (29.7) | 87.5 (35.1) | 91.9 (33.8) | 4.9 (1.6) | 4.7 (2) | 5 (1.9) |
| Juices | – | – | 267 | 330 | 350 | 186.7 (57) | 190.5 (54.8) | 193 (52.8) | 8.8 (3.3) | 9.1 (3.1) | 9.4 (2.9)* |
| Nectars | | | 135 | 158 | 229 | 190 (45.7) | 184.8 (46.9) | 190.4 (50.2) | 10 (2.4) | 9.5 (2.8) | 9.6 (3.3) |
| | – | – | 5 | 23 | 39 | 135 (107.5) | 141.8 (75) | 145.4 (68.7) | 7.8 (6.6) | 6.8 (4.3) | 5.9 (4.1) |
| | ✓ | – | 114 | 124 | 178 | 203.3 (29.1) | 198.8 (28.6) | 204.6 (33.7) | 10.6 (1.6) | 10.4 (1.8) | 10.8 (2.1) |
| | – | ✓ | 10 | 6 | 5 | 104.7 (5.4) | 116.5 (23.1) | 91.6 (29.5) | 5.6 (0.4) | 5.2 (0.7) | 4.6 (1.6) |
| | ✓ | ✓ | 6 | 5 | 7 | 125.5 (29.1) | 117.2 (16) | 105.7 (14.7) | 6.5 (1.6) | 6 (0.6) | 5.7 (0.4) |
| Energy drinks | | | 65 | 110 | 136 | 174.2 (72.9) | 135.6 (92.9) | 141.5 (88.9)** | 9.6 (4.2) | 7.4 (5.5) | 7.7 (5.2)** |
| | – | – | 2 | 2 | 1 | 240 | 203 (52.3) | 224 | 13 | 11.4 (2.3) | 12.5 |
| | ✓ | – | 52 | 62 | 79 | 203.8 (28.7) | 203.1 (31.4) | 197.5 (40.6) | 11.4 (1.6) | 11.3 (1.8) | 11 (2.3) |
| | – | ✓ | 8 | 36 | 35 | 11.5 (4.7) | 10.6 (7.2) | 12.2 (6.1) | 0 (0) | 0 (0) | 0 (0) |
| | ✓ | ✓ | 3 | 10 | 21 | 72.7 (7.2) | 153.9 (45.6) | 142.4 (70.1) | 3.7 (0.4) | 8.3 (2.7) | 7.8 (3.9) |
| Sports drinks | | | 21 | 22 | 37 | 85.1 (30.9) | 74.2 (41.5) | 81.3 (38.3) | 3.9 (1.4) | 3.5 (2.1) | 3.7 (2) |
| | ✓ | – | 3 | 8 | 12 | 102.7 (26.5) | 104.5 (16.3) | 103.1 (24.1) | 5.1 (1.2) | 5.3 (0.7) | 5.2 (0.9) |
| | – | ✓ | 2 | 5 | 7 | 2.5 (3.5) | 4.6 (2.6) | 13.7 (21.5) | 0 (0) | 0 (0) | 0 (0) |
| | ✓ | ✓ | 16 | 9 | 18 | 92.1 (12.3) | 86 (13.7) | 93.1 (14.4) | 4.2 (0.4) | 3.9 (0.2) | 4.2 (0.4) |

*N, number of products; SD, standard deviation.*
*$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. Data for 2017 and 2019 from Hafner et. al (26).*

sale-weighted proportion of beverages with LNCS increased from 10.8 to 18.2%. Sales increased both for beverages with added only LNCS (from 4.1 to 8.8%) as well as for beverages with added LNCS and sugar (from 6.7 to 9.4%). The largest market share represented beverages with added sugar, but their share fell from 81.5% in 2017 to 72.2% in year 2020, when notably increased sales of beverages with LNCS were observed. Beverages without added sugar and LNCS, despite representing a large proportion of the available beverages, represented only 7.7% of the volume sales-market in 2017, and 9.6% in 2020.

We also observed that the energy and sugar content of the beverages also changed in the last few years (**Table 2**). Taking the whole 2020 dataset into account, the average content of energy and sugar in the non-alcoholic beverages was 140.1 kJ and 7.3 g/100 mL, respectively. The EV and TSC therefore significantly fell from 2017 (152.2 kJ and 8 g of sugar per 100 mL; $p < 0.001$) (**Table 2**). This reduction is the result of both increased use of LNCS and a reduction in TSC in other beverages. The results showed that both the EV and the TSC were considerably reduced when the sugars were partially replaced by LNCS, while the difference was even more pronounced when only LNCS were used for sweetening. Encouragingly, a decrease in the EV and TSC was also observed in beverages with only added sugar ($p < 0.01$). The amount of TSC also dropped slightly in beverages with only added LNCS ($p < 0.05$). In individual categories, considerable changes were observed in soft drinks, which, due to their abundance, contributed the most to the changes in the overall sample. The mean EV and TSC in soft drinks decreased both in the whole category ($p < 0.001$), and also specifically in soft drinks with only added sugar ($p < 0.01$). In other categories, we observed a decrease in the mean EV and TSC for energy drinks ($p < 0.01$), which was mostly due to the increased use of LNCS, as after segmentation based on added LNCS and added sugar, no differences were observed in the four segments. A slight increase in TSC was also observed in juices ($p < 0.05$), which could mean that the supply of juices from sweeter fruits is on the rise. Juices were also the category with the highest EV and TSC.

The food matching method was used to provide insights into reformulation practices in specific products. Analyses was undertaken with $N = 859$ non-alcoholic beverages in the 2020 dataset, for which matches were found in previous datasets. For 680 (79.2%) products, no change in TSC was observed. Out of 179 products with changed compositions, 144 (16.8%) showed a reduction and 35 (4.1%) an increase in TSC (**Figure 1**). It should be noted that 31 (3.6%) products with reduced TSC also contained LNCS, and 20 (2.3%) did not contain LNCS in previous years. Most products with reduced sugar content were in the category of soft drinks ($N = 107$); this trend was particularly notable in fruit drinks. Interestingly, the average TSC in sugar-reduced reformulated beverages was quite high (7.4 g per 100 mL, in comparison to the overall 2020 average of 7.3 g per 100 mL). The mean sugar reduction for products without LNCS was −1.1 g per 100 ml, and −3.3 g for products with added LNCS. Products reformulated with increased TSC had on average 1.0 g more sugar per 100 ml. Interestingly, we found four energy drinks that did not contain LNCS in previous datasets, where LNCS were added in 2020, but their sugar content remained the same.

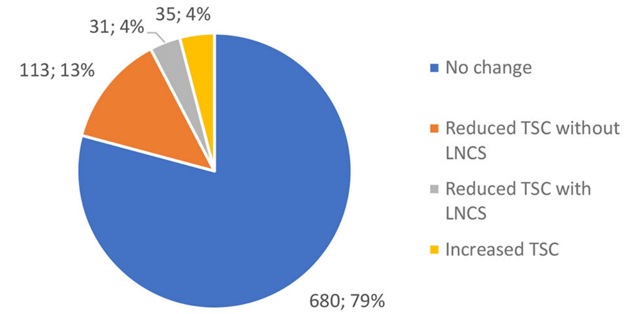


**FIGURE 1 |** Changes in the total sugar content (TSC) in beverages available in 2020, which were also found in the 2017/2019 dataset ($N = 859$; LNCS–low- and/or no-calorie sweeteners).

## Prevalence of Different LNCS and Their Combinations

An overview of the use of specific sweeteners showed no significant differences in comparison with the previous data collection in 2019 (**Figure 2**). However, some interesting trends identified in a previous study (26) continued in 2020. We observed a further increase in the use of sucralose, and a consequently lower use of most other sweeteners. Consistent with the results from 2019, the most common sweetener in 2020 remained acesulfame K ($N = 191$; 57.3%), followed by sucralose ($N = 139$; 41.7%) and aspartame ($N = 102$; 30.6%).

In beverages with LNCS, the use of multiple sweeteners at the same time is still prevalent (65%), while the use of only one LNCS (35%) is limited to the use of steviol glycosides ($N = 59$), sucralose ($N = 51$), and saccharin ($N = 6$) (**Supplementary Table 1**). Even though steviol glycosides and saccharin were used as single LNCS, they were always combined with added sugar. Meanwhile, sucralose was used as single LNCS without added sugar in 16 products. Drinks with LNCS most commonly contained a mix of two LNCS ($N = 128$, 38.4%). The most common combination was acesulfame K and sucralose ($N = 54$), followed by the combination of acesulfame K and aspartame ($N = 42$). Interestingly when mixing three or more LNCS, acesulfame K was always present in the mixture. Even for products containing only two LNCS, products that did not contain acesulfame K were in the minority ($N = 26$), and were usually combined with added sugar or fruit juices/concentrates. The results also suggested that as the number of sweeteners increases, the frequency of added sugar decreases.

# DISCUSSION

This study showed that the market for non-alcoholic beverages has changed rapidly in recent years. While a previous study found the first signs of growing use of LNCS in non-alcoholic beverages in Slovenia (26), this trend was clear when the latest data were taken into account. Our results showed a significant increase in the availability of beverages with LNCS, from 13.2% in 2017

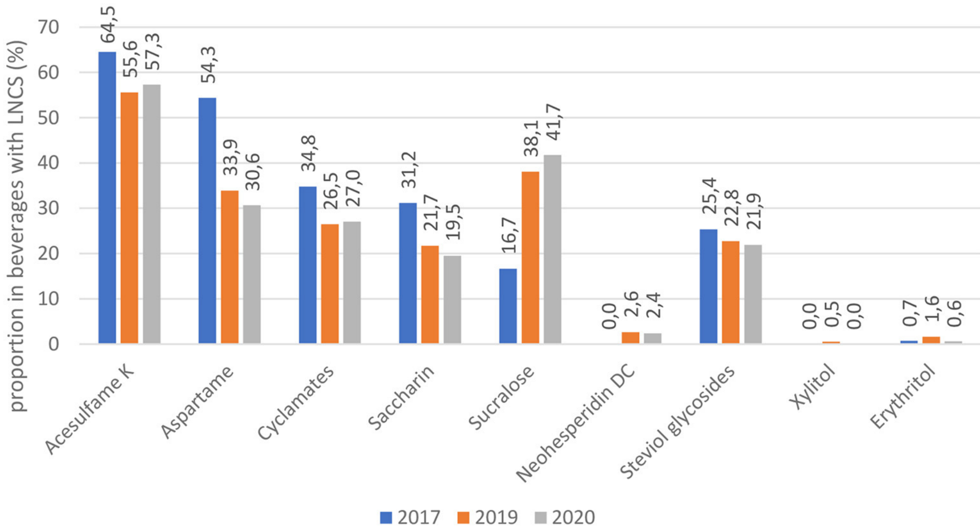


**FIGURE 2 |** Frequency of different low and no-calorie sweeteners (LNCS) in beverages with LNCS [data for 2017 and 2019 from Hafner et. al (26)].

and 15.5% in 2019, to 20.2% in 2020. In the period of 2017–2019 the changes mostly affected energy drinks; use of LNCS increased significantly. This trend stopped in 2020. On the other hand, in 2020 there was a marked growth in soft drinks with LNCS (from 19.6% in 2019 to 26.7% in 2020; $p < 0.01$), which is the most widespread category of non-alcoholic beverages in the Slovenian market.

We should note that the frequency of the use of LNCS in non-alcoholic beverages in Slovenia (20.2%) is still lower than in most other countries where it has been investigated (16, 37, 38). However, given the observed rapid changes, Slovenia is quickly approaching the proportion of beverages with LNCS in the US (23%; 2015–2017) (38) and Hong Kong (25%; 2019) (39). Even more frequent use of LNCS in beverages was reported in Spain (39%; 2013) (37), but the latter study was conducted with a very different methodological approach, using food consumption survey data. Very recently, a study with a methodological approach similar to our study was also conducted in Spain (15), but because only foods with LNCS were investigated, the proportion of use of LNCS was not calculated.

In our study, growth in the use of LNCS occurred particularly in the availability of beverages with both added sugar and LNCS. We also observed that increased sales-volume market-share for beverages with LNCS reflected the increase in the availability of such products. Increased sales for beverages with LNCS were also reported in a US study, which highlighted, that from 2002 to 2018, sales of products containing both added sugar and LNCS increased by almost 30% (40). The current study, as well as others (26, 38), observed that partially replacing sugar with LNCS can help reduce energy by almost half, while using only LNCS for sweetening leads to an even greater reduction. The increased supply of beverages with LNCS, therefore, impacted the overall mean EV and TSC, which decreased significantly from 2017 ($p < 0.001$). Encouragingly, the reduction of the average sugar content in the available non-alcoholic beverages was also influenced by the reduction of sugars in beverages sweetened only with added sugar ($p < 0.01$). This particularly applied to sugar-sweetened soft drinks.

Results of analyses of product-specific reformulations showed that 144 (16.8%) out of 859 matched products reduced their sugar content. Interestingly, a minority of these beverages were reformulated with the help of LNCS ($N = 31$; 3.6%). This is a promising result, since the Slovenian nutrition policy programme (28) encourages reformulation without LNCS, to reduce the preference for sweetness in the population. At the same time, it should be pointed out that products that used LNCS reduced their sugar content much more (−3.3 g per 100 ml) compared to those that did not use them (−1.1 g per 100 ml). However, the reformulated products still had a fairly high average sugar content (7.4 g per 100 ml), which was similar to the overall average (7.3 g), indicating that reformulation is particularly focused on products with a higher sugar content. Most of the reformulated beverages were soft drinks ($N = 107$), more precisely fruit drinks, which is the group of beverages that has been the most reformulated in Portugal in the past decade (41). It is also important to note that we also observed cases of increased TSC ($N = 35$), and those where added LNCS was not accompanied by lower TSC ($N = 4$). In such cases, the sweetness was increased without any health benefits, which is in clear conflict with public health goals.

Among the specific LNCS, the use of acesulfame K predominated in 2020, followed by sucralose and aspartame, which is aligned with the 2019 results. Most beverages contained multiple LNCS, while sole LNCS occurred mostly in beverages with sucralose and steviol glycosides, and occasionally saccharin. During recent years, the use of LNCS has shifted toward the increased use of sucralose and the decreased use of aspartame. Reduced occurrence of aspartame has also been observed in Portugal (41) and the US (40). Even though aspartame has been re-evaluated for safety (42), its use still remains controversial (43). This raises doubts in consumers, which is why manufacturers have replaced aspartame with novel, less notorious LNCS, such as sucralose and steviol glycosides (40). The use of sucralose is also increasing in Slovenia. In the US, the intake of sucralose increased by over 30% in a 16 year period, which could also be due to the increased supply of prepackaged products containing sucralose (40). Sucralose is relatively new to the market. Its use has rapidly increased because of its sucrose-like taste and good stability (44). Due to these properties, it appears both alone and in combination with other sweeteners. Increased use of sucralose has aroused the interest of researchers, who have begun to focus on its impact on health. Even though sucralose is currently considered safe, some studies indicate possible adverse effects on glucose metabolism, even in low amounts (15% of Acceptable Daily Intake, ADI) (45), and increased cardiovascular risk (46). A combination of carbohydrates and sucralose in beverages was highlighted as particularly risky (47). Interestingly, our study showed that use of steviol glycosides is slowly stagnating; they are only in fifth place among the most common sweeteners, used in 21.9% of beverages with LNCS. On the contrary, a considerable increase in the use of this sweetener has been reported from other countries (40, 41). In Chile, steviol glycosides are the second most used LNCS in the food supply, which has raised concerns that ADI value could be exceeded in some vulnerable groups, such as children (16), but a subsequent study indicated that this is not the case (48). A Portuguese study highlighted steviol glycosides as most commonly present in iced teas (41), which are a popular choice for children (49). Therefore, careful monitoring of this is crucial. In Slovenia, the frequency of use of steviol glycosides is currently relatively low, but given the rapid changes in LNCS use, this could change in the coming years. Steviol glycosides are commonly perceived by consumers as a natural sweetener, and are therefore rarely mixed with other LNCS (16, 50). However, the use of mixes of LNCS is still predominant in beverages. For the first time, we also reviewed LNCS combinations. We found out that combination of two LNCS is the most common, with acesulfame K and sucralose emerging as the most frequent blend. The same result was also reported in Spain, across the whole food supply (15). Blends can intensify the sweet taste of individual LNCS and prevent an unpleasant aftertaste (50). Our results indicated that with the increased number of LNCS, the frequency of added sugar has decreased. Therefore, beverages with only one LNCS in most cases also contained added sugars, while beverages with a blend of five LNCS did not contain any. This suggests that blends of LNCS could help to notably reduce sugar content in beverages.

It should be noted that in Slovenia beverages account for the largest share of sugar sold in the country (27), and are also the biggest contributor to free sugar intake (32% in adolescents, 34% in adults, and 31% in the elderly) (8). However, a reduction in TSC in beverages has occurred much slower than in some other countries, where sugar taxes have been implemented. For example, in Portugal, 1 year after tax implementation in 2017, 50% of soft drinks above the taxation level (8 g sugar/100 ml) reduced their TSC below this limit (51), and now only 15% exceed it (41). Meanwhile, in Slovenia (2020), about one third ($N = 310$; 34.5%) of soft drinks have a TSC above 8 g/100 ml. Although similar benefits of taxation have been seen in other countries, such reformulations might considerably increase the use of LNCS (52), and this could affect health risk analyses. At the same time, LNCS maintain or even intensify the sweetness of drinks, which hinders the main public health message to reduce the preferences for sweetness in our diets. Policy approaches for lowering the TSC in beverages are desirable, however, attention should be paid to possible excessive substitution of sugar with LNCS.

## Strengths and Limitations

As the main strength of this study, we should highlight the representativeness of the sample, which included beverages from all major retailers, representing a vast majority of the Slovenian food market. A repeated cross-sectional approach with the use of the same methodology and three time points allowed us to make meaningful comparisons and identify changes in the food supply. It is also noteworthy that there are only a few countries in Europe where the infrastructure enables such studies, which is why this study provided important insights to better understand the common European market. Some study limitations should be also mentioned. First, the data collection in 2020 included one discount retailer which was not included in monitoring in 2017/2019. We carefully checked that this did not have a major effect on the study results. In the whole 2020 dataset LNCS were present in 20.2% beverages; after exclusion of products found only at the additional retailer ($N = 103$), it was 19.8%. Another limitation is that sales data were not available for all beverages, however we covered 68% and 61% of the 2017 and 2020 sample, respectively. To exclude the possibility of the error due to the missing sales data, the offer of products included in sale-weighting were compared with the whole sample; no notable differences were observed between both samples. A limitation of the present study is also that all the information was extracted from food labels, which may differ from the actual chemical compositions of the beverages. We also only investigated the use of (declared) LNCS, and not their quantity, as this information is not indicated on the label. Finally, we used EAN barcodes for food matching in the analyses of product-specific reformulations. In case that the product has changed its EAN barcode, we were unable to match it with the previous formulation.

# CONCLUSIONS

This study explored the use of LNCS in non-alcoholic beverages in the 2020 edition of Slovenian branded foods database, for comparison with previous data. We showed that the use of LNCS in beverages increased for more than half (+53%)–from 13.2% in year 2017 to 20.2% in year 2020, with even more notable growth in soft drinks (for 59%–from 16.8% in 2017 to 26.7% in year 2020). Increased availability of beverages with LNCS also reflected in even higher increase in sales volumes; market-share of beverages with LNCS increased for 69%–from 10.8% in 2017 to 18.2% in year 2020. While study results also indicated some changes in the content of energy and sugars in both reformulated and newly launched beverages, most beverages in the Slovenian food supply still have very high energy/sugar content. To achieve public-health goals, more efficient reformulation activities are needed. Further monitoring of the composition of beverages in the food supply is needed also to assess the efficiency of such activities.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

IP: conceptualization, and manuscript writing—review and editing. EH: data analyses, formal analysis, writing—original draft preparation. IP and EH: methodology. All authors have read and agreed to the published version of the manuscript.

## FUNDING

Data collection for this study was supported by the national research programme Nutrition and Public Health (P3-0395, funded by the Slovenian Research Agency) and research project L3-9290, funded by the Ministry of Health of Republic of Slovenia and Slovenian Research Agency. The study was conducted within the Food Nutrition Security Cloud project (FNS-Cloud), which received funding from the European Union's Horizon 2020 Research and Innovation programme (H2020-EU.3.2.2.3.—A sustainable and competitive agri-food industry) under grant agreement no. 863059. Information and views in this report do not necessarily reflect the official opinion or position of the European Union. Neither European Union institutions and bodies, nor any person acting on their behalf, may be held responsible for the use that may be made of the information contained herein.

## ACKNOWLEDGMENTS

We thank the retailers for granting access to their stores to collect data for the study. We also acknowledge the support of Maša Hribar, Sanja Krušič, Hristo Hristov, Živa Lavriša, Anita Kušar, and Katja Žmitek (Nutrition Institute, Ljubljana, Slovenia) in conducting the food supply studies, and to students from the Biotechnical Faculty (University of Ljubljana) and BIC (Ljubljana) for their help with the data collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.778178/full#supplementary-material

## REFERENCES


1. Fujimaru T. Park Jh, Fau-Lim J, Lim J. Sensory characteristics and relative sweetness of tagatose and other sweeteners. *J Food Sci*. (2012) 77:S323–8. doi: 10.1111/j.1750-3841.2012.02844.x
2. Sylvetsky AC, Rother KI. Trends in the consumption of low-calorie sweeteners. *Physiol Behav*. (2016) 164:446–50. doi: 10.1016/j.physbeh.2016.03.030
3. Tahmassebi JF, BaniHani A. Impact of soft drinks to health and economy: a critical review. *Eur Arch Paediatr Dent*. (2020) 21:109–17. doi: 10.1007/s40368-019-00458-0
4. World Health Organization. *Diet, Nutrition and the Prevention of Chronic Diseases—Report of the Joint WHO/FAO Expert Consultation*. Geneva, Switzerland: World Health Organization (2003)
5. World Health Organization. *Guideline: Sugars Intake for Adults and Children*. Geneva, Switzerland: World Health Organization (2015).
6. Drewnowski A, Rehm CD. Consumption of added sugars among US children and adults by food purchase location and food source. *Am J Clin Nutr*. (2014) 100:901–7. doi: 10.3945/ajcn.114.089458
7. Azaïs-Braesco V, Sluik D, Maillot M, Kok F, Moreno LA. A review of total & added sugar intakes and dietary sources in Europe. *Nutr J*. (2017) 16:6. doi: 10.1186/s12937-016-0225-2
8. Zupanič N, Hristov H, Gregorič M, Blaznik U, Delfar N, Koroušić Seljak B, et al. Total and free sugars consumption in a slovenian population representative sample. *Nutrients*. (2020) 12:1729. doi: 10.3390/nu12061729
9. Ferruzzi MG, Tanprasertsuk J, Kris-Etherton P, Weaver CM, Johnson EJ. Perspective: the role of beverages as a source of nutrients and phytonutrients. *Adv Nutr*. (2020) 11:507–23. doi: 10.1093/advances/nmz115
10. Lobstein T. *Reducing Consumption of Sugar-Sweetened Beverages to Reduce the Risk of Childhood Overweight and Obesity*. London, World Health Organization (2014) Available onile at: https://www.who.int/elena/titles/bbc/ssbs_childhood_obesity/en/.
11. Pell D, Mytton O, Penney TL, Briggs A, Cummins S, Penn-Jones C, et al. Changes in soft drinks purchased by British households associated with the UK soft drinks industry levy: controlled interrupted time series analysis. *BMJ*. (2021) 372:n254. doi: 10.1136/bmj.n254
12. Bandy LK, Scarborough P, Harrington RA, Rayner M, Jebb SA. Reductions in sugar sales from soft drinks in the UK from 2015 to 2018. *BMC Med*. (2020) 18:20. doi: 10.1186/s12916-019-1477-4
13. Sánchez-Romero LM, Canto-Osorio F, González-Morales R, Colchero MA, Ng S-W, Ramírez-Palacios P, et al. Association between tax on sugar sweetened beverages and soft drink consumption in adults in Mexico: open cohort longitudinal analysis of health workers cohort study. *BMJ*. (2020) 369:m1311. doi: 10.1136/bmj.m1311
14. Popkin BM, Hawkes C. Sweetening of the global diet, particularly beverages: patterns, trends, and policy responses. *Lancet Diabetes Endocrinol*. (2016) 4:174–86. doi: 10.1016/S2213-8587(15)00419-2
15. de Lourdes Samaniego-Vaesken M, González-Fernández B, Partearroyo T, Urrialde R, Varela-Moreiras G. Updated database and trends of declared low- and no-calorie sweeteners from foods and beverages marketed in Spain. *Front Nutr*. (2021) 8:670422. doi: 10.3389/fnut.2021.670422
16. Sambra V, Lopez-Arana S, Caceres P, Abrigo K, Collinao J, Espinoza A, et al. Overuse of non-caloric sweeteners in foods and beverages in chile: a threat to consumers' free choice? *Front Nutr*. (2020) 7:68. doi: 10.3389/fnut.2020.00068
17. Dusemund B, Gilbert J, Gott D, Kenigswald H, König J, Lambré C, et al. Food additives and nutrient sources added to food: developments since the creation of EFSA. *EFSA J*. (2012) 10:s1006. doi: 10.2903/j.efsa.2012.s1006
18. Serra-Majem L, Raposo A, Aranceta-Bartrina J, Varela-Moreiras G, Logue C, Laviada H, et al. Ibero?american consensus on low- and no-calorie sweeteners: safety, nutritional aspects and benefits in food and beverages. *Nutrients*. (2018) 10:818. doi: 10.3390/nu10070818
19. Ruiz-Ojeda FJ, Plaza-Diaz J, Saez-Lara MJ, Gil A. Effects of sweeteners on the gut microbiota: a review of experimental studies and clinical trials. *Adv Nutr*. (2019) 10(Suppl. 1):S31–48. doi: 10.1093/advances/nmy037

20. Zhang X, Li X, Liu L, Hong F, Zhao H, Chen L, et al. Dose–response association between sugar- and artificially sweetened beverage consumption and the risk of metabolic syndrome: a meta-analysis of population-based epidemiological studies. *Public Health Nutr.* (2020) 24:3892–904. doi: 10.1017/S1368980020003614
21. de Koning L, Malik VS, Rimm EB, Willett WC, Hu FB. Sugar-sweetened and artificially sweetened beverage consumption and risk of type 2 diabetes in men. *Am J Clin Nutr.* (2011) 93:1321–7. doi: 10.3945/ajcn.110.007922
22. Hunter SR, Reister EJ, Cheon E, Mattes RD. Low Calorie Sweeteners Differ in Their Physiological Effects in Humans. *Nutrients.* (2019) 11:2717. doi: 10.3390/nu11112717
23. Pang MD, Goossens GH, Blaak EE. The Impact of Artificial Sweeteners on Body Weight Control and Glucose Homeostasis. *Front Nutr.* (2021) 7:333. doi: 10.3389/fnut.2020.598340
24. Dewinter L, Casteels K, Corthouts K, Van de Kerckhove K, Van der Vaerent K, Vanmeerbeeck K, et al. Dietary intake of non-nutritive sweeteners in type 1 diabetes mellitus children. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* (2016) 33:19–26. doi: 10.1080/19440049.2015.1112039
25. Garavaglia MB, Rodríguez García V, Zapata ME, Rovirosa A, González V, Flax Marcó F, et al. Non-nutritive sweeteners: children and adolescent consumption and food sources. *Arch Argent Pediatr.* (2018) 116:186–91. doi: 10.5546/aap.2018.eng.186
26. Hafner E, Hribar M, Hristov H, Kušar A, Žmitek K, Roe M, et al. Trends in the use of low and no-calorie sweeteners in non-alcoholic beverages in Slovenia. *Foods.* (2021) 10:387. doi: 10.3390/foods10020387
27. Zupanič N, Hribar M, Fidler Mis N, Pravst I. Free sugar content in pre-packaged products: does voluntary product reformulation work in practice? *Nutrients.* (2019) 11:2577. doi: 10.3390/nu11112577
28. National Program on Nutrition and Health Enhancing Physical Activity 2015–2025, (2015).
29. Dunford EK, Mhurchu CN, Huang L, Vandevijvere S, Swinburn B, Pravst I, et al. A comparison of the healthiness of packaged foods and beverages from 12 countries using the health star rating nutrient profiling system, 2013-2018. *Obes Rev.* (2019) 20:107–15. doi: 10.1111/obr.12879
30. Dunford E, Webster J, Metzler AB, Czernichow S, Ni Mhurchu C, Wolmarans P, et al. International collaborative project to compare and monitor the nutritional composition of processed foods. *Eur J Prev Cardiol.* (2012) 19:1326–32. doi: 10.1177/1741826711425777
31. Neal B, Sacks G, Swinburn B, Vandevijvere S, Dunford E, Snowdon W, et al. Monitoring the levels of important nutrients in the food supply. *Obesity Reviews.* (2013) 14:49–58. doi: 10.1111/obr.12075
32. Rayner M, Vandevijvere S. *INFORMAS Protocol: Food Labelling Module*. Auckland: The University of Auckland (2017).
33. NUTRIS. *Podatkovna Baza CLAS Kot Orodje za Vrednotenje Sprememb na Podrocju Ponudbe Predpakiranih Živil*. Sloveniji (2021). Available onilne at: https://www.nutris.org/clas/.
34. Pivk Kupirovič U, Miklavec K, Hribar M, Kušar A, Žmitek K, Pravst I. Nutrient Profiling Is Needed to Improve the Nutritional Quality of the Foods Labelled with Health-Related Claims. *Nutrients.* (2019) 11:287. doi: 10.3390/nu11020287
35. European Commission. *Regulation (EC) No 1333/2008 of the European Parliament and of the Council of 16 December 2008 on food additives*. Official Journal of the European Union (2008.)
36. Bernstein JT, Christoforou AK, Weippert M, L'Abbé MR. Reformulation of sugar contents in Canadian prepackaged foods and beverages between 2013 and 2017 and resultant changes in nutritional composition of products with sugar reductions. *Public Health Nutr*. (2020) 23:2870–8. doi: 10.1017/S1368980020001159
37. Samaniego-Vaesken ML, Ruiz E, Partearroyo T, Aranceta-Bartrina J, Gil A, Gonzalez-Gross M, et al. Added Sugars and Low- and No-Calorie Sweeteners in a Representative Sample of Food Products Consumed by the Spanish ANIBES Study Population. *Nutrients.* (2018) 10:1265. doi: 10.3390/nu10091265
38. Dunford EK, Taillie LS, Miles DR, Eyles H, Tolentino-Mayo L, Ng SW. Non-nutritive sweeteners in the packaged food supply-an assessment across 4 countries. *Nutrients.* (2018) 10:257. doi: 10.3390/nu10020257
39. BYS, Coyle DH, Dunford EK, Wu JHY, Louie JCY. The use of non-nutritive and low-calorie sweeteners in 19,915 local and imported pre-packaged foods in Hong Kong. *Nutrients*. (2021) 13:1861. doi: 10.3390/nu13061861
40. Dunford EK, Miles DR, Ng SW, Popkin B. Types and amounts of nonnutritive sweeteners purchased by us households: a comparison of 2002 and 2018 nielsen homescan purchases. *J Acad Nutr Diet.* (2020) 120:1662–71. doi: 10.1016/j.jand.2020.04.022
41. Silva PD, Cruz R, Casal S. Sugars and artificial sweeteners in soft drinks: a decade of evolution in Portugal. *Food Control.* (2021) 120:107481. doi: 10.1016/j.foodcont.2020.107481
42. EFSA Panel on Food Additives and Nutrient Sources added to Food (ANS). Scientific Opinion on the re-evaluation of aspartame (E 951) as a food additive. *EFSA J.* (2013) 11:3496-759.
43. Czarnecka K, Pilarz A, Rogut A, Maj P, Szymańska J, Olejnik Ł, et al. Aspartame—true or false? Narrative review of safety analysis of general use in products. *Nutrients.* (2021) 13:1957. doi: 10.3390/nu13061957
44. Shankar P, Ahuja S, Sriram K. Non-nutritive sweeteners: Review and update. *Nutrition.* (2013) 29:1293–9. doi: 10.1016/j.nut.2013.03.024
45. Romo-Romo A, Aguilar-Salinas CA, Brito-Córdova GX, Gómez-Díaz RA, Almeda-Valdes P. Sucralose decreases insulin sensitivity in healthy subjects: a randomized controlled trial. *Am J Clin Nutr.* (2018) 108:485–91. doi: 10.1093/ajcn/nqy152
46. Risdon S, Battault S, Romo-Romo A, Roustit M, Briand L, Meyer G, et al. Sucralose and cardiometabolic health: current understanding from receptors to clinical investigations. *Adv Nutr.* (2021) 12:1500–13. doi: 10.1093/advances/nmaa185
47. Dalenberg JR, Patel BP, Denis R, Veldhuizen MG, Nakamura Y, Vinke PC, et al. Short-term consumption of sucralose with, but not without, carbohydrate impairs neural and metabolic sensitivity to sugar in humans. *Cell Metab.* (2020) 31:493–502. doi: 10.1016/j.cmet.2020.01.014
48. Martínez X, Zapata Y, Pinto V, Cornejo C, Elbers M, van der Graaf M, et al. Intake of non-nutritive sweeteners in chilean children after enforcement of a new food labeling law that regulates added sugar content in processed foods. *Nutrients.* (2020) 12:1594. doi: 10.3390/nu12061594
49. Smith LH, Baumker E. Sugar-sweetened beverage behaviors of tennessee school children: how do parent- and child-report compare on school days and nonschool days? *J Spec Pediatr Nurs.* (2019) 24:e12231. doi: 10.1111/jspn.12231
50. O'Brien-Nabors L. *Alternative Sweeteners*. 3 ed. New york, NY:Marcel Dekker (2001).p. 541.
51. Goiana-da-Silva F, Morais Nunes A, Miraldo M, Bento A, Breda J, Ferreira Araújo F. Using pricing policies to promote public health: the sugar sweetened beverages taxation experience in portugal. *Acta Med Port*. (2018) 31:191–5. doi: 10.20344/amp.10222
52. Popkin BM, Ng SW. Sugar-sweetened beverage taxes: lessons to date and the future of taxation. *PLoS Med.* (2021) 18:e1003412. doi: 10.1371/journal.pmed.1003412


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 23 November 2021
doi: 10.3389/fnut.2021.773260

# Use Database to Evaluate the Prevalence of Hunger Among Adolescents in Brazil

*Ana Laura Benevenuto de Amorim*[1,2], *José Raimundo Sousa Ribeiro Junior*[1], *Helida Ventura Barbosa Gonçalves*[1,2] *and Daniel Henrique Bandoni*[1]*

[1] Nutrition and Food Service Research Center, Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil,
[2] Interdisciplinary Program of Health Sciences, Federal University of São Paulo (UNIFESP), Santos, Brazil

Food insecurity and malnutrition have become serious problems in many countries. In recent years, Brazil has experienced an increase in the prevalence of food insecurity and hunger. However, there is limited information on the status of these issues, and food security assessments are only performed as household measures. Therefore, the use of available databases is essential to expand information and support decision-making in the fight against food insecurity. This study aimed to evaluate the relationship between reports of hunger among adolescents and their sociodemographic characteristics. We used data from the 2015 National School Health Survey. The main variable of interest was obtained from responses to the following question: "Over the past 30 days, how often have you gone hungry because you did not have enough food at home?". The responses were separately gathered from those who reported going hungry and those who did not. Socioeconomic characteristics were evaluated simultaneously. For statistical analysis, a Pearson chi-square test and multiple analyses were performed using Poisson regression models. A total of 101,888 adolescents were evaluated. The variable used to measure hunger was associated with maternal education, internet access, and fruit intake. The results showed a positive association between adolescents who reported going hungry and women, black and indigenous adolescent students living in households with more than five people, adolescents not living with their father, and adolescents planning to work or not knowing what they will do after completing the ninth grade. The results demonstrate that it is possible to use secondary data with a single question to assess, monitor, and provide insights into how food security impacts the sociodemographic groups differently.

**Keywords: hunger, adolescent, food insecurity, database, food access, nutrition, social determinants of health**

OPEN ACCESS

***Edited by:***
*Alessandra Durazzo, Council for Agricultural Research and Economics, Italy*

***Reviewed by:***
*Naia Ortelan, Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS), Brazil*
*Ana Leite, Centro Hospitalar de Vila Nova de Gaia, Portugal*

***Correspondence:**
*Daniel Henrique Bandoni*
*danielbandoni@gmail.com*

***Specialty section:***
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

***Received:*** *09 September 2021*
***Accepted:*** *20 October 2021*
***Published:*** *23 November 2021*

***Citation:***
*Amorim ALB, Ribeiro Junior JRS, Gonçalves HVB and Bandoni DH (2021) Use Database to Evaluate the Prevalence of Hunger Among Adolescents in Brazil. Front. Nutr. 8:773260. doi: 10.3389/fnut.2021.773260*

## INTRODUCTION

Adolescence is the stage of life marked by intense physical, cognitive, and emotional development. Thus, access to a well-balanced diet (one that is sufficient, varied, and complete) is even more relevant during this period (1, 2). Studies have presented the relationship between eating habits and the socioeconomic level of adolescents; youths belonging to more favoured economic classes have unhealthy eating habits, while those from poorer families consume rice and beans more often. During adolescence, nutritional problems that have occurred since childhood can be minimised, and healthy eating habits and lifestyles can be formed (3–7).

Food and Nutritional Security can be defined as a concept that "exists when all people have, at all times, physical, social, and economic access to safe food, consumed in sufficient quantity and quality, that meets their nutritional needs and food preferences in an environment with adequate sanitation and health services, allowing a healthy and active life" (8). There are several methodologies and indicators that can be used to measure food security; however, limited information is available for the assessment of food insecurity (FI) and hunger in specific population groups (9).

Food insecurity and hunger are significant problems in Brazil, and the number of people who experience food insecurity (FI) has increased in recent years, from 37.5 million (18.3%) households in 2013 to 43.1 million (20.6%) households in 2017 and 2018 (10). Data on the status of FI and hunger during the coronavirus disease-2019 (COVID-19) pandemic demonstrate the worsening of this scenario in Brazilian households, with intense growth in the prevalence of hunger (from 5.9 to 9%) and FI (from 36.7 to 55.2%) in the Brazilian population (11).

The prevalence of FI is higher in Brazilian households with children and adolescents than in households without children and adolescents (10). However, limited studies have evaluated the prevalence of hunger among adolescents and their associated factors. Previous studies have shown that FI, especially moderate and severe hunger, in this age group affects physical and emotional development and is related to high rates of school absenteeism and excessive weight (2, 12, 13).

Due to the inter-sectoral, cross-sectional, and multicausal nature of food and nutrition security (FNS), its understanding depends on the association among various aspects of reality. Low income is one of the most important socioeconomic determinants of FI and hunger. Other demographic and socioeconomic factors, such as colour/race, age, household location, access to water and other public utilities, living conditions, number of people per household, level of education, and employment status, are associated with an increased risk of FI (14–17).

Thus, it is fundamental to determine how many adolescents experience hunger. This study aimed to help close this gap and evaluate the relationship between hunger and sociodemographic characteristics among Brazilian adolescents based on the data from the National School Health Survey (PeNSE, 2015).

## METHODS

The National Survey of School Health (PeNSE) was first conducted in 2009 through an agreement with the Brazilian Institute of Geography and Statistics and the Ministry of Health with support from the Ministry of Education, and was designed to monitor the risk factors and protect the health and lifestyle behaviours of students in Brazilian schools. This study used data from the third edition of the PeNSE conducted in 2015, which investigated the behavioural risk and health protection factors in a sample of students on their ninth year of primary education in public and private schools throughout Brazil (18). This sample of adolescents adequately represented the youth across Brazil, including all 27 federative units. The survey was approved by the National Commission of Research Ethics (n. 1.006.467), Conep (18).

### Dependent Variable

The key variable of interest was the self-perception of hunger, which was identified based on the following question: "How often during the last 30 days have you gone hungry because you do not have enough food at home?" The response categories were never, rarely, sometimes, most of the time, and always.

For analysis, responses to how often the student was hungry were grouped into the following two categories:

1. Students who reported not going hungry: The responses considered for this category were never and rarely.
2. Students who reported going hungry: The responses considered for this category were sometimes, most of the time, and always.

To evaluate data on the prevalence of hunger obtained by asking a single question, a multiple correspondence analysis was performed to assess the relationship between the major variable of interest and consumption dimension of FI and income proxy. To determine the consumption dimension of FI, the frequency of fruit consumption measure was used, never, 1–3 days in a week, and $\geq$ 4 days in a week. To measure income proxy, the following variables were assessed:

1. Maternal education, divided into five categories, incomplete primary education (<7 years of education), primary education (8 years of education), secondary education (9 to 11 years of education), higher level of education (>11 years of education), and unknown (due to high level of missing data [24.7%], this category was retained).
2. Internet access at home (yes or no).

### Independent Variables

To evaluate the relationship between self-perceived hunger and sociodemographic factors, the independent variables were grouped into the following categories:

1. Geographic region of Brazil: North, Northeast, Southeast, South, and Midwest.
2. Place of residence: capital or non-capital cities.
3. School area: urban or rural.
4. School administrative status: public or private.
5. Sex: male or female.
6. Race/skin colour: White, Black, Asian, mixed race, or indigenous.
7. Age: 11–13, 14–15, or 16–19 years.
8. Lives with their mother: yes or no.
9. Lives with their father: yes or no.
10. Household size: up to four people, or five or more people.
11. Have any paid work: yes or no.
12. Plans after completing the ninth grade: keep studying fulltime, work fulltime, work and study, others, or do not know.


**FIGURE 1** | Flowchart of students enrolment in the National Survey of school Health (PeNSE), Brazil, 2015.

## Analysis

Data were analysed using descriptive statistics, and Pearson's chi-square test of association was performed to determine the outcome of self-perceived hunger. A $p$-value of $< 0.01$ was considered significant.

The correlation between hunger (outcome) and sociodemographic characteristics was tested using univariate and multiple Poisson regression models with robust variance, which estimated the crude and adjusted prevalence ratios and their respective 95% confidence intervals (CIs). Statistically significant variables ($p < 0.01$) and those with changes in prevalence ratios of at least 10% were included in the final model. Analyses were performed based on the design of the sample using the statistic package Stata 14.0.

## RESULTS

Among the 102,072 students interviewed, 101,888 (99.82%) answered the question on how often during the last 30 days they had gone hungry (**Figure 1**). Among them, 11.18% ($n = 11,740$) confirmed that they had experienced hunger.

The multiple correspondence analysis showed that self-perceived hunger was related to poor access to food and lower income levels. The first dimension explained 80% of the variation. We interpreted this dimension to confirm that self-perception of hunger can be used to assess hunger among students (**Figure 2**).

The following variables were significantly related to the hunger status of the students: geographic region, school administrative status, sex, race/skin colour, age, household size, and plans after completing the ninth grade (**Table 1**).

The adjusted models showed that adolescents studying in private schools had a lower prevalence ratio (PR) of hunger (PR:0.985; 95% CI:0.977–0.994), while those from the North (PR:0.983;95% CI:0.974–0.992) and Midwest regions (PR:0.994; 95% CI.983–1.005), women (PR: 1.024; 95% CI: 1.018–1.031), Black (PR: 1.019;95% CI: 1.008–1.031) and indigenous (PR: 1.023;95% CI: 1.005–1.042) adolescent students living in households with more than five people (PR: 1.02; 95% CI: 1.013–1.028), adolescents not living with their father (PR: 1.019; 95% CI: 1.011–1.025), and adolescents planning to work (PR: 1.036; 95% CI: 1.019–1.052) or knowing their plans after completing primary school (PR: 1.025; 95% CI: 1.011–1.04) had a higher prevalence of hunger (**Table 2**).

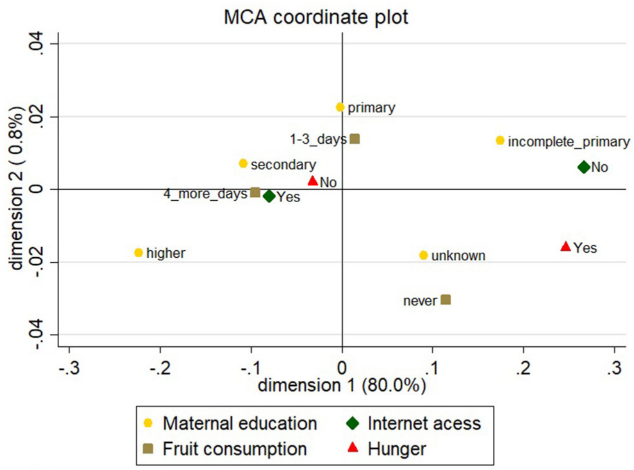


**FIGURE 2** | Multiple Correspondence plot.

## DISCUSSION

The results of this study demonstrate that reports of hunger among Brazilian adolescents have exceeded 10%, which is directly related to sociodemographic factors, allowing to determine the most vulnerable populations. To the best of our knowledge, this is the first study using a representative national sample to explore the reports of hunger among this population. The results of this study are consistent with those of previous national research conducted in 2017 and 2018, showing that 7.3% of the population aged 5–17 years experienced hunger (10).

The self-perception of adolescents used to report hunger in this study was shown to be related to measures that represent the scarcity of resources. The variables selected for correspondence analysis concur with those used in other studies (19). FI-driven food changes are accompanied by decreases in the variety of food intake, which can be expressed by consumption of fruits (20, 21). Maternal education has stronger effects than paternal education at the level of income and the cognitive performance of children, and is associated with lower prevalence of childhood undernutrition (22, 23). It is well established that internet access is correlated with income (24) and educational achievement (25).

Therefore, demonstrating the reliability of using a population survey database with a representative sample is essential to assess

**TABLE 1 |** Sociodemographic characteristics according to hunger status, Brazil, 2015.

| Variable | Student status | | | |
|---|---|---|---|---|
| | Not reported hungry | | Reported hungry | |
| | *n* | (%) | *n* | (%) |
| **Region*** | | | | |
| North | 20,504 | 86.40 | 3,393 | 13.60 |
| Northeast | 32,312 | 88.45 | 3,961 | 11.55 |
| Southeast | 16,124 | 89.67 | 161 | 10.33 |
| South | 8,753 | 88.95 | 1,075 | 11.05 |
| Midwest | 12,455 | 88.17 | 1,701 | 11.83 |
| **Place of residence** | | | | |
| Capital city | 45,301 | 88.84 | 5,795 | 11.16 |
| Non-capital city | 44,847 | 88.81 | 5,945 | 11.19 |
| **School area** | | | | |
| Urban | 83,274 | 89.00 | 10,356 | 11.00 |
| Rural | 6,874 | 86.89 | 1,384 | 13.11 |
| **School type** | | | | |
| Public | 71,107 | 88.35 | 9,877 | 11.65 |
| Private | 19,041 | 91.58 | 1,863 | 8.42 |
| **Gender*** | | | | |
| Male | 43,945 | 89.92 | 5,220 | 10.08 |
| Female | 46,203 | 87.78 | 6,520 | 12.22 |
| **Race/Colour*** | | | | |
| White | 30,384 | 89.91 | 3,343 | 10.09 |
| Black | 11,126 | 86.68 | 1,683 | 13.32 |
| Yellow | 4,014 | 87.77 | 560 | 12.23 |
| Brown | 41,382 | 88.86 | 5,479 | 11.14 |
| Indigenous | 3,149 | 86.45 | 666 | 13.55 |
| **Age (years)*** | | | | |
| ≤13 | 15,692 | 90.70 | 1,543 | 9.30 |
| 14–15 | 64,277 | 89.26 | 8,075 | 10.74 |
| ≥16 | 10,179 | 82.83 | 2,122 | 17.17 |
| **Live with the mother** | | | | |
| Yes | 80,115 | 88.95 | 10,212 | 11.05 |
| No | 10,005 | 87.69 | 1,522 | 12.31 |
| **Live with the father*** | | | | |
| Yes | 56,653 | 89.71 | 6,860 | 10.29 |
| No | 33,421 | 87.27 | 4,864 | 12.73 |
| **Household size*** | | | | |
| Up to 4 people | 52,840 | 89.97 | 5,876 | 10.03 |
| 5 or more people | 37,292 | 87.21 | 5,863 | 12.79 |
| **Have any paid work** | | | | |
| Yes | 11,118 | 88.09 | 1,716 | 11.91 |
| No | 78,979 | 88.94 | 10,019 | 11.06 |
| **Plan to do after completing 9th grade*** | | | | |
| Study full-time | 11,971 | 89.81 | 1,467 | 10.19 |
| Work full-time | 5,844 | 85.35 | 972 | 14.65 |
| Work and study | 60,972 | 89.43 | 7,458 | 10.57 |
| Other and don't know | 11,296 | 86.47 | 1,838 | 13.53 |

**p* < 0.01, chi-square test.*

**TABLE 2 |** Results of Poisson regression model analysis of the association between self-perceived hunger and sociodemographic variables, Brazil, 2015.

| | Crude Prevalence Ratio | CI 95% | | Adjusted Prevalence Ratio* | CI 95% | |
|---|---|---|---|---|---|---|
| **Region** | | | | | | |
| North | 1 | | | 1 | | |
| Northeast | 0.982* | 0.972 | 0.992 | 0.983* | 0.974 | 0.992 |
| Southeast | 0.971* | 0.961 | 0.982 | 0.982* | 0.971 | 0.992 |
| South | 0.978* | 0.966 | 0.989 | 0.989 | 0.977 | 1.001 |
| Midwest | 0.984 | 0.973 | 0.996 | 0.994 | 0.983 | 1.005 |
| **Gender** | | | | | | |
| Male | 1 | | | 1 | | |
| Female | 1.019* | 1.013 | 1.026 | 1.024* | 1.018 | 1.031 |
| **School type** | | | | | | |
| Public | 1 | | | 1 | | |
| Private | 0.971* | 0.963 | 0.979 | 0.985* | 0.977 | 0.994 |
| **Race/Colour** | | | | | | |
| White | 1 | | | 1 | | |
| Black | 1.029* | 1.018 | 1.041 | 1.019* | 1.008 | 1.031 |
| Yellow | 1.019 | 1.000 | 1.039 | 1.012 | 0.994 | 1.031 |
| Brown | 1.009 | 1.000 | 1.018 | 1.001 | 0.992 | 1.010 |
| Indigenous | 1.031* | 1.013 | 1.050 | 1.023* | 1.005 | 1.042 |
| **Age (years)** | | | | | | |
| ≤13 | 1 | | | 1 | | |
| 14–15 | 1.013 | 1.004 | 1.022 | 1.010 | 1.003 | 1.021 |
| ≥16 | 1.072* | 1.057 | 1.087 | 1.061* | 1.047 | 1.075 |
| **Household size** | | | | | | |
| Up to 4 people | 1 | | | 1 | | |
| 5 or more people | 1.025* | 1.018 | 1.033 | 1.020* | 1.013 | 1.028 |
| **Live with the father** | | | | | | |
| Yes | 1 | | | 1 | | |
| No | 1.022* | 1.015 | 1.029 | 1.019* | 1.011 | 1.025 |
| **Plan to do after completing 9th grade** | | | | | | |
| Study full-time | 1 | | | 1 | | |
| Work full-time | 1.041* | 1.024 | 1.057 | 1.036* | 1.019 | 1.052 |
| Work and study | 1.003 | 0.993 | 1.014 | 0.999 | 0.989 | 1.009 |
| Other and don't know | 1.030* | 1.016 | 1.045 | 1.025* | 1.011 | 1.040 |

**Adjusted according to other sociodemographic variables: sex, school type, race, have any paid work, region, age, household size, living with the father, and plans after completing the ninth grade. *p < 0.01.*

the prevalence of hunger in a specific population (adolescents). Additionally, the PeNSE is performed at regular intervals (3–4 years); thus, it is possible to determine the status of hunger in this population and, in the future, to access the impact of the pandemic, providing important indicators of the prevalence of hunger and FI in this population.

The data from the PeNSE indicated that reports of hunger were more frequent among public school students than among private school students, similar to the findings of the Brazilian Food Insecurity Scale for public and private school adolescent students in Brazilian capitals (1), reinforcing the importance of the National School Feeding Program (Programa Nacional de Alimentação Escolar-PNAE), which may offer principal daily free meals to many students. However, limited studies have examined the factors associated with the consumption of food offered in school feeding programs. Locatelli et al. (26) observed that consumption of the school lunch offered by the PNAE is influenced by the same determinants, which were also found in this study.

The PNAE is currently designed to contribute to biopsychosocial development, growth, learning, school performance, and formation of healthy food habits through education, food and nutrition activities, and provision of meals, which meet the school nutritional requirements during the scholastic year. Moreover, one of the guidelines of the program is to ensure food and nutrition security (27). Because of the importance of this policy in ensuring food and nutrition security, school meals should be provided during school holidays and vacations, especially for the most vulnerable populations (28).

In terms of race and skin colour, black and indigenous students were the most vulnerable to hunger. A similar result was found in other studies that showed significantly higher prevalence of hunger in households with at least one non-White adolescent (29, 30). The prevalence of hunger was also higher among adolescents who did not live with their fathers, emphasising how difficult it is for women to ensure FNS in their households. Some studies have reported a higher prevalence of moderate-to-severe hunger in households where the head of the family was a woman (31, 32). The PeNSE data also indicated a positive relationship between hunger and female students, demonstrating the greater vulnerability of women even before reaching adulthood, and reinforcing the importance of policies dedicated to them.

Students residing in the North and Midwest regions of the country had the highest prevalence of self-reported hunger. However, in the literature, studies on FI in Brazilian regions have found the highest prevalence of hunger (moderate and severe FI) in the North and Northeast regions because of socioeconomic inequalities that have existed in these regions for years and the absence of public policies that guarantee basic rights to health, food, education, and sanitation (10, 32). A possible explanation for the result in the Midwest region was that the proportion of indigenous students (3.88%) in this area was greater than that in the South (2.26%) and Southeast regions (2.86%), based on the sample evaluated.

Existing evidence shows that the prevalence of hunger is higher among rural households (32, 33); however, the PeNSE only considered the area in which schools were located. Therefore, students living in rural areas may have access to schools in urban areas due to the availability of public-school transportations.

Unlike most studies, this study reflects hunger at the individual level and not at the household level; therefore, we could not find a relationship between FNS and the age of adolescents in the literature. However, the prevalence of hunger was greater among older ninth-grade students (aged ≥ 16 years). This fact may be justified by several vulnerability factors that are associated with hunger, since most students in the ninth grade are aged between 14 and 15 years.

In addition to the relationship between inequality and FNS, which has a direct association with access to food (29), some studies have found the same between FNS and poor nutrition. In Brazil, higher prevalence of underweight has been found among underprivileged children and adolescents. Existing data have also indicated that Black and mixed-race women with low income and low levels of education have higher prevalence of obesity and low stature than more privileged women. In addition, Black and mixed-race individuals are more likely to have low-income status. In contrast, higher-income White men with higher levels of education are more likely to be overweight and obese (34). These data reinforce the need for public policies to promote social equality.

This study has some limitations. PeNSE only contains a single question related to FI and hunger; it asks the students about the frequency of going hungry due to not having sufficient food at home over the past 30 days. This method of assessing hunger could explain the prevalence ratios and CI close to 1, which were significant in multiple analysis due to the large sample size. Although the PeNSE question is insufficient to determine household FI, affirmative responses should be read at the same time as evidence of FI. Additionally, a Brazilian study demonstrated the validity of two-item FI screening to identify families at risk of FI and used a question very similar to that used in this study: "During the past 3 months, has a minor in your household gone hungry and not eaten because there was not enough money to buy food?" (35).

To address the issue of hunger, it is important to evaluate the food system based on five pillars: availability, accessibility, utilisation, stability, and sustainability (36, 37). Hunger is a complex social phenomenon that cannot be hidden through various euphemisms, such as "severe FI" or "very low food security," and food security will not be achieved using simple methods, such as genomics-based breeding of crops (38).

Childhood and adolescence are directly affected by the impact of austerity policies. Considering the current political situation in Brazil and the dismantling of public policies that fight against hunger and guarantee food and nutrition security, we would like to emphasise the importance of the PNAE and the need for investment in policies that further equality among sexes, races, and permanence in school given the direct relationship between these determinants and reports of hunger.

The COVID-19 pandemic has further aggravated the economic, educational, and public health crisis in Brazil, as the country was one of the most severely affected countries (39) because of the negligence of the federal government. Moreover, public schools remained closed in 2020 and part of 2021, which may have exacerbated hunger and social inequality among students.

## FUTURE CHALLENGES

The results of this study address the self-perception of hunger among Brazilian adolescents. There is a paucity of studies investigating the prevalence of FI and hunger in specific populations, such as adolescents, since most studies have assessed FI in households. Regarding existing policies to help end hunger among adolescents, it is very important to understand the associated factors. Strengthening school feeding programs is essential to tackle FI among students, and the implementation of FI screening in this population is important to identify schools with higher proportions of adolescents experiencing hunger, seeking to increase equity in the PNAE.

Future research can validate this simple tool to evaluate hunger in this specific context, which can be crucial in improving hunger and FI screening. The impact of the COVID-19 pandemic on hunger among students should be further investigated, especially with the closing of schools and reduction of access to food offered in this environment.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analysed in this study. This data can be found here: https://www.ibge.gov.br/estatisticas/sociais/educacao/9134-pesquisa-nacional-de-saude-do-escolar.html.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Brazilian National Commission of Research Ethics–Conep. Written informed consent to participate in this study was provided by the participant's legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

DB conceptualised the idea. AA and DB performed the analyses. AA and JR wrote the first draught of the manuscript. HG and DB extensively and critically reviewed the manuscript. All the authors have approved the submission of the manuscript.

## REFERENCES


1. Coelho SE dos AC, Vianna RP de T, Segall-Correa AM, Peres-Escamilla R, Gubert MB. Insegurança alimentar entre adolescentes brasileiros: um estudo de validação da escala curta de insegurança alimentar. *Rev Nutr*.(2015) 28:385–95. doi: 10.1590/1415-52732015000400005
2. Pérez-Escamilla R, Vianna RP de T. Food insecurity and the behavioral and intellectual development of children: a review of the evidence. *J Appl Res Child*. (2012) 3:93. Available online at: https://digitalcommons.library.tmc.edu/childrenatrisk/vol3/iss1/9/
3. da Veiga GV, Sichieri R. Correlation in food intake between parents and adolescents depends on socioeconomic level. *Nutrition Res*. (2006) 26:517–23. doi: 10.1016/j.nutres.2006.09.002
4. de Sousa JG, Lima LR, Fernandes CRS, dos Santos, GM. Atividade física e hábitos alimentares de adolescentes escolares: Pesquisa Nacional de Saúde do Escolar (PENSE), 2015. *Rev Brasi Nutr Esportiva*. (2019) 13:87–93.
5. Salvador CCZ, Kitoko PM, Gambardella AMD. Nutritional status of children and adolescents: factors associated to overweight and fat accumulation. *J Hum Growth Dev*. (2014) 24:313. doi: 10.7322/jhdg.88969
6. Rodrigues RDRM, de Souza BDSN, Cunha DB, Estima CDCP, Sichieri R, Yokoo EM. Associação entre tempo de exposição à tela e variação de ingestão alimentar entre adolescentes em idade escolar de Niterói/Rio de Janeiro, Brasil. *CadernosSaúdeColetiva*. (2020) 28:24–33.
7. Sarlio-Lahteenkorva S, Lahelma E. The association of body mass index with social and economic disadvantage in women and men. *Int J Epidemiol*. (1999) 28:445–9. doi: 10.1093/ije/28.3.445
8. Committee on World Food Security (CFS). *Coming to Terms With Terminology: Food Security Nutrition Security Food Security and Nutrition Food and Nutrition Security*. Rome: FAO (2012). Available online at: http://www.fao.org/3/MD776E/MD776E.pdf
9. Peng W, Dernini S, Berry EM. (2018). Coping with food insecurity using the sociotype ecological framework. *Front Nutr*. (2018) 5:107. doi: 10.3389/fnut.2018.00107
10. Instituto Brasileiro de Geografia e Estatística (IBGE). *Pesquisa de Orçamentos Familiares 2017-2018: análise da segurança alimentar no Brasil*. Brasil (2020). Available online at: https://biblioteca.ibge.gov.br/visualizacao/livros/liv101749.pdf
11. Rede Brasileira de. Pesquisa em Soberania e Segurança Alimentar (Rede PENSSAN). *VIGISAN: Inquérito Nacional sobre Insegurança Alimentar no Contexto da Pandemia da COVID-19 no Brasil Rio de Janeiro*. (2021). Available online at: http://olheparaafome.com.br/ (accessed August, 2021)
12. Kac G, Velásquez-Melendez G, Schlüssel MM, Segall-Côrrea AM, Silva AA, Pérez-Escamilla R. Severe food insecurity is associated with obesity among Brazilian adolescent females. *Public Health Nutr*. (2012) 15:1854–60. doi: 10.1017/S1368980011003582
13. Tamiru D, Argaw A, Gerbaba M, Ayana G, Nigussie A, Belachew T. Household food insecurity and its association with school absenteeism among primary school adolescents in Jimma zone, Ethiopia. *BMC Public Health*. (2016) 16:802. doi: 10.1186/s12889-016-3479-x
14. Bezerra TA, Olinda RA de, Pedraza DF. Insegurança alimentar no Brasil segundo diferentes cenários sociodemográficos. *CienSaude Colet*. (2017) 22:637–51. doi: 10.1590/1413-81232017222.19952015
15. Cotta RMM, Machado JC. Programa Bolsa Família e segurança alimentar e nutricional no Brasil: revisão crítica da literatur. *Rev PanamSalud Publica*. (2013) 33:54–60. doi: 10.1590/S1020-49892013000100008
16. Galesi LF, Quesada KR, Oliveira MRM de. Indicadores de Segurança Alimentar e Nutricional. *Rev Simbio-Logias*. (2009) 2:221–30. Available online at: https://www.ibb.unesp.br/#!/ensino/departamentos/educacao/revista-simbio-logias/
17. Santos TG dos, Silveira JAC da, Longo-Silva G, Ramires EKNM, Menezes RCE de. Tendência e fatores associados à insegurança alimentar no Brasil: pesquisa nacional por amostra de domicílios 2004. *Cad Saude Publica*. (2018) 34:4. doi: 10.1590/0102-311x00066917
18. Instituto Brasileiro de Geografia e Estatística (IBGE). *Pesquisa nacional de saúde do escolar : 2015*. Brasil (2016). Available online at: https://biblioteca.ibge.gov.br/visualizacao/livros/liv97870.pdf
19. Dhurandhar EJ. 2016. The food-insecurity obesity paradox: a resource scarcity hypothesis. *Physiol Behav*. (2016) 162:88–92. doi: 10.1016/j.physbeh.2016.04.025
20. Turnbull O, Homer M, Ensaff H. Food insecurity: its prevalence and relationship to fruit and vegetable consumption. *J Hum Nutr Diet*. (2021) 34:849–57. doi: 10.1111/jhn.12866
21. de Araújo ML, de Deus Mendonça R, Lopes Filho JD, Lopes ACS. Association between food insecurity and food intake. *Nutrition*. (2018) 54:54–9. doi: 10.1016/j.nut.2018.02.023
22. Chevalier A, Harmon C, O'Sullivan V, Walker I. The impact of parental income and education on the schooling of their children. *IZA J Labor Econ*. (2013) 2:1–22. doi: 10.1186/2193-8997-2-8
23. Carneiro P, Meghir C, Parey M. Maternal education, home environments, and the development of children and adolescents. *J Eur Econ Assoc*. (2013) 11:123–60. doi: 10.1111/j.1542-4774.2012.01096.x

24. Hilbert M. When is cheap, cheap enough to bridge the digital divide? modeling income related structural challenges of technology diffusion in Latin America. *World Dev*. (2010) 38:756–70. doi: 10.1016/j.worlddev.2009.11.019
25. Wainer J, Vieira P, Melguizo T. The association between having access to computers and Internet and educational achievement for primary students in Brazil. *Comput Educ*. (2015) 80:68–76. doi: 10.1016/j.compedu.2014.08.007
26. Locatelli NT, Canella DS, Bandoni DH. Fatores associados ao consumo da alimentação escolar por adolescentes no Brasil: resultados da PeNSE 2012. *Cad Saude Publica*. (2017) 33:4. doi: 10.1590/0102-311x00183615
27. Brasil. Lei no 11.947. *Presidência da República - Casa Civil. Dispõe sobre o atendimento da alimentação escolar*. Brasil (2009). Available online at: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/lei/l11947.htm
28. Amorim ALB de, Ribeiro Junior JRS, Bandoni DH. Programa Nacional de Alimentação Escolar: estratégias para enfrentar a insegurança alimentar durante e após a COVID-19. *Rev AdmPública*. (2020) 54:1134–45. doi: 10.1590/0034-761220200349
29. Panigassi G, Segall-Corrêa AM, Marin-León L, Pérez-Escamilla R, Sampaio M de FA, Maranha LK. Insegurança alimentar como indicador de iniqüidade: análise de inquérito populacional. *Cad Saude Publica*. (2008) 24:2376–84. doi: 10.1590/S0102-311X2008001000018
30. Guerra LD da S, Espinosa MM, Bezerra ACD, Guimarães LV, Lima-Lopes MA. Insegurança alimentar em domicílios com adolescentes da amazônia legal Brasileira: prevalência e fatores associados. *CadSaude Publica*. (2013) 29:335–48. doi: 10.1590/S0102-311X2013000200020
31. Bittencourt L de S, Santos SMC. dos, Pinto E de J, Aliaga MA, Ribeiro-Silva R de C. Factors associated with food insecurity in households of public school students of Salvador city, Bahia, Brazil. *J Heal PopulNutr*. (2013) 31:471–9. doi: 10.3329/jhpn.v31i4.20002
32. Instituto Brasileiro de. Geografia e Estatística IBGE. *Pesquisa Nacional por Amostra de Domíclios - Segurança Alimentar Brasil*. (2014). Available online at: https://biblioteca.ibge.gov.br/visualizacao/livros/liv91984.pdf
33. Morais D de C, Dutra LV, Franceschini S. do CC, Priore SE. Insegurança alimentar e indicadores antropométricos, dietéticos e sociais em estudos brasileiros: uma revisão sistemática. *CienSaude Colet*. (2014) 19:1475–88. doi: 10.1590/1413-81232014195.13012013
34. Canella DS, Duran AC, Claro RM. Malnutrition in all its forms and social inequalities in Brazil. *Public Health Nutr*. (2020) 23:s29–38. doi: 10.1017/S136898001900274X
35. Poblacion A, Segall-Corrêa AM, Cook J, Taddei JA de AC. Validity of a 2-item screening tool to identify families at risk for food insecurity in Brazil. *Cad Saude Publica*. (2021) 37:e00132320. doi: 10.1590/0102-311x00132320
36. Berry EM, Dernini S, Burlingame B, Meybeck A, Conforti P. Food security and sustainability: can one exist without the other? *Public Health Nutr*. (2015) 18:2293–302. doi: 10.1017/S136898001500021X
37. Smetana SM, Bornkessel S, Heinz V. A path from sustainable nutrition to nutritional sustainability of complex food systems. *Front Nutr*. (2019) 6:39. doi: 10.3389/fnut.2019.00039
38. Mir RR, Kumar A, Pandey MK, Isobe SN. Achieving nutritional security and food safety through genomics-based breeding of crops. *Front Nutr*. (2021) 8:638845. doi: 10.3389/fnut.2021.638845
39. Jaime PC. Pandemia de COVID19: implicações para (in)segurança alimentar e nutricional. *CienSaude Colet*. (2020) 25:2504. doi: 10.1590/1413-81232020257.12852020

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

METHODS
published: 25 November 2021
doi: 10.3389/fnut.2021.777246

# Nutrition Status of Children, Teenagers, and Adults From National Health and Nutrition Surveys in Mexico From 2006 to 2020

*Teresa Shamah-Levy* [1*†‡], *Lucia Cuevas-Nasu* [1†‡], *Martín Romero-Martínez* [1†‡], *Ignacio Méndez Gómez-Humaran* [2†‡], *Marco Antonio Ávila-Arcos* [1†‡] and *Juan A. Rivera* [3†‡]

[1] *Evaluation and Survey Center, Center for Evaluation and Survey Research, National Institute of Public Health, Cuernavaca, Mexico,* [2] *Center for Mathematics Research, Unit Aguascalientes, Aguascalientes, Mexico,* [3] *National Institute of Public Health, Cuernavaca, Mexico*

**Background:** Population-level health and nutrition surveys provide critical anthropometric data used to monitor trends of the prevalence of under nutrition and overweight in children under 5 years old, and overweight and obesity in the population over 5 years of age.

**Objective:** Analyze the children malnutrition and overweight and obesity in children, teenagers and adults through the National Health and Nutrition Surveys information available from public databases.

**Materials and Methods:** Comparable anthropometric data was gathered by five Mexican National Health and Nutrition Surveys (in Spanish, ENSANUT). In pre-school-age children, under nutrition status was identified through underweight (Z-score below −2 in weight-for-age), stunting (chronic malnutrition) (Z-score below −2 for length/height-for-age), or wasting (Z-score below −2, for weight-for-length/height); overweight status was defined as a body mass index (BMI, kg/m$^2$) for age over +2. For school-age children and adolescents, a Z-score BMI between +1 and +2 deviations was defined as overweight, and between +2 and +5.5 as obesity. In adults (≥20 years of age), overweight status was classified as a BMI between 25.0 and 29.9, and obesity as ≥30.

**Results:** The anthropometric data presented derives from the databases of five survey years of the Mexican National Health and Nutrition Survey: 2006, 2012, 2016, 2018, and 2020. They include a total of 210,915 subjects with complete anthropometric data (weight, length/height) distributed on five survey moments; subjects were categorized by age group: pre-school-age children (*n* = 25,968), school-age children (*n* = 42,255), adolescents (*n* = 39,275), and adults (*n* = 103,417). Prevalence of malnutrition by indicator was calculated: in pre-school-age children: low height- and weight-for-age, low

OPEN ACCESS

**Edited by:**
*Alessandra Durazzo, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Shaonong Dang, Xi'an Jiaotong University, China*
*Mostafa Waly, Sultan Qaboos University, Oman*

***Correspondence:**
*Teresa Shamah-Levy*
*tshamah@insp.mx*

† *These authors have contributed equally to this work*

**‡ORCID:**
*Teresa Shamah-Levy*
*orcid.org/0000-0002-1357-720X*
*Lucia Cuevas-Nasu*
*orcid.org/0000-0001-9890-3563*
*Martín Romero-Martínez*
*orcid.org/0000-0002-2413-2089*
*Ignacio Méndez Gómez-Humaran*
*orcid.org/0000-0002-3908-0867*
*Marco Antonio Ávila-Arcos*
*orcid.org/0000-0002-5698-8666*
*Juan A. Rivera*
*orcid.org/0000-0003-2586-4908*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** *15 September 2021*
**Accepted:** *25 October 2021*
**Published:** *25 November 2021*

**Citation:**
*Shamah-Levy T, Cuevas-Nasu L, Romero-Martínez M, Gómez-Humaran IM, Ávila-Arcos MA and Rivera JA (2021) Nutrition Status of Children, Teenagers, and Adults From National Health and Nutrition Surveys in Mexico From 2006 to 2020. Front. Nutr. 8:777246. doi: 10.3389/fnut.2021.777246*

weight-for-height, and overweight; and in school-age children, adolescents, and adults, the indicators calculated were overweight and obesity.

**Conclusions:** Results demonstrate the importance of maintaining systematic, reliable, and timely national anthropometric data in the population, in order to detect and track trends and to form the basis of nutrition-related public policy.

**Keywords: anthropometry, weight, height, body mass index, databases, national surveys, Mexico**

## INTRODUCTION

Reducing the global burden of malnutrition is a fundamental component of the United Nations Sustainable Development Goals, which include monitoring indicators of under nutrition, overweight, and obesity.

Mexico has a history of over three decades conducting health and nutrition surveys as part of a nationwide monitoring system. The National Health Surveys System in Mexico includes a series of multi-thematic surveys on health and nutrition topics; surveys are probabilistic and nationally representative. The results of these surveys have been key for evaluating the performance of the Mexican health system, providing precise, detailed, and representative information about the status of health and nutrition of the population and for planning health and nutrition policies.

The first national health and nutrition surveys were conducted by the Secretariat of Health (SoH) or the National Public Health Institute (INSP) and the operatives of the Health surveys were different from those of the nutrition surveys which only included mothers and children (1, 2).

In 2006, the nutrition and health surveys were combined into a single Mexican National Health and Nutrition Survey (in Spanish, ENSANUT), which has since been applied four times, all of them conducted by INSP using similar methods and sampling procedures in order to allow comparison: 2006 (3), 2012 (4), 2016 at "*Medio Camino (MC)*" (halfway point) (5), and 2018–2019 (6). The surveys, with the exception with the one collected in 2016, were conducted every 6 years, at the end of the Federal Administrations.

Beginning in 2020, with the support of the Health Secretariat, ENSANUT is now a continuous survey, with data collection every year, so that within a 5-year time window a sample is created which is representative at the national level, urban/rural level, and most importantly, by state level. Notably, the 2020 ENSANUT obtained data aimed at updating the general landscape on the frequency, distribution, and trends of health and nutrition status indicators and their determinants, but in addition, it includes a specific COVID-19 component which measured SARS-CoV-2 antibody seroprevalence in a population subsample, which allows an estimation of the proportion of the population with possible exposure to the coronavirus infection. In addition, the information collected provides a closer look at experiences and changes in behavior, food security, diet, physical activity, and healthcare seeking behavior of the Mexican population stemming from the COVID19 pandemic and the measures adopted to contain it, including the way in which individuals have dealt with the confinement period and to what extent they have adopted mitigation measures.

In light of the poverty and inequality persistent in Mexico, one critical factor driving the creation of the Mexican nutrition surveys was the need to monitor the nutritional status of children under 5 years old. For three decades, these surveys have documented national trends in under nutrition for children under five—recognized since the 1980's as an immense public health challenge in the country—offering data disaggregated by age, sex, geographical region, and socioeconomic level. On the other hand, an epidemic of overweight and obesity in all age groups in Mexico has also been documented in the last three decades, in addition to other serious public health problems in the population (7–9). The ENSANUT 2021 is currently underway and will generate information on progress made and challenges which remain, as well as supporting the identification of health and nutrition priorities for the coming years and necessary strategies given the ongoing COVID-19 epidemic in Mexico.

The goal of this study was to review the anthropometric data available through national databases in Mexico, which are key to tracking nutritional status indicators including under nutrition, overweight and obesity in the population. These indicators may be used to inform and craft nutrition-specific public health policy based on transparent, reliable, updated, and timely information in Mexico.

## MATERIALS AND METHODS

The ENSANUT is a probabilistic household-level survey, which assigns to each Mexican household a known value $>0$ as probability of selection. The ENSANUT is a stratified cluster survey, with strata defined by population size of the locality of interest: rural (1–2,499 inhabitants), urban (2,500–99,999 inhabitants), or metropolitan ($\geq$100,000 inhabitants). Primary Sampling Units (PSU) of ENSANUT were geographical areas defined by INEGI (National Institute of Geography and Statistics): ENSANUT 2018 used clusters of the Master Sampling Frame of INEGI for surveys, and the rest of ENSANUT used the Basic Geographical Areas (AGEB). ENSANUT 2016 and ENSANUT 2020 were designed to make inferences on regions: ENSANUT 2020 (nine regions) and ENSANUT 2016 (four regions); and ENSANUT 2006, 2012, and 2018 were designed to make inferences on the States of México (32). As an example of parameter used to estimate the sample size, we present the sample size of ENSANUT 2018. For the planning of ENSANUT 2018, a deff $= 2.0$ was used, a value that was estimated from the experience in surveys carried out by INSP and INEGI. The sample size by state was calculated with the formula: $n = Z^2 \frac{p(1-p)}{\delta^2 RK}$ Deff.

where $n$ = sample size in households, p = proportion to estimate, Z = Quantile 97.5% of a unit normal distribution (Z = 1.96), δ = is the semi-amplitude of the confidence interval, R = expected response rate (85% in dwellings, 85% in adults and adolescents, 88% in schoolchildren, and 90% in pre-school), K = number of individuals expected to obtain from the interest group (pre-school K = 0.24, schoolchildren K = 0.26, adolescents K = 0.48, adults K = 0.99). Based on the assumed parameters, a sample size of 1,580 homes per state was chosen. The sample size will allow estimating prevalences of 10% with the following semi-amplitudes: 2.54% in adults, 3.57% in adolescents, 5.14% in schoolchildren, and 4.91% in preschool.

In general, the ENSANUT allocates sample sizes by State, seeking to ensure that all households have a similar probability of selection and selects clusters with probabilities proportional to their population.

Further details of sampling procedures and data results are discussed in previous articles and are publicly available through the website www.ensanut.insp.mx, which allows users to download data and survey instruments.

The ENSANUT disaggregates the population by four age groups: pre-school-age children (0–4 years), school-age children (5–9 years), adolescents (10–19 years), and adults (≥20 years). Selection procedures differ by age group. One adolescent is selected to represent the adolescents of a household. For pre-school-age and school-age children, the selection procedure has evolved over time; in 2006 one child between zero and 9 years of age was selected for each household; ENSANUT 2012 and 2016 selected both one pre-school-age and one school-age child per household, and since ENSANUT 2018 all pre-school-age children are included and only one school-age child is selected. From ENSANUT 2006 to 2018, only one adult per household was selected; however, in the 2020 survey one adult might be selected from each of the following age groups: 20–34, 35–49, and 50 years and above. Each person selected from a household was interviewed and his or her anthropometric measurements were taken by a group of trained specialists following a standardized protocol which guaranteed comparability between measurements of different surveys.

Anthropometric data was standardized in body weight and height measurements according to international standards (10, 11). In the case of the pre-school-age population, weight-for-height, and weight- and height-for-age indexes were constructed, which were then converted to Z-scores according to World Health Organization (WHO) standards (12). Participants were classified as under nutrition status was identified through underweight (Z-score below −2 in weight-for-age), stunting (chronic malnutrition) (Z-score below −2 for length/height-for-age), or wasting (Z-score below −2, for weight-for-length/height); overweight status was defined as a body mass index (BMI, kg/m$^2$) for age over +2. In the school-age and adolescent groups, BMI (kg/m$^2$) was calculated by age and sex and categorized according to WHO guidelines (13) as overweight when BMI Z-score was between +1 and +2, and obese when Z-score was between +2 and +5.5. For adults, WHO classification standards were also applied: overweight was considered as BMI of 25.0–29.9, and obesity as BMI ≥30 (14).

## Ethics Statement

ENSANUT protocols were approved by the committees of Ethics, Biosecurity and Research of the National Institute of Public Health of Mexico, and all participants signed an informed consent form. The data contained in the databases complies with the guidelines established by the Ethics Committee and national laws regarding confidentiality and transparency. Under no circumstances data that could reveal the identity of study participants is disclosed.

# RESULTS

## Relevance of Public Health Databases Use

Since 2006, anthropometric data has been one of the main outputs of the Mexican National Health and Nutrition Surveys, gathered across all survey years and for all age groups (pre-school-age to seniors) (**Table 1**). Related results have been specifically disaggregated, analyzed and reported within public reports, and have emerged as a topic of significant interest to the public health community, decision makers, and the general population. The corresponding datasets have provided information on the nutrition status of the population across survey years, and perhaps most importantly, through a transversal comparable design, have also revealed trends which have alerted about the increasing trends of overweight and obesity, which can be characterized as epidemics; at the same time it has informed about the decline of stunting in general and its persistence in some marginalized subgroups.

In regards to data quality, in addition to the standardization process required by personnel responsible of the anthropometric measurements, data was recorded using a software which required that at least two duplicate measurements be taken, allowing data collection to proceed only if the measurements differed by no more than 200 g for weight in children and 400 g in adults, and 4 mm for height for both age groups; this was intended to minimize errors related to recording, dictation, or measurement. Finally, the software also addressed sampling and data-integrity issues, allowing data collection only from those individuals previously selected in the sampling process and guaranteeing through unique identification codes that every anthropometric measurement link to its counterparts from the other survey sections (e.g., household characteristics, health indicators, or feeding practices).

Since all data obtained in the surveys by the National Institute of Public Health received public funding, datasets from all survey years may be downloaded freely from the ENSANUT website (https://ensanut.insp.mx/). For each survey year, a link reading "*Descargar bases*" (Download datasets) leads the user to data gathered, ordered by survey topic and population. Popular data processing systems such as SPSS, STATA, or the non-proprietary CSV format are the ideal options for dataset download, or alternatively, the paper version of the survey instrument in PDF and a catalog in Excel may be downloaded from the same

**TABLE 1 |** Anthropometric information available in databases in Mexico, by age group and survey year.

| SURVEY | Format(s) | Description | Measurement | Age groups | | | | | | | | | | Representativity | Extra variables/info |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Preschool 0–4 | | School 5–11 | | Teenagers 12–19 | | Adults ≥20 | | Total | | | |
| | | | | *n* | *N* | *n* | *N* | *n* | *N* | *n* | *N* | *n* | *N* | | |
| 2006 | SPSS (*.sav) | One single SPSS file including all the population groups for the survey | Weight<br>Height | 7,725<br>7,725 | 9,441,669<br>9,441,669 | 15,165<br>15,165 | 15,885,880<br>15,885,880 | 14,794<br>14,794 | 18,962,586<br>18,962,586 | 33,785<br>33,785 | 59,011,558<br>59,011,558 | 71,469<br>71,469 | 103,301,693<br>103,301,693 | Urban/Rural localities inside the four regions of the country | Includes waist circumference for adult population |
| 2012 | SPSS (*.sav), STATA (*.dta), CSV | One single SPSS file including all the population groups for the survey, additional file including survey dates and a processed version of the dataset that include variables with score-z and BMI computations | Weight<br>Height | 10,887<br>10,853 | 10,966,681<br>10,966,681 | 16,484<br>16,476 | 16,565,012<br>16,565,012 | 14,214<br>14,207 | 18,308,611<br>18,308,611 | 38,267<br>38,218 | 69,245,519<br>69,245,519 | 79,733<br>79,733 | 115,085,824<br>115,085,824 | Urban/ Rural inside any of the 32 states that conform the country | Includes blood pressure, calf, waist and hip circumference, toe to knee distance, and demi-span |
| 2016 | SPSS (*.sav), STATA (*.dta), CSV | This data corresponds to the mid-point 2016 survey, one file includes all the age-groups | Weight<br>Height | 2,028<br>2,020 | 11,104,791<br>11,031,753 | 2,391<br>2,391 | 11,602,871<br>11,602,871 | 3,431<br>3,430 | 23,045,978<br>22,997,366 | 8,435<br>8,413 | 70,888,336<br>70,760,153 | 16,285<br>16,254 | 116,641,975<br>116,392,142 | Urban/Rural localities inside the four regions of the country | Includes blood pressure, calf, arm, waist and hip circumference, toe to knee distance, and demi-span |
| 2018 | CSV | One single CSV file including all the population groups for the survey | Weight<br>Height | 3,776<br>3,998 | 9,736,969<br>9,737,540 | 6,268<br>6,268 | 11,004,874<br>11,004,874 | 5,672<br>5,672 | 22,904,289<br>22,904,289 | 13,053<br>13,012 | 61,793,694<br>61,587,914 | 28,769<br>28,950 | 105,439,826<br>105,234,617 | Urban/Rural inside some selected states of the country * | Includes blood pressure, calf, arm, waist and hip circumference, toe to knee distance, and demi-span, this is the only survey conducted along with the National Statistics and Informatics Institute (INEGI) |
| 2020 | SPSS (.sav), STATA (.dta), CSV | One single file including all the population groups for the survey | Weight<br>Height | 1,594<br>1,594 | 10,097,632<br>10,097,632 | 1,955<br>1,955 | 15,132,188<br>15,132,188 | 1,172<br>1,172 | 17,545,092<br>17,545,092 | 9,989<br>9,989 | 83,869,561<br>83,869,561 | 14,710<br>14,710 | 126,644,473<br>126,644,473 | Urban/ Rural inside any of the 32 states that conform the country | Includes blood pressure and a post-sampling generated variable to expand to the 2020 census reported population. This is the first edition from the "continuous" mode of the survey and was conducted during COVID-19 pandemic |

*Source: ENSANUT 2006, 2012, 2016, 2018–19, and 2020.*

**TABLE 2 |** The complete structure of one of the anthropometric tables (ENSANUT 2018).

**TABLE NAME: Antropometría**

| ID | Variable name | Question (English) | Type | Length | Valid codes | Value labels | |
|---|---|---|---|---|---|---|---|
| | | | | | | Numeric | Label |
| **1** | **UPM** | **Primary sampling unit** | **C** | **5** | **00001…03938** | | |
| **2** | **VIV_SEL** | **Selected residential house** | **C** | **2** | **01…25** | | |
| **3** | **HOGAR** | **Household number** | **C** | **1** | **1…4** | | |
| **4** | **NUMREN** | **Family member id** | **C** | **2** | **01…19** | | |
| 5 | PESO1_1 | Now I will weight (Name) Kg/g first Measurement | C | 7 | 002.500…222.220, 222.222,b | 222.222 | Not weighted |
| 6 | PESO1_2 | Now I will weight (Name) Kg/g second Measurement | C | 7 | 002.500… 222.220,b | | |
| 7 | P2 | Clothes | C | 1 | 1.0.0.4,9,b | 1 | Light |
| | | | | | | 2 | Thick |
| | | | | | | 3 | Without clothes |
| | | | | | | 4 | Not weighted |
| | | | | | | 9 | Don't know |
| 8 | P3 | Weight result | C | 1 | 1.0.0.3,9,b | 1 | Without problem |
| | | | | | | 2 | Physical problem |
| | | | | | | 3 | Did not cooperate |
| | | | | | | 9 | Don't know |
| 9 | TALLA4_1 | Now I will measure the height of (name) cm first measurement | C | 5 | 045.5…196.2,222.2,b | 222.2 | Not measured |
| 10 | TALLA4_2 | Now I will measure the height of (name) cm first measurement | C | 5 | 045.5…196.4,b | | |
| 11 | P5 | Height result | C | 1 | 1.0.0.3,9,b | 1 | Without problem |
| | | | | | | 2 | Physical problem |
| | | | | | | 3 | Did not cooperated |
| | | | | | | 9 | Don't know |
| 12 | P6 | Are you … | C | 1 | 1.0.0.4, 9,b | 1 | Pregnant? |
| | | | | | | 2 | Breastfeeding? |
| | | | | | | 3 | Pregnant and breastfeeding? |
| | | | | | | 4 | None of the previous |
| | | | | | | 9 | Don't know |
| 13 | P7_1 | How many months of pregnancy? | C | 2 | 01…09, 99,b | 99 | Not specified |
| 14 | CIRCUNFERENCIA8_1 | Now I will measure the waist of (name) first measurement | C | 5 | 011.4…187.8,222.2,b | 222.2 | Not measured |
| 15 | CIRCUNFERENCIA8_2 | Now I will measure the waist of (name) second measurement | C | 5 | 011.3…187.1,b | | |
| 16 | P9 | Result of waist measurement | C | 1 | 1.0.0.3,9,b | 1 | Without problem |
| | | | | | | 2 | Physical problem |
| | | | | | | 3 | Did not cooperate |
| | | | | | | 9 | Don't know |
| 17 | P10 | Did (name) lose weight recently (in the last 3 months)? | C | 1 | 1.0.0.3, 8, 9,b | 1 | No weight lost |
| | | | | | | 2 | Lost between 1 and 3 kg |
| | | | | | | 3 | Lost more than 3 kg |
| | | | | | | 8 | Not responding |
| | | | | | | 9 | Don't Know |

*(Continued)*

**TABLE 2 |** Continued

| TABLE NAME: Antropometría | | | | | | | |
|---|---|---|---|---|---|---|---|
| **ID** | **Variable name** | **Question (English)** | **Type** | **Length** | **Valid codes** | **Value labels** | |
| | | | | | | **Numeric** | **Label** |
| 18 | P11 | (Name) suffered any limb amputation? | C | 1 | 1,3.0.0.6,b | 1 | Yes, upper limbs and is able to walk |
| | | | | | | 2 | Yes, upper limbs and is unable to walks |
| | | | | | | 3 | Yes, lower limbs |
| | | | | | | 4 | No, and he/she can stand up |
| | | | | | | 5 | No and he/she can only be seated |
| | | | | | | 6 | No, and is bedridden |
| 19–25 | These are the same variables as P5–P11 but only for individuals over 60 years old, that can stand. | | | | | | |
| 26 | TALLAPIE17_1 | Now I will measure the distance between the heel and knee of (name), cm | C | 5 | 021.0…158.0,222.2,b | 222.2 | Not measured |
| 27 | P18 | Heel-Knee measurement result | C | 1 | 1.0.0.3,b | 1 | Without problem |
| | | | | | | 2 | Physical problem |
| | | | | | | 3 | Did not cooperate |
| 28 | CIRCPANTORRILLA19_1 | Now I will measure the thigh circumference of (Name) cm | C | 5 | 016.6…89.9,222.2,b | 222.2 | Not measured |
| 29 | P20 | Thigh circumference measurement result | C | 1 | 1.0.0.3,b | 1 | Without problem |
| | | | | | | 2 | Physical problem |
| | | | | | | 3 | Did not cooperate |
| 30–32 | These are the same variables as P14–P16 but only for individuals over 60 years old, that can stand. | | | | | | |
| 33 | HEMIENVERGADURA23_1 | Now I will measure the distance from the chest to the medium fingertip of (name) cm | C | 5 | 070.0…095.1,222.2,b | 222.2 | Not measured |
| 34 | P24 | Half arm-span measurement result | C | 1 | 1.0.0.3,b | 1 | Without problem |
| | | | | | | 2 | Physical problem |
| | | | | | | 3 | Did not cooperate |
| 35 | MEDIABRAZO25_1 | Now I will measure the mid-arm circumference of (name) cm | C | 5 | 022.3…037.8,222.2,b | 222.2 | Not measured |
| 36 | P26 | Mid-arm circumference measurement result | C | 1 | 1.0.0.3,b | 1 | Without problem |
| | | | | | | 2 | Physical problem |
| | | | | | | 3 | Did not cooperate |
| 37 | P27_1_1 | Now I will measure the blood pressure of (name) first measurement, systolic | C | 3 | 059…255,b | | |
| 38 | P27_1_2 | Now I will measure the blood pressure of (name) second measurement, systolic | C | 3 | 040…193,222,b | 222 | Not measured |
| 39 | P27_2_1 | Now I will measure the blood pressure of (name) first measurement, diastolic | C | 3 | 056…256,b | | |
| 40 | P27_2_2 | Now I will measure the blood pressure of (name) second measurement, diastolic | C | 3 | 000…191,b | | |
| 41 | P28 | Blood pressure, time of measurement (hours: min) | C | 5 | 01:00…23:59 | | |

*(Continued)*

**TABLE 2** | Continued

| **TABLE NAME: Antropometría** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **ID** | **Variable name** | **Question (English)** | **Type** | **Length** | **Valid codes** | **Value labels** | |
| | | | | | | **Numeric** | **Label** |
| 42 | P29 | Arm used to measure blood pressure | C | 1 | 1.0.0.3,b | 1 | Left arm |
| | | | | | | 2 | Right arm |
| | | | | | | 3 | Not measured |
| 43 | P30 | Blood pressure measurement result | C | 1 | 1.0.0.4,b | 1 | Without problem |
| | | | | | | 2 | Physical problem |
| | | | | | | 3 | Did not cooperate |
| | | | | | | 4 | Denial |
| **System Variables** | | | | | | | |
| 44 | EDAD | Age in years | C | 3 | 000…111 | 000 | Under 1 year old |
| 45 | EDAD_MESES | Age in months | C | 4 | 0…1337 | | |
| 46 | SEXO | Gender | C | 1 | 1,2 | 1 | Male |
| | | | | | | 2 | Female |
| 47 | ENT | State code | C | 2 | 01.0.0.32 | | |
| 48 | DOMINIO | Locality domain | C | 1 | 1,2 | 1 | Urban |
| | | | | | | 2 | Rural |
| 49 | ALTITUD | Altitude in meters over sea level | C | 4 | 0…3241 | | |
| 50 | REGION | Region of the country | C | 1 | 1…4 | 1 | North |
| | | | | | | 2 | Center |
| | | | | | | 3 | Mexico city |
| | | | | | | 4 | South |
| 51 | EST_DIS | Design strata | C | 3 | 001…301 | | |
| 52 | UPM_DIS | Design psu | C | 5 | 00001…03938 | | |
| 53 | ESTRATO | Sociodemographic strata | C | 1 | 1…4 | | |
| 54 | F_ANTROP | Original weight | C | 6 | 141…132198 | | |
| 55 | F_ANTROP_INSP | Calibrated weight | C | 6 | 0…132198 | | |
| 56 | GPO_INSP | Age group | C | 1 | 0…4 | 1 | Pre-school children |
| | | | | | | 2 | School children |
| | | | | | | 3 | Teenagers |
| | | | | | | 4 | Adults |
| 57 | DIAS | Age in days | C | 5 | 9…38542 | | |

website for each survey topic. SPSS and STATA files, in addition to the provided reference index, are labeled by variable and value, and missing values are declared. All data files include, in addition to information about the section, the key variables used at household and individual level in order to enable file merges with other ENSANUT datasets on topics such as disease, household characteristics, blood samples, healthcare service usage, and others. Another important feature is that the public datasets include geographic variables which allow all surveys to be identified by municipality (county); notably, although the sampling methodology identifies both households and street blocks, due to confidentiality laws the only data available to the public is at county level. Finally, all sampling variables are included in the datasets: design strata, PSU, and weights which allow the calculation of confidence intervals for and population parameters. Datasets also include auxiliary variables commonly used in anthropometric calculations or to filter observations, such as gender, age, birth, survey date, pregnancy, disability, and clothing characteristics in the case of pre-school-age children.

## Database Structure Description

In the **Table 2** the complete structure of one of the anthropometric tables (ENSANUT 2018) is presented as an example. The included data for every variable is: name, data type, width, and an English-translated version of the question, it is important to mention that this last field (in Spanish) corresponds exactly to the question in the PDF of the questionnaire as well as the data entry application. In the documentation, the range of valid and missing values are also included. In those variables that applies a translated list of value labels (codes) is also provided.

The ENSANUT data files always begin with the key variables (ID: 1–4) that uniquely identify every observation (households or individuals), this fields allow all the merge operations with other survey sections, following (ID: 5–16) the anthropometric variables can be found, including the repeated measurements and additional control variables like the clothes type, measurements result for every indicator, and pregnancy. The variables ID: 17 and 18 are used to register weight loss and limb amputations. The variables corresponding to 19–25 IDs are the same as 5–11 but adapted to individuals of 60 and more years old. The variables (ID: 26–36) are gathered to estimate the actual height using proximal indicators like heel-knee height or mid arm span for people that are unable to stand, here, some body composition variables like thigh or mid-arm circumferences are located. After this, the table contains the blood pressure variables (ID: 37–43) and finally, a group of variables entitled "System variables" (ID:

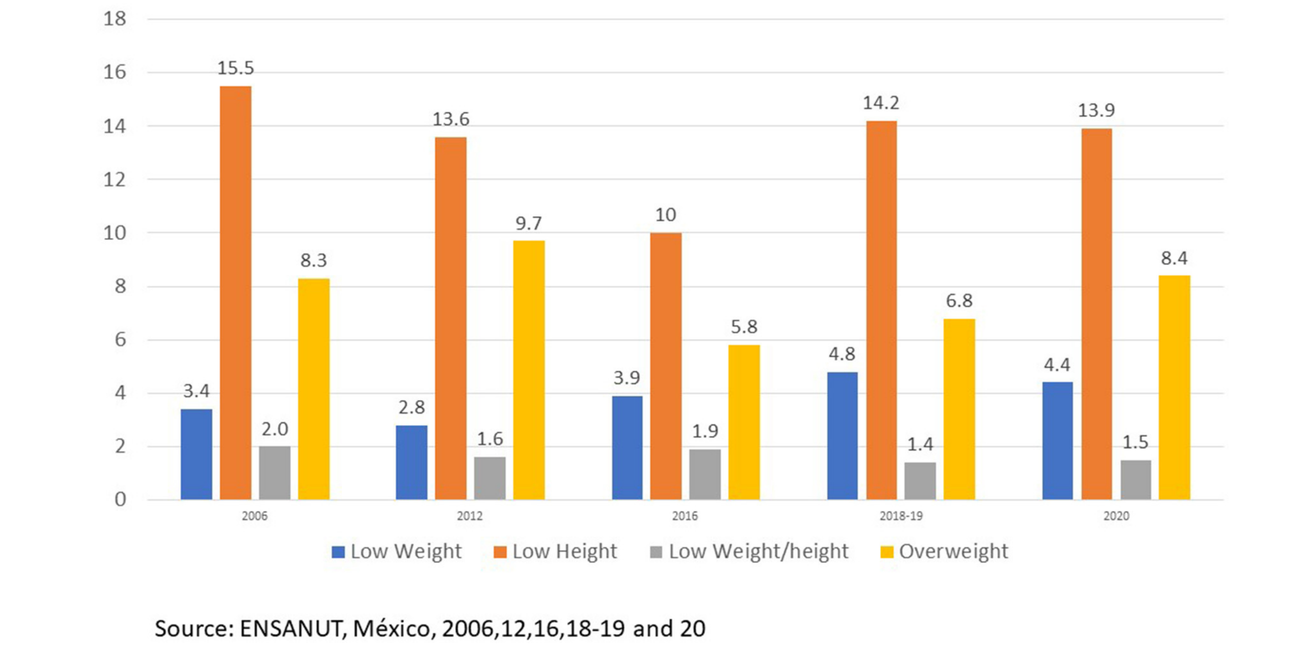


**FIGURE 1 |** Prevalences of nutrition status in children <5 years age by year survey in Mexico. *Source: ENSANUT 2006, 2012, 2016, 2018–19, and 2020*.

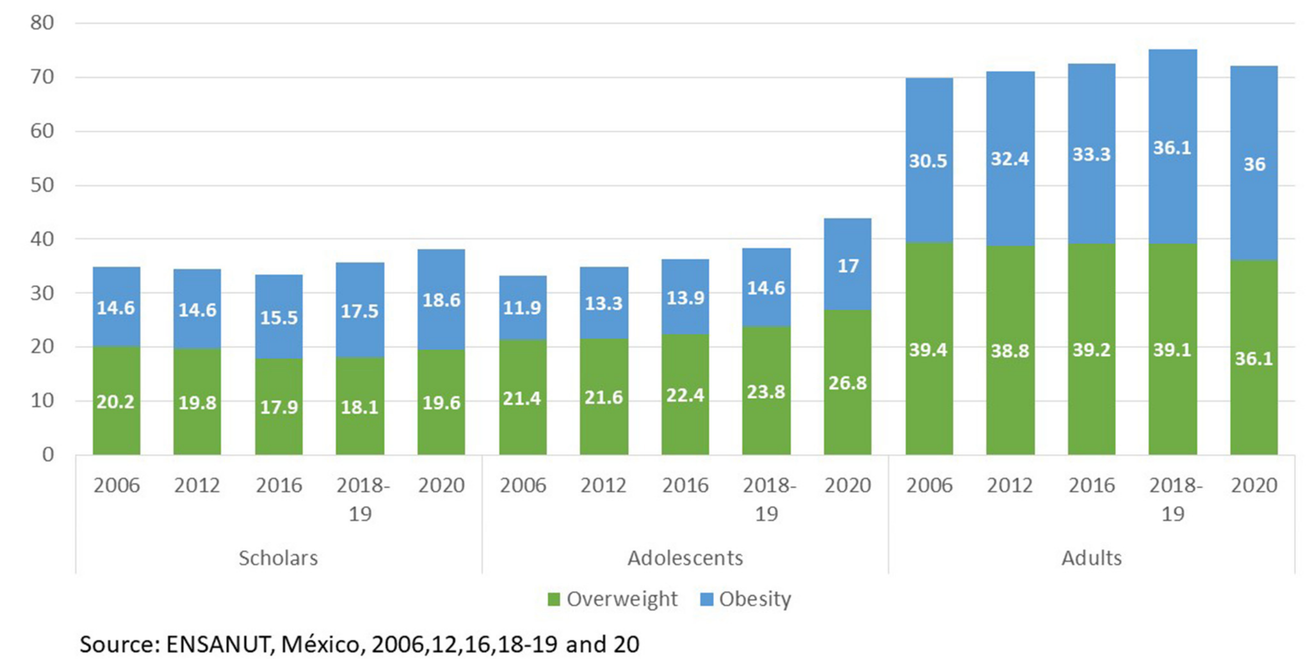


**FIGURE 2 |** Prevalences of overweight and obesity by age group and year survey in Mexico. *Source: ENSANUT 2006, 2012, 2016, 2018–19, and 2020*.

44–57) that include geographic variables (state code, altitude), analysis variables often used in the crosstabs creation of the surveys (gender, age groups, Country region, type of locality, etc.); in this section are also included, all the sampling related variables (strata, PSU, weights) used to estimate variances.

## Database Use

Prevalences of malnutrition: low weight, stunting, and wasting in children under 5 years old and overweight and obesity for all the selected population were obtained. **Figure 1** shows that the stunting prevalence in 2006 was 15.5% then decreased to 10% in 2016 and finally increased up to 13.9% in 2020. This, constitutes one of the main public health problems in population under 5 years old in Mexico (15, 16).

**Figure 2** shows the trends of overweight and obesity prevalences among the population groups. For school age and teenagers excessive weight is above 30% whereas in adult population this cipher exceeds 70%.

Back in 2006, with the results obtained through ENSANUT variables analysis. Mexico was declared at national alert due to the weight excess in the population. This triggered the development of National public policies aimed to immediately address this problem (7, 8).

## DISCUSSION

This study presents the anthropometric data available at a national level for inhabitants of Mexico. ENSANUT can explore information by characteristics such as area of residence (urban or rural), geographical region, state, age group, and sex. The findings highlight the availability of comparable, consistent, transparent, and high-quality nutrition data which can be used to support local-level analyses.

High-quality and up-to-date anthropometric data serves as the basis for the evaluation of combined nutritional status indexes of weight, age, and height, and helps to make the difference between high-quality surveys and those with low-quality data. In the context of surveys like the ENSANUT which are applied periodically, the validity, and robustness of the data collected allows estimation of the changes over time of the prevalence of nutritional status indicators of the Mexican population (see **Supplemental Material**) (17).

Previous studies have demonstrated that investments toward improving the quality of data collection procedures improves the quality of anthropometric data (18). However, variations in survey quality are also related to contextual factors which present challenges to successful implementation: for example, conflict, political instability, or geographically isolated populations.

Among the strengths of this study is that explore data from the different ENSANUTS; it is worth noting that to our knowledge, Mexico is one of the first countries in Latin America to make this data freely available to the public. The methodological alignment between the distinct ENSANUT survey all analysis and comparison across five time points, which may be further explored through characteristics such as area type (urban or rural), or ecological variables from robust and reliable sources. Another strength of this study is the probabilistic design, which allows estimations from a sample representative at the national, regional and area (urban or rural) level, and comparisons among previous national health and nutrition surveys with similar methodologies. This contributes to the overarching goal of track trends, inform public policy makers, and contribute to the reformulation of actions seeking to prevent and treat malnutrition and its outcomes among different age groups in Mexico.

Limitations of our study include the sample size of anthropometric data on child malnutrition in the ENSANUT-MC 2016, since although this "halfway point" survey shared the same sampling design as other survey years, its focus was to identify overweight and obesity. The same should be noted for survey year 2020, considering that sample size for states should be completed in 2024.

## CONCLUSION

In conclusion, the periodic collection and availability of population-level anthropometric data is a key component of a system which monitors and evaluates national nutrition trends and allows comparison at the local, national, regional, and global level. The availability of this data implies a nationwide effort which serves to: provide critical data to determine the magnitude, distribution, and trends in population nutrition; maintain systematic, reliable, and timely information based on multiple points of analysis; and inform the formulation and evaluation of national public policy around nutrition and associated factors.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of the National Public Health Institute. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

TS-L, JR, and LC-N: conceptualization and investigation. MR-M, IG-H, and MÁ-A: methodology. MÁ-A: analysis. TS-L: writing—original draft. MR-M, LC-N, and JR: writing—review and editing. All authors contributed to the development of this manuscript and read and approved the final version.

## FUNDING

All the National Health and Nutrition Surveys were funded by the Ministry of Health of Mexico, and in 2018 also by the Geographic and Statistic Institute of Mexico (INEGI).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.777246/full#supplementary-material

The supplemental materials include anthropometric information available by survey year and age group, converted to prevalences, and trends over time (2006–2020).

**Supplementary Table 1 |** Nutritional status of the population of children under 5 years of age, as evaluated by weight-for-height/length, height/length-for-age, weight-for-age, and presence of obesity.

**Supplementary Table 2 |** Prevalence of overweight and obesity in school-age children (5–11 years).

**Supplementary Table 3 |** Prevalence of overweight and obesity in adolescents (12–19 years).

**Supplementary Table 4 |** Prevalence of overweight and obesity in adults (≥20 years).

## REFERENCES

1. Sepúlveda-Amor J, Angel Lezana M, Tapia-Conyer R, Luis Valdespino J, Madrigal H, Kumate J.. Nutritional status of pre-school children and women in Mexico: results of a probabilistic national survey. *Gac Med Mex.* (1990) 126:207–24.
2. Resano-Pérez E, Méndez-Ramírez I, Shamah-Levy T, Rivera JA, Sepúlveda-Amor J. Methods of the National Nutrition Survey 1999. *Salud Publica Mex.* (2003) 45:S558–64. doi: 10.1590/S0036-36342003001000012
3. Palma O, Shamah T, Franco A, Olaiz G, Mendez I. *Methodology. National Health and Nutrition Survey 2006.* Cuernavaca: National Public Health Institute (2006). p. 19–33.
4. Romero-Martínez M, Shamah-Levy T, Franco-Nuñez A, Villalpando S, Cuevas-Nasu L, Gutiérrez JP, et al. National Health and Nutrition Survey 2012: design and coverage. *Salud Publica Mex.* (2013) 55:S332–40.
5. Romero-Martínez M, Shamah-Levy T, Cuevas-Nasu L, Méndez Gómez-Humarán I, Gaona-Pineda EB, Gómez-Acosta LM, et al. Methodological design of the National Mid-Stage Health and Nutrition Survey 2016. *Salud Publica Mex.* (2017) 59:299–305. doi: 10.21149/8593
6. Romero-Martínez M, Shamah-Levy T, Vielma-Orozco E, Heredia-Hernández O, Mojica-Cuevas J, Cuevas-Nasu L, et al. National Health and Nutrition Survey 2018-19: methodology and perspectives. *Salud Publica Mex.* (2019) 61:917–23. doi: 10.21149/11095
7. Shamah-Levy T, Cuevas-Nasu L, Méndez-Gómez Humarán I, Morales-Ruán C, Valenzuela-Bravo DG, Gaona-Pineda EB, et al. Prevalencia y predisposición a la obesidad en una muestra Nacional de niños y adolescentes en México. *Salud Publica Mex.* (2020) 62:725–33. doi: 10.21149/11552
8. Barquera S, Hernández-Barrera L, Trejo-Valdivia B, Shamah T, Campos-Nonato I, Rivera-Dommarco J. Obesidad en México, prevalencia y tendencias en adultos. Ensanut 2018–19. *Salud Publica Mex.* (2020) 62:682–92. doi: 10.21149/11630
9. Shamah-Levy T, Romero-Martínez M, Cuevas-Nasu L, Méndez Gómez-Humaran I, Antonio Avila-Arcos M, Rivera-Dommarco JA. The Mexican National Health and Nutrition Survey as a basis for public policy planning: overweight and obesity. *Nutrients.* (2019) 11:1727. doi: 10.3390/nu11081727
10. Lohman T, Roche A, Martorell R. Anthropometric standarization reference manual. Champlaign, IL: Human Kinetics (1988) 1–177.
11. Habicht JP. Standardization of anthropometric methods in the field. *PAHO Bull.* (1974) 76:375–84.
12. *WHO Anthro for Personal Computers, version 3.2.2, 2011: Software for Assessing Growth and Development of the World'S Children*. Geneva: WHO (2010). Available online at: http://www.who.int/childgrowth/software/en/ (accessed September 2, 2021).
13. de ONIS M, Onayango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ.* (2007) 85:660–7. doi: 10.2471/BLT.07.043497
14. WHO Expert Committee on Physical Status: the Use and Interpretation of Anthropometry (1993: Geneva, Switzerland) and World Health Organization. *Physical Status: The Use of and Interpretation of Anthropometry, Report of a WHO Expert Committee. World Health Organization*. Available online at: https://apps.who.int/iris/handle/10665/37003 (accessed September 2, 2021).
15. Cuevas-Nasu L, Shamah-Levy T, Hernández-Cordero SL, et al. Trends of malnutrition in Mexican children under five years from 1988 to 2016: analysis of five national surveys. *Salud Publica Mex.* (2018) 60:283–90. doi: 10.21149/8846
16. Cuevas-Nasu L, García-Guerra A, González-Castell LD, Morales-Ruan MC, Méndez-Gómez HI, Gaona-Pineda EB, et al. Magnitude and trends of malnutrition, and stunting associated factors among children under five years old in Mexico, Ensanut 2018-19. *Salud Publica Mex.* (2021) 63:339–49. doi: 10.21149/12193
17. Perumal N, Namaste S, Qamar H, Aimone A, Bassani DG, Roth DE. Anthropometric data quality assessment in multisurvey studies of child growth. *Am J Clin Nutr.* (2020) 112(Suppl 2):806S−15S. doi: 10.1093/ajcn/nqaa162
18. Leidman E, Mwirigi LM, Maina-Gathigi L, Wamae A, Imbwaga AA, Bilukha OO. Assessment of anthropometric data following investments to ensure quality: Kenya Demographic Health Surveys case study, 2008 to 2009 and 2014. *Food Nutr Bull.* (2018) 39:406–19. doi: 10.1177/0379572118783181

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 03 December 2021
doi: 10.3389/fnut.2021.777364

# Vitamin D Biofortification of Pork May Offer a Food-Based Strategy to Increase Vitamin D Intakes in the UK Population

*Holly R. Neill[1], Chris I. R. Gill[1], Emma J. McDonald[2], W. Colin McRoberts[3] and L. Kirsty Pourshahidi[1]**

[1] Nutrition Innovation Centre for Food and Health, School of Biomedical Sciences, Ulster University, Coleraine, United Kingdom, [2] Devenish Nutrition Ltd., Belfast, United Kingdom, [3] Agri-Food and Biosciences Institute, Belfast, United Kingdom

OPEN ACCESS

**Edited by:**
Alessandra Durazzo,
Council for Agricultural Research and Economics, Italy

**Reviewed by:**
Armin Zittermann,
Heart and Diabetes Center North Rhine-Westphalia, Germany
Susheel Kumar Sharma,
ICAR Research Complex for NEH Region, India

***Correspondence:**
L. Kirsty Pourshahidi
k.pourshahidi@ulster.ac.uk

**Specialty section:**
This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition

**Received:** 15 September 2021
**Accepted:** 03 November 2021
**Published:** 03 December 2021

**Citation:**
Neill HR, Gill CIR, McDonald EJ, McRoberts WC and Pourshahidi LK (2021) Vitamin D Biofortification of Pork May Offer a Food-Based Strategy to Increase Vitamin D Intakes in the UK Population. Front. Nutr. 8:777364. doi: 10.3389/fnut.2021.777364

Hypovitaminosis D is prevalent worldwide, with many populations failing to achieve the recommended nutrient intake (RNI) for vitamin D (10–20 μg/day). Owing to low vitamin D intakes, limited exposure to ultraviolet-B (UVB) induced dermal synthesis, lack of mandatory fortification and poor uptake in supplement advice, additional food-based strategies are warranted to enable the UK population to achieve optimal vitamin D intakes, thus reducing musculoskeletal risks or suboptimal immune functioning. The aims of the current study were to (1) determine any changes to vitamin D intake and status over a 9-year period, and (2) apply dietary modeling to predict the impact of vitamin D biofortification of pork and pork products on population intakes. Data from the UK National Diet and Nutrition Survey (Year 1–9; 2008/09–2016/17) were analyzed to explore nationally representative mean vitamin D intakes and 25-hydroxyvitamin D (25(OH)D) concentrations ($n$ = 13,350). Four theoretical dietary scenarios of vitamin D pork biofortification were computed (vitamin D content +50/100/150/200% vs. standard). Vitamin D intake in the UK population has not changed significantly from 2008 to 2017 and in 2016/17, across all age groups, 13.2% were considered deficient [25(OH)D < 25 nmol/L]. Theoretically, increasing vitamin D concentrations in biofortified pork by 50, 100, 150, and 200%, would increase vitamin population D intake by 4.9, 10.1, 15.0, and 19.8% respectively. When specifically considering the impact on gender and age, based on the last scenario, a greater relative change was observed in males (22.6%) vs. females (17.8%). The greatest relative change was observed amongst 11–18 year olds (25.2%). Vitamin D intakes have remained stable in the UK for almost a decade, confirming that strategies are urgently required to help the population achieve the RNI for vitamin D. Biofortification of pork meat provides a proof of concept, demonstrating that animal-based strategies may offer an important contribution to help to improve the vitamin D intakes of the UK population, particularly adolescents.

**Keywords: cholecalciferol, 25-hydroxyvitamin D (25(OH)D), National Diet and Nutrition Survey, dietary modeling, bio-addition, meat, feed supplementation, UVB radiation**

## INTRODUCTION

Substantial evidence exists to suggest hypovitaminosis D is prevalent globally and, assuming minimal sunlight exposure, many also fail to achieve the recommended nutrient intake (RNI) for vitamin D (10–20 μg/day) (1–3). Within the UK, one in five adults do not meet the RNI of 10 μg/day which is currently advised by the Scientific Advisory Committee on Nutrition (4) to reduce the risk of musculoskeletal diseases (5, 6). Therefore, a large proportion of the population have suboptimal 25-hydroxyvitamin D (25(OH)D) concentrations (a reliable and robust blood marker of vitamin D status). The exact 25(OH)D concentration classified as deficient is disputed and ranges from <25 to <50 nmol/L (4, 5, 7). The reason for suboptimal vitamin D status is multifactorial, owed predominantly to limited natural food sources, lack of mandatory fortification, poor implementation of supplement advice, especially amongst low socioeconomic groups, and insufficient dermal synthesis from ultraviolet-B (UVB) exposure (8, 9). Thus, additional sustainable food-based strategies are urgently warranted to enable populations to achieve adequate vitamin D intakes (10).

Fortification and biofortification (also referred to as "bio-addition") are popular initiatives to help alleviate vitamin deficiencies globally; the former being widely accepted by consumers nowadays and a plethora of research confirming its effectiveness (11–13). Vitamin D fortification in particular is well-established as a method to improve consumer vitamin D intakes and subsequently increase circulating 25(OH)D concentrations (14–16). Many have argued the benefits and feasibility of vitamin D fortification policies, considering both the efficacy and safety of various vehicles and scenarios (16–20). Countries having implemented mass vitamin D fortification policies observed increased 25(OH)D concentrations and reduced rates of deficiency in their populations (21–23). However, within the UK, vitamin D fortification is applied on a voluntary basis, with the exception of infant formula. On a population level, mandatory biofortification and fortification presents as the only method to increase 25(OH)D concentrations within the general population.

Biofortification refers to the endogenous increase of nutrients during the growth phase, either in animal-based or plant-based foods. Evidence has shown the plausibility of increasing the vitamin D content in commodity animal-based foods such as pork, beef, chicken and eggs, as well as mushrooms (24). Amongst animal foods, vitamin D biofortification can be achieved by either UVB radiation and/ or feed alteration. Whilst oily fish such as salmon, herring, sardines and mackerel are one of the richest sources of vitamin $D_3$ (∼3–19 μg/100 g) (25), in general, its consumption remains unpopular in the UK owing to a variety of factors including taste preferences, perceived cost and sustainability concerns (26). Currently, meat and meat products remain the main contributor (30%) to vitamin D intakes in the UK adult (19–64 years) diet (27), despite relatively lower vitamin D contents per 100 g (∼0.1–1.9 μg/100 g) (25). In particular, pork is popular amongst the UK population and is the most widely consumed meat worldwide (28). Typically, pig feed in the UK contains the maximum permitted concentration of 50 μg (2,000 IU) vitamin D/kg owing to animal health benefits; albeit this is considered substandard by many, and revaluation has been proposed (29). Thus, owing to dietary supplementation restrictions, UVB exposure is warranted to further elevate vitamin D concentrations in pork meat. As such, enriching common foods, especially meat and meat products, may complement the diet of at-risk populations.

Dietary mathematical modeling can predict the impact of vitamin D biofortification implementation and has been previously conducted to assess differing vitamin D fortification scenarios (17–19, 30, 31). However, to the authors knowledge, this is the first study to investigate the theoretical impact of biofortified meat.

Therefore, the aims of the current study were (1) quantify vitamin D intake and status over a 9-year period and; (2) use dietary modeling scenarios to predict how biofortification of pork meat could improve vitamin D intakes across the UK population.

## MATERIALS AND METHODS

### Study Population

Participants from the UK National Diet and Nutrition Survey Rolling Program (NDNS) Years 1–9 (2008/09–2016/17) dataset were used in the present manuscript. Details regarding the design, participant selection, recruitment process and data collection of NDNS are reported in full elsewhere (32–36). In short, jointly funded by Public Health England and the Food Standards Agency, NDNS is a UK-wide continuous cross-sectional survey. Fieldwork began in 2008 and provides quantitative comprehensive information regarding diet, nutritional status, sociodemographic characteristics, lifestyle, and physical activity levels from a nationally representative sample of the general UK population aged 1.5 years and older. Using household post-code details from the Post-code Address File, participants were stratified and randomly recruited to take part. All food and drink consumption were estimated by participants using a self-reported 3 or 4-day food diary. Nutrient intakes were then quantified using McCance and Widdowson's The Composition of Foods Integrated Dataset (CoFID). Following written consent, fasting blood samples were collected by venepuncture and transported in a cool box to a local processing field laboratory within 2 h of blood collection. All samples were centrifuged at 2,000 g for 20 min at 4°C and aliquoted. Survey years 1–6 (2008–2014) used competitive chemiluminescence immunoassay (CLIA) to determine plasma and serum 25(OH)D, whilst Years 7–9 (2014–2017) measured serum 25(OH)D by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). Quality controls were included within each batch of samples and National Institute of Standards and Technology (NIST) confirmed accuracy. Involvement in blood sampling was lower than other aspects of NDNS observations and therefore, participant numbers are inconsistent between vitamin D intake and status. Ethical approval was obtained from Oxfordshire Research Ethics Committee (Ref. No. 07/H0604/113) and data were made available from the UK Data Archives (37).

## Data Modeling

As meat and meat products are one of the main contributors (18–31%) to vitamin D intakes within the UK population (27), pork and pork products were selected for predictive dietary modeling (see full list of food codes in **Supplementary Material**). The theoretical percentage increases from biofortification were selected based upon prior on-farm UVB exposure biofortification work in pigs (38, 39). Owing to varying concentrations of vitamin D in pork and pork products, biofortification increases were considered on a percentage basis. As such, in addition to current standard vitamin D content (baseline, 0%), four scenarios were examined to determine the effect of pork biofortification: 50, 100, 150, and 200% increase in total vitamin D concentrations in pork. Within the NDNS dataset, vitamin D is defined as total vitamin D, including both forms of vitamin $D_3$ and $D_2$. Composite dishes which did not include a significant proportion of pork were not included. Scenarios including and excluding supplements were both explored, as well as subgroup analyses of sex and age groups (1.5–3, 4–10, 11–18, 19–64, 65+ years).

## Data Analysis

Analysis of all data were performed using the Statistical Package for the Social Science for Windows (IBM SPSS Statistics version 25, Chicago IL, USA). All values are expressed as mean ± standard deviation (SD), unless otherwise specified. Descriptive statistics were used to present participant characteristics, mean vitamin D intakes and 25(OH)D concentrations. Normality tests were conducted for all data using Kolmogorov-Smirnov testing and, where necessary, data was log transformed. Independent *t*-tests and one-way analysis of variance (ANOVA) with *post-hoc* Tukey test were performed to calculate subgroup analysis in sex, age and season for vitamin D intakes and 25(OH)D concentrations. Current baseline total vitamin D intakes were calculated per participant and then expressed as a daily average based on the number of completed food diary days. The relevant pork food codes were identified from NDNS datasets (**Supplementary Material**) and a SPSS syntax was created to calculate the new vitamin D content of these foods based on our four scenarios (+50, 100, 150, and 200%). These values were applied to individual food intake data to create four new total vitamin D intakes for each participant. This data was again divided by the number of completed food diary days. One-way within-subjects repeated measures ANOVA and *post-hoc* pairwise comparisons were conducted to identify significance differences in daily vitamin D intakes between the various dietary modeling scenarios. Relative percentage change was calculated as below.

$$\textit{Relative percentage change} = \frac{\textit{vitamin D intake from biofortified pork} - \textit{vitamin D intake from standard pork}}{\textit{vitamin D intake from standard pork}} \, x\,100$$

Values of $p < 0.05$ were deemed statistically significant throughout and results displayed in tabular and graphical form.

**TABLE 1** | Participant characteristics from the UK National Diet and Nutrition Survey (NDNS) Years 1–9 (2008–2017).

| | All (*n* = 13,350) | Male (*n* = 6,161) | Female (*n* = 7,189) | *P*-value* |
|---|---|---|---|---|
| Age (y) | 30 ± 24 | 28 ± 24 | 32 ± 24 | <0.001 |
| Weight (kg) | 58.5 ± 27.3 | 59.9 ± 30.1 | 57.3 ± 24.7 | NS |
| Height (m) | 1.53 ± 0.24 | 1.56 ± 0.27 | 1.51 ± 0.22 | <0.001 |
| BMI (kg/m$^2$) | 23.5 ± 6.5 | 23.0 ± 6.2 | 24.0 ± 6.7 | <0.001 |
| Supplement user (%) | 21.6 | 19.0 | 23.9 | <0.001 |

*Data is presented as mean ± standard deviation, unless otherwise specified.*
**P-value difference within rows between male and female participants; independent samples t-test on log transformed data, where required. Significance set at p < 0.05 throughout. UK, United Kingdom; n, number of participants; y, years; kg, kilograms; m, meters; BMI, body mass index; NS, not significant.*

# RESULTS

The study included 13,350 participants (46% males, 54% females), ranging in age from 1.5 to 96 years and residing in England (57.6%), Northern Ireland (13.6%), Scotland (15.5%), and Wales (13.3%). Participant characteristics are summarized in **Table 1**. The majority of participants never smoked (56.0%) while 21.9 and 21.2% either currently smoked or were ex-smokers, respectively.

## Vitamin D Intakes

Vitamin D mean intakes have not changed significantly between 2008 to 2017 in the UK population when considering diet alone or in combination with supplement intake ($p > 0.05$; **Figures 1A,B**; **Supplementary Material**). Including supplemental intake, 95.8% of participants failed to achieve the recommendation of 10 μg/day. The mean vitamin D intake for those below the RNI was 2.76 ± 1.99 μg/day. Participants consuming 10 μg or above (4.2%) daily vitamin D had mean intakes of 19.35 ± 24.29 μg/day. When considering diet alone, males reported a significantly higher mean vitamin D intake compared to females over all 9 years combined (M 2.66 ± 1.99 μg/day and F 2.30 ± 1.66 μg/day, $p < 0.05$) as well as each individual survey year. However, females reported greater mean daily vitamin D intakes in comparison with males when both diet and supplements were included overall years combined (M 3.42 ± 4.42 μg/day and F 3.50 ± 7.59 μg/day, $p < 0.05$). **Figures 2A,B** and **Supplementary Material** outline intakes from age groups in each survey year. In general, those aged 65 years and over consumed the highest amounts of vitamin D, whilst the lowest was most commonly observed in those aged 1.5–3 years. Some significant differences were observed in 4–10 years (2014/15 vs. 2016/17) and 19–64 years (2008/09 vs. 2015/16) from diet alone, and 1.5–3 years (2009/10 vs. 2015/16 and 2016/17) from diet and supplementation.

## Dietary Modeling Scenarios

Across all participants, our modeling scenarios demonstrated that a 5, 10, 15, or 20% increase in population vitamin

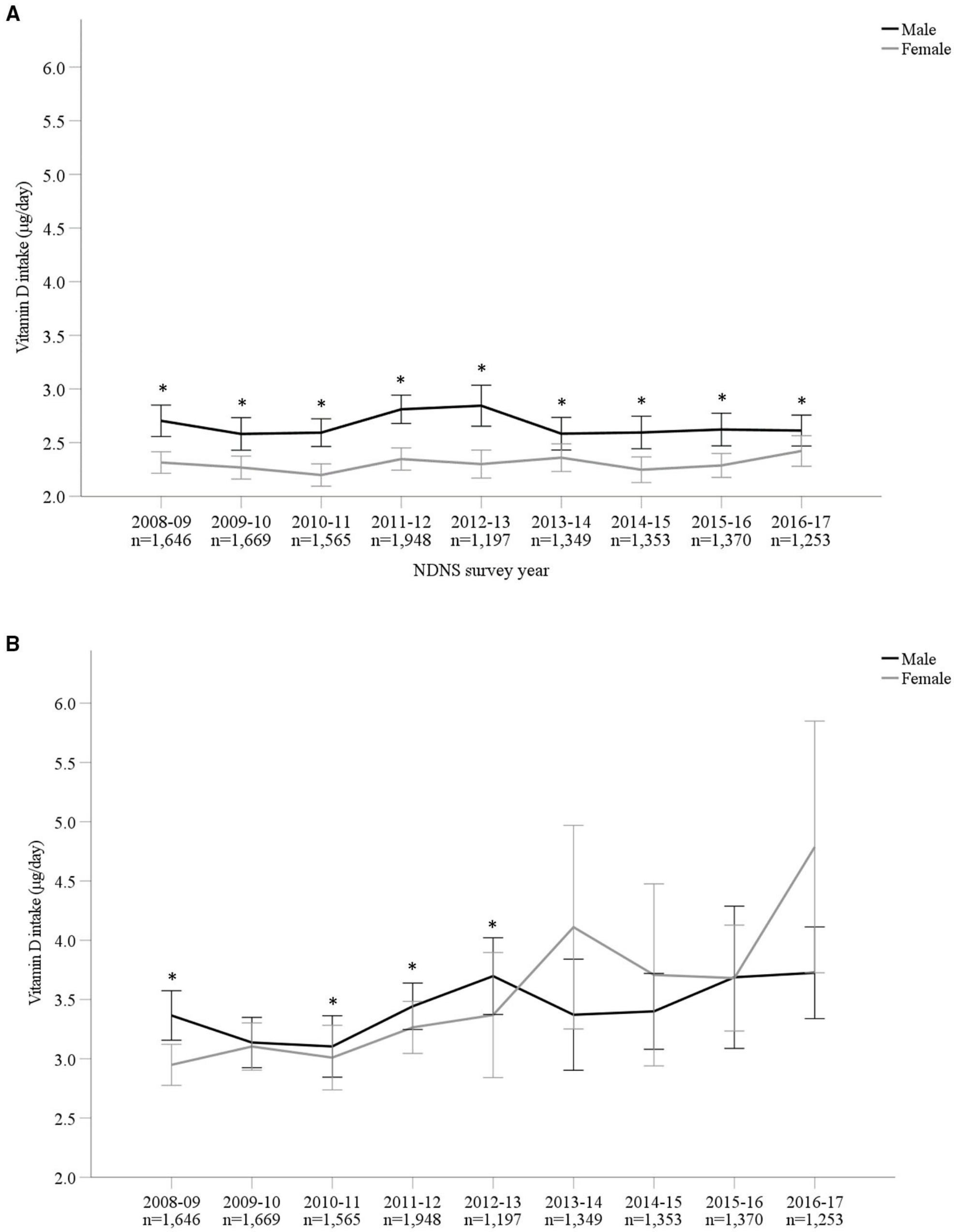


**FIGURE 1 | (A,B)** Vitamin D intake (μg/day) from diet alone **(A)** and in combination with supplements **(B)** from Years 1–9 (2008–2017) of the UK National Diet and Nutrition Survey (NDNS). Data is presented as mean (95% CI). *Denotes significant difference ($p < 0.05$) between male and female participants; independent samples *t*-test using log transformed data. No significant difference ($p > 0.05$) between survey years in total group or within each gender; one-way ANOVA tests using log transformed data. UK, United Kingdom; *n*, number of participants; y, years; CI, confidence interval.

D intake was achievable if the concentrations in biofortified pork were elevated by 50, 100, 150, and 200%, respectively (**Table 2**). Considering the 200% increase scenario, a greater relative change was observed in males (22.6%) compared to females (17.8%) (**Table 2**), and although older adults (65 years and over) had significantly greater vitamin D intakes compared to other age categories (3.28 ± 2.27 μg/day), this age group observed the smallest relative increase from the dietary modeling scenarios (14.3%) (**Table 2**). This may be owed to fish and fish dishes also substantially contributing to vitamin D intakes in older adults (**Supplementary Material**). The greatest relative change was observed amongst 11–18 years, where 200% vitamin D biofortification of pork and pork products would result in a 25.2% increase in

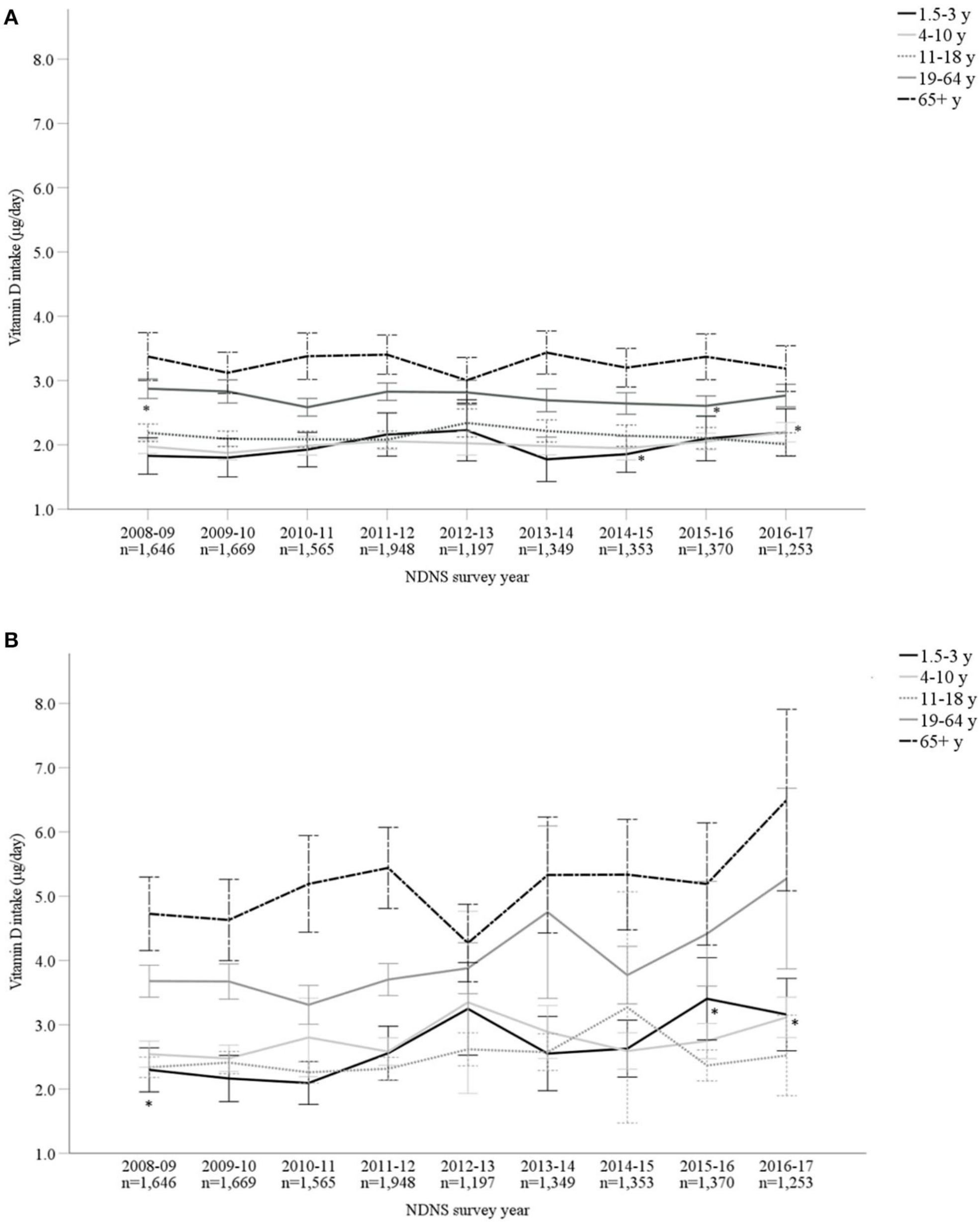


**FIGURE 2 | (A,B)** Vitamin D intake (µg/day) from diet alone **(A)** and in combination with supplements **(B)**, split by age categories, from Years 1–9 (2008–2017) of the UK National Diet and Nutrition Survey (NDNS; total $n = 13,350$). Data is presented as mean (95% CI). *Denotes significant difference ($p < 0.05$) between survey years within age group; one-way ANOVA tests using log transformed data. For *post-hoc* (Tukey) tests between ages groups in the same survey year, see **Supplementary Material**. UK, United Kingdom; *n*, number of participants; y, years; CI, confidence interval.

mean daily vitamin D intakes. The range of vitamin D intakes from 50, 100, 150, and 200% scenarios were <0.01–22.42, <0.01–22.66, <0.01–22.90, and <0.01–23.14 µg/day, respectively. In general, dietary modeling showed a significant difference between each age cohort, with the exception of 4–10 years and 11–18 years ($p > 0.05$). Significant increases in vitamin D daily intakes were evident from current baseline values for each of the four modeled changes ($p < 0.05$).

## 25(OH)D Concentrations

Overall, mean 25(OH)D concentrations across all 9 years were 46.8 ± 21.2 nmol/L (M 46.5 ± 20.6 nmol/L and F 47.1 ± 21.6 nmol/L). The lowest and highest mean 25(OH)D concentrations

**TABLE 2 |** Theoretical mean vitamin D intakes (μg/day), split by gender and age ranges, from diet alone of UK population in response to varying increases in vitamin D concentration of pork and pork products ($n = 13,350$).

| Population | | Vitamin D intake (μg/day) | | | | |
|---|---|---|---|---|---|---|
| | | Vitamin D increases in pork | | | | |
| | | 0% | 50% | 100% | 150% | 200% |
| **All** | Mean ± SD | $2.47 \pm 1.83^{a}$ | $2.59 \pm 1.86^{b}$ | $2.72 \pm 1.90^{c}$ | $2.84 \pm 1.94^{d}$ | $2.96 \pm 2.00^{e}$ |
| $n = 13,350$ | Relative change (%) | N/A | 4.9 | 10.1 | 15.0 | 19.8 |
| | Maximum (μg/day) | 22.19 | 22.42 | 22.66 | 22.90 | 23.14 |
| **Male** | Mean ± SD | $2.66 \pm 1.99^{a}$ | $2.81 \pm 2.03^{b}$ | $2.96 \pm 2.07^{c}$ | $3.11 \pm 2.13^{d}$ | $3.26 \pm 2.21^{e}$ |
| $n = 6,161$ | Relative change (%) | N/A | 5.6 | 11.3 | 16.9 | 22.6 |
| **Female**$^{\dagger}$ | Mean ± SD | $2.30 \pm 1.66^{a}$ | $2.40 \pm 1.68^{b}$ | $2.51 \pm 1.70^{c}$ | $2.61 \pm 1.73^{d}$ | $2.71 \pm 1.77^{e}$ |
| $n = 7,189$ | Relative change (%) | N/A | 4.3 | 9.1 | 13.5 | 17.8 |
| **1.5–3 y** | Mean ± SD | $1.97 \pm 1.90^{a,z}$ | $2.05 \pm 1.90^{b,z}$ | $2.12 \pm 1.91^{c,z}$ | $2.20 \pm 1.93^{d,z}$ | $2.28 \pm 1.94^{e,z}$ |
| $n = 1,173$ | Relative change (%) | N/A | 4.1 | 7.6 | 11.7 | 15.7 |
| **4–10 y** | Mean ± SD | $2.01 \pm 1.19^{a,y}$ | $2.12 \pm 1.22^{b,y}$ | $2.23 \pm 1.26^{c,y}$ | $2.35 \pm 1.30^{d,y}$ | $2.46 \pm 1.36^{e,y}$ |
| $n = 2,554$ | Relative change (%) | N/A | 5.5 | 10.9 | 16.9 | 22.4 |
| **11–18 y** | Mean ± SD | $2.14 \pm 1.38^{a,y}$ | $2.27 \pm 1.42^{b,y}$ | $2.41 \pm 1.48^{c,y}$ | $2.54 \pm 1.55^{d,y}$ | $2.68 \pm 1.63^{e,y}$ |
| $n = 2,821$ | Relative change (%) | N/A | 6.1 | 12.6 | 18.7 | 25.2 |
| **19–64 y** | Mean ± SD | $2.74 \pm 1.99^{a,x}$ | $2.88 \pm 2.02^{b,x}$ | $3.01 \pm 2.07^{c,x}$ | $3.14 \pm 2.13^{d,x}$ | $3.28 \pm 2.19^{e,x}$ |
| $n = 5,223$ | Relative change (%) | N/A | 5.1 | 9.9 | 14.6 | 19.7 |
| **65+ y** | Mean ± SD | $3.28 \pm 2.27^{a,w}$ | $3.40 \pm 2.29^{b,w}$ | $3.51 \pm 2.31^{c,w}$ | $3.63 \pm 2.34^{d,w}$ | $3.75 \pm 2.37^{e,w}$ |
| $n = 1,579$ | Relative change (%) | N/A | 3.7 | 7.0 | 10.7 | 14.3 |

*Data is presented as mean ± standard deviation, unless otherwise specified. Values not sharing a common superscript letter in rows (a, b, c, d, e) are significantly different ($p < 0.001$) between modeling scenarios; one-way repeated measures ANOVA using log transformed data. $^{\dagger}$Denotes all mean intakes are significantly different between males and females; independent t-test. Values not sharing a common superscript letter in columns (z, y, x, w) are significantly different ($p < 0.05$) between age ranges of the same modeling scenario; one-way ANOVA with Tukey test using log transformed data. UK, United Kingdom; n, number of participants; SD, standard deviation; y, years; N/A, not applicable.*

were observed in 2009–2010 (44.5 ± 19.4 nmol/L) and 2016–2017 (51.7 ± 22.5 nmol/L), respectively. The most recent data (for 2016/17) reported significantly increased 25(OH)D concentrations compared to 2009–2014 ($p < 0.05$). However, when split by age categories, those aged 1.5–3 years and 11–18 years reported no significant difference across all 9 survey years ($p > 0.05$) (**Supplementary Material**). For both males and females, vitamin D status significantly varied across seasons, except 2013/14 in males, with the highest concentration observed during late summer months (July to September) and lowest in late winter (January to March) (**Figures 3A,B**; **Supplementary Material**). **Figure 4** presents the percentage of participants classified as vitamin D deficient based on different cut-off levels. In 2016/17, when including all age groups, 45.6, 19.3, and 13.2% presented 25(OH)D concentrations deemed insufficient (<50, 30, and 25 nmol/L, respectively; **Supplementary Material**). Across all nine survey years, 15.6% of participants had 25(OH)D concentrations <25 nmol/L (mean 18.9 ± 5.1 nmol/L) (4). This increased to 24.4 and 58.5% when considering <30 nmol/L (mean 22.0 ± 5.8 nmol/L) and <50 nmol/L (mean 32.3 ± 10.5 nmol/L) as the deficiency threshold classified by the U.S. Institute of Medicine and The Endocrine Society, respectively (5, 7).

# DISCUSSION

For almost a decade, mean vitamin D intakes have remained suboptimal and stagnant in the UK. Only 17% of adults (19–64 years) reported supplementing with vitamin D in 2016/17 and a larger proportion of females take a vitamin D supplement compared to men (20 vs. 13%) (40) thus, it would appear that there has been limited implementation following updated vitamin D supplement advice in 2016 (4). The COVID-19 pandemic, however, has heightened public interest and awareness of the importance of vitamin D (41) and it may be postulated that this has resulted in a greater uptake in supplementation which may be reflected in future data. Nevertheless, the current findings confirm that vitamin D intakes are unacceptably low and strongly supports the need for mandatory biofortification as an additional food-based solution to offer a diet-focused approach in alleviating vitamin D deficiency.

The current paper, for the first time, demonstrates that increasing the vitamin D content in pork and pork products, without altering consumers habitual diet, resulted in significant, albeit modest, increases to the vitamin D intakes within the UK population. It is important to be cognizant that this theoretical change was based on achievable increases in vitamin D that

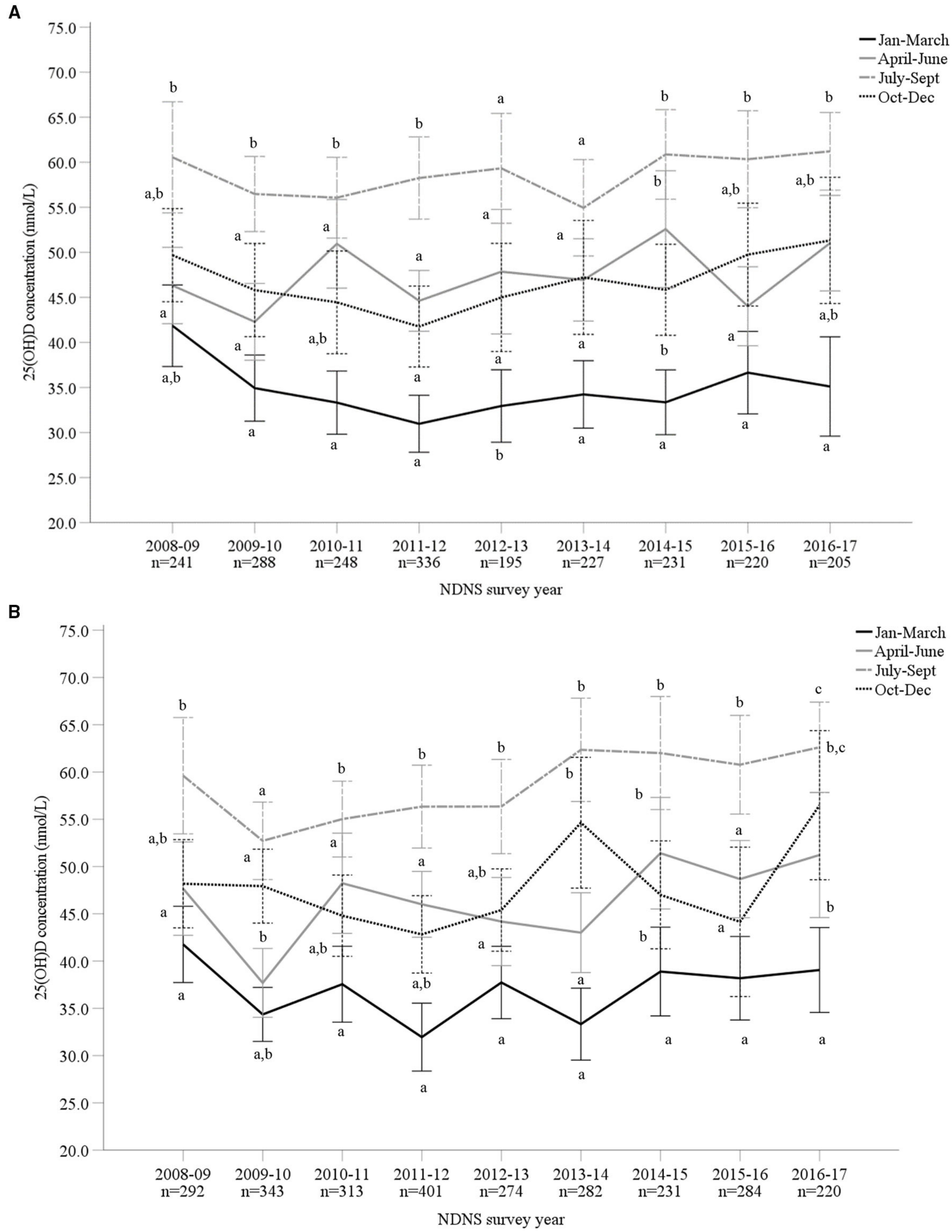


**FIGURE 3 | (A,B)** Vitamin D status [25-hydroxyvitamin D (25(OH)D) nmol/L] of male (**A**; $n = 2,191$) and female (**B**; $n = 2,640$) adults aged 19–64 years from Years 1–9 (2008–2017) of the UK National Diet and Nutrition Survey (NDNS; total $n = 4,831$). Data is presented as mean (95% CI). Values not sharing a common superscript letter (a, b, c) are significantly different ($p < 0.05$) between seasons in each survey year; one-way ANOVA and *post-hoc* (Tukey) tests. 25(OH)D concentration data from standardized liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). UK, United Kingdom; $n$, number of participants; CI, confidence interval.

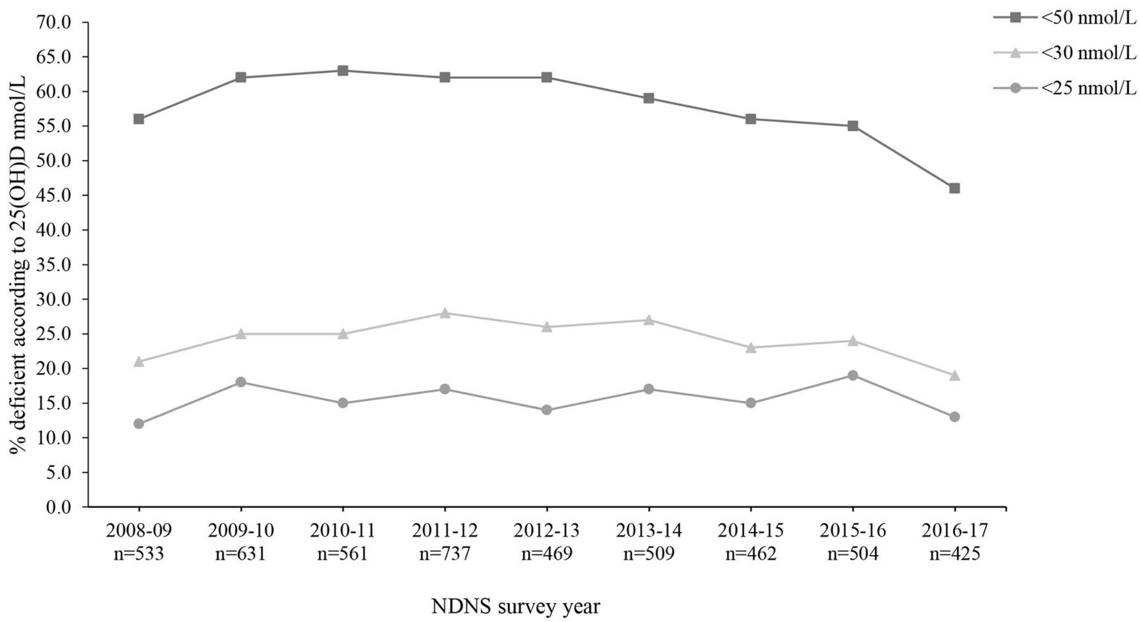


**FIGURE 4 |** Percentage (%) of participants from the UK National Diet and Nutrition Survey (NDNS) classified as vitamin D deficient based on The Endocrine Society and EFSA (<50 nmol/L), US Institute of Medicine (<30 nmol/L) and SACN (<25 nmol/L) cut-off values for 25(OH)D concentrations ($n$ = 4,831). UK, United Kingdom; EFSA, European Food Safety Authority; US, United States; SACN, Scientific Advisory Committee on Nutrition (UK); 25(OH)D, 25-hydroxyvitamin D; $n$, number of participants who provided blood sample.

can be reasonably expected using natural biofortification, for example UVB radiation and enriched animal feed, rather than traditional exogenous fortification (38, 39). Previous research has theorized the potential impact of vitamin D fortification in ready-to-eat cereals (RTEC), milk, bread, plain yogurt, cheese, eggs, crackers, and wheat flour (17–19, 30, 31). In these studies, as expected, predicted exogenous fortification produced higher theoretical intakes compared to those reported within the present study. Fortification of semi-skimmed cows' milks theoretically increased median vitamin D intakes from 2.3 to 6.1 μg/day (17) whilst fortifying both milk and bread resulted in ∼70% of Irish individuals meeting the recommended 10 μg/day (19). Notably, endogenous vitamin D biofortification offers an increased challenge as greater natural inter and intra-variability exists in end-point vitamin D concentrations in meat, compared to traditional fortification practices whereby there is rigid control on the quantity of vitamer being added to the foodstuff during processing. However, biofortification may be perceived as a more natural food-based strategy and presents vast opportunities for both the food industry and populations.

Additionally, improvements within the current paper were only modeled based on altering only one food type (pork) within a main food group (meat). Even greater improvements would be expected if this practice was implemented across a wider animal food portfolio to account for dietary diversity and include, for example, chicken, beef and eggs, of which on-farm evidence provides proof of concept (24, 42–44). Regardless of the biofortification food carrier, it would be short-sighted and ill-advised to focus on a single commodity as this would undoubtedly exclude a proportion of the population who identify as low or non-consumers. Of note, owing to the rise in vegetarian and vegan dietary trends, motivated by environmental and health concerns (45), coupled with the prevalence of religious groups vulnerable to vitamin D deficiency such as veiled Moslem women living in Europe, additional plant-based strategies must also be considered to maximize the benefits of biofortification. Human randomized controlled trials (RCTs) have investigated the efficacy of biofortified mushrooms and bread baked with UV-treated yeast. Whilst bread had poor bioavailability, mushrooms may be an alternative biofortification food vehicle to increase 25(OH)$D_2$ concentrations amongst non-meat consumers (24). Future research should further explore bioavailability of UV-treated yeast to confirm these findings. Despite vitamin $D_3$ being considered more effective than vitamin $D_2$ (46, 47), and 25(OH)$D_2$ often increasing at the apparent expense of 25(OH)$D_3$, for vegans and vegetarians who may have limited vitamin $D_3$ in their diet, total 25(OH)D concentrations should increase with vitamin $D_2$-biofortified products. Owing to limited viable non-animal biofortified foods, vegetarian and vegan consumers may be more likely to benefit from fortified foods such as breakfast cereals (18, 48), wheat flour/bread (19, 31, 48, 49) and fruit juice (48, 50). Nonetheless, meat remains a popular staple component for a large proportion of the UK population (40). Therefore, it is advantageous for those who choose to include pork and pork products in their diet to have access to high-quality, nutrient-dense meat to aid in reaching nutritional recommendations. The marketed nutritional benefit of vitamin D biofortified pork, however, should ensure the avoidance of a "halo effect" or positivity bias (51–53), whereby the enhanced micronutrient content may encourage consumers to contradict the recommendation to reduce red and processed meat intake.

Future vitamin D dietary modeling should explore the proportion of population below the RNI and exceeding the tolerable upper intake level (UL) in various scenarios using both animal and plant-based foods, combined with the inclusion and exclusion of varying supplemental intakes. This evaluation will inform how to achieve the goal of having nearly the entire population above the recommended intake without exceeding the UL. Future dietary modeling research may also explore environmental or individual factors associated with suboptimal 25(OH)D concentrations such as educational status, socioeconomic status and genetics (54, 55). Importantly, there is a need to ensure those of lower socioeconomic status are not further disadvantaged by biofortified products being a premium price or having an image of exclusivity. Moreover, prior to the implementation of vitamin D biofortification, there should be a thorough assessment of its efficacy, feasibility, production costs for farmer and food industry, as well as consideration for regulatory aspects (48).

Notably, biofortification would only be effective if it is part of a mandatory vitamin D fortification program by the government, with both strategies recognized as the only measure to improve vitamin D status in the general population. The United States, Canada and Finland have effectively implemented mass vitamin D food fortification policies (16, 22, 23) which should act as a benchmark for the UK. If enforced, revision of supplement guidance may be required to assess risk of toxicity and reduce the likelihood of exceeding the 50–100 μg/day (2,000–4,000 IU/day) UL for adults and children (4, 48). Within the current modeling scenario however, safety concerns are low and no participant exceeded the UL following vitamin D biofortification of pork and pork products. This differs to supplement intake which can pose risk of toxicity depending on the quantity and frequency of consumption. Importantly, pork meat biofortification holds a degree of biological regulatory control, compared to traditional fortification. Similar to humans, in UVB-exposed pigs the CYP24A1 enzyme is induced which acts as a feedback mechanism to prevent vitamin D toxicity (56).

Owing to stable vitamin D intakes, it is unsurprising that, in general, 25(OH)D concentrations also reflected this outcome. Considering all age groups, some significant increases in 25(OH)D concentrations were observed in 2016/17 compared to earlier years which is encouraging; however, this was not consistent when split by age groups. Nevertheless, as hypothesized, significant changes were observed between seasons owed to variation in UVB radiation during warmer months (April to September). Vitamin D is a fat-soluble prohormone and therefore some bodily stores remain prior to the expected wintertime nadir, explaining the gradual decline to the lowest status observed in early spring. The consistency of insufficient vitamin D status further highlights the need for additional strategies. Owing to such seasonal variations, calculations could be performed based on the predicted timeline for on-farm UV biofortification duration, slaughter and pork processing to identify pork sold during September to April and thus, sector biofortification to the seasons where suboptimal status is most prevalent. However, this may present industrial challenges and complications which may not be realistic in the commercially relevant or widescale context. If mandatory pork biofortification was implemented throughout the year, and at the optimized level (200% increase in vitamin D content), the contribution of cutaneous synthesis during summer months may subsequently decrease to account for greater dietary intakes. Priority should first focus on increasing year-round vitamin D status in the majority of the UK population by a combination of biofortification, fortification and supplemental strategies.

NDNS provides critically important national food intake data however, it is not without limitations. Food diaries, used within the current study, are inherently flawed owing to high prevalence of underreporting. Healthy or unhealthy bias can be present (57) and is often associated with age, body mass index (BMI, $kg/m^2$), socioeconomic status and ethnicity (58). Results from doubly labeled water (DLW), a well-recognized method to measure energy expenditure in free-living individuals, suggest that energy intake reported by participants has been underreported (energy intake: total energy expenditure = 0.73). A detailed overview of DLW results from NDNS subsamples have been described elsewhere (35).

Future research should model 25(OH)D concentrations combined with total vitamin D (vitamin $D_3$ + vitamin $D_2$). Results may otherwise under-represent the true impact of biofortified pork, with an even greater increase hypothesized if 25(OH)D was independently considered. This is owed to the natural presence of 25(OH)D in pork meat (59, 60), further increases following biofortification practices (24, 38) and suggestions this metabolite may be five times more potent than parental vitamin D to increase circulating 25(OH)D concentrations (25, 61). Although changes to 25(OH)D concentrations were not within the scope of the present study, depending upon absorptive capacity and baseline status, an increase of 1 μg of vitamin $D_3$ has been reported to equate to an approximate increase of 0.7–1.0 nmol/L of serum 25(OH)D concentrations (62). Undoubtedly, there is a need for acute and chronic human RCTs to explore the bioavailability, bioaccessibility and real-life application of the modeling scenarios within this study to confirm the impact of vitamin D biofortified pork meat on circulating 25(OH)D concentrations; thus, potentially reducing musculoskeletal risks or suboptimal immune functioning.

## CONCLUSION

Evidently, there is no panacea for hypovitaminosis D; rather, an integrated strategy is urgently required to reduce prevalence rates. Mandatory biofortification may offer an easily implemented strategy to help bridge the gap between current vitamin D intakes and those recommended by government guidelines in order to limit the risk of musculoskeletal diseases. As meat and meat products provide a sizeable contribution to vitamin D intakes in UK diets, even without biofortification, it is sensical to improve the vitamin D content within these food groups to benefit the UK population and, in particular, at-risk subgroups and supplement non-users. Due to the prominence of meat in UK diets, a lower resistance to uptake would be anticipated and potentially allow for greater impact. Nevertheless, a combined approach using a range of biofortified food products is required to ensure the number of non-consumers is limited and the widest proportion of the population can be reached, particularly those susceptible to lower 25(OH)D concentrations. With a clear need to prioritize and strengthen initiatives which will help in alleviating vitamin D deficiency, mandatory biofortification of pork may offer a modest, yet vitally important contribution to increasing intakes, particularly in adolescents.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: NatCen Social Research and MRC Elsie Widdowson Laboratory (37).

## AUTHOR CONTRIBUTIONS

HN, CG, and LP designed the research. HN combined the database, performed analyses, and prepared the manuscript. All authors contributed to interpreting the results, read and approved the final manuscript.

## FUNDING

This work was funded as part of a Department for the Economy (DfE) Co-operative Awards in Science and Technology (CAST) PhD studentship, supported by Devenish Nutrition Limited.

## ACKNOWLEDGMENTS

The authors wish to thank the participants of the National Diet and Nutrition Survey whose contribution enabled the present work.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.777364/full#supplementary-material

## REFERENCES

1. Amrein K, Scherkl M, Hoffmann M, Neuwersch-Sommeregger S, Köstenberger M, Tmava Berisha A, et al. Vitamin D deficiency 20: an update on the current status worldwide. *Eur J Clin Nutr.* (2020) 74:1498–513. doi: 10.1038/s41430-020-0558-y
2. Hilger J, Friedel A, Herr R, Rausch T, Roos F, Wahl DA, et al. A systematic review of vitamin D status in populations worldwide. *Br J Nutr.* (2014) 111:23–45. doi: 10.1017/S0007114513001840
3. Spiro A, Buttriss JL. Vitamin D: An overview of vitamin D status and intake in Europe. *Nutr Bull.* (2014) 39:322–50. doi: 10.1111/nbu.12108
4. Scientific Advisory Committee on Nutrition (SACN). *Vitamin D and Health* (2016). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/537616/SACN_Vitamin_D_and_Health_report.pdf (accessed August 19, 2019).
5. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP, et al. Evaluation, treatment, and prevention of vitamin D deficiency: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab.* (2011) 96:1911–30. doi: 10.1210/jc.2011-0385
6. Wolff AE, Jones AN, Hansen KE. Vitamin D and musculoskeletal health. *Nat Clin Pract Rheumatol.* (2008) 4:580–8. doi: 10.1038/ncprheum0921
7. Institute of Medicine. Overview of vitamin D. In: Ross AC, Taylor CL, Yaktine AL, et al., editors. *Dietary Reference Intakes for Calcium and Vitamin D.* Washington, DC: National Academies Press (2011). Available online at: www.ncbi.nlm.nih.gov/books/NBK56061/ (accessed December 14, 2020).
8. Black LJ, Jacoby P, Nowson CA, Daly RM, Lucas RM. Predictors of vitamin D-containing supplement use in the Australian population and associations between dose and serum 25-hydroxyvitamin D concentrations. *Nutrients.* (2016) 8:356. doi: 10.3390/nu8060356
9. Greene-Finestone LS, Langlois KA, Whiting SJ. Characteristics of users of supplements containing vitamin D in Canada and associations between dose and 25-hydroxvitamin D. *Appl Physiol Nutr Metab.* (2013) 38:707–15. doi: 10.1139/apnm-2012-0316
10. Hayes A, Cashman KD. Food-based solutions for vitamin D deficiency: putting policy into practice and the key role for research. *Proc Nutr Soc.* (2017) 76:54–63. doi: 10.1017/S0029665116000756
11. Grønborg IM, Tetens I, Christensen T, Andersen EW, Jakobsen J, Kiely M, et al. Vitamin D-fortified foods improve wintertime vitamin D status in women of Danish and Pakistani origin living in Denmark: a randomized controlled trial. *Eur J Nutr.* (2020) 59:741–53. doi: 10.1007/s00394-019-01941-6
12. Keats EC, Neufeld LM, Garrett GS, Mbuya MNN, Bhutta ZA. Improved micronutrient status and health outcomes in low- and middle-income countries following large-scale fortification: evidence from a systematic review and meta-analysis. *Am J Clin Nutr.* (2019) 109:1696–708. doi: 10.1093/ajcn/nqz023
13. Beal T, Massiot E, Arsenault JE, Smith MR, Hijmans RJ. Global trends in dietary micronutrient supplies and estimated prevalence of inadequate intakes. *PLoS ONE.* (2017) 12:e0175554. doi: 10.1371/journal.pone.0175554
14. Black LJ, Seamans KM, Cashman KD, Kiely M. An updated systematic review and meta-analysis of the efficacy of vitamin D food fortification. *J Nutr.* (2012) 142:1102–8. doi: 10.3945/jn.112.158014
15. Itkonen ST, Erkkola M, Lamberg-Allardt CJE. Vitamin D fortification of fluid milk products and their contribution to vitamin D intake and vitamin D status in observational studies-a review. *Nutrients.* (2018) 10:1054. doi: 10.3390/nu10081054
16. Pilz S, Marz W, Cashman KD, Kiely ME, Whiting SJ, Holick MF, et al. Rationale and plan for vitamin D food fortification: a review and guidance paper. *Front Endocrinol.* (2018) 9373:e00373. doi: 10.3389/fendo.2018.00373
17. Weir RR, Johnston M, Lowis C, Fearon AM, Stewart S, Strain JJ, et al. Vitamin D3 content of cows' milk produced in Northern Ireland and its efficacy as a vehicle for vitamin D fortification: a UK model. *Int J Food Sci Nutr.* (2021) 72:447–55. doi: 10.1080/09637486.2020.1837743
18. Calame W, Street L, Hulshof T. Vitamin D serum levels in the UK population, including a mathematical approach to evaluate the impact of vitamin D fortified ready-to-eat breakfast cereals: application of the NDNS database. *Nutrients.* (2020) 12:1868. doi: 10.3390/nu12061868
19. McCourt A, McNulty BA, Walton J, O'Sullivan A. Efficacy and safety of food fortification to improve vitamin D intakes of older adults. *Nutrition.* (2020) 75:76110767. doi: 10.1016/j.nut.2020.110767
20. Cashman KD, Kazantzidis A, Webb AR, Kiely M. An integrated predictive model of population serum 25-hydroxyvitamin D for application in strategy

development for vitamin D deficiency prevention. *J Nutr.* (2015) 145:2419–25. doi: 10.3945/jn.115.217968

21. Ikonen H, Lumme J, Seppälä J, Pesonen P, Piltonen T, Järvelin M, et al. The determinants and longitudinal changes in vitamin D status in middle-age: a Northern Finland Birth Cohort 1966 study. *Eur J Nutr.* (2021) 60:4541–53. doi: 10.1007/s00394-021-02606-z
22. Jääskeläinen T, Itkonen ST, Lundqvist A, Erkkola M, Koskela T, Lakkala K, et al. The positive impact of general vitamin D food fortification policy on vitamin D status in a representative adult Finnish population: evidence from an 11-y follow-up based on standardized 25-hydroxyvitamin D data. *Am J Clin Nutr.* (2017) 105:1512–20. doi: 10.3945/ajcn.116.151415
23. Raulio S, Erlund I, Männistö S, Sarlio-Lähteenkorva S, Sundvall J, Tapanainen H, et al. Successful nutrition policy: improvement of vitamin D intake and status in Finnish adults over the last decade. *Eur J Public Health.* (2017) 27:268–73. doi: 10.1093/eurpub/ckw154
24. Neill HR, Gill CIR, McDonald EJ, McRoberts WC, Pourshahidi LK. The future is bright: Biofortification of common foods can improve vitamin D status. *Crit Rev Food Sci Nutr.* (2021) 1–17. doi: 10.1080/10408398.2021.1950609. [Epub ahead of print].
25. Public Health England. *McCance and Widdowson's Composition of Foods Integrated Dataset* (2021). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/971018/McCance_Widdowsons_Composition_of_Foods_Integrated_Dataset_2021.xlsx (accessed July 5, 2021).
26. Public Health England. *NDNS: Results from Years 9 to 11 (Combined) – Statistical Summary* (2020). Available online at: https://www.gov.uk/government/publications/ndns-results-from-years-9-to-11-2016-to-2017-and-2018-to-2019/ndns-results-from-years-9-to-11-combined-statistical-summary#contents (accessed December 18, 2020).
27. Public Health England. *NDNS Results from Years 7 and 8 (Combined): Data Tables* (2018). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/699242/NDNS_yr_7_to_8_statistics.xlsx (accessed August 12, 2021).
28. Food and Agriculture Organization of the United Nations, (FAO). *Sources of Meat* (2014). Available online at: http://www.fao.org/ag/againfo/themes/en/meat/backgr_sources.html#:~:text=Pork%20is%20the%20most%20widely,35%25%20and%2022%25%20respectively (accessed December 18, 2020).
29. European Food Safety Authority. *Opinion of the Scientific Panel on Additives and Products or Substances Used in Animal Feed (FEEDAP) on the Evaluation of Safety and Efficacy of "Hy•D" (Calcifediol), based on 25-hydroxylcholecalciferol/25-hydroxy-pre-cholecalciferol, as feed additive in accordance with Council Directive 70/524/EEC.* Available online at: https://www.efsa.europa.eu/en/efsajournal/pub/224 (accessed October 8, 2021).
30. Grønborg I, Tetens I, Ege M, Christensen T, Andersen E, Andersen R. Modelling of adequate and safe vitamin D intake in Danish women using different fortification and supplementation scenarios to inform fortification policies. *Eur J Nutr.* (2019) 58:227–32. doi: 10.1007/s00394-017-1586-9
31. Allen RE, Dangour AD, Tedstone AE, Chalabi Z. Does fortification of staple foods improve vitamin D intakes and status of groups at risk of deficiency? A United Kingdom modeling study. *Am J Clin Nutr.* (2015) 102:338–44. doi: 10.3945/ajcn.115.107409
32. Public Health England. *NDNS: Results from Years 9 to 11 (Combined) – Appendices* (2020). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943627/NDNS_results_from_years_9_to_11_combined_appendices__2_.zip (accessed December 18, 2020).
33. Public Health England. *NDNS Results from Years 7 and 8 (Combined): Appendices* (2018). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/700933/NDNS_appendices.zip (accessed December 18, 2020).
34. Public Health England. *NDNS Results from Years 5 and 6 Combined: Appendices and Tables* (2016). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/551420/NDNS_Years_5_and_6_appendices_and_tables.zip (accessed December 18, 2020).
35. Public Health England. *National Diet and Nutrition Survey: Appendices and Tables* (2014). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/594406/NDNS_Appendices_and_tables.zip (accessed December 18, 2020).
36. Bates B, Lennox A, Prentice A, Bates C, Page P, Nicholson S, et al. *National Diet and Nutrition Survey Results from Years 1, 2, 3 and 4 (combined) of the Rolling Programme (2008/2009–2011/2012)* (2014). p. 41–55. Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/594361/NDNS_Y1_to_4_UK_report_full_text_revised_February_2017.pdf. (accessed July 5, 2021).
37. NatCen Social Research, MRC Elsie Widdowson Laboratory. *National Diet and Nutrition Survey Years 1-9, 2008/09-2016/17.* 15th ed. UK Data Service. SN: 6533 (2019). Available online at: http://doi.org/10.5255/UKDA-SN-6533-18 (accessed April 02, 2020).
38. Neill HR, Gill CIR, McDonald EJ, McRoberts WC, Rosbotham EJ, Boland R, et al. Improving vitamin D content in pork meat by UV bio-enrichment. *Proc Nutr Soc.* (2021) 80:2639. doi: 10.1017/S0029665121002639
39. Larson-Meyer DE, Ingold BC, Fensterseifer SR, Austin KJ, Wechsler PJ, Hollis BW, et al. Sun exposure in pigs increases the vitamin D nutritional quality of pork. *PLoS ONE.* (2017) 12:e0187877. doi: 10.1371/journal.pone.0187877
40. Public Health England. *NDNS: Results from Years 9 to 11 (Combined) – Data Tables* (2020). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943121/NDNS_from_years_1_to_9_data_tables.zip (accessed December 14, 2020).
41. Hamulka J, Jeruszka-Bielak M, Górnicka M, Drywień ME, Zielinska-Pukos MA. Dietary supplements during COVID-19 outbreak. Results of Google trends analysis supported by PLifeCOVID-19 Online Studies. *Nutrients.* (2020) 13:54. doi: 10.3390/nu13010054
42. Barnkob LL, Argyraki A, Jakobsen J. Naturally enhanced eggs as a source of vitamin D: a review. *Trends Food Sci Technol.* (2020) 102:62–70. doi: 10.1016/j.tifs.2020.05.018
43. Duffy SK, Kelly AK, Rajauria G, O'Doherty JV. Biofortification of meat with vitamin D. *CAB Rev.* (2018) 13:1–11. doi: 10.1079/PAVSNNR201813045
44. Guo J, Lovegrove JA, Givens DI. 25(OH)D3-enriched or fortified foods are more efficient at tackling inadequate vitamin D status than vitamin D3. *Proc Nutr Soc.* (2018) 77:282–91. doi: 10.1017/S0029665117004062
45. Medawar E, Huhn S, Villringer A, Veronica Witte A. The effects of plant-based diets on the body and the brain: a systematic review. *Transl Psychiatry.* (2019) 9:226–17. doi: 10.1038/s41398-019-0552-0
46. Tripkovic L, Lambert H, Hart K, Smith CP, Bucca G, Penson S, et al. Comparison of vitamin D2 and vitamin D3 supplementation in raising serum 25-hydroxyvitamin D status: a systematic review and meta-analysis. *Am J Clin Nutr.* (2012) 95:1357–64. doi: 10.3945/ajcn.111.031070
47. Wilson LR, Tripkovic L, Hart KH, Lanham-New SA. Vitamin D deficiency as a public health issue: using vitamin D2 or vitamin D3 in future fortification strategies. *Proc Nutr Soc.* (2017) 76:392–9. doi: 10.1017/S0029665117000349
48. Moyersoen I, Devleesschauwer B, Dekkers A, Verkaik-Kloosterman J, De Ridder K, Vandevijvere S, et al. A novel approach to optimize vitamin D intake in Belgium through fortification based on representative food consumption data. *J Nutr.* (2019) 149:1852–62. doi: 10.1093/jn/nxz119
49. Jayaratne N, Hughes MCB, Ibiebele TI, van den Akker S, van der Pols, Jolieke C. Vitamin D intake in Australian adults and the modeled effects of milk and breakfast cereal fortification. *Nutrition.* (2013) 29:1048–53. doi: 10.1016/j.nut.2013.02.011
50. Biancuzzo RM, Young A, Bibuld D, Cai MH, Winter MR, Klein EK, et al. Fortification of orange juice with vitamin D(2) or vitamin D(3) is as effective as an oral supplement in maintaining vitamin D status in adults. *Am J Clin Nutr.* (2010) 91:1621–6. doi: 10.3945/ajcn.2009.27972
51. Ikonen I, Sotgiu F, Aydinli A, Verlegh PWJ. Consumer effects of front-of-package nutrition labeling: an interdisciplinary meta-analysis. *J Acad Mark Sci.* (2020) 48:360–83. doi: 10.1007/s11747-019-00663-9
52. van buul VJ, Brouns, FJPH. Nutrition and health claims as marketing tools. *Crit Rev Food Sci Nutr.* (2015) 55:1552–60. doi: 10.1080/10408398.2012.754738
53. Volkova E, Ni Mhurchu C. The influence of nutrition labeling and point-of-purchase information on food behaviours. *Curr Obes Rep.* (2015) 4:19–29. doi: 10.1007/s13679-014-0135-6

54. Heald AH, Anderson SG, Scargill JJ, Short A, Holland D, Khan A, et al. Social disadvantage is associated with lower vitamin D levels in older people and there is no surrogate for its measurement. *Gerontol Geriatr Med*. (2017) 3:7843. doi: 10.1177/2333721417697843
55. Naugler C, Zhang J, Henne D, Woods P, Hemmelgarn BR. Association of vitamin D status with socio-demographic factors in Calgary, Alberta: an ecological study using Census Canada data. *BMC Public Health*. (2013) 13:316. doi: 10.1186/1471-2458-13-316
56. Bikle D. Vitamin D: production, metabolism, and mechanisms of action. In: Feingold KR, Anawalt B, Boyce A, et al., editors. *Endotext*. South Dartmouth (MA): MDText.com, Inc. Available online at: https://www.ncbi.nlm.nih.gov/books/NBK278935/ (accessed December 20, 2020).
57. Maki KC, Slavin JL, Rains TM, Kris-Etherton PM. Limitations of observational evidence: implications for evidence-based dietary recommendations. *Adv Nutr*. (2014) 5:7–15. doi: 10.3945/an.113.004929
58. Gibney M, Allison D, Bier D, Dwyer J. Uncertainty in human nutrition research. *Nat Food*. (2020) 1:247–9. doi: 10.1038/s43016-020-0073-2
59. Schmid A, Walther B. Natural vitamin D content in animal products. *Adv Nutr*. (2013) 4:453–62. doi: 10.3945/an.113.003780
60. Clausen I, Jakobsen J, Leth T, Ovesen L. Vitamin D3 and 25-hydroxyvitamin D3 in raw and cooked pork cuts. *J Food Compos Anal*. (2003) 16:575–85. doi: 10.1016/S0889-1575(03)00064-4
61. Cashman KD, Seamans KM, Lucey AJ, Stöcklin E, Weber P, Kiely M, et al. Relative effectiveness of oral 25-hydroxyvitamin D3 and vitamin D3 in raising wintertime serum 25-hydroxyvitamin D in older adults. *Am J Clin Nutr*. (2012) 95:1350–6. doi: 10.3945/ajcn.111.031427
62. Alshahrani F, Aljohani N. Vitamin D: deficiency, sufficiency and toxicity. *Nutrients*. (2013) 5:3605–16. doi: 10.3390/nu5093605


**Conflict of Interest:** EJM was employed by the industrial partner Devenish Nutrition Ltd.

The authors declare that this study received funding from Devenish Nutrition Ltd., in the form of a CAST PhD Studentship awarded to the lead author. The funder had the following involvement in the study: review and approval of the final manuscript.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 09 December 2021
doi: 10.3389/fnut.2021.753059

# Development of a Vitamin K Database for Commercially Available Food in Australia

*Claire R. Palmer[1,2], Henrietta Koch[2], Sujata Shinde[3], Lauren C. Blekkenhorst[1,3], Joshua R. Lewis[1,3,4], Kevin D. Croft[2], Jonathan M. Hodgson[1,3] and Marc Sim[1,3]**

[1] Institute for Nutrition Research, School of Health and Medical Sciences, Edith Cowan University, Perth, WA, Australia, [2] School of Biomedical Sciences, The University of Western Australia, Perth, WA, Australia, [3] Medical School, Faculty of Health and Medical Sciences, The University of Western Australia, Perth, WA, Australia, [4] Centre for Kidney Research, Children's Hospital at Westmead, School of Public Health, Sydney Medical School, The University of Sydney, Sydney, NSW, Australia

Vitamin K content of foods is known to vary substantially by geographical location. In Australia, no Vitamin K database of food exists, thereby creating ambiguity when trying to develop national dietary intake guidelines. This investigation aimed to develop a Vitamin K database for commonly consumed foods that are commercially available in Australian supermarkets. The Vitamin K1 (phylloquinone; PK) and K2 (menaquinone; MK4, MK7) content of 60 foods known to contain Vitamin K were assessed (e.g., vegetables fruits, oils, animal products, dairy and fermented foods). A liquid chromatography with tandem mass spectrometry (LCMS/MS) method was developed and used to measure PK and MKs in different foods with an improved chromatographic separation and detection of Vitamin K's and their analogs. The LOD and LOQ for PK and MK4 was 0.1, 0.5 ng/ml and 0.5, 1.0 ng/ml, respectively. The majority foods contained detectable PK (53/60), about half contained MK4 (31/60), and few contained MK7 (3/60). PK was highest in green leafy vegetables, with moderate amounts in oils. Highest MK4 content was in chicken eggs and meat products such as ham and chicken. This database enables nutritional epidemiologist to estimate dietary Vitamin K intake, especially in Australian cohorts, for a range of health outcomes.

**Keywords: food analysis, food composition, food database, Vitamin K1, Vitamin K2, phylloquinone, menaquinone**

OPEN ACCESS

**Edited by:**
*Massimo Lucarini,*
*Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Jun Wu,*
*Harbin Engineering University, China*
*Jolanta Kumirska,*
*University of Gdansk, Poland*

***Correspondence:**
*Marc Sim*
*marc.sim@ecu.edu.au*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** *04 August 2021*
**Accepted:** *05 November 2021*
**Published:** *09 December 2021*

**Citation:**
*Palmer CR, Koch H, Shinde S, Blekkenhorst LC, Lewis JR, Croft KD, Hodgson JM and Sim M (2021) Development of a Vitamin K Database for Commercially Available Food in Australia. Front. Nutr. 8:753059. doi: 10.3389/fnut.2021.753059*

## INTRODUCTION

Vitamin K refers to a group of fat-soluble vitamins best known for their role in blood coagulation. Other biological processes that Vitamin K has been implicated include blood calcium regulation, vascular anti-calcification, and bone metabolism (1). There are two main forms of Vitamin K; Vitamin $K_1$ (phylloquinone; PK) and Vitamin $K_2$ (menaquinones; MK). Phylloquinone is most abundant in green leafy vegetables and their oils, but is also present in smaller concentrations in the majority of food groups such as fruit, meat, and dairy products (2). In contrast to PK, the MKs are a group of isoprenologs where the side chain varies by the number of isoprenoid units ranging from four to thirteen repeats (MK4 to MK13; see **Supplementary Figure 1**) (3). Despite having analogous structures, the origins of menaquinones differ. MK4 is synthesized from PK by animals, and thus is obtained in the diet from animal products (3). All other MK are synthesized by anaerobic bacteria and thus are found in fermented foods, such as cheese (4).

Presently, there are several databases listing the PK content in a range of foods, especially vegetables and fruits. The most extensive of these databases has been developed by the United States Department of Agriculture (USDA) that has recently been updated (5). The USDA National Nutrient Database has recently been combined with the United Kingdom Composition of Foods Integrated Dataset (COFID) to provide an update to Vitamin K values in the Irish Food Composition Database (6). Although MK4 content of some food are listed, the USDA database does not include other MKs (MK5-12). Generally, such MKs are found in much lower quantities in meats and cheese compared to MK4, with the exception of MK7 known to be high in fermented foods such as Natto (7). Besides the USDA, MK reference ranges for food products are also limited to a few individual investigations (4, 7–10).

There are numerous methods to measure both PK and MK in biological matrices, as reviewed previously (11, 12). However, there is presently no standardized method to quantitate PK and MK in food. This has likely contributed toward a lack of detailed information on the Vitamin K content of food, which is reported to vary by regions including Europe, Asia and the United States of America by up to 50% (13). In Australia, a Vitamin K database for commonly consumed foods does not exist. This can limit nutritional epidemiology when researchers seek to investigate the potential health benefits of Vitamin K (PK and/or MK) for a range of health outcomes in Australian cohorts. Since MKs are estimated to constitute about 10% of total dietary Vitamin K intake, with the majority (up to 40%) of dietary MK attributed to MK4 (3, 14), the aim of this investigation was to develop a preliminary Vitamin K database, assessing PK and MK4 for commonly consumed commercially available foods in Australia.

# METHODS

## Database Creation

Food items were primarily selected based on the foods assessed in a commonly used Australian food frequency questionnaire (FFQ) developed by the Cancer Council of Victoria (DQES V2) (15, 16). This FFQ is designed to cover dietary intake over a period of 12 months, a timeframe thought to represent the "usual" diet. From the 101 foods and beverages (including alcohol) recorded in the FFQ, this was subsequently condensed to 60 food items which were known to contain PK and MK4 or MK7, based on previous work (8–10, 17). The main categories of food groups considered included vegetables ($n = 20$), fruits ($n = 3$), oils ($n = 4$), animal products ($n = 16$), dairy ($n = 14$) and fermented foods ($n = 3$). This list also included food items unique to Australia [kangaroo, yeast extract spread (Vegemite)]. It was beyond the scope of the current investigation to explore all other MKs in the selected foods. As previously highlighted, MKs are estimated to constitute about 10% of total dietary Vitamin K intake, with the majority (up to 40%) of dietary MK attributed to MK4 (3, 14). Hence, we were primarily concerned with assessing MK4 content in food. Nevertheless, we also quantified MK7 which is known to be present in food such as fermented vegetables and dairy.

Three leading local supermarket franchises were visited to obtain one sample (or brand) of food item from each store. This resulted in up to three samples per food item obtained for analysis, depending on availability. Vegetable, fruit, and meat products were processed in a food processor (Multichopper, Sunbeam Australia) prior to storage. The food samples were stored in 5 ml freezer tubes at −80°C until extraction. Every food item obtained was subsequently analyzed in duplicate. For example, if three separate brands of full fat milk were obtained, a total of six samples would have been analyzed. Once analyzed, the median value was calculated to provide an estimate of the content of Vitamin K in the analyzed food item.

### Chemicals and Reagents

Chemicals used for extraction were of HPLC grade and solvents used for chromatography were of LCMS grade. Chemicals: dichloromethane (VWR International Ltd, Tingalpa, QLD), absolute ethanol (VWR International Ltd, Tingalpa, QLD), ammonium formate (Sigma Aldrich, Castle Hill, NSW), formic acid (Univar, Sydney, NSW) n-hexane (Fisher Scientific, Loughborough, England), methanol (Fisher Scientific, Loughborough, England), isopropanol (Fisher Scientific, Loughborough, England) and diethyl ether (Sigma Aldrich, Castle Hill, NSW) were all used as received.

Deuterium-labeled Vitamin K standards: PK-$d_7$ (5,6,7,8-$d_4$,2-methyl-$d_3$), MK4-$d_7$ (MK4)-(5,6,7,8-$d_4$,2-methyl-$d_3$) and MK7-$d_7$ (MK7)-(5,6,7,8-$d_4$,2-methyl-$d_3$) were purchased from Sigma Aldrich (St Louis, USA), each with chemical purity of ≥95%, and ≥98 atom % deuterium. Non-deuterated PK and MK4 were purchased from Sigma Aldrich (St Louis, USA). Non-deuterated MK7 was unavailable to our lab.

### Vitamin K Extraction

The method of Vitamin K extraction was adapted from prior published methods (18–20). All procedures were performed under yellow light to reduce the photo-oxidation of Vitamin K by UV light. Approximately 0.2 g of the food was weighed into 15 ml screw-capped polyethylene centrifuge tubes and spiked with PK-$d_7$ (10 ng), MK4-$d_7$ (10 ng) and MK7-$d_7$ (15 ng). Proteins were denatured with ethanol (1 ml), followed by a 1 min wait period. Extraction was performed with hexane (2 ml) and Millipore $H_2O$ (1 ml). The samples were vortexed (1 min) and shaken (3 min), followed by further agitation in an ultrasound bath (10 min), and gyratory mixer (20 min). Meat samples were also sonicated for 30 s, to further homogenize the sample (Branson Ultrasonics Corporation, Sonicator 150). Samples were centrifuged at 1,800×g for 10 min at 25°C. The upper organic layer was removed and purified on 3 ml silica columns (Agilent Technologies Bond Elute, Mulgrave, VIC) according to Tarvainen et al. (21). Since Vitamin K is fat soluble, foods with a high fat content used a high capacity column (Agilent Technologies, Bond Elute silica, 10 g, 60 ml, 120 μm, Mulgrave, VIC), subsequently, extraction volumes were doubled, and the columns were preconditioned with 60 ml diethyl ether, washed with 60 ml hexane, and eluted with 54 ml of hexane: diethyl ether (3.5: 96.5%). The purified solution was collected into glass tubes then dried under nitrogen at 50°C. The residue was dissolved in dichloromethane (20 μl), dried under nitrogen, and then heated for 10 mins at 60°C to remove any residual solvent. Samples were then reconstituted in isopropanol (200 μl) and placed in amber vials for analysis by LCMS/MS.

### Liquid Chromatography With Tandem Mass Spectrometry (LCMS/MS) Parameters

The extracted phylloquinone and menaquinones were analyzed on a Thermo Scientific TSQ Quantum Ultra Triple Quadrupole mass spectrometer equipped with a heated ESI source. It was operated in the positive ion mode and connected to an Accela Autosampler system. Chromatographic separation was performed on a reverse-phase Accucore PFP HPLC column (2.6 μm, 100 × 2.1 mm). The mobile phase consisted of methanol containing 0.1% formic acid (solvent A) and 5 mM ammonium formate with 0.1% formic acid (solvent B). Column and tray temperature was maintained at 40 and 35°C, respectively. The run time for the LC method was 6 mins and the solvent gradient conditions were as follows, 90% A at 0 time, isocratic at 90% A for 1 mins, 100% A at 1.01 mins and held at 100% A to 5 mins, then reduced to 90% A at 5.01 mins and held at 90% A for 6 mins to equilibrate to starting conditions. The flow rate was 0.5 ml/min, with a 10 μl injection volume. The mass spectrometer was used in the multiple reaction monitoring mode with argon as the collision gas to detect PK, PK-$d_7$, MK4, MK4-$d_7$, MK7 and MK7-$d_7$, with MS parameters listed in **Table 1**. For quantitative analysis, ratios of area of peak for PK and MK4, to their deuterated standards at the respective RTs, to known amount of added internal standards, PK-$d_7$ and MK4-$d_7$ were used to calculate concentration in μg and expressed as per 100 g of food matrix. The concentration of MK7 in food samples was determined by measuring the area under the peak of non-deuterated precursor ion, transition *m/z* 650 > product ion *m/z* 187 based on the RT of the deuterated internal standard, MK7-$d_7$, transition *m/z* 657.1 > *m/z* 194.2 (**Supplementary Figure 2**). The ratio of area of peaks for MK7 and MK7-$d_7$, assuming a 1:1 response, to a known amount of MK7-$d_7$ (15 ng) added during the analysis was used to calculate the concentration of MK7 in μg and expressed as per 100 g of food matrix.

**TABLE 1** | MRM transitions and RT for deuterated and non-deuterated phylloquinone (PK), menaquinone-4 (MK4), and menaquinone-7 (MK7).

| | Precursor ion *m/z* | Product ion *m/z* | Collision Energy eV | RT mins |
|---|---|---|---|---|
| PK | 451.0 | 187.0 | 25 | 1.90 |
| PK-$d_7$ | 458.4 | 194.2 | 23 | 1.91 |
| MK4 | 445.0 | 187.0 | 20 | 1.53 |
| MK4-$d_7$ | 452.0 | 194.0 | 25 | 1.57 |
| MK7 | 650.0 | 187.0 | 26 | - |
| MK7-$d_7$ | 657.1 | 194.2 | 28 | 2.30 |

*RT, retention time.*

### Method Validation

Linearity, limit of detection (LOD), and limit of quantification (LOQ) were determined for the vitamin PK and MK4 by using incrementally diluted calibration standards ranging from 0.01 to 500 ng/ml. Linear regression analysis was conducted using the ratio of the K vitamin to deuterated internal standard (PK-d7 or MK4-d7) as a function of concentration of the K vitamin. The $R^2$ for both PK and MK4 was 0.9987 (**Supplementary Figure 3**). LOD and LOQ was determined in accordance with the US Food and Drug Administration guidelines (22). The LOD and LOQ for PK was 0.1 and 0.5 ng/ml. respectively. The LOD and LOQ of MK4 was 0.5 and 1 ng/ml, respectively. Precision was determined by calculating the % CV of PK and MK4 concentrations in standard solutions and was repeated over three consecutive days to determine intra- and inter-assay variability. Intra and inter-assay variability for phylloquinone was 10.6 and 12.8%, respectively. Intra and inter-assay variability for MK4 was 3.9 and 10.1%, respectively.

Due to the wide range of food matrices investigated precision was not determined in all matrices. As such, a simulated matrix was used to determine recovery. Recovery was determined for PK and MK4 by spiking a low-fat milk matrix with known amounts of PK, PK-d7, MK4, and MK4-d7. In two tubes a known amount of the standards was added at the start of the extraction process and in three separate tubes the same amount of standard was added at the end of the extraction process, with the recovery determined by comparing the measured concentrations of PK and MK4 in samples spiked before extraction to those spiked following purification. Recovery for PK was 41.3% and for MK4 was 43.4%; the recovery for PK-d7 and MK4-d7 was 37.0 and 40.2%, respectfully, resulting in approximately equal ratios. MK7 was not included in these experiments due to lack of a MK7 standard.

# RESULTS

## Food Database

The PK, MK4, and MK7 content for the individual foods analyzed are displayed in **Table 2.** All PK and MK4 measurements were higher than the LOQ. The selective reaction monitoring chromatograms of PK, MK4 and MK-7 in selected foods (including spinach, cheese and natto, respectively) are presented in **Supplementary Figure 2**. The majority of assessed foods contained detectable PK (53/60), just over half contained MK4 (31/60), and few contained MK7 (3/60).

Phylloquinone was detected in all vegetables, being highest in green leafy vegetables such as spinach (median 263.0 μg/100 g)

**TABLE 2 |** Median (range) values of the phylloquinone and menaquinone content of commonly consumed foods obtained from Australian supermarkets.

| Food item | Phylloquinone (μg/100 g) | MK4 (μg/100 g) | MK7 (μg/100 g) |
|---|---|---|---|
| **Vegetables** | | | |
| Spinach | 262.9 (244.7–286.2) | ND | ND |
| Kale | 128.5 (92.5–272.6) | ND | ND |
| Cabbage | 70.4 (28.6–94.4) | ND | ND |
| Broccoli | 67.9 (50.9–92.9) | ND | ND |
| Carrot | 36.4 (10.0–67.5) | ND | ND |
| Green beans | 32.5 (29.1–46.7) | ND | ND |
| Cucumber | 26.4 (17.1–32.0) | ND | ND |
| Lettuce | 25.7 (10.3–43.2) | ND | ND |
| Peas | 22.7 (17.9–31.4) | ND | ND |
| Kidney beans | 21.0 (5.4–29.7) | ND | ND |
| Zucchini | 18.3 (12.8–24.7) | ND | ND |
| Celery | 17.3 (13.6–20.8) | ND | ND |
| Cauliflower# | 16.4 (12.5–20.1) | ND | ND |
| Pumpkin | 14.8 (3.3–19.3) | ND | ND |
| Capsicum | 9.55 (8.2–14.1) | ND | ND |
| Tofu# | 7.0 (5.9–8.1) | ND | ND |
| Tomato | 6.0 (4.4–7.3) | ND | ND |
| Bean shoots* | 5.4 (5.1–5.7) | ND | ND |
| Potato | 0.5 (0.2–0.7) | ND | ND |
| **Fruits and nuts** | | | |
| Avocado | 23.6 (16.2–31.4) | ND | ND |
| Cashews | 12.3 (5.4–14.0) | ND | ND |
| Pear | 1.8 (1.4–2.7) | ND | ND |
| Apple | 1.4 (1.2–2.2) | ND | ND |
| **Oils** | | | |
| Canola oil | 75.0 (64.6–121.4) | ND | ND |
| Margarine | 69.2 (53.4–80.1) | 6.41 (3.9–8.1) | 3.5 (0.0–37.8) |
| Olive oil | 43.6 (38.6–45.1) | ND | ND |
| Butter | 5.10 (3.5–5.6) | 24.69 (30.9–30.4) | ND |
| **Animal products** | | | |
| Egg | 3.34 (1.9–4.4) | 32.61 (22.5–37.9) | ND |
| Ham | ND | 28.78 (22.1–45.7) | ND |
| Chicken | 0.3 (0.2–0.6) | 26.4 (18.5–31.8) | ND |
| Salami | 0.3 (0.0–0.9) | 18.18 (16.0–42.0) | ND |
| Pork | ND | 15.9 (8.5–19.1) | ND |
| Beef | 0.6 (0.4–1.6) | 15.6 (5.9–21.6) | ND |
| Beef mince | 1.2 (0.5–1.8) | 13.7 (12.2–17.8) | ND |
| Bacon | ND | 12.72 (9.5–13.0) | ND |
| Lamb | 0.4 (0.3–0.5) | 11.6 (7.8–14.9) | ND |
| Sausage (beef) | 0.3 (0.2–0.7) | 3.0 (2.1–9.7) | ND |
| Barramundi | 0.2 (0.0–0.4) | 2.16 (1.5–2.7) | ND |
| Tinned salmon | 0.4 (0.3–0.7) | 2.0 (1.4–4.4) | ND |
| Tinned tuna | ND | 1.4 (1.1–2.0) | ND |
| Kangaroo* | 0.17 (0.13–0.21) | 1.1 (1.0–1.2) | ND |
| Veal* | 0.1 (0.1–0.2) | 1.0 (0.9–1.1) | ND |
| Snapper# | ND | 0.8 (0.7–1.1) | ND |
| **Dairy** | | | |
| Thickened cream | ND | 20.1 (18.2–29.8) | 0.9 (0.4–3.1) |
| Sour cream | 3.0 (2.6–4.4) | 12.72 (10.9–16.4) | ND |
| Brie | 1.84 (1.1–2.9) | 9.34 (3.7–11.3) | ND |
| Cheddar | 1.1 (0.8–1.2) | 5.5 (4.6–5.9) | ND |
| Parmesan | 0.7 (0.4–1.0) | 5.4 (3.3–8.0) | ND |
| Cream cheese | 1.3 (0.7–1.9) | 5.0 (3.4–10.8) | ND |
| Low fat cheddar | 0.7 (0.5–0.9) | 4.1 (3.4–4.7) | ND |
| Ice cream | 2.3 (0.0–5.3) | 4.07 (1.8–6.7) | ND |
| Greek yogurt | 0.6 (0.0–1.2) | 3.5 (0.0–5.1) | ND |
| Full fat milk | 1.3 (1.1–2.7) | 1.21 (1.0–1.8) | ND |
| Yogurt | 0.3 (0.2–0.5) | 1.0 (0.5–3.0) | ND |
| Cottage cheese | 0.4 (0.3–0.5) | 0.97 (0.6–1.1) | ND |
| Reduced fat milk | 0.21 (0.0–0.3) | ND | ND |
| Skim milk | ND | ND | ND |
| **Fermented foods** | | | |
| Sauerkraut* | 7.0 (3.2–10.7) | ND | ND |
| Natto* | 6.2 (6.0–6.4) | ND | 81.6 (68.8–94.5) |
| Yeast extract* spread (vegemite) | 2.34 (1.9–2.8) | (1.4–8.6) | ND |

*(Continued)*

**TABLE 2 |** Continued

**Indicates mean value; ND, not detected. Three separate samples were analyzed for all food items apart from those indicated by # or *, where two or a single food item was analyzed, respectively. Each sample of food was also assessed in duplicate. Where only one food sample was analyzed (in duplicate), the mean and range has been presented instead.*

and kale (median 128.5 μg/100 g) (**Figure 1**). Phylloquinone was of lower abundance (<25 μg/100 g) in fruit. MK4 and MK7 were not detected in any non-fermented vegetable or fruit. Japanese fermented soybeans (Natto), contained the highest content of MK7 (mean 81.6 μg/100 g).

Vegetable oils contained moderate amounts of PK. Noteworthy, margarine contained moderate amounts of PK, MK4 and in some samples MK7. Butter contained smaller amounts of PK (5.1 μg/100 g), but a high abundance of MK4 (24.7 μg/100 g).

The majority of animal products, with exception of reduced fat and skim milk, contained MK4 (**Figure 2**). The highest MK4 content was in chicken eggs (32.61 μg/100 g) and meat products such as ham (28.8 μg/100 g) and chicken (26.4 μg/100 g). MK4 content was low in fish products (<5.0 μg/100 g) and kangaroo meat (1.1 μg/100 g). The PK content of all animal products was minimal (<6.0 μg/100 g). Finally, only a small amount of MK7 was detected in thickened cream (0.9 μg/100 g).

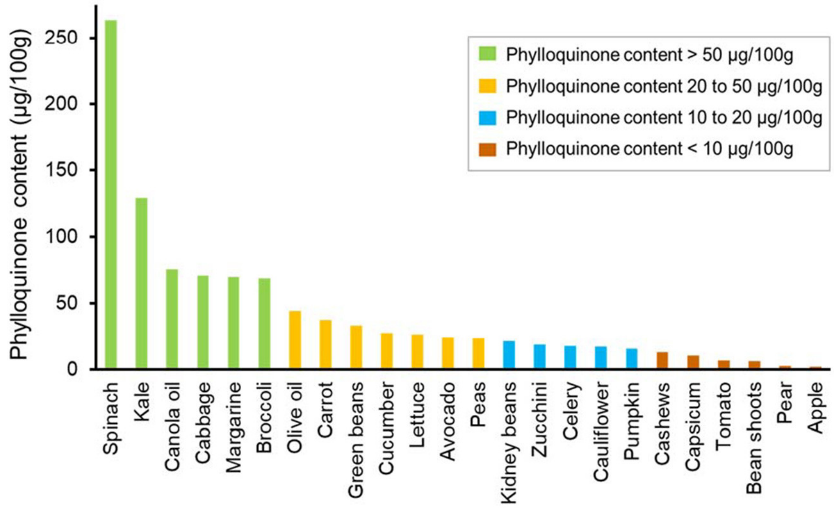


**FIGURE 1 |** Median phylloquinone content of individual vegetables, vegetable oils, and fruits from highest to lowest.

## DISCUSSION

For the first time, we present an Australian food composition database for Vitamin K. Specifically, we present the PK, MK4 and MK7 content of 60 commonly consumed commercially available foods. Food items analyzed comprised a wide range of food groups known to provide the majority of dietary Vitamin K1 and K2, including vegetables, oils, meat, and dairy. We were able to accomplish this by adopting an LCMS/MS method to assess the Vitamin K content of these food items. Findings of this work may help with the revaluation of Vitamin K intake guidelines in Australia where an adequate intake of 60 and 70 µg/day for men and women, respectively, is proposed (23), which are substantially lower than the USA (120 µg/day for men, 90 µg/day for women) (24).

Similar to previous work, vegetables were found to be a major source of PK. We report that specific vegetables including spinach and kale as the richest sources of PK (263, 129 µg/100 g, respectively). Noteworthy, the PK content of these vegetables differed substantially compared to previous work. For example, Schurgers and Vermeer report that PK in spinach and kale to be 387 and 817 µg/100 g. Alternatively, the USDA report these measurements to be 483 and 390 µg/100 g, respectively. Compared to these studies, PK content in spinach we measured varied between 33 and 45%. For kale, the PK content was approximately 6 times lower compared to Schurgers and Vermeer (8), and 3 times lower compared to the USDA. Nevertheless, it has been suggested that the average PK content in "green vegetables" to be in the range of 100–750 µg/100 g (14). Although canola oil and margarine were shown to contain a high-moderate amount of PK, these foods are typically consumed in quantities less than 100 g/day. Therefore, such foods may not be a large contributor toward overall dietary Vitamin K intake. Collectively, these results highlight that the importance of accounting for regional differences in PK content of food to reduce inaccuracies when trying to estimate PK intake for specific populations. Noteworthy, the Vitamin K content of fruits are reported to be low (25, 26), a finding we also observed in apples and pears in this study. Although it would have been ideal to assess the Vitamin K content of more fruits, due to limited resources, we selected some of the most commonly consumed fruits previously shown to contain Vitamin K.

Others including ourselves have now demonstrated that the MK4 in foods also differ between regions (13). Specifically, MK4 content of beef cuts from the USA (1.1–9.3 µg/100 g) was substantially lower compared to Japan (15.0 ± 7.0 µg/100 g) (9, 18). Egg yolk MK4 content was also 4 times greater in Japan compared to the USA (64 vs. 15.5 µg/100 g, respectively). In comparison, our measurements in whole egg and beef indicated a MK4 content of 33 and 16 µg/100 g, respectively. Such regional differences may be attributed to differences in food production including the use of menadione in animal feeds (13). Hard and soft cheese are also known to be dietary sources of MKs (4, 13) that provide between 4 and 10 µg/100 g (13). These findings are consistent with our results where MK4 content of brie, cheddar and parmesan were 9, 6 and 5 µg/100 g, respectively. Although not assessed, it is worth highlighting that up to 15% of variability has been reported in the MK content (MK-6 to MK-10) of semi-hard cheese varieties from three different European countries, including France, Poland and Denmark (4). Similar to previous work, MK7 was not detected (or found in very small amounts) in cheese varieties (13), whilst margarine was found to contain almost double the amount of MK7 compared to MK4 (6 vs. 11 µg/100 g, respectively). In line with previous work, natto was found to be a rich source of MK7 (82 µg/100 g). However, this was substantially lower compared to the MK7 content previously measured (902–998 µg/100 g) (9, 13, 21).

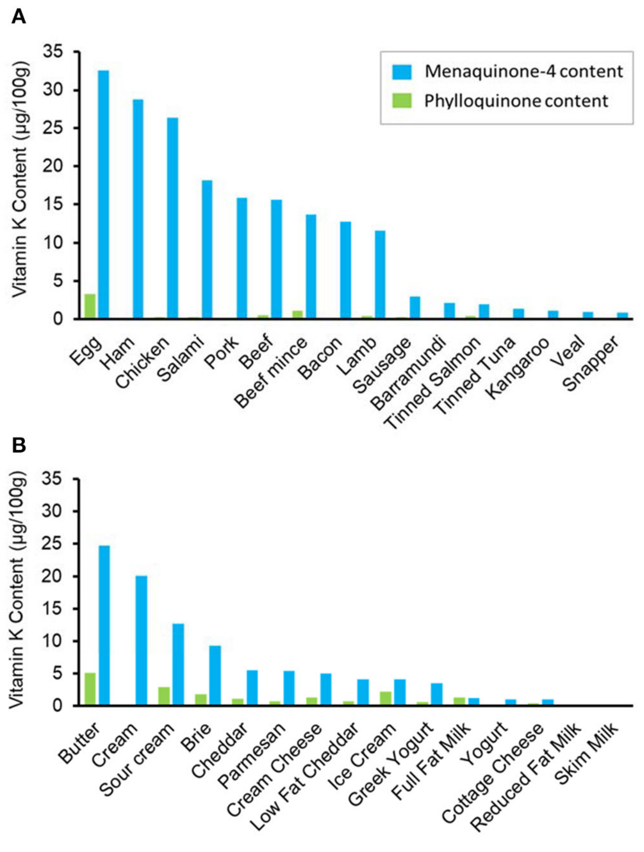


**FIGURE 2 |** Median menaquinone-4 (blue) and phylloquinone (green) content of **(A)** meat products and **(B)** dairy products. Presented from highest to lowest menaquinone content.

In order to create food databases, a validated method of measuring PK and MKs simultaneously, with accuracy and high throughput, is required. This is further complicated as the Vitamin K content in foods are typically low, in conjunction with the added complexity of extraction from the food matrix (27). We have previously provided an overview of the current methods for analyzing Vitamin K compounds in food (12). An emerging method to measure Vitamin K in foods is a HPLC-MS method developed by Karl et al. (20), which has high versatility and may be used to measure PK and all MKs in a variety of forms such as food, serum and feces. A combination of HPLC and gas chromatography-MS (GC-MS) has been used with deuterium-labeled internal standards to accurately measure PK in serum (28). However, there are limitations in using GC for measuring Vitamin K, as high temperatures are needed to volatise Vitamin K compounds (>300°C) (29). Therefore, HPLC is the preferred method to separate Vitamin K compounds for analysis. Previously, reverse-phase HPLC with fluorescent detection or HPLC–MS, with K1 or deuterated-K1 as the internal standard have been used to determine PK and MK from food (4, 20). Most recently, a whole range of fermented foods were assessed for their PK and MK content using a newly developed ultra-high performance liquid chromatography–atmospheric pressure chemical ionization tandem mass spectrometric (UHPLC-APCI-MS/MS) method (21). They reported a LOD and LOQ for PK and MK4 of 1.0, 3.2 pg and 6.4, 21.2 pg, respectively. In the present study, we used HPLC method coupled to heated electrospray ionization tandem mass spectrometry (H-ESI-MS/MS) to measure PK and MK in different foods with an improved chromatographic separation and detection of Vitamin K's and their analogs with the LOD and LOQ for PK and MK4 as 1 and 5 pg (0.1 and 0.5 ng/ml) and 5 and 10 pg (0.5 and 1 ng/ml), respectively. These values are comparable to that reported by Tarvainen et al. (21). In comparison, we have developed an efficient LCMS/MS method with a shorter chromatography run time for the quantification of PK and MK4 with the use of deuterated internal standards. Specifically, when adopting a modified LCMS/MS method here (30), we recorded faster elution of PK and MK4 retention times (1.90 and 1.53 mins), compared to previous work analyzing fermented foods (5.52 and 4.66 mins) (21) as well as human serum and plasma (12.20 and 10.41 mins) (19).

In conclusion, we have provided the Vitamin K (PK, MK4 and MK7) content of commonly consumed commercially available food products in Australia. The LCMS/MS method enabled the quantification of PK, MK4, and MK7 in these food products with acceptable LOD and LOQ and inter and intra-assay precision. In conjunction with other Vitamin K food database, this data can be used by nutritional epidemiologist seeking to quantify dietary intake, especially in Australian cohorts, for a range of health outcomes.

## DATA AVAILABILITY STATEMENT

Within reasonable request, the raw data supporting the conclusions of this article will be made available by the authors.

## AUTHOR CONTRIBUTIONS

MS, LB, JL, KC, and JH conceived and designed the study. CP and HK performed the analysis. SS and KC did the LCMS/MS method development. CP and MS prepared the manuscript. CP had the primary responsibility for the final content. All authors reviewed, read, and approved the final manuscript.

## FUNDING

This research was supported by a Royal Perth Hospital Research Foundation Springboard Grant. The salary of MS is supported by a Royal Perth Hospital Research Foundation Career Advancement Fellowship (CAF00/2020). The salary of LB is supported by a National Health and Medical Research Council (NHMRC) of Australia Emerging Leadership Investigator Grant (ID: 1172987) and a National Heart Foundation of Australia Post-Doctoral Research Fellowship (ID: 102498). The salary of JL is supported by a National Heart Foundation of Australia Future Leader Fellowship (ID: 102817). The salary of JH is supported by a National Health and Medical

Research Council of Australia Senior Research Fellowship (ID: 1116973). None of the funding agencies had any role in the conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.753059/full#supplementary-material

## REFERENCES


1. DiNicolantonio JJ, Bhutani J, O'Keefe JH. The health benefits of Vitamin K. *Open Heart.* (2015) 2:e000300. doi: 10.1136/openhrt-2015-000300
2. Thane CW, Paul AA, Bates CJ, Bolton-Smith C, Prentice A, Shearer MJ. Intake and sources of phylloquinone (Vitamin K1): variation with socio-demographic and lifestyle factors in a national sample of British elderly people. *Br J Nutr.* (2002) 87:605–13. doi: 10.1079/BJN2002583
3. Shearer MJ, Newman P. Recent trends in the metabolism and cell biology of Vitamin K with special reference to Vitamin K cycling and MK-4 biosynthesis. *J Lipid Res.* (2014) 55:345–62. doi: 10.1194/jlr.R045559
4. Manoury E, Jourdon K, Boyaval P, Fourcassie P. Quantitative measurement of Vitamin K2 (menaquinones) in various fermented dairy products using a reliable high-performance liquid chromatography method. *J Dairy Sci.* (2013) 96:1335–46. doi: 10.3168/jds.2012-5494
5. United States Department of Agriculture Agricultural Research Service. *USDA Food and Nutrient Database for Dietary Studies 2017–2018*. Food Surveys Research Group Home Page (2020). Available online at: http://www.ars.usda.gov/nea/bhnrc/fsrg. (accessed July 1, 2021).
6. Kingston C, Kehoe L, McNulty BA, Nugent AP, Cashman KD, Flynn A, et al. Updating of the Irish Food Composition Database for Vitamin K1 and Vitamin K2. *Proc Nutr Soc.* (2020) 79:OCE1. doi: 10.1017/S0029665119001459
7. Vermeer C, Raes J, Van't Hoofd C, Knapen MH, Xanthoulea S. Menaquinone content of cheese. *Nutrients.* (2018) 10:446. doi: 10.3390/nu10040446
8. Schurgers LJ, Vermeer C. Determination of phylloquinone and menaquinones in food. *Pathophysiol Haemost Thromb.* (2000) 30:298–307. doi: 10.1159/000054147
9. Kamao M, Suhara Y, Tsugawa N, Uwano M, Yamaguchi N, Uenishi K, et al. Vitamin K content of foods and dietary Vitamin K intake in Japanese young women. *J Nutr Sci Vitaminol.* (2007) 53:464–70. doi: 10.3177/jnsv.53.464
10. Fu X, Harshman SG, Shen X, Haytowitz DB, Karl JP, Wolfe BE, et al. Multiple Vitamin K forms exist in dairy foods. *Curr Dev Nutr.* (2017) 1:e000638. doi: 10.3945/cdn.117.000638
11. Zhang Y, Bala V, Mao Z, Chhonker YS, Murry DJ. A concise review of quantification methods for determination of Vitamin K in various biological matrices. *J Pharm Biomed Anal.* (2019) 169:133–41. doi: 10.1016/j.jpba.2019.03.006
12. Palmer CR, Blekkenhorst LC, Lewis JR, Ward NC, Schultz CJ, Hodgson JM, et al. Quantifying dietary Vitamin K and its link to cardiovascular health: a narrative review. *Food Funct.* (2020) 11:2826–37. doi: 10.1039/C9FO02321F
13. Walther B, Karl JP, Booth SL, Boyaval P. Menaquinones, bacteria, and the food supply: the relevance of dairy and fermented food products to Vitamin K requirements. *Adv Nutr.* (2013) 4:463–73. doi: 10.3945/an.113.003855
14. Schurgers L, Geleijnse J, Grobbee D, Pols H, Hofman A, Witteman J, et al. Nutritional intake of vitamins K1 (phylloquinone) and K2 (menaquinone) in the Netherlands. *J Nutr Environ Med.* (1999) 9:115–22. doi: 10.1080/13590849961717
15. Keogh JB, Lange K, Syrette J. Comparative analysis of two FFQ. *Public Health Nutr.* (2010) 13:1553–8. doi: 10.1017/S1368980010000066
16. Hebden L, Kostan E, O'Leary F, Hodge A, Allman-Farinelli M. Validity and reproducibility of a food frequency questionnaire as a measure of recent dietary intake in young adults. *PLoS ONE.* (2013) 8:e75156–e75156. doi: 10.1371/journal.pone.0075156
17. Westenbrink S, Jansen-van der Vliet M. *Dutch Food Composition Table.* (2006). Available online at: https://www.rivm.nl/en/dutch-foodcomposition-database (accessed March 30, 2019).
18. Elder SJ, Haytowitz DB, Howe J, Peterson JW, Booth SL. Vitamin k contents of meat. dairy, and fast food in the US Diet. *J Agric Food Chem.* (2006) 54:463–7. doi: 10.1021/jf052400h
19. Gentili A, Cafolla A, Gasperi T, Bellante S, Caretti F, Curini R, et al. Rapid, high performance method for the determination of Vitamin K1, menaquinone-4 and Vitamin K1 2,3-epoxide in human serum and plasma using liquid chromatography-hybrid quadrupole linear ion trap mass spectrometry. *J Chromatogr A.* (2014) 1338:102–10. doi: 10.1016/j.chroma.2014.02.065
20. Karl JP, Fu X, Dolnikowski GG, Saltzman E, Booth SL. Quantification of phylloquinone and menaquinones in feces, serum, and food by high-performance liquid chromatography-mass spectrometry. *J Chromatogr B.* (2014) 963:128–33. doi: 10.1016/j.jchromb.2014.05.056
21. Tarvainen M, Fabritius M, Yang B. Determination of Vitamin K composition of fermented food. *Food Chem.* (2019) 275:515–22. doi: 10.1016/j.foodchem.2018.09.136
22. U.S. Department of Health and Human Services. *Food and Drug Administration. Bioanalytical Method Validation. Guidance for Industry.* Rockville, MD (2018). Available online at: https://www.fda.gov/files/drugs/published/Bioanalytical-Method-Validation-Guidance-for-Industry.pdf (accessed July 1, 2021).
23. National Health and Medical Research Council. *Nutrient Reference Values for Australia and New Zealand: Including Recommended Dietary Intakes*. Report 1864962372. Canberra: National Health and Medical Research Council (2006).
24. Institute of Medicine. *Dietary Reference Intakes for Vitamin, A., Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc*. Washington, DC: National Academy Press (2001).
25. Dismore ML, Haytowitz DB, Gebhardt SE, Peterson JW, Booth SL. Vitamin K content of nuts and fruits in the US diet. *J Am Diet Assoc.* (2003) 103:1650–2. doi: 10.1016/j.jada.2003.09.028
26. United States Department of Agriculture. *National Nutrient Database for Standard Reference Release.* USDA (2015).
27. Zhang Y, Zhou W-E, Yan J-Q, Liu M, Zhou Y, Shen X, et al. A review of the extraction and determination methods of thirteen essential vitamins to the human body: an update from 2010. *Molecules.* (2018) 23:1484. doi: 10.3390/molecules23061484
28. Dolnikowski GG, Sun Z, Grusak MA, Peterson JW, Booth SL. HPLC and GC/MS determination of deuterated Vitamin K (phylloquinone) in human serum after ingestion of deuterium-labeled broccoli. *J Nutr Biochem.* (2002) 13:168–74. doi: 10.1016/S0955-2863(01)00210-8
29. Koivu-Tikkanen T. *Determination of Phylloquinone and Menaquinones in Foods by HPLC*. Finland: Department of Applied Chemistry and Microbiology, University of Helsinki (2001).

30. Doyle RM, Cherkassky A. *LC-MS/MS Quantitative Analysis of the Vitamin K's and their Metabolites in Serum for Research Use*. PO64920ENO. Waltham, MA: Thermo Scientific (2017).

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 15 December 2021
doi: 10.3389/fnut.2021.778369

# Heterogeneity in Meat Food Groups Can Meaningfully Alter Population-Level Intake Estimates of Red Meat and Poultry

***Lauren E. O'Connor[1]*, Kirsten A. Herrick[1], Ruth Parsons[2] and Jill Reedy[1]***

*[1] Risk Factor Assessment Branch, Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health, Rockville, MD, United States, [2] Information Management Services, Inc., Rockville, MD, United States*

OPEN ACCESS

**Edited by:**
*Alessandra Durazzo,*
*Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Ali Imran,*
*Government College University Faisalabad, Pakistan*
*Daniela Assumpção,*
*State University of Campinas, Brazil*

***Correspondence:**
*Lauren E. O'Connor*
*leoconnor24@gmail.com*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** *16 September 2021*
**Accepted:** *11 November 2021*
**Published:** *15 December 2021*

**Citation:**
*O'Connor LE, Herrick KA, Parsons R and Reedy J (2021) Heterogeneity in Meat Food Groups Can Meaningfully Alter Population-Level Intake Estimates of Red Meat and Poultry. Front. Nutr. 8:778369. doi: 10.3389/fnut.2021.778369*

Heterogeneity in meat food groups hinders interpretation of research regarding meat intake and chronic disease risk. Our objective was to investigate how heterogeneity in red meat (RM) and poultry food groups influences US population intake estimates. Based on a prior systematic review, we created an ontology of methods used to estimate RM [1= unprocessed RM; 2 (reference)= unprocessed RM + processed RM; 3= unprocessed RM + processed RM + processed poultry; and 4=unprocessed RM + processed RM + processed poultry + chicken patties/nuggets/tenders (PNT)] and three for poultry [A=unprocessed poultry; B= unprocessed poultry + PNT; C (reference)= unprocessed poultry + processed poultry + PNT). We applied methods to 2015–18 National Health and Nutrition Examination Survey data to estimate RM and poultry intake prevalence and amount. We estimated and compared intakes within RM and within poultry methods *via* the NCI Method for individuals ≥2 years old (*n* = 15,038), adjusted for age, sex, and race/Hispanic origin. We compared the population percentage that exceeded age- and sex-specific RM and poultry allotments from the Dietary Guidelines for Americans recommended eating patterns. The percent that consumed RM ranged from 47 ± 1.2% to 75 ± 0.8% across methods and mean amount ranged from 10.5 ± 0.28 to 18.2 ± 0.35 lean oz-equivalents/week; 38 ± 1.2% to 71 ± 0.7% and 9.8 ± 0.35 to 13.3 ± 0.35 lean oz-equivalents/week across poultry methods. Estimates for higher, but not lower, intake percentiles differed across RM methods. Compared to the reference, Method 1 was ≥3.0 oz-equivalents/week lower from 20th-70th percentiles, ≥6.0 oz-equivalents/week lower from 75th-90th percentiles, and ≥9.0 oz-equivalents/week lower for the 95th percentile. Method 4, but not Method 3, was ≥3.0 oz-equivalents/week higher than the reference from 50 to 95th percentiles. The population percentage that exceeded allotments was 27 ± 1.8% lower for Method 1, 9 ± 0.8% higher for Method 3, and 14 ± 0.9% higher for

Method 4 compared to the reference. Differences were less pronounced for poultry. Our analysis quantifies the magnitude of bias introduced by heterogeneous meat food group methodology. Explicit descriptions of meat food groups are important for development of dietary recommendations to ensure that research studies are compared appropriately.

**Keywords: dietary assessment, standardization, food groups, nutrition surveillance, nutrition epidemiology, U.S. populations**

## INTRODUCTION

Dietary guidance in the U.S. emphasizes adoption of food group-based dietary patterns to meet nutrient needs and prevent chronic disease risk (1). Heterogeneity in research questions, study design, sample populations, and dietary assessment methodologies precludes adoption of standardized food groups or food group lexicons (2) by nutrition researchers. Dietary assessment tools, such as food frequency questionnaires and 24-h dietary recalls, suit different research purposes and collect varying levels of detailed data on each reported food group (3). The level of detail of the dietary assessment tool(s) employed dictates how food groups can be subsequently operationalized. Even with comprehensive data collected at the individual food level, there is a lack of standardized definitions of food groups across public health and research organizations to guide researchers and the public (4, 5). These two factors contribute heterogeneity to how researchers operationalize food groups across research studies, influencing scientists' and policy makers' ability to collate and translate research into food-based dietary pattern recommendations.

The 2015–20 and 2020–25 Dietary Guidelines for Americans (DGA) scientific advisory committees noted that heterogeneity in food groups was most prominent in research about meat intake and chronic disease risk (6, 7). A systematic review showed that meat terminology used throughout chronic disease literature as well as the foods included within meat food groups differed within and between observational and experimental nutrition research studies (8). A challenge in assessing meat subgroups is that dietary assessment tools and database do not disaggregate processed meat into processed red meat and processed poultry inhibiting researchers' ability to create accurate and detailed red meat and poultry food groups. For example, this has led to processed red meat and processed poultry being omitted, thus researchers estimate intakes of unprocessed red meat and unprocessed poultry only. Or researchers have grouped all processed meat, inclusive of processed red meat and processed poultry, with unprocessed red meat, hence the "red and processed meat" food group commonly used in the literature (8). Further, about 25% of researchers don't include any description of how red meat and poultry food groups are operationalized which provide no indication of potential misclassification (9–11). One would hypothesize that variations in analytical decisions of how to operationalize food groups would meaningfully influence intake estimates because each method represents a distinct and unique food group. Therefore, the objective of this analysis was to assess how intakes of red meat and poultry and the proportion of the population below, within, and above allotments from the 2020–25 DGA recommended eating patterns differs based on the method used to operationalize red meat and poultry food groups. This analysis will aid understanding of the degree to which misclassification within meat food groups influences population-level intakes.

## METHODS

### Study Design

We used data from the 2015–16 and 2017–18 National Health and Nutrition Examination Survey (NHANES) which is conducted by the U.S. Centers for Disease Control and Prevention's National Center for Health Statistics (NCHS). NHANES uses a multistage, complex, probability sample to release health and nutrition data every 2 years that is representative of the non-institutionalized U.S. population (12). Participants are recruited for a household interview and a physical examination conducted in the NHANES Mobile Examination Centers (MEC). Survey design and analytical weighting procedures are described in detail previously (13).

### Ethics

All NHANES protocols are approved by the NCHS Research Ethics Review Board (14, 15). Participants aged ≥18 years provide written consent, and written consent is provided by a parent or guardian for those aged <18 years. Additional assent is obtained for those 7–17 years.

### Demographic Data Collection

Self-reported demographic data are collected during the at-home interview (16, 17). Demographic variables relevant to this analysis were age (≥2 years old), gender (male, female), race and Hispanic origin (Non-Hispanic White, Non-Hispanic-Black, Non-Hispanic Asian, and Hispanic), family income to poverty ratio (PIR; ≤ or >130% which is the cut off for the Supplemental Nutrition Assistance Program), educational attainment for participants 19 years old (high school or less, more than high school), and head of household educational attainment for participants 2–18 years old (high school or less and more than high school).

### Dietary Intake Data Collection

Dietary data are available from NHANES *via* a joint effort between NCHS and the US Department of Agriculture (USDA) and are referred to as the What We Eat In America (WWEIA) (18) component. Self-reported dietary data are collected in the MEC *via* trained interview administered 24-h recalls using the USDA's computer assisted Automated Multiple Pass Method (18).

Participants are asked to recall what foods, beverages, and dietary supplements they consumed the prior day and the amount they consumed. Participants ≥12 years old completed the 24-h dietary recall on their own, participants 6–11 years old were assisted by a parent or guardian, and participants ≤5 years old had a parent or guardian proxy complete the 24-h recall. A second 24-h dietary recall is conducted *via* telephone 3–10 days later. Each food and beverage reported in a 24-h dietary recall is coded to correspond to a food code in the USDA's Food and Nutrient Database for Dietary Studies (FNDDS) (19). Each food code subsequently links to the Food Patterns Equivalents Database (FPED) which disaggregates food code components into servings sizes [ounce-equivalents (oz.-eq), cup equivalents, teaspoon equivalents, or grams] of 37 distinct food pattern components (7 of which are related to meat intake) used to model the food patterns recommended in the DGA (20). Each food code is also linked to a WWEIA food category and subcategory which describes the food code "as consumed," e.g., "bacon" or "burgers" (21).

## Red Meat and Poultry Food Groups

There is heterogeneity in food group definitions across public health and research organizations (4, 5). Therefore, it is important for us to describe the definitions used for this analysis. No public health or research organization defines and describes all meat food groups needed for our purposes, so we relied on three resources: the FPED (2), the DGA (1, 22), and the American Meat Science Association (AMSA) Lexicon (23). Overall, the term "meat" refers to "skeletal muscle and associated tissues derived from mammalian, avian, reptilian, amphibian, and aquatic species harvested for human consumption" (23). The FPED defines (1) "meat," i.e., red meat, as "beef, veal, pork, lamb, and game meat; excludes organ meat and cured meat;" (2) poultry as "chicken, turkey, Cornish hens, duck, goose, quail, and pheasant (game birds); excludes organ meat and "cured meat;" and (3) cured meat as "frankfurters, sausages, corned beef, cured ham and luncheon meat that are made from beef, pork, or poultry." The "cured meat" FPED variable encompasses most processed meat consumed in the US (2). Therefore, the term "processed" rather than "cured" will be used throughout the manuscript to be consistent with DGA terminology (22, 24). The AMSA Lexicon was used to identify additional types of processed meat, other than cured, that could be estimated using FNDDS data. This resulted in additionally including chicken patties, nuggets, and tenders as processed poultry products because they are considered further processed by AMSA and are reasonably estimated in FNDDS using the WWEIA categories. By default, meats that are not processed will be referred to as "unprocessed." The gram weight of solid fats present in meat above 2.63 grams is allocated to the solid fat FPED gram weight rather than meat (21, 22). Therefore, all red meat and poultry food groups operationalized in our analysis are in lean meat ounce-equivalents (oz-eq).

We used a systematic review of meat terminology to build an ontology of common methods in which "red meat" and "poultry" food groups were operationalized by researchers in the nutrition and chronic disease literature (8). A challenge in assessing red meat and poultry intake is that dietary assessment tools and available dietary databases do not disaggregate processed meat groups into processed red meat and processed poultry. Therefore, researchers generally defaulted to one of the following described methodological decisions when assessing red meat and/or poultry intake (8). First, researchers may exclude processed meat completely and assess only unprocessed red meat or unprocessed poultry (2, 25, 26) which underestimates true red meat and poultry intakes. Second, researchers may classify all processed meat as processed red meat, i.e., "red and processed meat" (27–29), which overestimates red meat and underestimates poultry. A third option is to disaggregate processed meat into processed red meat and processed poultry and reaggregate with unprocessed red meat and unprocessed poultry, respectively. Yet, dietary assessment methods and available dietary databases rarely allow for this option. For our analysis, we used the NCI Processed Meat Categories SAS program (30) to disaggregate the processed meat FPED variable into processed red meat and processed poultry. This program also disaggregates chicken patties, nuggets, and tenders from the unprocessed poultry FPED because these are considered processed by some definitions (23). In brief, this program text-mines descriptive data for all food codes in FNDDS that contain a processed meat FPED component. The details of the code are previously described (30). In summary, for this analysis we compared four methods of operationalizing a red meat food group and three methods of operationalizing a poultry food group which are described in **Figure 1**.

## Red Meat and Poultry Allotments

To calculate allotment ranges, we first determined age- and sex-specific energy intake ranges (from sedentary to active) using the Institute of Medicine's estimated energy requirements for each population subgroup 2020–25 DGA (1). Those energy intake ranges were corresponded to red meat and poultry allotment amounts modeled in the recommended food patterns from the 2020–25 DGA (31) (**Supplementary Table 1.1**). The red meat categories in **Figure 1** were compared to the "meat" (i.e., red meat) allotment ranges and the poultry categories were compared to the poultry allotment ranges. Additional details about this analysis are described in the footnotes of **Supplementary Table 1.1**.

## Analytical Sample

Our analytical sample consists of all participants aged ≥2 years old who participated in the MEC examination of the 2015–16 and 2017–18 NHANES cycles ($n$ = 17,945). Participants were excluded if they did not provide at least one dietary recall deemed reliable by NCHS data reporting standards ($n$ = 2,907) (32), resulting in a final analytical sample of 15,038 participants. Of those participants, 84.4% had a reliable second dietary recall from which the data were used in statistical modeling of usual intakes described below. Unweighted examination response rates for participants in this age range were 58.7% in 2015–16 and 48.8% in 2017–18 (33). Enhanced weighting adjustments are applied by NCHS to limit potential response rate biases for population subgroups (34).

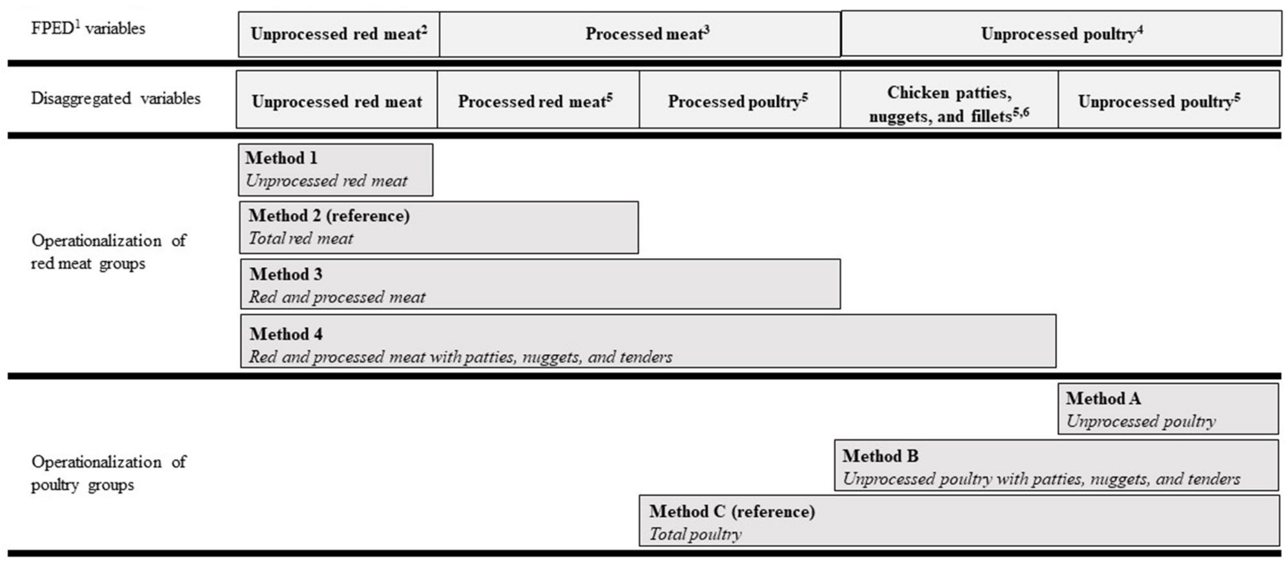


**FIGURE 1 |** Ontology of red meat and poultry food group methods used in nutrition research. [1]The US Department of Agriculture's Food Patterns Equivalents Database, available here: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/. [2]"Meat" FPED (PF_meat), defined as beef, veal, pork, lamb, and game meat; excludes organ meat and cured meat. [3]"Cured meat" FPED (PF_curedmeat), defined as frankfurters, sausages, corned beef, cured ham and luncheon meat that are made from beef, pork, or poultry. [4]"Poultry" FPED (PF_poult), defined as chicken, turkey, Cornish hens, duck, goose, quail, and pheasant (game birds); excludes organ meat and cured meat. [5]Disaggregated from FPED using the Processed Meat Categories SAS program from the US National Cancer Institute, available here: https://epi.grants.cancer.gov/ProcessedMeatCategories/. [6]Considered "further processed" by American Meat Science Association Lexicon, available here: https://dl.sciencesocieties.org/publications/mmb/articles/2/3/1.

## Statistical Analysis

Population characteristics and intake prevalence (i.e., the percent of the population who reported consuming red meat or poultry) were estimated *via* survey commands in SAS Version 9.4 (SAS Institute Inc. Cary, NC, USA). For each method described in **Figure 1**, we used The NCI Method (35, 36) to estimate usual intake (1) distributions of red meat and poultry and (2) proportions of the population whose intakes were below, within, or above red meat and poultry allotment ranges from the 2020–25 DGA recommended eating patterns. We chose the two-part model (i.e., for episodic food consumption) because >10% of the population had zero intakes of all red meat and poultry food groups listed in **Figure 1** (37). Integerized balanced repeated replication (BRR) weights were used to account for the day of the week that the 24-h recall was conducted, differential weighting for subpopulations, and the multistage complex sampling design of NHANES (36). We calculated 32 BRR weights by all 60 post-stratification combinations of age, sex, and race/ethnicity consistent with the NHANES sampling methods and used 0.3 for the Fay method which correlates with a perturbation factor of 70% (38). Weekly red meat and poultry allotments from the 2020-25 DGA were divided by 7 to be incorporated into The NCI Method SAS macros. The results are presented on a weekly basis in which estimated means and standard errors were multiplied back by 7 to be consistent with the DGA food patterns (1). Estimates were adjusted for age, gender, and race/Hispanic origin.

We then conducted pairwise comparisons to investigate differences in estimates within red meat food group methods and within poultry food group methods to assess how operationalization of each method influences population-level estimates. We used *total red meat* (red meat Method 2) and *total poultry* (poultry Method 3) as the reference method because this is likely the estimate that is most representative of true red meat and poultry intakes. We will highlight total red meat and total poultry intake estimates throughout the results, as intake estimates of these two food groups are a novel contribution to the literature. Due to the large sample size and increased precision *via* measurement error correction *via* usual intake modeling, even the most conservative Bonferroni correction ($P < 0.00007$) resulted in most comparisons being statistically significant. Therefore, we will highlight effect sizes rather than the Bonferroni corrected $P$ values throughout the results (39). Comparing intake amounts, we note meaningful effect sizes between methods as a difference of ≥3, ≥6, or ≥9 lean oz-eq/week. Three lean oz-eq is a recommended serving size of meat, so these effect sizes can be interpreted as a difference of 1, 2, or 3 servings/week between methods. When comparing the percent of population below, within, or above allotment ranges, we note meaningful effect sizes between methods as a difference of ≥10% and ≥20%. All analyses accounted for the complex survey design of NHANES and were weighted using day one dietary intake sample weights to account for oversampling, non-response, and post-stratification.

# RESULTS

## Prevalence of Consumption and Mean Intake

Our sample is nationally representative of the U.S. population aged ≥2 years old and is described in **Table 1**. Total red meat (i.e., unprocessed and processed red meat) is consumed in 70 ± 0.8% of the population, with a mean total red meat intake of 14.0 ± 0.35 lean oz-eq/week. Total poultry (i.e., unprocessed and processed poultry) is consumed in 71 ± 0.7% of the population, with a mean total poultry intake of 13.3 ± 0.35 lean oz-eq/week. Individuals <19 years old are consuming more total poultry than red meat, and individuals ≥19 years old consume more total red meat than total poultry (**Table 2**).

**TABLE 1** | Demographic characteristics of individuals from a representative sample of the US population.

| Characteristic | Age 2–18 years | Age ≥19 years | Age ≥2 years |
|---|---|---|---|
| Total | 5,037 | 10,001 | 15,038 |
| Sex | | | |
| Male | 2,508 (51%) | 4,850 (48%) | 7,358 (49%) |
| Female | 2,529 (49%) | 5,151 (52%) | 7,680 (51%) |
| Race and Hispanic origin | | | |
| Non-Hispanic White | 1,566 (51%) | 3,467 (63%) | 5,033 (60%) |
| Non-Hispanic Black | 1,135 (13%) | 2,238 (11%) | 3,373 (12%) |
| Non-Hispanic Asian | 438 (4%) | 1,172 (6%) | 1,610 (6%) |
| Hispanic | 1,477 (24%) | 2,688 (16%) | 4,165 (18%) |
| **PIR** | | | |
| <130% | 1,834 (31%) | 2,723 (21%) | 4,557 (24%) |
| ≥130% | 2,772 (69%) | 6,210 (79%) | 8,982 (76%) |
| Educational attainment | | | |
| High school or less | 1,645 (71%) | 4,295 (37%) | 5,940 (41%) |
| More than high school | 492 (29%) | 5,453 (63%) | 5,945 (59%) |

*Data presented as unweighted number of participants in each sample strata and weighted percent of population is presented in parentheses. Data source: US Centers for Disease Control and Prevention/National Center for Health Statistics, NHANES 2015–2018.*

The prevalence of those ≥2 years old reporting consumption of red meat ranges from 47 ± 1.2% to 75 ± 0.8% of the population and mean intake amounts range from 10.5 ± 0.28 to 18.2 ± 0.35 lean oz-eq/week (**Table 2**), depending on which method is operationalized (see ontology in **Figure 1**). The prevalence of those ≥2 years old reporting poultry intake ranges from 38 ± 1.0% to 71 ± 0.7% and mean intake amounts range from 9.8 ± 0.35 to 13.3 ± 0.35 lean oz-eq/week (**Table 2**) depending on which method is operationalized (see ontology in **Figure 1**). Weekly mean intake estimates for each age and sex subgroup are shown in **Supplementary Tables 1.3**, **1.4**.

### Red Meat

Intake of total red meat (Method 2, and the reference for all comparisons) for individuals aged ≥2 years was 13.6 ± 0.32 lean oz-eq/week at the 50th percentile and 31.0 ± 1.04 lean oz-eq/week at the 95th percentile. Intake increased from Method 1 to 2 (i.e., the progression from unprocessed red meat to total red meat which additionally included processed red meat) by ≥3.0 lean oz-eq/week from the 20–70th percentile, by ≥6.0 lean oz-eq/week from the 75–90th percentile, and by ≥9.0 lean

**TABLE 2** | Prevalence and amount of self-reported intake of red meat and poultry in the US estimated *via* different methods.

| Characteristic | Age 2–18 years *n* = 5,037 | Age ≥19 years *n* = 10,001 | Age ≥2 years *n* = 15,038 |
|---|---|---|---|
| **Intake prevalence of red meat (%)** | | | |
| Method 1: Unprocessed red meat | 40 ± 1.3 | 49 ± 1.3 | 47 ± 1.2 |
| Method 2: Total red meat | 67 ± 1.3 | 70 ± 0.8 | 70 ± 0.8 |
| Method 3: Red and processed meat | 70 ± 1.3 | 73 ± 0.8 | 72 ± 0.8 |
| Method 4: Red and processed meat with patties, nuggets, and tenders | 76 ± 1.2 | 75 ± 0.8 | 75 ± 0.8 |
| **Mean intake amount of red meat (lean oz-eq/week)** | | | |
| Method 1: Unprocessed red meat | 7.0 ± 0.35 | 11.2 ± 0.35 | 10.5 ± 0.28 |
| Method 2: Total red meat | 9.8 ± 0.35 | 15.4 ± 0.35 | 14.0 ± 0.35 |
| Method 3: Red and processed meat | 12.6 ± 0.35 | 18.2 ± 0.35 | 16.8 ± 0.35 |
| Method 4: Red and processed meat with patties, nuggets, and tenders | 14.7 ± 0.47 | 18.9 ± 0.42 | 18.2 ± 0.35 |
| **Intake prevalence of poultry groups (%)** | | | |
| Method A: Unprocessed poultry | 32 ± 1.3 | 40 ± 1.1 | 38 ± 1.0 |
| Method B: Unprocessed poultry with patties, nuggets, and tenders | 43 ± 1.7 | 43 ± 1.1 | 43 ± 1.0 |
| Method C: Total poultry | 75 ± 1.2 | 70 ± 0.8 | 71 ± 0.7 |
| **Mean intake amount of red meat (lean oz-eq/week)** | | | |
| Method A: Unprocessed poultry | 7.0 ± 0.42 | 10.5 ± 0.35 | 9.8 ± 0.35 |
| Method B: Unprocessed poultry with patties, nuggets, and tenders | 9.1 ± 0.49 | 11.2 ± 0.35 | 10.5 ± 0.35 |
| Method C: Total poultry | 11.2 ± 0.56 | 14.0 ± 0.35 | 13.3 ± 0.35 |

*Weighted mean ± standard error of the mean. See **Figure 1** for ontology of red meat and poultry method. Data source: US Centers for Disease Control and Prevention/National Center for Health Statistics, NHANES 2015–2018, day 1 dietary recall data.*

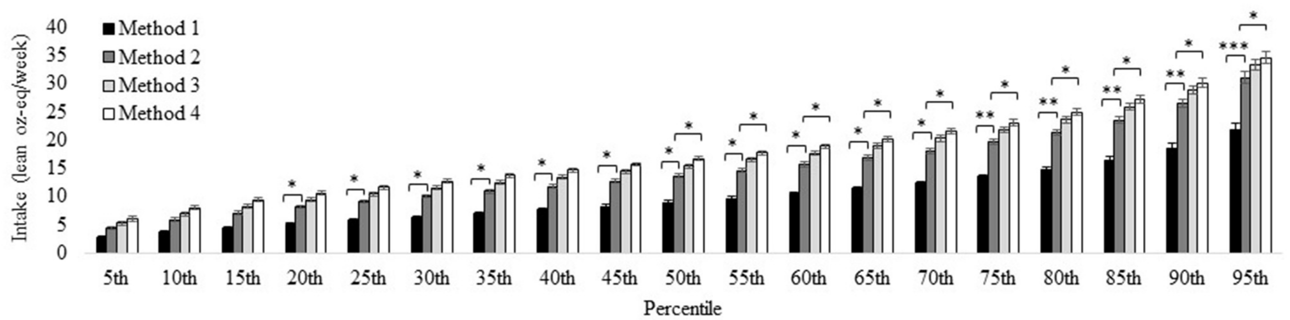


**FIGURE 2 |** Distribution of red meat intake for the U.S. population aged ≥2 years estimated *via* different methods. Method 1: Unprocessed red meat, includes beef, veal, pork, lamb, and game meat; excludes organ meat and processed meat. Method 2 (reference): Total red meat, includes unprocessed red meat and processed red meat. Method 3: Red and processed meat, includes unprocessed red meat, processed red meat, and processed poultry. Method 4: Red and processed meat, additionally including chicken patties, nuggets, and tenders. See **Figure 1** for further descriptions of each method. Results are shown as mean ± SEM within each percentile estimated *via* the NCI Method for usual dietary intakes. *Difference from Method 1 estimate (effect size) is ≥ 3.0 oz-eq per week; **difference from Method 1estimate is ≥ 6.0 oz-eq servings per week; ***difference from Method 1 estimate is ≥ 9.0 oz-eq per week. Data source: US Centers for Disease Control and Prevention/National Center for Health Statistics, NHANES 2015-2018, day 1 and 2 dietary recall data.

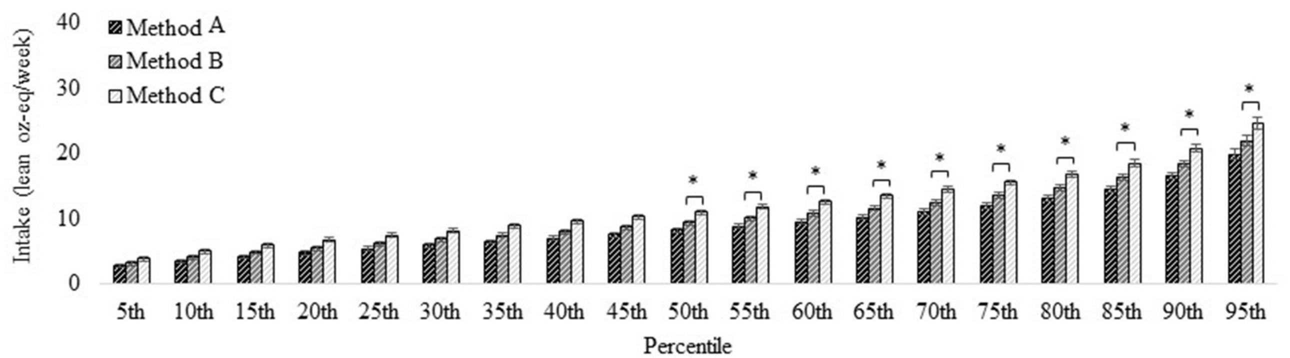


**FIGURE 3 |** Distribution of poultry intake for the U.S. population aged ≥2 years estimated *via* different methods. Method A: Unprocessed poultry, includes chicken, turkey, Cornish hens, duck, goose, quail, and pheasant (game birds); excludes organ meat and cured meat; additionally excludes chicken patties, nuggets, and tenders. Method B: Unprocessed poultry, includes chicken, turkey, Cornish hens, duck, goose, quail, and pheasant (game birds); excludes organ meat and cured meat and includes chicken patties, nuggets, and tenders. Method C (reference): Total poultry, includes unprocessed poultry, processed poultry, and chicken patties, nuggets, and tenders. See **Figure 1** for further descriptions of each method. Results are shown as mean ± SEM within each percentile estimated *via* the NCI Method for usual dietary intakes. *Difference from Method 1 estimate (effect size) is ≥ 3.0 oz-eq per week. Data source: US Centers for Disease Control and Prevention/National Center for Health Statistics, NHANES 2015–2018, day 1 and 2 dietary recall data.

oz-eq/week for the 95th percentile (**Figure 2**). The differences between intakes of Method 1 (unprocessed red meat) and Method 2 (total red meat) were largest at the higher end of the intake distribution for males of all age categories. Intake estimates using Method 2 vs. Method 3 (i.e., the progression from total red meat to "red and processed meat" which additionally includes processed poultry) were within one 3 lean oz-eq serving/week across the distribution and for each sex and age subgroup. Intake estimates using Method 4 (i.e., additional inclusion of chicken patties, nuggets, and tenders in "red and processed meat") from the 50th (16.7 ± 0.34 lean oz-eq/week) to 95th (34.6 ± 1.05 lean oz-eq/week) percentile differed from Method 2 (total red meat) by ≥3.0 lean oz-eq/week. See **Supplementary Tables 1.5**, **1.6** for age- and sex-specific intakes using each red meat method.

## Poultry

Intake of total poultry (Method C and the reference for all comparisons) for individuals aged ≥2 years was 11.0 ± 0.37 lean oz-eq/week at the 50th percentile and 24.6 ± 0.86 lean oz-eq/week at the 95th percentile. Estimates using Method A were ≥3.0 lean oz-eq/week less than Method C from the 50–95th percentile (i.e., the progression from unprocessed poultry, excluding chicken patties, nuggets, and tenders to total poultry inclusive of chicken patties, nuggets, and tenders as well as processed poultry products). Estimates using Method B did not differ from Method C (**Figure 3**), in which the difference between the two methods is the additional inclusion of processed poultry products in Method C (total poultry). Differences between methods were similar across all age and sex subgroups (**Supplementary Tables 1.7**, **1.8**).

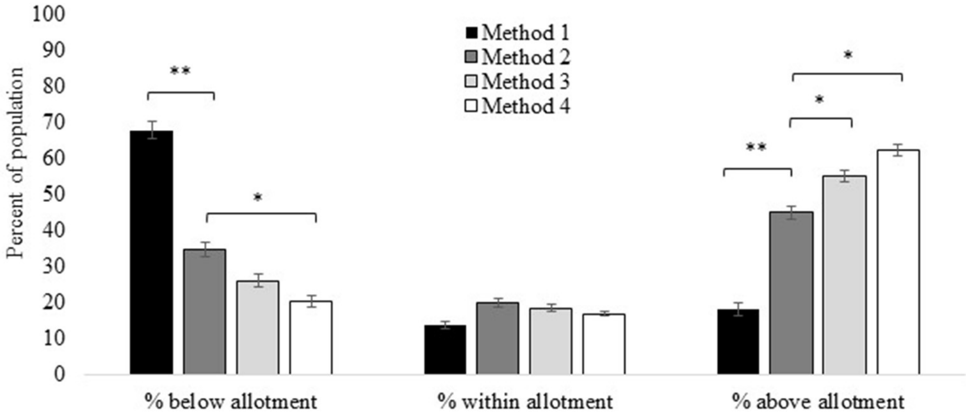


**FIGURE 4 |** Red meat intakes compared to allotment ranges in the US 2020–25 Dietary Guidelines for Americans recommended eating patterns estimated *via* different methods. Method 1: Unprocessed red meat, includes beef, veal, pork, lamb, and game meat; excludes organ meat and processed meat. Method 2 (reference): Total red meat, includes unprocessed red meat and processed red meat. Method 3: Red and processed meat, includes unprocessed red meat, processed red meat, and processed poultry. Method 4: Red and processed meat, additionally including chicken patties, nuggets, and tenders. See **Figure 1** for further descriptions of each method. Results are shown as mean ± SEM *via* the NCI Method for usual dietary intakes. *Difference from Method 1 estimate (effect size) is ≥10%; **difference from Method 1 estimate is ≥20%. Data source: US Centers for Disease Control and Prevention/National Center for Health Statistics, NHANES 2015-2018, day 1 and 2 dietary recall data.

## Comparison of Intakes to Allotment Ranges From the 2020–25 DGA Eating Patterns

### Red Meat

We compared intake amounts of individuals ≥2 years old to the corresponding red meat allotment ranges from the 2020–25 DGA recommended eating patterns. For total red meat (Method 2 and the reference for all comparisons), 35 ± 2.0% of the population was below, 20 ± 1.1% was within, and 45 ± 1.8% was above their age- and sex-specific red meat allotment ranges.

The percentage of the population that had intakes below their allotment range decreased by 33 ± 2.6% from Method 1 to Method 2 (i.e., the progression from unprocessed red meat to total red meat which includes processed red meat; **Figure 4**; **Supplementary Table 1.9**). The percentage was similar between Method 3 and Method 2 (i.e., the progression from total red meat to "red and processed meat" which additionally includes processed poultry) but was 14 ± 0.9% lower when using Method 4 (i.e., additional inclusion of chicken patties, nuggets, and tenders in "red and processed meat") vs. Method 2 (total red meat). The percent of the population whose intakes were within their age- and sex-specific allotment range did not differ by ≥10% based on which red meat method was used (**Figure 4**; **Supplementary Table 1.10**). The percent of the population who had intakes above their age- and sex-specific allotment was 27 ± 1.8% lower when using Method 1 (unprocessed red meat) vs. Method 2 (total red meat), was 10 ± 0.7% higher when using Method 3 (red and processed meat) vs. Method 2 (total red meat), and 17 ± 1.1% higher when using Method 4 (red and processed meat plus chicken patties, nuggets, and tenders) vs. Method 2 (total red meat).

Most age and sex subgroups consumed amounts of total red meat (Method 2; reference) that were within red meat allotment ranges from the 2020–25 DGA eating patterns (**Figure 5**). The exceptions are males aged 19–30, 31–50, 51–70, and 71+ years who consumed amounts above the allotment ranges. Intakes were below the age- and sex-specific red meat allotment ranges when using Method 1 (unprocessed red meat). The exceptions were males aged 31–50 whose mean intakes were within the allotment range, and males aged 51–70 whose intakes were just above the allotment range. Results were similar when using Method 3 (red and processed meat) and Method 2 (total red meat), except that the inclusion of processed poultry in Method 3 pushed intakes of males aged 14–18 years above allotment ranges. When chicken patties, nuggets, and filets were additionally included in Method 4, mean intakes of males aged 2–4, 5–8, and 9–13 years were pushed above allotment ranges. See **Supplementary Table 1.3** for age- and sex-specific mean intakes for each red meat method.

### Poultry

We compared intake of individuals ≥2 years old to the corresponding poultry allotment ranges from the 2020–25 DGA recommended eating patterns. Using total poultry (Method C and reference for all comparisons), 36 ± 2.6% of the population was below, 18 ± 0.9% was within, and 47 ± 2.6% was above age- and sex-specific allotment ranges (**Figure 6**). The percentage of the population that had intakes below their allotment range was 12 ± 1.1% higher with Method B (unprocessed poultry, including chicken tenders, nuggets, and patties) than Method C (total poultry). The percentage of the population that had intakes below their allotment range was ≥20% higher using Method A (unprocessed poultry, excluding chicken tenders, nuggets, and patties) vs. Method C (**Figure 6**; **Supplementary Table 1.12**). There was no difference between poultry methods for the percent within allotment ranges, except a lower percent of males and females aged 5–8 years (**Supplementary Table 1.13**) were within

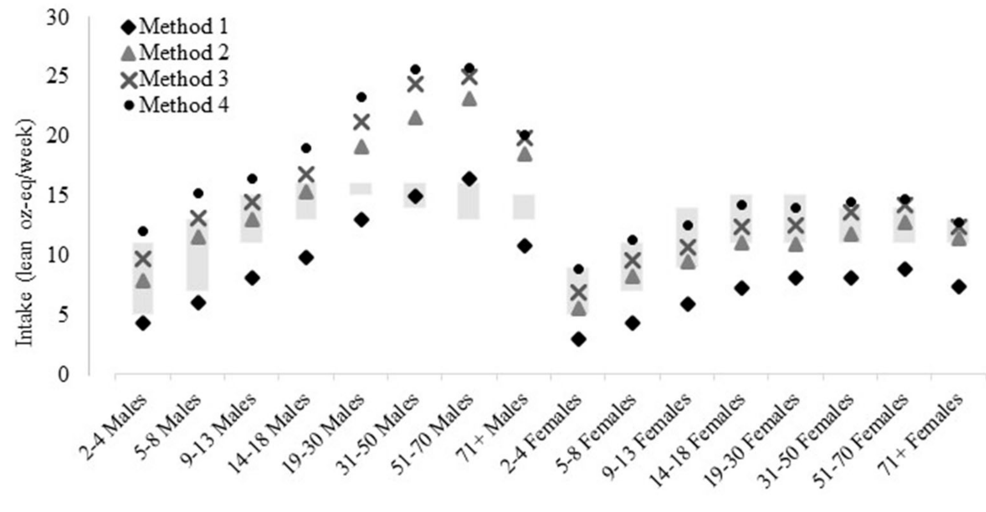


**FIGURE 5 |** Red meat intakes compared to allotment ranges in the US 2020-25 Dietary Guidelines for Americans recommended eating patterns, by age and sex, estimated *via* different methods. Method 1: Unprocessed red meat, includes beef, veal, pork, lamb, and game meat; excludes organ meat and processed meat. Method 2 (reference): Total red meat, includes unprocessed red meat and processed red meat. Method 3: Red and processed meat, includes unprocessed red meat, processed red meat, and processed poultry. Method 4: Red and processed meat, additionally including chicken patties, nuggets, and tenders. See **Figure 1** for further descriptions of each method. Results are shown as mean estimated *via* the NCI Method for usual dietary intakes. Data source: US Centers for Disease Control and Prevention/National Center for Health Statistics, NHANES 2015-2018, day 1 and 2 dietary recall data.

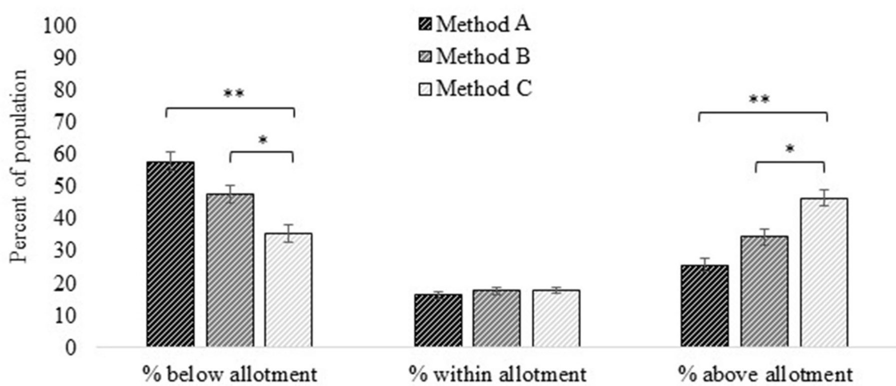


**FIGURE 6 |** Poultry intakes compared to allotment ranges in the US 2020–25 Dietary Guidelines for Americans recommended eating patterns estimated *via* different methods. Method A: Unprocessed poultry, includes chicken, turkey, Cornish hens, duck, goose, quail, and pheasant (game birds); excludes organ meat and cured meat; additionally excludes chicken patties, nuggets, and tenders. Method B: Unprocessed poultry, includes chicken, turkey, Cornish hens, duck, goose, quail, and pheasant (game birds); excludes organ meat and cured meat and includes chicken patties, nuggets, and tenders. Method C (reference): Total poultry, includes unprocessed poultry, processed poultry, and chicken patties, nuggets, and tenders. See **Figure 1** for further descriptions of each method. Results are shown as mean ± SEM within each percentile estimated *via* the NCI Method for usual dietary intakes. *Difference from Method 1 estimate (effect size) is ≥10%; **difference from Method 1 estimate is ≥20%. Data source: US Centers for Disease Control and Prevention/National Center for Health Statistics, NHANES 2015-2018, day 1 and 2 dietary recall data.

allotment ranges when using Method A vs. Method C. Method C differed from Method B by 12 ± 0.8% and differed from Method A by ≥20% when estimating the percent above the allotment range. The largest differences were between Method A and Method C for the younger age and sex subgroups (**Supplementary Table 1.14**).

Mean intake of total poultry (Method C) is within the age- and sex-specific poultry allotment ranges from the DGA recommended eating patterns for almost all age and sex subgroups (**Figure 7**). The exceptions were males aged 19–30 and 31–50 years and females aged 19–30 and 31–50 years who were above their allotment ranges and males and females aged ≥71 years who were below allotment ranges. A similar pattern was seen for Method B (unprocessed poultry including chicken tenders, nuggets, and patties), but females aged 19–30 and 31–50 years were within allotment range and females 51–70 years

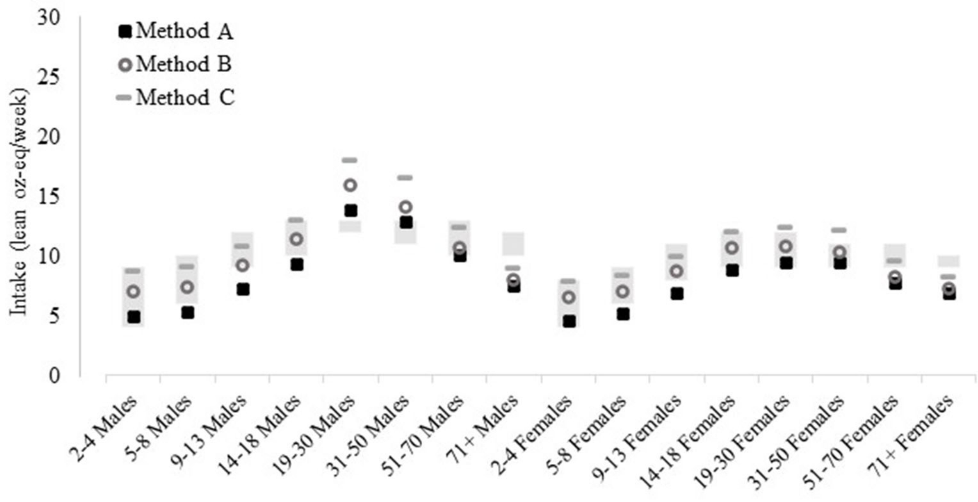


**FIGURE 7 |** Poultry intakes estimated by various poultry food group methods compared to allotment ranges in the US 2020–25 Dietary Guidelines for Americans recommended eating patterns, by age and sex, estimated *via* different methods. Method A: Unprocessed poultry, includes chicken, turkey, Cornish hens, duck, goose, quail, and pheasant (game birds); excludes organ meat and cured meat; additionally excludes chicken patties, nuggets, and tenders. Method B: Unprocessed poultry, includes chicken, turkey, Cornish hens, duck, goose, quail, and pheasant (game birds); excludes organ meat and cured meat and includes chicken patties, nuggets, and tenders. Method C (reference): Total poultry, includes unprocessed poultry, processed poultry, and chicken patties, nuggets, and tenders. See **Figure 1** for further descriptions of each method. Results are shown as mean estimated *via* the NCI Method for usual dietary intakes. Data source: US Centers for Disease Control and Prevention/National Center for Health Statistics, NHANES 2015–2018, day 1 and 2 dietary recall data.

were below with Method B. When chicken patties, nuggets, and tenders are further excluded in Method A, males and females aged 5–8, 9–13, and 14–18 were below allotment ranges.

## DISCUSSION

Heterogeneity in meat terminology and operationalization of meat food groups represents a methodological challenge in elucidating chronic disease implications of meat consumption (4, 6–8). The objective of this analysis was to demonstrate and quantify an example of how heterogeneity in meat food groups impacts population-level estimates of red meat and poultry intakes. Our results were supportive of our hypotheses that variation in meat food groups would meaningfully influence population-level intake estimates. In particular, decisions to include or exclude processed products from red meat and poultry food groups meaningfully changed weekly intake estimates of US populations. This methodological decision also meaningfully changed the percent of individuals who are below, within, and above red meat and poultry allotment ranges from the 2020–25 DGA recommended eating patterns. For example, most age and sex subgroups are below the allotment red meat range when processed red meat products are excluded but are within or substantially above (e.g., males aged 19–70 years old) the allotment range when processed red meat products are included. Further, misclassifying processed poultry as processed red meat, i.e., using the broad "red and processed meat" food group, exacerbates these differences. This analysis demonstrated that misclassification of processed meat and how processed meats are defined within meat food groups meaningfully influence population-level intake estimates of red meat and poultry.

In addition to a methodological demonstration, our analysis also provides novel intake estimates of total lean red meat and total lean poultry for the US population. The USDA routinely estimates unprocessed lean red meat and unprocessed lean poultry intake in the US (2, 40). Our estimate of total lean red meat and total lean poultry further incorporated processed products using the NCI's Processed Meat Categories method (30), adding novel insight into meat intake behaviors of the US population. Our estimates of total lean red meat intake were 3.0–9.0 lean oz-eq higher than estimates of unprocessed lean red meat across percentiles of consumption, and intake differences between total lean red lean meat and unprocessed lean red meat increased as consumption amounts increased. This emphasizes the importance of including processed red meat in lean red meat food groups to describe population intakes. The differences in intake between total lean red meat and unprocessed lean red meat were greatest for males aged 19–70 years old, who tend to be the highest consumers of animal-based protein sources in the US (1, 40). Differences across distributions were less pronounced for poultry estimates. Further, our estimates for total lean red meat ($14.0 \pm 0.35$ lean oz-eq/week) and total lean poultry ($13.3 \pm 0.35$ lean oz-eq/week) mean intakes were within one oz-eq/week of one another and had similar prevalence of consumption ($\sim$70% of individuals $\geq$2 yeas reported consumption). Adults $\geq$19 years consumed $\sim$1.5 lean oz-eq/week more total lean red meat than total poultry, but adolescents <19 years consumed $\sim$1.5 lean oz-eq/week more total poultry than total lean red meat. True estimates of total lean red meat and total poultry are likely higher, and potentially more divergent than 1.5 oz-eq/week, because some red meat products (such as bacon or pork chops) and poultry products (fried chicken or retaining the skin) are higher in fat (41). Future methodological development is needed to additionally include the non-lean gram weight of total red meat and total poultry to obtain more accurate population-level intake estimates. This is of particular relevance because high-fat meats are associated with various chronic disease end points, as summarized during the 2020–25 DGA development process (42).

We compared mean intakes of age and sex population subgroups to the age- and sex-specific red meat and poultry allotment ranges from the 2020–25 DGA recommended eating patterns. The allotment ranges in the recommended eating patterns result from food pattern modeling conducted in the evidence review portion of the DGA process (43). The foods chosen to create a representative composite of "red meat" in the food pattern modeling and resulting allotment ranges include both processed and unprocessed red meat. Similarly, the "poultry" composite contains unprocessed and processed poultry. Based on these models, one would expect that total red meat and total poultry would be the most appropriate food group for comparison to the allotment ranges. Yet, the FPED, which is designed to represent the 37 different components of the recommended eating patterns, does not provide total red meat or total poultry variables. Therefore, there is a disconnect between the food pattern modeling and the available FPEDs. Total lean red meat and total lean poultry can be reasonably estimated by linking the Processed Meat Categories program to the FPED, as we demonstrated in this analysis. Our results showed that most population subgroups consumed total red meat amounts within allotment ranges (except males aged 19–70 who were above allotment ranges by up to 7 lean oz-eq/week). Comparatively, excluding processed red meat, we found that intakes of unprocessed red meat were below red meat allotment ranges for most age and sex subgroups, and within allotment ranges for males 31–70 years old. Therefore, inclusion vs. exclusion of processed red meat when comparing intakes to allotment ranges results in very different conclusions about red meat intake behaviors. Notably, it is the processed portion of total red meat and total poultry that seems to be pushing certain population subgroups beyond their age- and sex-specific allotment ranges. Most population subgroups in the US consume total red meat and total poultry amounts that are within their age- and sex-specific allotment ranges, except that males 19–70 years consume too much red meat and males and females aged 19–50 consume too much poultry.

A push for methodological transparency in food group operationalization needs to be met with greater consensus on how food groups are defined. For example, the 2020–25 DGA do not provide an explicit definition of processed meat, but recommend that meat and poultry be consumed as "fresh, frozen, or canned, and in lean forms (e.g., chicken breast or ground turkey) versus processed meats (e.g., hot dogs, sausages, ham, luncheon meats)" (1). The prior 2015–2020 DGA defined processed meat as "all meat or poultry products preserved by smoking, curing, salting, and/or the addition of chemical preservatives" including "all types of meat or poultry sausages (bologna, frankfurters, luncheon meats and loaves, sandwich spreads, viennas, chorizos, kielbasa, pepperoni, salami, and summer sausages), bacon, smoked or cured ham or pork shoulder, corned beef, pastrami, pig's feet, beef jerky, marinated chicken breasts, and smoked turkey products" (22). Other organizations, such as AMSA, consider chicken nuggets to be processed as well (23). Our results demonstrate that nuances in processed meat definitions are a major driver of heterogeneity in meat food groups. This likely has differential implications across life stages that should be considered in research design. For example, classification of chicken patties, nuggets, and tenders as either unprocessed or processed poultry alters poultry intake estimates by $\sim$2 lean oz-eq/week for males and females aged 5–18 years old, but not other age groups. Consensus on processed meat definitions is also important for public health messaging. Processed meat intake is recognized as a cancer risk factor (44–46) and has also been associated with increased risk for cardiovascular disease and type 2 diabetes (47). Processed meat is generally higher in sodium and saturated fat content than unprocessed meat (4). Also, processed meat is commonly cured and cooked with high heat, which increases concentrations of N-nitroso, heterocyclic amine, and polycyclic aromatic carbons (48). There is also potential for high heme-iron intake to increase risk for disease, although this is debated (49, 50). These nutritive and non-nutritive compounds are proposed mechanistic links between processed meat consumption and some cancer types (51, 52). Public health organizations recommend consumers avoid or limit processed meat intake to reduce risk for chronic diseases (1), particularly cancer (44). Yet, it is difficult to communicate this message to the general public without a standardized definition of which meat products are considered processed and should be avoided.

For our analysis, we created an ontology of commonly used methods of estimating red meat and poultry food groups from a rigorous and comprehensive systematic review of nutrition research (8). We compared four methods of operationalizing red meat groups and three methods of operationalizing poultry groups within one nationally representative sample of the US population. Yet, our ontology is certainly not an exhaustive list of how researchers operationalize these groups. For example, some researchers classify "sandwich meats" as unprocessed red meat, yet most sandwich meats (i.e., deli meats) would be considered processed by publicly available definitions as well as within the FPED (4). A second example is the use of the term "white meat" which researchers may use to describe a variety of poultry-containing food groups, sometimes including pork or lean and fatty fish (8, 53). Further, we used NCI's Processed Meat Categories SAS program to disaggregate processed meat into processed red meat and processed poultry. This program has been shown to potentially overestimate processed red meat and underestimate processed poultry by $\sim$10–15% (30). Therefore, in our analysis, the differences between red meat methods may be overestimated and the differences within poultry methods may be underestimated. Our analysis is strengthened by using 24-h recall data and the USDA's databases which provide a comprehensive and detailed assessment of participant intake. Although not an exhaustive list of red meat and poultry food group methods, our results serve as a demonstrative example of why capturing nuances in food groups is quantifiably important in estimating intakes.

## CONCLUSION

Unsurprisingly, different methods of operationalizing red meat and poultry food groups resulted in different population-level intake estimates, simply because they are distinct and unique food groups. The current analysis quantifiably demonstrates the magnitude of potential bias induced by heterogeneous meat food group methods and for which meat subtypes and population subgroups the bias may be most prominent. Future research is needed to understand if the degree of misclassification in red meat and poultry food groups is meaningful enough to influence associations between red meat and poultry intakes and chronic disease risk, particularly cancer. Understanding this bias becomes important during development of public health dietary recommendations to ensure that studies can be grouped and compared appropriately. This work highlights that it may not be appropriate to compare studies that exclude processed products to studies that include processed products in red meat and poultry estimates because, as the current results demonstrate, these are distinct food groups in which intake estimates can differ by up to 9 lean oz-eq/week. The ultimate meat food group method employed depends largely on the research question and the dietary assessment tool used. However, promoting clear and transparent descriptions of meat food group terminology and methodology will provide clarity for researchers and public health professionals when creating and disseminating evidence-based public recommendations.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.cdc.gov/nchs/nhanes/index.htm.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the National Center for Health Statistics. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

LEO, KAH, and JR designed the research. LEO conducted the research, wrote the paper with editorial assistance from all authors, and has primary responsibility for final content. LEO and RP analyzed the data. All authors have read and approved the final manuscript.

## ACKNOWLEDGMENTS

We would like to acknowledge Lisa Kahle at IMS Inc. for contributions to programming for the usual intake analyses. We also thank all NHANES participants for their willingness to participate and provide valuable data.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.778369/full#supplementary-material

## REFERENCES

1. –2025 Dietary Guidelines for Americans. Ninth *edition, 2020. United States Department of Agriculture and United States Department of Health and Human Services.* (2020). Available online at: https://www.dietaryguidelines.gov/sites/default/files/2020-12/Dietary_Guidelines_for_Americans_2020-2025.pdf (accessed June 4, 2021).
2. Food Patterns Equivalents Database 2017-2018. *Methodology and User Guide. Food Surveys Research Group, Agricultural Research Service, United States Department of Agriculture.* (2021). Available online at: https://www.ars.usda.gov/ARSUserFiles/80400530/pdf/fped/FPED_1718.pdf (accessed June 4, 2021).
3. Diet Assessment Primer Roadmap. Dietary Assessment Primer. National Cancer Institute, National Institutes of Health. Available online at: https://dietassessmentprimer.cancer.gov/roadmap.html (accessed June 4, 2021).
4. Gifford CL, O'Connor LE, Campbell WW, Woerner DR, Belk KE. Broad and inconsistent muscle food classification is problematic for dietary guidance in the U.S. *Nutrients.* (2017) 9:1027. doi: 10.3390/nu9091027
5. Mathews R, Chu Y. Global review of whole grain definitions and health claims. *Nutr Rev.* (2020) 78:98–106. doi: 10.1093/nutrit/nuz055
6. Guidelines Advisory Committee. *Scientific Report of the 2020 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Agriculture the Secretary of Health Human Services. U.S. Department of Agriculture, Agricultural Research Service, Washington, DC.* (2020). Available online at: https://www.dietaryguidelines.gov/2020-advisory-committee-report (accessed June 4, 2021).
7. Dietary Guidelines Advisory Committee. *Scientific Report of the 2015 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Agriculture the Secretary of Health Human Services. U.S. Department of Agriculture, Agricultural Research Service, Washington, DC.* (2015). Available online at: https://health.gov/sites/default/files/2019-09/Scientific-Report-of-the-2015-Dietary-Guidelines-Advisory-Committee.pdf. 4 June 2021 (accessed June 4, 2021).
8. O'Connor LE, Gifford CL, Woerner DR, Sharp JL, Belk KE, Campbell WW. Dietary meat categories and descriptions in chronic disease research are substantively different within and between experimental and observational studies: a systematic review and landscape analysis. *Adv Nutr.* (2020) 11:41–51. doi: 10.1093/advances/nmz116
9. Isanejad M, LaCroix AZ, Thomson CA, Tinker L, Larson JC, Qi Q, et al. Branched-chain amino acid, meat intake and risk of type 2 diabetes in the women's health initiative. *Br J Nutr.* (2017) 117:1523–30. doi: 10.1017/S0007114517001568
10. Cross AJ, Peters U, Kirsh VA, Andriole GL, Reding D, Hayes RB, et al. A prospective study of meat and meat mutagens and prostate cancer risk. *Cancer Res.* (2005) 65:11779–84. doi: 10.1158/0008-5472.CAN-05-2191
11. Haring B, Gronroos N, Nettleton JA, von Ballmoos MC, Selvin E, Alonso A. Dietary protein intake and coronary heart disease in a large community based cohort: results from the Atherosclerosis Risk in Communities (ARIC) study [corrected]. *PLoS ONE.* (2014) 9:e109552. doi: 10.1371/journal.pone.0109552
12. National Health and Nutrition Examination Survey. *National Center for Health Statistics, Centers for Disease Control and Prevention.* (2021). Available

online at: https://www.cdc.gov/nchs/nhanes/index.htm (accessed June 4, 2021).

13. NHANES Survey Methods and Analytical Guidelines. *National Center for Health Statistics, Centers for Disease Control and Prevention.* (2021). Available online at: https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx#sample-design (accessed June 4, 2021).
14. National Center for Health Statistics. *Centers for Disease Control and Prevention.* (2021). Available online at: https://wwwn.cdc.gov/nchs/data/nhanes/2015-2016/manuals/2016_Interviewer_Procedures_Manual.pdf (accessed June 4, 2021).
15. Interviewer Procedures Manual, National Health and Nutrition Examination Survey. 2017-18. *National Center for Health Statistics, Centers for Disease Control and Prevention.* (2017). Available online at: https://wwwn.cdc.gov/nchs/data/nhanes/2017-2018/manuals/2017_Interviewer_Procedures_Manual.pdf (accessed June 4, 2021).
16. -2018 Data Documentation, Codebook, and Frequencies, Demographic Variables and Sample Weights (DEMO_J). *National Health and Nutrition Examination Survey. 2020. National Center for Health Statistics, Centers for Disease Control and Prevention.* (2021). Available online at: https://wwwn.cdc.gov/Nchs/Nhanes/2017-2018/DEMO_J.htm (accessed June 4, 2021).
17. -2016 Data Documentation, Codebook, and Frequencies, Demographic Variables and Sample Weights (DEMO_I). *National Health and Nutrition Examination Survey. 2017. National Center for Health Statistics, Centers for Disease Control and Prevention.* (2021). Available online at: https://wwwn.cdc.gov/Nchs/Nhanes/2015-2016/DEMO_I.htm (accessed June 4, 2021).
18. Moshfegh AJ, Rhodes DG, Baer DJ, Murayi T, Clemens JC, Rumpler WV, et al. The US department of agriculture automated multiple-pass method reduces bias in the collection of energy intakes. *Am J Clin Nutr.* (2008) 88:324–32. doi: 10.1093/ajcn/88.2.324
19. Food and Nutrient Database for Dietary Studies. *Food Surveys Research Group, Agricultural Research Service, United States Department of Agriculture.* (2021). Available online at: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds/ (accessed June 4, 2021).
20. Food Patterns Equivalents Database. *Food Surveys Research Group, Agricultural Research Service, United States Department of Agriculture.* (2021). Available online at: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fped-overview/ (accessed June 4, 2021).
21. What are WWEIA Food Categories? *Food Surveys Research Group, Agricultural Research Service, United States Department of Agriculture.* (2021). Available online at: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/dmr-food-categories/ (accessed June 4, 2021).
22. -2020 Dietary Guidelines for Americans. Eighth *edition, 2015. United States Department of Agriculture and United States Department of Health and Human Services.* (2021). Available online at: https://health.gov/sites/default/files/2019-09/2015-2020_Dietary_Guidelines.pdf (accessed June 4, 2021).
23. Seman DL, Boler DD, Carr C, Dikeman ME, Owens CM, Keeton JT, et al. Meat Science Lexicon. *Meat Muscle Biol.* (2018) 2(3). doi: 10.22175/mmb2017.12.0059
24. Glossary of the 2015-2020 Dietary Guidelines for Americans. Eighth *edition, 2015. United States Department of Agriculture and United States Department of Health and Human Services.* (2021). Available online at: https://health.gov/sites/default/files/2019-09/2015-2020_Dietary_Guidelines.pdf (accessed June 4, 2021).
25. Abete I, Romaguera D, Vieira AR, Lopez de Munain A, Norat T. Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality: a meta-analysis of cohort studies. *Br J Nutr.* (2014) 112:762–75. doi: 10.1017/S000711451400124X
26. Aune D, De Stefani E, Ronco A, Boffetta P, Deneo-Pellegrini H, Acosta G, et al. Meat consumption and cancer risk: a case-control study in Uruguay. *Asian Pac J Cancer Prev.* (2009) 10:429–36.
27. Cho E, Chen WY, Hunter DJ, Stampfer MJ, Colditz GA, Hankinson SE, et al. Red meat intake and risk of breast cancer among premenopausal women. *Arch Intern Med.* (2006) 166:2253–9. doi: 10.1001/archinte.166.20.2253
28. Cross AJ, Freedman ND, Ren J, Ward MH, Hollenbeck AR, Schatzkin A, et al. Meat consumption and risk of esophageal and gastric cancer in a large prospective study. *Am J Gastroenterol.* (2011) 106:432–42. doi: 10.1038/ajg.2010.415
29. Azadbakht L, Esmaillzadeh A. Red meat intake is associated with metabolic syndrome and the plasma C-reactive protein concentration in women. *J Nutr.* (2009) 139:335–9. doi: 10.3945/jn.108.096297
30. O'Connor LE, Wambogo EA, Herrick KA, Reedy J. A standardized assessment of processed red meat and processed poultry intake in the us population aged ≥ 2 years using NHANES. *J Nutr.* (2021) nxab316. doi: 10.1093/jn/nxab316. [Epub ahead of print].
31. Parker JD TM, Makram T, Malec DJ, Beresovsky V, Carroll M, Gonzalez JF, et al. "Vital and health statistics," in *National Center for Health Statistics Data Presentation Standards for Proportions, 2017. National Center for Health Statistics, Centers for Disease Control and Prevention.*
32. Dietary Interview—Individual Foods, First Day (DR1IFF_J), 2017-2018 Data Documentation, Codebook, and Frequencies, National Health and Nutrition Examination Survey. *National Center for Health Statistics, Centers for Disease Control and Prevention.* Available online at: https://wwwn.cdc.gov/Nchs/Nhanes/2017-2018/DR1IFF_J.htm (accessed June 4, 2021).
33. Unweighted Response. *Rates for NHANES 2017-2018 by Age and Gender, National Center for Health Statistics, Centers for Disease Control and Prevention.* Available online at: https://wwwn.cdc.gov/nchs/data/nhanes3/ResponseRates/NHANES-2017-2018-Response-Rates-508.pdf (accessed June 4, 2021).
34. Fakhouri TH, Martin CB, Chen TC, Akinbami LJ, Ogden CL, Paulose-Ram R, et al. An investigation of nonresponse bias and survey location variability in the 2017-2018 National Health and Nutrition Examination Survey. Vital and Health statistics. Series 2. *Data Eval Meth Res.* (2020) 185:1–36.
35. Usual Dietary Intakes: The NCI Method. *Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health.* Available online at: https://epi.grants.cancer.gov/diet/usualintakes/method.html (accessed June 4, 2021).
36. Herrick KA, Rossen LM, Parsons R, Dodd KW. Estimating usual dietary in take from national health and nut rition examination survey data using the national cancer institute method. *Vital Health Stat.* (2018) 178:1–63.
37. Usual Dietary Intakes: Details of the Method. *Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health.* Available online at: https://epi.grants.cancer.gov/diet/usualintakes/details.html (accessed June 4, 2021).
38. Chen TC, Clark J, Riddles MK, Mohadjer LK, Fakhouri TH. National Health and Nutrition Examination Survey, 2015– 2018: sample design and estimation procedures. Available online at: https://www.cdc.gov/nchs/data/series/sr_02/sr02-184-508.pdf
39. Sorkin JD, Manary M, Smeets PAM, MacFarlane AJ, Astrup A, Prigeon RL, et al. A guide for authors and readers of the American Society for Nutrition Journals on the proper use of *P* values and strategies that promote transparency and improve research reproducibility. *Am J Clin Nutr.* (2021) 114:1280–5. doi: 10.1093/ajcn/nqab223
40. Zeng L, Ruan M, Liu J, Wilde P, Naumova EN, Mozaffarian D, et al. Trends in processed meat, unprocessed red meat, poultry, and fish consumption in the United States, 1999-2016. *J Acad Nutr Diet.* (2019) 119:1085–98 e12. doi: 10.1016/j.jand.2019.04.004
41. Pereira PM, Vicente AF. Meat nutritional composition and nutritive role in the human diet. *Meat Sci.* (2013) 93:586–92. doi: 10.1016/j.meatsci.2012.09.018
42. Part D. Chapter 8: Dietary Patterns. *Guidelines Advisory Committee. 2020. Scientific Report of the 2020 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Agriculture the Secretary of Health Human Services. U.S. Department of Agriculture, Agricultural Research Service, Washington, DC.* (2021). Available online at: https://www.dietaryguidelines.gov/sites/default/files/2020-07/PartD_Ch8_DietaryPatterns_first-print.pdf (accessed June 4, 2021).
43. Food Pattern Modeling: Ages 2 Years and Older. *2020 Dietary Guidelines Advisory Committee. Food Pattern Modeling Report. Office of Nutrition Guidance and Analysis, Center for Nutrition Policy and Promotion, United States Department of Agriculture.* (2021). Available

online at: https://www.dietaryguidelines.gov/sites/default/files/2020-07/FoodPatternModeling_Report_2YearsandOlder.pdf (accessed June 4, 2021).
44. Limit Consumption of Red and Processed Meat. *American Institute for Cancer Research.* (2021). Available online at: https://www.aicr.org/cancer-prevention/recommendations/limit-consumption-of-red-and-processed-meat/ (accessed June 4, 2021).
45. Red meat and Processed Meat Consumption. *Online Summary of Trends in US Cancer Control Measures. Cancer Trends Progress Report. National Cancer Institute, National Institutes of Health.* (2021). Available online at: https://progressreport.cancer.gov/prevention/red_meat (accessed June 4, 2021).
46. Red Meat and Processed Meat. *IARC Monograph, Volume 114. International Agency for Research on Cancer, World Health Organization.* (2021). Available online at: https://monographs.iarc.who.int/wp-content/uploads/2018/06/mono114.pdf (accessed June 4, 2021).
47. Micha R, Michas G, Mozaffarian D. Unprocessed red and processed meats and risk of coronary artery disease and type 2 diabetes–an updated review of the evidence. *Curr Atheroscler Rep.* (2012) 14:515–24. doi: 10.1007/s11883-012-0282-8
48. Sinha R, Cross A, Curtin J, Zimmerman T, McNutt S, Risch A, et al. Development of a food frequency questionnaire module and databases for compounds in cooked and processed meats. *Mol Nutr Food Res.* (2005) 49:648–55. doi: 10.1002/mnfr.200500018
49. Han M, Guan L, Ren Y, Zhao Y, Liu D, Zhang D, et al. Dietary iron intake and risk of death due to cardiovascular diseases: a systematic review and dose-response meta-analysis of prospective cohort studies. *Asia Pac J Clin Nutr.* (2020) 29:309–21.
50. Seiwert N, Heylmann D, Hasselwander S, Fahrer J. Mechanism of colorectal carcinogenesis triggered by heme iron from red meat. *Biochim Biophys Acta Rev Cancer.* (2020) 1873:188334. doi: 10.1016/j.bbcan.2019.188334
51. Sinha R, Peters U, Cross AJ, Kulldorff M, Weissfeld JL, Pinsky PF, et al. Meat, meat cooking methods and preservation, and risk for colorectal adenoma. *Cancer Res.* (2005) 65:8034–41. doi: 10.1158/0008-5472.CAN-04-3429
52. Cross AJ, Sinha R. Meat-related mutagens/carcinogens in the etiology of colorectal cancer. *Environ Mol Mutagen.* (2004) 44:44–55. doi: 10.1002/em.20030
53. Cross AJ, Leitzmann MF, Gail MH, Hollenbeck AR, Schatzkin A, Sinha R. A prospective study of red and processed meat intake in relation to cancer risk. *PLoS Med.* (2007) 4:e325. doi: 10.1371/journal.pmed.0040325

**Conflict of Interest:** RP is employed by Information Management Services, Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 16 December 2021
doi: 10.3389/fnut.2021.794592

Check for updates

# Evolution of Nutritional Habits Behaviour of Spanish Population Confined Through Social Media

*Miguel Mariscal-Arcas[1]*, Sonia Delgado-Mingorance[2], Borja Saenz de Buruaga[2], Alba Blas-Diaz[1], Jose Antonio Latorre[3], Manuel Martinez-Bebia[3], Nuria Gimenez-Blasi[4], Javier Conde-Pipo[1], Leticia Cantero[1], Alejandro Lopez-Moro[1] and Maria Jose Jimenez-Casquet[1]*

*[1] Department of Nutrition and Food Science, University of Granada, Granada, Spain, [2] Master of Food and Fit, University of Granada, Granada, Spain, [3] Department of Food Technology, Nutrition and Food Science, University of Murcia, Murcia, Spain, [4] Department of Nutrition and Dietetics, Faculty of Health Sciences, Catholic University of Avila, Avila, Spain*

OPEN ACCESS

***Edited by:***
*Massimo Lucarini,*
*Council for Agricultural Research and Economics, Italy*

***Reviewed by:***
*Muhammad Sohail Afzal,*
*University of Management and Technology, Pakistan*
*Santosh Kumar Swain,*
*Siksha O Anusandhan University, India*

****Correspondence:***
*Miguel Mariscal-Arcas*
*mariscal@ugr.es*

***Specialty section:***
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

***Received:*** *13 October 2021*
***Accepted:*** *18 November 2021*
***Published:*** *16 December 2021*

***Citation:***
*Mariscal-Arcas M, Delgado-Mingorance S, Saenz de Buruaga B, Blas-Diaz A, Latorre JA, Martinez-Bebia M, Gimenez-Blasi N, Conde-Pipo J, Cantero L, Lopez-Moro A and Jimenez-Casquet MJ (2021) Evolution of Nutritional Habits Behaviour of Spanish Population Confined Through Social Media. Front. Nutr. 8:794592. doi: 10.3389/fnut.2021.794592*

**Introduction:** In Spain, on 14 March 2020, a state of alarm is declared to face the health emergency situation caused by the COVID-19 coronavirus, limiting the freedom of movement of people. The Spanish population is confined.

**Objective:** With this situation, "NUTRITIONAL HEALTH IS NOT CONFINED" arises a research project that seeks to promote nutritional education based on the pattern of the Mediterranean diet (MD) using new computer technologies. It is about providing the population with the information of general interest about the promotion of a healthy diet through social networks and analysing the impact of its dissemination, in the form of a longitudinal intervention study of the Spanish nutritional evolution during confinement, with a daily survey format, and it is intended to assess food consumption during the period of confinement. Materials and methods: In total, 936 participants were asked every day. Short publications were published every day based on the scientific evidence (FAO, WHO, AECOSAN) through social media such as Instagram, accompanied by a questionnaire of 11 questions (yes/no) where it was intended to assess the evolution of daily consumption.

**Results and Discussion:** The diffusion through social media has allowed to have a greater reach of the population. We observed that mood throughout confinement generally improves. There are certain eating habits from the MD that are well established in the daily diet of our population, such as the consumption of fruits, vegetables, legumes, dairy products, and eggs. It seems that enjoying good health is a growing concern in pandemic situations, which is why inappropriate behaviours such as "snacking" between meals or the consumption of processed foods such as snacks, industrial pastries, soft drinks, and sweets are avoided, increasing the amount of healthy food such as meat and fish. This study opens up future avenues of research promoting MD and implements new cohort nutritional databases, especially about young adult people, who are adept at navigating digital spaces and therefore using social media.

**Keywords: social media, database, nutrition methodology, population confined, diet**

## INTRODUCTION

Throughout history, there have been various catastrophes such as wars, terrorist attacks, floods, earthquakes, and pandemics, which highlight the vulnerability of the population and require the urgent mobilisation of interventions (1). In Spain, on 14 March 2020, a state of alarm is declared to face the health emergency situation caused by the COVID-19 coronavirus, limiting the freedom of movement of people. The Spanish population is confined. Faced with this situation and with a project already started on how social networks influence the dissemination of healthy aspects based on the Mediterranean diet (MD), "NUTRITIONAL HEALTH IS NOT CONFINED" (#nutritionalhealthisnotconfined) arises. A scientific initiative for the transfer of research aims to make confinement more bearable by promoting nutritional education through the daily publication in a digital blog, of posts based on scientific evidence, of different aspects related to a healthy diet. The dissemination of content on social media presents a series of advantages with respect to the paper format, such as immediate communication and instantaneous assessment of the reactions of readers, and also sharing knowledge of any field reaching more niches of population (2). That is why this format is chosen with the aim of reaching as many people as possible without biases of age, profession, interest in a particular healthy life, or level of training or special knowledge, only the employment of the social network chosen for the dissemination of #nutritionalhealthisnotconfined (3).

This blog would be based on the Mediterranean pattern for its great association with the idea of health and quality of life, and also for its proven beneficial effects on cardiovascular diseases and even prevention of some types of cancer. One of the components that characterise this eating pattern is olive oil, whose main fatty acid is oleic acid. Because it is also the main source of lipids, the MD has a high monounsaturated or saturated fatty acid (MUFA/SFA) ratio compared to other countries in the world that use animal-type fats, with the corresponding advantages at the cardiovascular level (4). In addition, the high consumption of foods of plant origin such as legumes, fruits (as a usual dessert), vegetables, mushrooms, and nuts produces a great contribution of fibre, vitamins, and minerals at the same time that they provide a large amount of water, drink par excellence in the Mediterranean and fundamental in our diet. In addition, it establishes a daily consumption of bread and cereals such as pasta and rice, preferably whole grains, and also dairy products (mainly yoghourt and cheeses) since they are a source of protein of high biological value, of vitamins and minerals (calcium, phosphorus), and of the well-known probiotics capable of improving the balance of the intestinal microbiota (5). As a source of protein, the Mediterranean pattern defends animal protein, but prioritising fish and eggs over meat. The MD not only focuses on food, but also defends a healthy lifestyle, which is why it emphasises the need to perform physical activity every day adapted to our abilities to maintain good health (6, 7).

As usually happens with any natural catastrophe, both physical and psychological consequences remain in the population that must be analysed and treated (1). When disaster strikes, everyday life patterns and normal behaviour are disrupted. In the case of confinement, one can speak of loneliness and social isolation, which in recent years the World Health Organisation (WHO) had classified as a challenge for global public health. Increased stress and anxiety due to the obligation to be locked up, due to the loss of routine activities, due to the lack of information about the situation experienced or the uncertainty of what will happen in the near future, attached at fear of possible contagion, and also the employment and economic situation of each person, can be some of the most common manifestations (8). In addition, they can affect not only on a psychological level, but also on a physical level, compromising our immune system due to a chronic inflammatory reaction (9). Maintaining the mental health of citizens has been a challenge these days, which is why telephone medical consultations and online psychological counselling services have been enabled. In this case, communication and social media have done a very positive job: being able to communicate and even see loved ones and friends who are far away or even read news about how to cope with confinement (10).

For all this, the ultimate goal of "NUTRITIONAL HEALTH IS NOT CONFINED" (#nutritionalhealthisnotconfined) is proposed as a measure to help the population during confinement, where to explain the properties of the foods that make up the MD, some original recipes how to consume them, their health benefits, and even how they can affect our declared mood. Social media and new technologies can be the good tools (3). Therefore, once the scientific literature has been reviewed, and the insufficiency of studies and the inconsistency of results have been proven, the main objective of this applied research work is to provide the population with information of general interest on the promotion of a healthy diet through social media and analyse the impact of its dissemination, in the form of a longitudinal intervention study of the Spanish nutritional evolution during confinement. In a daily survey format, designed and validated by authors (3, 6, 7, 11–14), it is intended to assess food consumption during the quarantine period. Databases and information achieved through this new format of study could be useful to inform the population through social media about forms of healthy eating based on the MD, publicise not so widespread foods or different ways of consuming them and also their benefits in the body, describe possible mood changes during confinement and how to cope with them through diet, explain different aspects related to a healthy life style and physical activity practise, carry out daily surveys to assess the way the population feeds during quarantine and also making an assessment of the declared daily mood and body perception of the participants, and analysing how all of the above is disseminated through something newer than paper format (magazines, newspapers, and television) such as social media and how it this new format impacts the population.

## MATERIALS AND METHODS

The format of "NUTRITIONAL HEALTH IS NOT CONFINED" (#nutritionalhealthisnotconfined) was used through publications of short duration and free dissemination based on the transfer knowledge of the authors using international recommendations FAO/WHO (15–18) and scientific evidence, disseminating this information and asking everyday through Instagram social media. Daily questionnaire included written text and publications in video format. Each daily questionnaire was accompanied by 11 questions with dichotomous answer options (yes/no) where it was intended to describe and assess the evolution of daily consumption of water, alcohol, vegetable products, animal products, processed and ultraprocessed, and those related to nutritional habits as snacking between meals or other important as changes in daily body perception, changes in mood perception and the influence that each daily questionnaire may had had on the subject.

## Study Population

In total, 936 subjects were freely recruited and surveyed each day during the quarantine state occurred in Spain from March to July 2020. We used the official social media (Instagram) of a Spanish winter sports federation (@fadiandalucia) as representative sample of the physically active Spanish population who have more than 1,000 people federated in Spain and more than 2,200 followers. We also used the official social media (Instagram) of scientific knowledge transfer academic platform designed by our nutritional research group in the Department of Nutrition and Food Science of the University of Granada, Spain call MM Health Science (@mmhealthscience) [section: "#lasaludnutricionalnoseconfina" (#nutritionalhealthisnotconfined)] (3) with more than 820 followers. The validity and reliability of this tool used in this study were obtained through the research work published by Mariscal-Arcas (3, 6, 7, 11–14). The age of the subjects who voluntarily participated in the study was between 17 and 65 years, excluding any subject under 17 years of age from the study. Only followers of one of these two official social networks were allowed to participate to control the population. The study was approved by the Research Ethics Committee of the Andalusian Public Health Service, Spain (reference number. 0756-N-20). The research protocol was carried out in accordance with the Declaration of Helsinki for Human Studies of the World Medical Association, with strict respect for the confidentiality of the information in accordance with Organic Law 15/1999, of 13 December, on the protection of personal data in all the processes of collection and treatment of the information obtained and Organic Law 3/2018, of 5 December, on the protection of personal data and guarantee of digital rights. The population under study is all followers of mentioned accounts, who participated voluntarily by consenting through a private request on their profiles, as represented in **Table 1**. To analyse dietary changes during confinement, previous consumption data for each food group according to the SENC Healthy Eating Guidelines have been used (19).

The material used for the development of the publications went from official database sources of information in the field of nutrition (FAO, WHO) and from scientific sources based on evidence (PubMed, Scopus, and Google Scholar). To evaluate the results, the percentage of affirmative responses in the survey is analysed, which associates it with the days on which the study is carried out. As a strength, this nutritional study through social media allows to have a wider scope of population including different points of Spain distant from each other and mostly young adult population, however, part of the limitations of the study lies in the insufficient scientific work that currently exists. Another limitation in the use of surveys on the Instagram platform is the impossibility of adding many extended questions.

## Statistical Analysis

SPSS version 22.0 (IBM. Chicago. IL) was used for the statistical analysis. A descriptive analysis was conducted to calculate means, standard deviations (SD), medians, maximum and minimum values, histograms, and normality tests, followed by the application of the Student's $t$-test, analysis of variance (ANOVA), Pearson's correlations. $p$-Value $< 0.05$ was considered statistically significant.

# RESULTS

## Usual Eating Habits

Before the state of alarm and confinement, the Spanish population, according to the SENC, has the following eating habits: 97% of people regularly consume meat, at levels above the recommendations for red and processed meats in 57% of cases, especially in those under the age of 35. In total, 95.5% of people usually consume fish, 98.8% eggs, and 88.6% dairy products. Only 30% meet the recommendations for fruit consumption and 21.3% for vegetables, although 55% consume at least one daily serving of whole fruit and 53.7% at least one serving of vegetables. A significantly higher proportion of women have adequate vegetable intakes. In total, 54% of people report usual intakes of pulses in line with recommendations, although 51% of those over 55 years of age report inadequate intakes for this food group, as for cereals and potatoes. A total of 89.6% of people report using virgin olive oil is an added fat in culinary preparations, whereas 20.4% of people also use olive oil and 16.8% sunflower oil. As regards the consumption of sweet and bakery products, 84% of respondents reported regular consumption, although 75% of men and 65% of women consumed them occasionally. Those under the age of 35 years have a higher proportion of inadequate consumption for this food group (40%), for savoury snacks (39%), and sugary drinks (10%) (19).

Analysing the result of the percentage of 'YES' responses for each of the responses in relation to the days of the study, trend lines were generated for each of the items asked. As can be seen in **Figure 1**, the trend in water consumption, which turned out to be quite high, is not falling below 90% except for specific days that are also concentrated in the first half of the study and then remaining around 100% quite consistently. However, analysing the trend line of alcohol consumption, as can be seen in **Figure 2**, it remained constant throughout the entire confinement, with a mean of "YES" responses not more than 35%. On the other hand, in the case of the answers to the question "Do you feel heavier than yesterday?", the trend line of the graph was slightly downwards from 30 to 40% (**Figure 3**). Regarding the "snacking" between hours, the situation was totally different, since there was a clear decrease as the days passed, decreasing

**TABLE 1 |** Population that constitutes the study by age groups and origin as a percentage ($n = 936$).

| | | @fadiandalucia % (*N*) | @mmhealthscience % (*N*) |
|---|---|---|---|
| Sex | Men | 62 (427.80) | 37 (91.02) |
| | Women | 38 (262.20) | 63 (154.98) |
| Age range (yrs) | 17–24 | 12 (76.90) | 20 (47.74) |
| | 25–34 | 28 (195.20) | 46 (115.16) |
| | 35–44 | 29 (200.10) | 21 (52.66) |
| | 45–54 | 20 (139.00) | 8 (20.68) |
| | 55–64 | 7 (49.30) | 8 (20.68) |
| | ≥65 | 2 (14.80) | 1 (3.46) |
| Origin | Granada | 40 (376.00) | 20 (186.10) |
| | Madrid | 18 (169.00) | 1.8 (17.22) |
| | Malaga | 2.2 (20.70) | 0 (0.00) |
| | Barcelona | 1.4 (13.80) | 0 (000) |
| | Seville | 1.4 (13.80) | 0 (0.00) |
| | Murcia | 0 (0.00) | 13.3 (124.60) |
| | Cangas | 0 (0.00) | 0.7 (7.38) |
| | Lorca | 0 (0.00) | 0.7 (7.38) |
| Declared co-morbidities | Respiratory diseases | 4.2 (39.31) | 4 (37.44) |
| | Cardiovascular diseases | 0 (0.00) | 1.6 (14.97) |
| | Diabetes | 1.9 (17.78) | 1 (9.36) |
| | Hypercholesterolemia | 2.3 (21.52) | 3 (28.08) |
| | Hypertension | 2 (18.72) | 2.1 (18.72) |
| | Allergies/intolerances | 1.4 (13.10) | 2.7 (25.27) |

| | | Mean (SD) | Min | Max | *P** | *P*** |
|---|---|---|---|---|---|---|
| Age (yrs) | Men | 34.78 (10.92) | 19.00 | 65.00 | 0.706 | |
| | Women | 35.40 (12.48) | 17.00 | 72.00 | | |
| | Total | 35.21 (12.00) | 17.00 | 72.00 | | |
| Weight before confinement (Kg) | Men | 80.43 (11.82) | 53.00 | 114.30 | 0.001 | |
| | Women | 63.64 (11.10) | 41.00 | 98.90 | | $R = 0.995$ $P = 0.001$ |
| | Total | 68.85 (13.72) | 41.00 | 114.30 | | |
| Weight at the end of confinement (Kg) | Men | 80.43 (11.82) | 53.00 | 114.30 | 0.001 | |
| | Women | 63.64 (11.10) | 41.00 | 98.90 | | |
| | Total | 68.85 (13.72) | 41.00 | 114.30 | | |
| Height (cm) | Men | 178.16 (6.94) | 168.00 | 196.00 | 0.001 | |
| | Women | 164.88 (6.13) | 148.00 | 181.00 | | |
| | Total | 169.00 (8.87) | 148.00 | 196.00 | | |
| BMI before confinement (Kg/m$^2$) | Men | 25.29 (3.15) | 18.56 | 36.57 | 0.001 | |
| | Women | 23.43 (4.10) | 15.70 | 38.67 | | $R = 0.993$ $P = 0.001$ |
| | Total | 24.01 (3.92) | 15.70 | 38.67 | | |
| BMI at the end of confinement (Kg/m$^2$) | Men | 25.43 (0.37) | 18.56 | 36.57 | 0.001 | |
| | Women | 23.39 (0.32) | 16.04 | 38.67 | | |
| | Total | 24.02 (0.25) | 16.04 | 38.67 | | |

**Student's t-test by gender. **Pearson's correlations with the Total variable.*

from approximately 50 to 30% (**Figure 4**). The consumption of foods of animal origin showed very different results from each other. In the case of dairy products (**Figure 5**), the percentage of YES responses to the question Did you consume any dairy or milk? remained around 85% throughout the confinement time with a slight upward trend. However, the consumption of animal protein sources such as meat and fish rose to 90–95% toward the second half of the study with several days of 100% affirmative responses. Regarding daily consumption of egg, the trend line remained practically straight around 50–55% of positive responses as shown in **Figure 6**. Foods such as fruits, vegetables, and legumes (**Figure 7**) were consumed in a

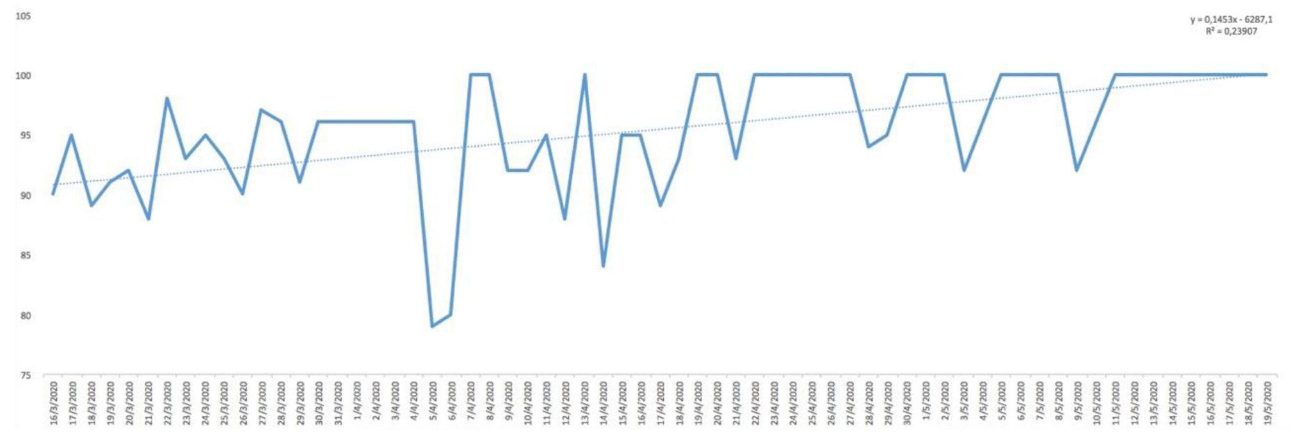


**FIGURE 1 |** YES percentage response analysis of did you drink water yesterday? Depending on the days of study.

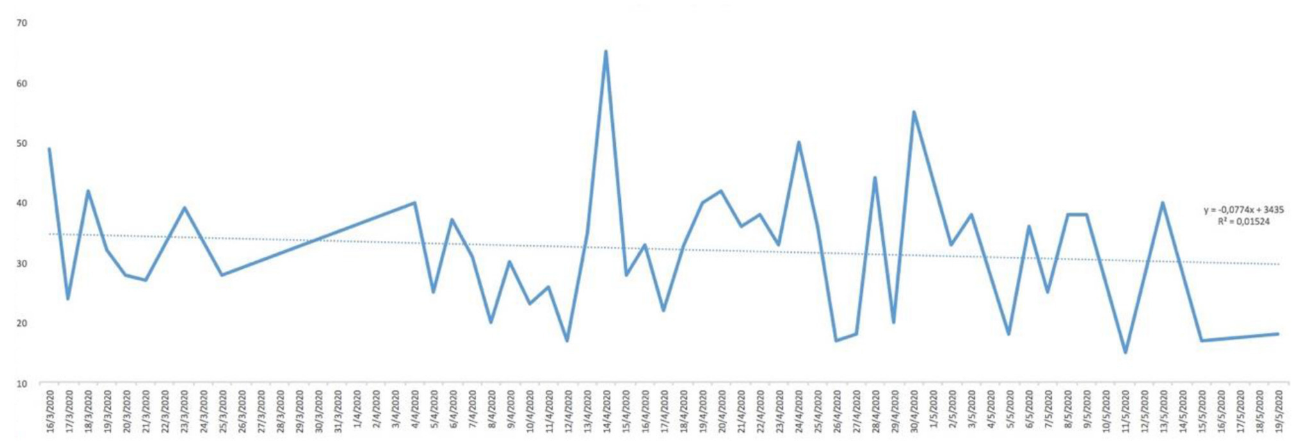


**FIGURE 2 |** YES percentage response analysis of did you drink alcohol? Depending on the days of study.

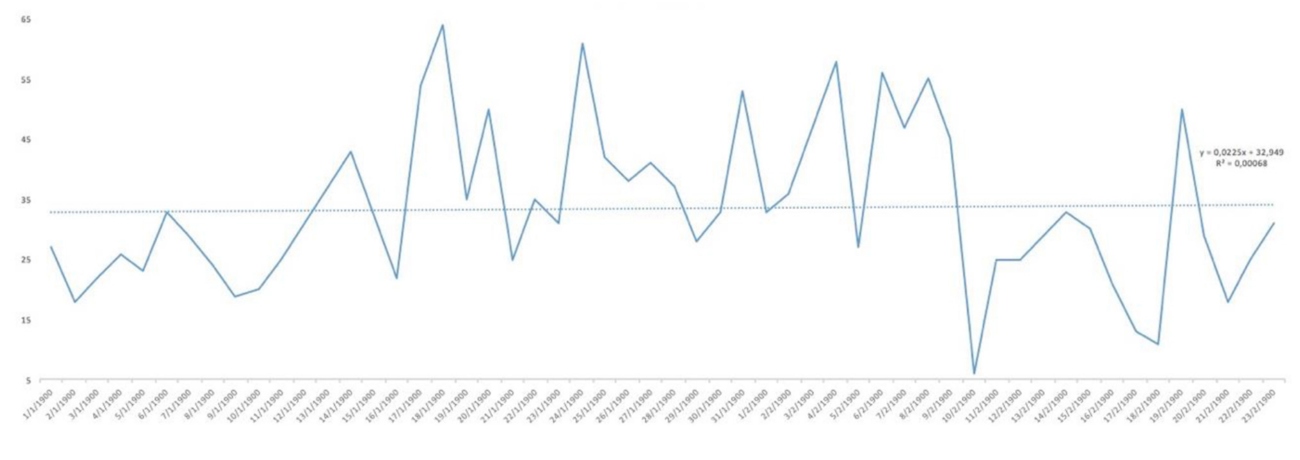


**FIGURE 3 |** YES percentage response analysis of Do you feel heavier than yesterday? Depending on the days of study.

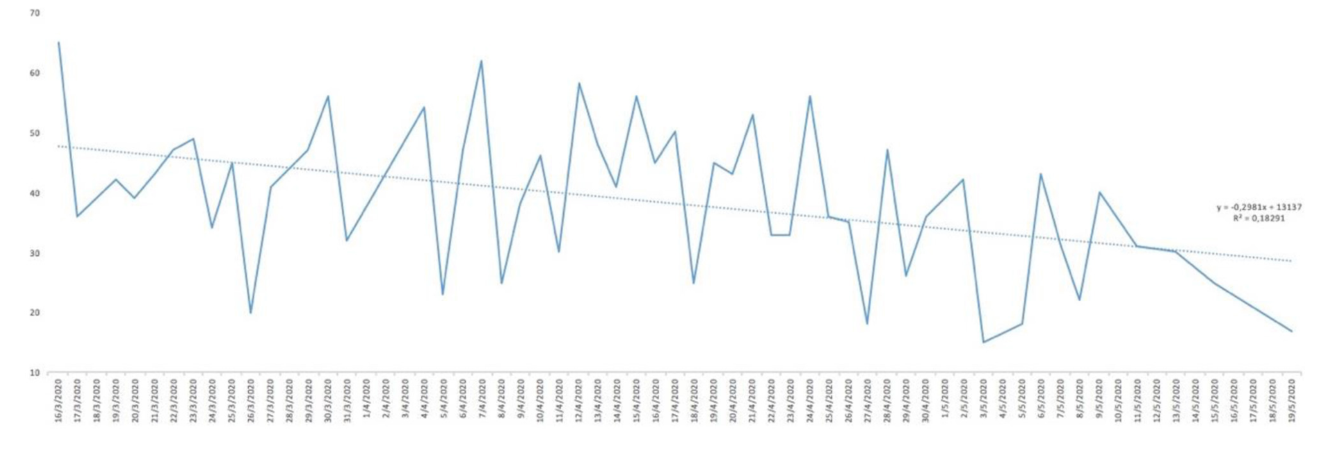


**FIGURE 4 |** YES percentage response analysis of did you sting between hours? Depending on the days of study.

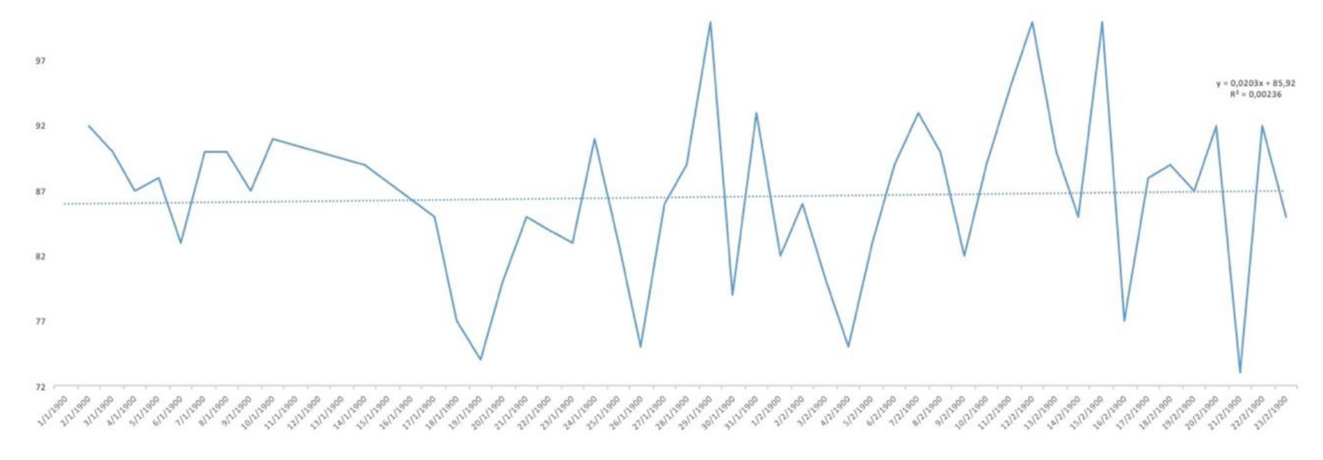


**FIGURE 5 |** YES percentage response analysis of did you consume any dairy or milk yesterday? Depending on the days of study.

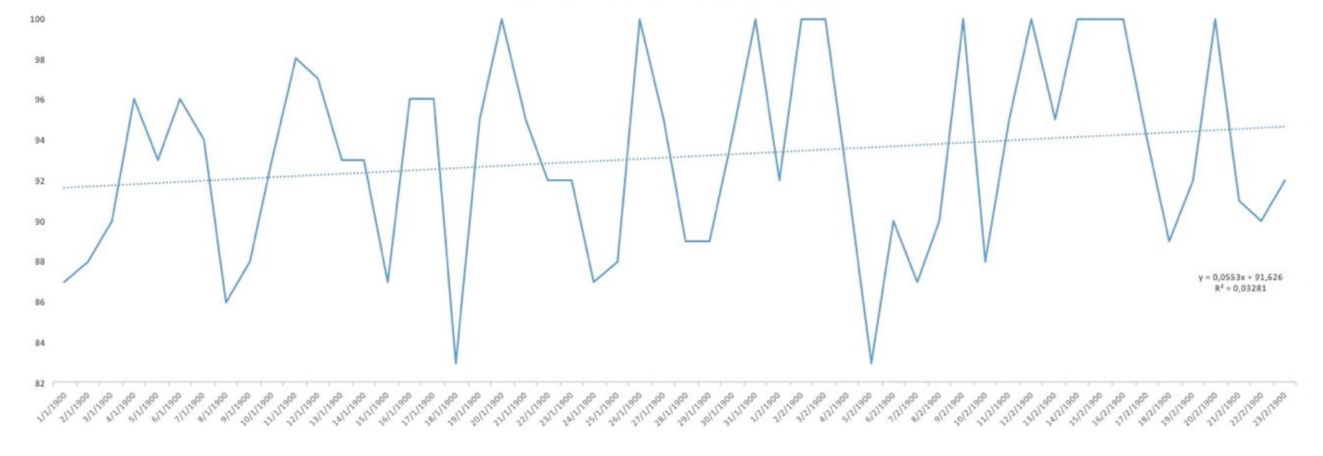


**FIGURE 6 |** YES percentage response analysis of were there eggs in any of your meals? Depending on the days of study.

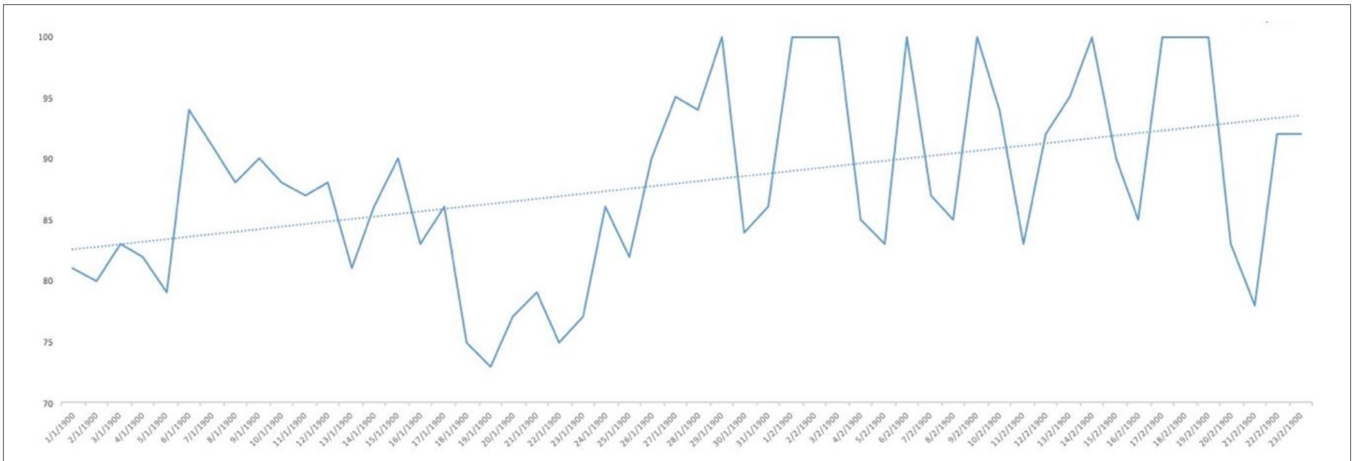


**FIGURE 7 |** YES percentage response analysis of did you eat fruit, vegetables or legumes yesterday? Depending on the days of study.

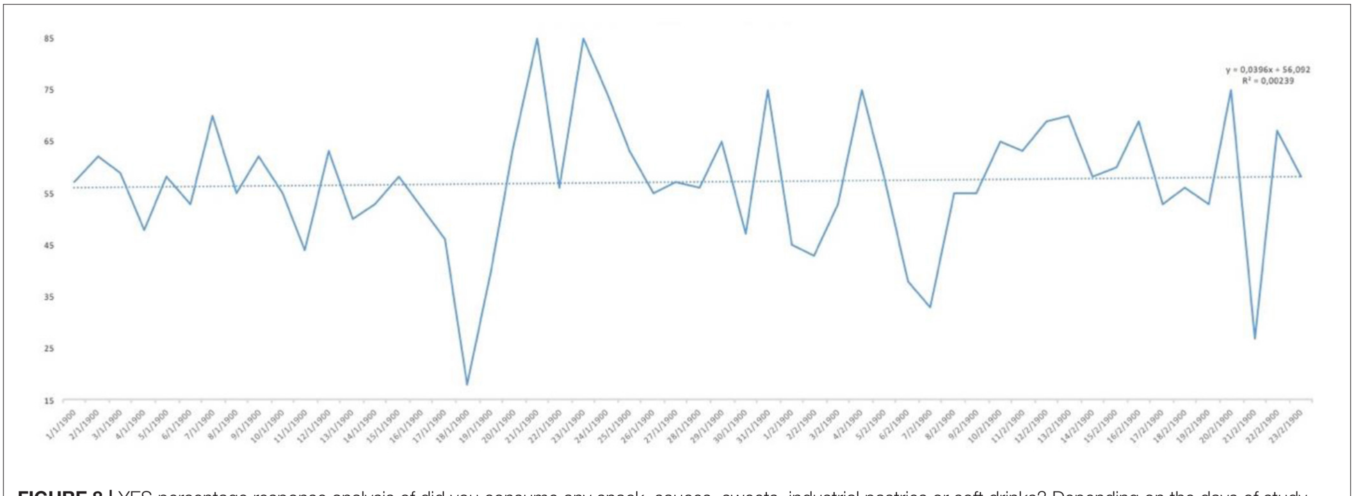


**FIGURE 8 |** YES percentage response analysis of did you consume any snack, sauces, sweets, industrial pastries or soft drinks? Depending on the days of study.

percentage above 90% throughout the entire quarantine, with a slight upward trend. However, the consumption of snacks, sauces, sweets, industrial pastries, or soft drinks fell from 55% to 35% as the progress of the study days (**Figure 8**).

In **Figure 9**, the affirmative responses regarding mood are reflected and, although we see large peaks up and down, the trend line was slightly upwards, reaching 55% positive responses. Finally, as to whether the posts published are useful for the current health of the population, as can be seen in **Figure 10**, the analysis of positive responses began to be high (80%), increasing as the days of confinement passed in general, although with great diversity in the responses as shown by the peaks in the graph.

In **Table 2**, descriptive statistical parameters are analysed, such as the minimum and maximum values of positive and negative responses in percentage, the mean and median of these responses, and the standard deviation (SD). The affirmative answer to the question "Do you feel heavier than yesterday?"(**Figure 11**) is, at some point, presented a maximum of 65%, while in the case of the negative answer, it was 85%. However, the mean of the positive answer corresponds only to 32.15% and 67.84% for the negative answer. The consumption of water showed results of more positive than negative responses. At some point, 100% affirmative responses were declared with a minimum of 79%. In this case, the mean and median correspond to 95% [mean = 95.05% (SD: 5.29) and median = 95%]. The "pecking" between hours declared more disparate responses with a maximum of 65% affirmative responses and a minimum of 15%. The mean and median are not coincident [mean 38.92% (SD: 12.21) and median 41%]. The negative response for alcohol consumption presented higher maximum (94%) and minimum (36%) than for the positive response. However, there were not very uniform results with the mean of 66.42% (SD: 13.65); it should be noted that the median does not have similar value as the mean. In the case of the consumption of fruit, vegetables, and legumes, the percentage of YES remained high throughout the study, since the minimum of positive responses was 83%, which means that only almost 20% in at some point declared not to consume fruit, vegetables, and legumes. The results regarding dairy consumption presented

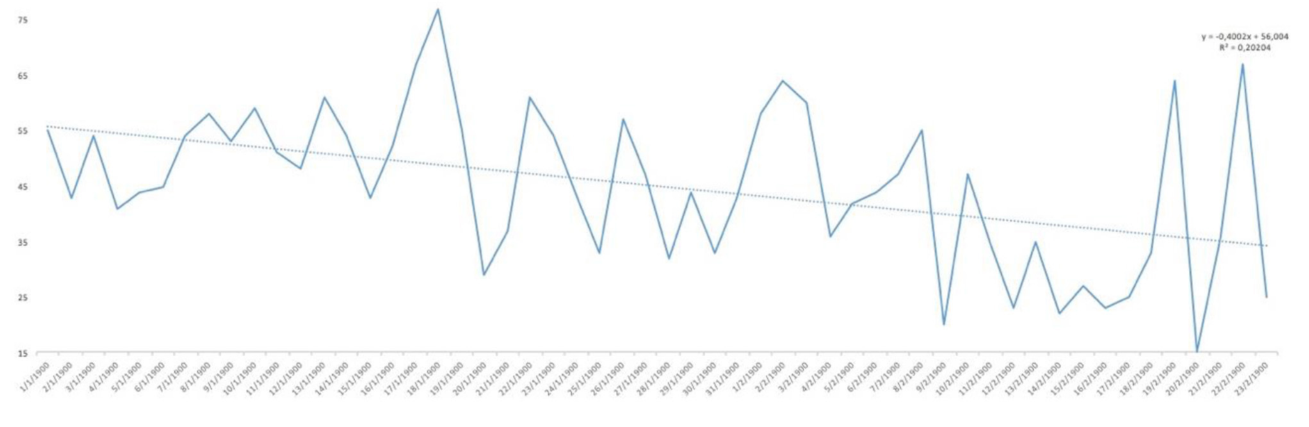


**FIGURE 9 |** YES percentage response analysis of does your mood improve? Depending on the days of study.

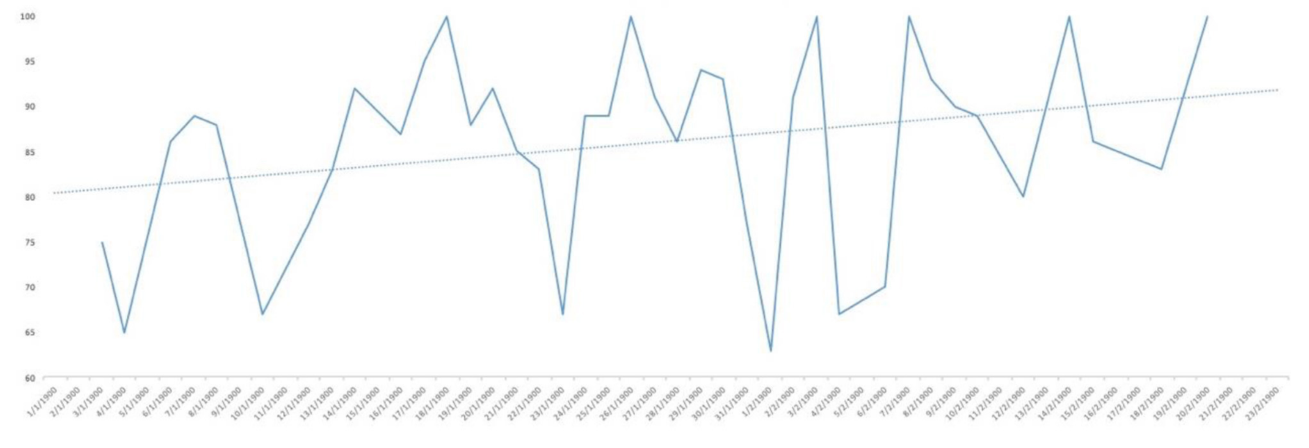


**FIGURE 10 |** YES percentage response analysis of has this post been useful for your health today? Depending on the days of study.

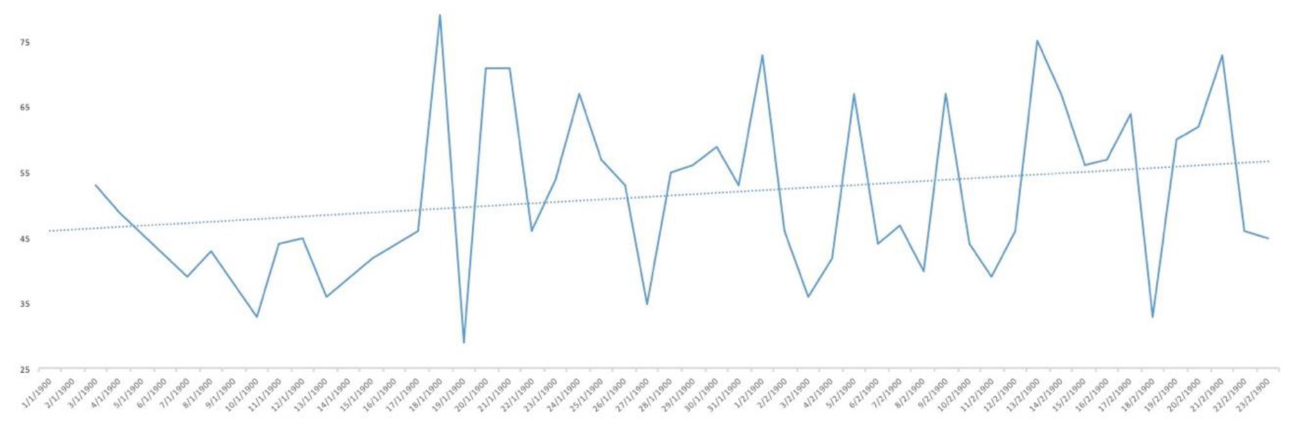


**FIGURE 11 |** YES percentage response analysis of did you eat any meat or fish? Depending on the days of study.

a maximum of 100% affirmative responses and a minimum of 73%. These responses were more uniform among the study participants showing a mean of 86.52% (SD: 6.46) and a median of 87% practically equal to the value of the mean. At some point, a maximum of 85% of the participants declared consuming eggs in their meals, but on the other hand, there was also a maximum of

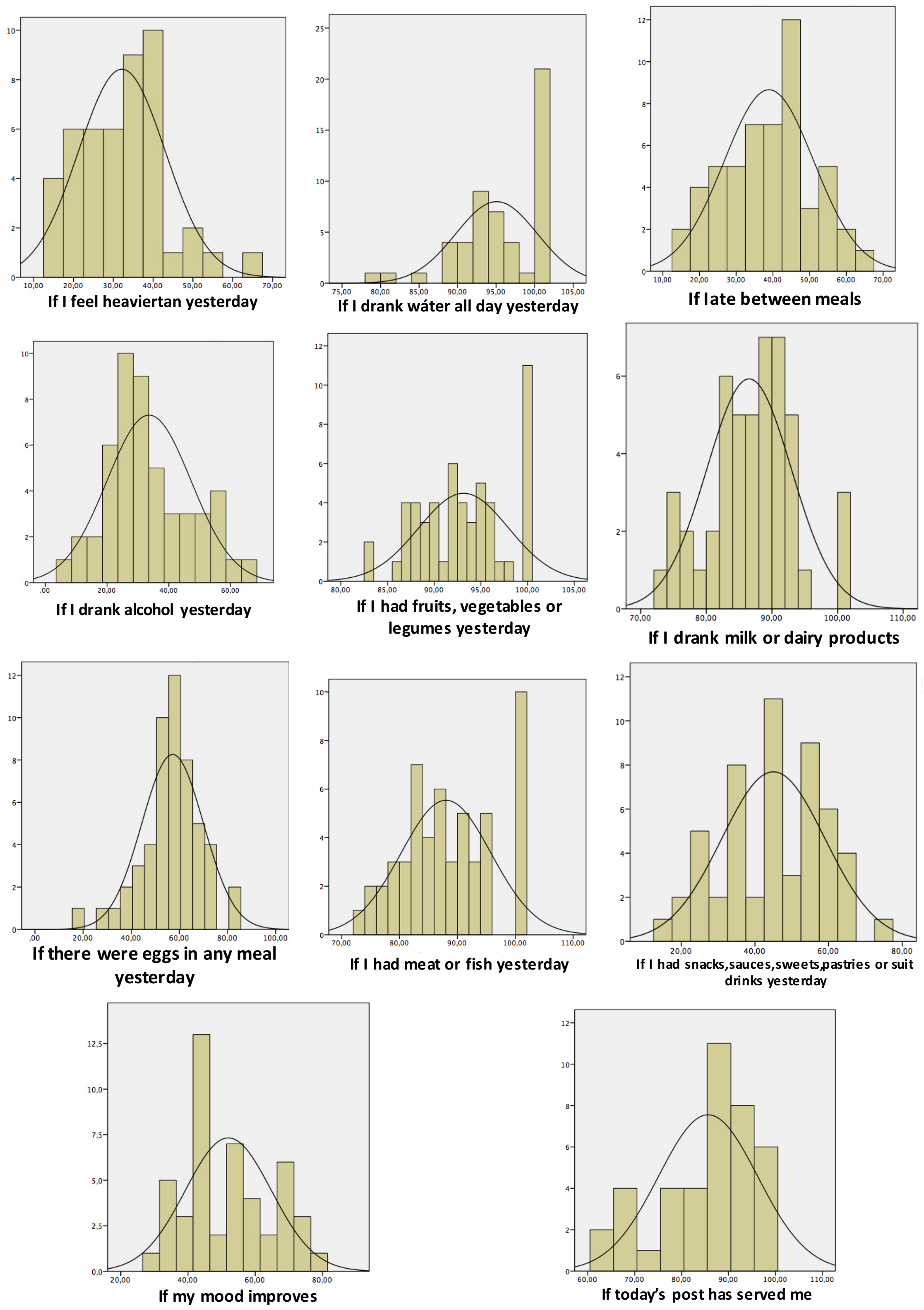


**FIGURE 12 |** Normality curves for each of the questions throughout the study period.

**TABLE 2 |** Percentages of responses for each of the study questions.

| | Minimum % | Maximum % | Mean % | Median SD | % |
|---|---|---|---|---|---|
| Yes I feel heavier than yesterday | 15..00 | 65.00 | 32.15 | 10.89 | 33.00 |
| I don't feel heavier than yesterday | 35.00 | 85.00 | 67.84 | 10.89 | 67.00 |
| Yes I drank water all day yesterday | 79.00 | 100.00 | 95.05 | 5.29 | 95.00 |
| I didn't drink water all day yesterday | 0.00 | 21.00 | 4.94 | 5.29 | 5.00 |
| Yes, I ate between hours yesterday | 15.00 | 65.00 | 38.92 | 12.21 | 41.00 |
| I did not eat between hours yesterday | 35.00 | 85.00 | 61.07 | 12.21 | 59.00 |
| Yes, I drank alcohol yesterday | 6.00 | 64.00 | 33.58 | 13.65 | 31.00 |
| I didn't drink alcohol yesterday | 36.00 | 94.00 | 66.42 | 13.65 | 69.00 |
| Yes I had fruit, vegetables, vegetables or legumes yesterday | 83.00 | 100.00 | 93.14 | 4.80 | 93.00 |
| I did not have fruit, vegetables, vegetables or legumes yesterday | 0.00 | 17.00 | 6.77 | 4.83 | 7.00 |
| Yes I had milk or some dairy yesterday | 73.00 | 100.00 | 86.52 | 6.46 | 87.00 |
| I did not drink milk or any dairy yesterday | 0.00 | 27.00 | 13.47 | 6.46 | 13.00 |
| Yes there were eggs at any meal yesterday | 18.00 | 85.00 | 57.18 | 12.79 | 57.00 |
| There were no eggs at any meal yesterday | 15.00 | 82.00 | 42.62 | 12.94 | 43.00 |
| Yes I had meat or fish yesterday | 73.00 | 100.00 | 88.03 | 7.77 | 87.00 |
| I did not have meat or fish yesterday | 0.00 | 25.00 | 11.74 | 7.56 | 12.50 |
| Yes I had snacks, sauces, sweets, pastries or soft drinks yesterday | 15.00 | 77.00 | 45.00 | 14.01 | 44.50 |
| I did not have snacks, sauces, sweets, pastries or soft drinks yesterday | 23.00 | 85.00 | 55.00 | 14.01 | 55.50 |
| Yes it improves my mood | 29.00 | 79.00 | 52.00 | 12.79 | 49.00 |
| Doesn't improve my mood | 21.00 | 71.00 | 48.00 | 12.79 | 51.00 |
| Yes, today's post has helped me | 63.00 | 100.00 | 85.65 | 10.56 | 88.00 |
| Today's post has not served me | 0.00 | 37.00 | 14.35 | 10.56 | 12.00 |

82% negative responses regarding the consumption of eggs and although mean and medians coincide, presenting a SD > 12 for both cases. The answer to the question Did you eat meat or fish yesterday? (**Figure 11**) is that the total of the responses leaned toward the YES, reaching 100% [mean = 88.03% (SD: 7.77)] in some cases compared to 25% in the case of negative responses [mean = 11.74% (SD: 7.56)]. However, for the consumption of snacks, sweets, industrial pastries, and soft drinks, the results were more controversial. Regarding maximum and minimum, both the affirmative and negative responses reached high values throughout the confinement time (positive: 77% and negative: 85%). For their part, the mean and median remain at almost identical values (45 and 44.5%, respectively, for YES and 55 and 55.5%, respectively, for NO) but the SD with respect to the mean is higher than in other answers (SD: 14.01). The state of mind of the study participants throughout the days, analysing **Table 2**, shows a maximum of 79% affirmative responses compared to a maximum negative 71%, so there was disparity in the results as evidenced by the SD: 12.79 from the mean. Concluding the analysis of **Table 2**, regarding the question Has today's post served you?, the affirmative responses throughout the days were more common (on 1 day of the study duration a 100% YES was reached) with a mean that is 85.65% (SD: 10.56) and median of 88%.

Normality tests: The analysis in the form of histograms for each of the positive responses of the survey establishes symmetric normality curves of normal distribution in all of them (**Figure 12**). Some of them are flatter, as in the case of the consumption of water or fruit, vegetables, and legumes whereas others such as the consumption of dairy or eggs appear more elongated.

*Pearson's correlation analysis between variables: **Table 3** refers to the correlation between the two variables analysed for each question: percentage of affirmative responses and the days of confinement. That is, it confirms whether the results presented in the previous graphs are significant or not, whether there is a correlation between them or not. Pearson's correlation coefficient was used with a statistical significance level of $p \leq 0.05$.* Only some type of correlation was observed for the responses on water consumption, snacking between meals, consumption of meat and fish, and consumption of snacks, sauces, sweet pastries, or soft drinks. The existing correlation for water consumption with respect to the days on which the study was $R = 0.489$ in the coefficient with a significance level of 97% ($p = 0.001$), for which a positive correlation is declared. As the days passed, the consumption of water increased. Regarding the "pecking" between hours, the opposite case occurs: a negative correlation is observed; As the days passed, their consumption decreased as indicated by a negative Pearson's coefficient ($R = -0.428$, $p = 0.001$). The same case is the consumption of snacks, sauces, sweets, industrial pastries, and soft drinks, obtaining a correlation of $R = -0.444$, $p = 0.001$. Of all the data analysed, the last one that presented correlation was the question related to the consumption of meat or fish, obtaining in this case a positive

**TABLE 3 |** Correlations in the evolution of the percentages of affirmative responses during the study period.

| | | Evolution of days |
|---|---|---|
| Yes I feel heavier than yesterday | Pearson's correlation | −0,123 |
| | *P* | 0.414 |
| Yes I drank water all day yesterday | Pearson's correlation | 0.489** |
| | *P* | 0.001 |
| Yes I ate between hours yesterday | Pearson's correlation | −0.428** |
| | *P* | 0.001 |
| Yes I drank alcohol yesterday | Pearson's correlation | 0.031 |
| | *P* | 0.832 |
| Yes I had fruit, vegetables, vegetables or legumes yesterday | Pearson's correlation | 0.174 |
| | *P* | 0.209 |
| Yes I had milk or some dairy yesterday | Pearson's correlation | 0.030 |
| | *P* | 0.839 |
| Yes there were eggs at any meal yesterday | Pearson's correlation | 0.047 |
| | *P* | 0.738 |
| Yes I had meat or fish yesterday | Pearson's correlation | 0.397** |
| | *P* | 0.003 |
| Yes I had snacks, sauces, sweets, pastries or soft drinks yesterday | Pearson's correlation | −0.444** |
| | *P* | 0.001 |
| Yes it improves my mood | Pearson's correlation | 0.235 |
| | *P* | 0.112 |
| Yes, today's post has helped me | Pearson's correlation | 0.284 |
| | *P* | 0.076 |

correlation ($R = 0.397$, $p = 0.003$), for which an increase in this type of food as the days of confinement progress.

## DISCUSSION

Digital nutrition promotion interventions provide an opportunity to address the public health issue of improving people's nutrition (3, 20). There are inherent elements of subjectivity in the interpretation of this case study (3, 21); however, it presents insight into how our subjects (audience) engaged with one science communication endeavour using digital platforms as social media. Zarnowiecki found a positive effect of the digital intervention on child nutrition across a range of dietary outcomes (20). Additionally, having reviewed and analysed the data obtained in Instagram, we consider these data to be of great use for future research related to the implementation of MD, and also to promote nutritional health through social media, applications, blogs, etc. (3, 22–24, 24). The main objective of this study was to provide the population with the information of general interest on the promotion of healthy eating through social media and analyse the impact of its dissemination, in addition to assessing the consumption of food during the period of confinement. The results suggest that a high percentage of the population considered the action positive as indicated by Mariscal-Arcas (3). The study carried out through social media such as *Instagram* was divided into two parts. At first, the profile of a Spanish winter sports federation (FADI, @fadiandalucia) was used; a profile with consolidated followers, which allowed the survey to have a greater reach with more than 1,000 participants compared to the more than 800 participants of the *@mmhealthscience* profile, created during the study period as a scientific knowledge transfer academic platform designed by our nutritional research group in the University of Granada, Spain (3): the age of the population, mostly young adult people (up to 55 years old), the majority profile in the use of social media. This population coincides with that used in a similar study by Mariscal-Arcas (3).

Analysing the results inferred in the study, we can deduce information about the food consumption of our population and also the state of mind during the situation experienced in the days of confinement. In general, it can be seen that the consumption of dairy products, fruits, vegetables, legumes, and eggs is quite high throughout the confinement and linear (it remains high at all times, except in the case of fruits, vegetables, and legumes that rises as the progress of study days). It could be because the study was carried out in a Spanish population, almost entirely Andalusian Region and Murcia Region, so that the MD is very present in daily life, and it is not affected by confinement (neither increasing nor decreasing). However, along the same lines as the Mediterranean pattern, it should be noted that the consumption of animal protein sources, such as meat and fish, increases throughout the quarantine period with no differences between gender as described by Tkachenko (25). Therefore, perhaps influenced by having more time to devote to cooking, a consequence of the restriction of the population's freedom of movement seems that more time has been spent making more sophisticated and different recipes, trying new foods, etc. This theory is maintained in other studies, carried out during the pandemic, such as the one carried out by Di Renzo in the Italian population. In that, the increase in the consumption of natural products to the detriment of processed ones, an increase in the preparation of meals, and a greater adherence to the Mediterranean pattern are defended although it was already high (26). It could also be due to a greater concern for health these days where the fact of maintaining an adequate state of health acquires special relevance. This theory of greater self-care for health would be supported by Luzi and Radaelli who speak of being overweight as the main risk factor not only because it increases the risk of infection and complications for obese people, but also a high prevalence of obese individuals within the population, and it may increase the chance of the emergence of a more virulent viral strain, prolong the shedding of the virus throughout the population, and could eventually increase the overall death rate from a flu pandemic (27).

It could also be due to a greater concern for health these days where the fact of maintaining an adequate state of health acquires special relevance. This theory of greater self-care for health would be supported by Luzi and Radaelli (27) who speak of being overweight as the main risk factor not only because it increases the risk of infection and complications for obese people, but also a high prevalence of obese individuals within the population, and it may increase the chance of the emergence of a more virulent viral strain, prolong the shedding of the virus throughout the population, and could eventually increase the overall death rate from a flu pandemic along these lines of increased health concern, one could speak of the reduction in alcohol consumption declared, with higher consumption in the middle of the period. According to the study carried out by Sidor (28), alcohol consumption in some people is increased due to the stress produced by prolonged stays at home, although it only accounts for 14% of the study population (28). However, in this case, no statistically significant correlation is observed according to Pearson's correlation analysis, therefore, confinement would have nothing to do with the consumption or not of alcoholic beverages. It is a stable habit in the population that it could be said that is not affected by the stress that a confinement situation can entail or the uncertainty of not knowing what will happen (work situation, evolution in the number of cases of infected by SARS-CoV-2) or the idea of drinking to escape. In addition, the consumption of 'junk food' such as snacks, industrial pastries, sweets, and soft drinks is also reduced over time and also 'snacking' between meals. In both cases, there is a significant negative correlation (97%), that is, as the days of confinement go by, the study population decreases the consumption of this type of products and also snacking, a theory confirmed by Di Renzo (26).

Water consumption does present an upward trend line throughout the confinement, although from the beginning, the percentage of YES responses remains high, with a daily average that does not fall below 90% with a positive correlation and a significance level of 97%, so the passage of days would be related to the increase in water consumption. We could also relate it to an increase in the subjects' concern for the state of health (19, 29). A recent review highlights that balanced nutrition can help maintain immunity, which is essential for the prevention and treatment of viral infections (30). We can conclude that the period of confinement due to the COVID-19 pandemic in Spain has induced changes in diet with a trend toward a higher consumption of fruit, vegetables, pulses, and fish and a lower consumption of bakery products, sweets, salty snacks, sugary drinks, and drinks with a high alcoholic content as other authors have found (19).

The reduction in physical activity to which the population has been subjected could have caused a greater sensation of heaviness on the part of the subjects; however, according to the data obtained from the results of the study, it does not seem to have influenced since it remains fairly stable throughout the days and does not present a significant correlation. The mood question of this study improves as the days go by, going from a daily average of 45% affirmative responses to 55% toward the end of confinement. It could be related to the end of the confinement situation (although the state of alarm is maintained, prohibitions are being lifted and they begin to let us go outside) although it could also be thought that it may be due to an adaptation to the new situation. In this last line, it is possible to establish a relationship with the state of imprisonment of criminals serving a sentence. Echeverri (31) analyses the psychological effects and the evolution of prisoners and establishes that, contrary to what was thought, as life in freedom is closer, the inmate manifests a greater conformity with the established social norms. The daily reality is imposed since the person who has been confined in prison for many years, as he sees the possibility of reintegrating himself into life in freedom, progressively adapts his behaviour to the social norms that he himself transgressed (31). However, the data referring to the correlation between both variables (passage of days and mood question) were not positive. According to Pearson's correlation analysis, there is no relationship between variables, so none of the theories outlined above would be valid. There is a good acceptance by the population regarding the daily post that is published since it began with 80% of YES responses with a slight rise over the days. Communicating nutrition information is a good strategy to reduce boredom, stress, and anguish caused by having to stay at home (3, 32); however, if the correlation between variables in **Table 3** is analysed, there is no significant correlation, so it can be said that publishing publications on a daily basis does not impact so much over time. Therefore, perhaps it would be more interesting to select good information with the intention of having a great impact but more sporadically.

The nutritional health information through the social media claim to improve the health status of the participants. Other studies have found that the influence of the confinement situation due to COVID-19 may have on the use of the Internet as a source of information on health-related issues in the world population. This situation of confinement could have affected the perception, the feeding, and the use of social media (3). The diffusion through social media has allowed to have a greater reach of the population, including different parts of Spain that are distant from each other and, for the most part, the young adult population. We observed that mood throughout confinement generally improves, although it is not related to the passage of time in confinement. There are certain eating habits from the MD that are well established in the daily diet of our population, such as the consumption of fruits, vegetables, legumes, dairy products, and eggs. It seems that enjoying good health is a growing concern in pandemic situations, which is why inappropriate behaviours such as "snacking" between meals or the consumption of processed foods such as snacks, industrial pastries, soft drinks, and sweets are avoided, increasing the amount of healthy food such as meat and fish. Future studies should examine that the effect these habits may have in quarantine situations. Finally, we conclude that the daily publication of posts is not as impressive sustained over time compared to if articles of higher quality of content are released more sporadically. This study, as indicated by Mariscal-Arcas (3), opens up future avenues of research promoting MD and implements new cohort nutritional databases, especially about young adult people, who are adept at navigating digital spaces and therefore using social media (24, 33–38).

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethics Committee of the Andalusian Public Health Service, Spain. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

The study was designed by MM-A. Data were collected and analysed by MM-A, SD-M, BS, AB-D, JL, MM-B, NG-B, JC-P, LC, AL-M, and MJ-C. Data interpretation and manuscript preparation were undertaken by MM-A, JC-P, LC, AL-M, and MJ-C. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the Andalusian Regional Government (Nutrition, Diet and Risks Assessment: AGR255), FEDER-ISCIII PI14/01040 and Consejería de Transformación Económica, Industria, Conocimiento y Universidades, Junta de Andalucía P18-RT-4247.

## ACKNOWLEDGMENTS

This paper will be part of SD-M's Master Thesis, being completed as part of the Food and Fit Master Program at the University of Granada, Spain. We express our gratitude for the cooperation we received from "Whats Creative Studio", "Sola Communication" and "MM Health Science" through social media.

## REFERENCES

1. Sever L, Balat A. Renal crisis in children during armed conflict. *Semin Nephrol.* (2020) 40:408–20. doi: 10.1016/j.semnephrol.2020.06.008
2. Alassiri SA, Alowfi AS. Public's Attitudes toward health information on twitter: a cross-sectional survey based on the Saudi population. *Cureus.* (2019) 11:e5863. doi: 10.7759/cureus.5863
3. Mariscal-Arcas M, Jimenez-Casquet MJ, Saenz de Buruaga B, Delgado-Mingorance S, Blas-Diaz A, Cantero L, et al. Use of social media, network avenues, blog and scientific information systems through the website promoting the Mediterranean Diet as a method of a health safeguarding. *Front Commun*. (2021) 6. doi: 10.3389/fcomm.2021.599661
4. Rodríguez-Palmero M. Efectos beneficiosos de la dieta mediterránea. *Offarm.* (2000) 19:104–9.
5. Zommiti M, Feuilloley MGJ, Connil N. Update of probiotics in human world: a nonstop source of benefactions till the end of time. *Microorganisms.* (2020) 8:E1907. doi: 10.3390/microorganisms8121907
6. Heras-Gonzalez L, Latorre JA, Martinez-Bebia M, Gimenez-Blasi N, Olea-Serrano F, Mariscal-Arcas M. Sugar consumption in schoolchildren from southern Spain and influence on the prevalence of obesity. *PLoS ONE.* (2020) 15:e0242602. doi: 10.1371/journal.pone.0242602
7. Giménez-Blasi N, Latorre Rodríguez JA, Martinez Bebia M, Olea-Serrano F, Mariscal Arcas M. Comparison of diet quality between young children and adolescents in the Mediterranean basin and the influence of life habits. *Nutr Hosp.* (2019) 36:387–93. doi: 10.20960/nh.2110
8. Cheikh Ismail L, Osaili TM, Mohamad MN, Al Marzouqi A, Jarrar AH, Abu Jamous DO, et al. Eating habits and lifestyle during COVID-19 lockdown in the United Arab Emirates: a cross-sectional study. *Nutrients.* (2020) 12:3314. doi: 10.3390/nu12113314
9. Di Santo SG, Franchini F, Filiputti B, Martone A, Sannino S. The effects of COVID-19 and quarantine measures on the lifestyles and mental health of people over 60 at increased risk of dementia. *Front Psychiatry.* (2020) 11:578628. doi: 10.3389/fpsyt.2020.578628
10. Courtet P, Olié E, Debien C, Vaiva G. Keep socially (but not physically) connected and carry on. *J Clin Psychiatry*. 2020 81:20com13370. doi: 10.4088/JCP.20com13370
11. Hamdan M, Monteagudo C, Lorenzo-Tovar ML, Tur JA, Olea-Serrano F, Mariscal-Arcas M. Development and validation of a nutritional questionnaire for the palestine population. *Public Health Nutr*. (2014) 17:2512–8. doi: 10.1017/S1368980013002711
12. Mariscal-Arcas M, Caballero-Plasencia ML, Monteagudo C, Hamdan M, Pardo-Vasquez MU, Olea-Serrano F. Validation of questionnaires to estimate adherence to the Mediterranean diet and life habits in older individuals in Southern Spain. *J Nutr Health Aging.* (2011) 15:739–43 doi: 10.1007/s12603-011-0136-z
13. Mariscal-Arcas M, Velasco J, Monteagudo C, Caballero-Plasencia MA, Lorenzo-Tovar ML, Olea-Serrano M. Comparison of methods to evaluate the quality of the Mediterranean diet in a large representative sample of young people in Southern Spain. *Nutr Hosp*. (2010) 25:1006–13. doi: 10.3305/nh.2010.25.6.4889
14. Rivas A, Romero M, Mariscal C, Monteagudo J, Hernandez F, Olea-Serrano. Validation of questionnaires for the study of food habits and bone mass. *Nutr Hosp*. (2009) 24:521–8.
15. FAO/WHO. *Preparation and Use of Food-Based Dietary Guidelines*. Geneva, Switzerland: Nutrition Programme-WHO (1996).
16. World Health Organization. *Global Action Plan for the Prevention and Control of NCDs 2013–2020*. Geneva, Switzerland: WHO (2013).
17. World Health Organization. *Sustainable Development Goals (SDGs)*. Geneva, Switzerland: WHO. Available online at: https://www.who.int/sdg/en/ (accessed Feburary 1, 2020).
18. FAO/WHO. *Fats and Fatty Acids in Human Nutrition: Report of an Expert Consultation.* Rome, Italy: FAO Food and Nutrition Paper No. 91; FAO (2010).
19. Aranceta-Bartrina J, Partearroyo T, López-Sobaler AM, Ortega RM, Varela-Moreiras G, Serra-Majem L, et al. Updating the food-based dietary guidelines for the Spanish population: the Spanish society of community nutrition (SENC) proposal. *Nutrients*. (2019) 11:2675 doi: 10.3390/nu111 12675
20. Zarnowiecki D, Mauch CE, Middleton G, Matwiejczyk L, Watson WL, Dibbs J, et al. Systematic evaluation of digital nutrition promotion websites and apps for supporting parents to influence children's nutrition. *Int J Behav Nutr Phys Act.* (2020) 17:17. doi: 10.1186/s12966-020-0915-1
21. Nieubuurt JT. Internet memes: leaflet propaganda of the digital age. *Front Commun.* (2020). 5:1–14. doi: 10.3389/fcomm.2020.5470
22. Hirasawa R, Saito K, Yachi Y, Ibe Y, Kodama S, Asumi M, et al. Quality of internet information related to the Mediterranean diet. *Public Health Nutr.* (2012) 15:885–93. doi: 10.1017/S1368980011002345
23. Gkouskou K, Markaki A, Vasilaki M, Roidis A, Vlastos I. Quality of nutritional information on the internet in health and disease. *Hippokratia.* (2011) 15:304–7.
24. Pavelle S, Wilkinson C. Into the digital wild: utilizing Twitter, Instagram, Youtube, and Facebook for effective science and environmental communication. *Front Commun.* (2020) 5:82. doi: 10.3389/fcomm.2020.575122
25. Tkachenko N, Chotvijit S, Gupta N, Bradley E, Gilks C, Guo W, et al. Google trends can improve surveillance of type 2 diabetes. *Sci Rep.* (2017) 7:4993. doi: 10.1038/s41598-017-05091-9
26. Di Renzo L, Gualtieri P, Pivari F, Soldati L, Attinà A, Cinelli G, et al. Hábitos alimenticios y cambios en el estilo de vida durante el encierro de COVID-19: una encuesta italiana. *Revista de medicina traslacional.* 18:229. doi: 10.1186/s12967-020-02399-5

27. Luzi L, Radaelli MG. Influenza and obesity: its odd relationship and the lessons for COVID-19 pandemic. *Acta Diabetol junio de.* (2020) 57:759–64. doi: 10.1007/s00592-020-01522-8
28. Sidor A, Rzymski P. Dietary choices and habits during COVID-19 lockdown: experience from Poland. *Nutrients.* (2020) 12:1657. doi: 10.3390/nu12061657
29. Pérez-Rodrigo C, Gianzo Citores M, Gotzone Hervás B, Ruiz Litago F, Casis Sáenz L, Aranceta-Bartrina J. Cambios en los hábitos alimentarios durante el periodo de confinamiento por la pandemia COVID-19 en España. *Rev Esp Nutr Comunitaria*. (2020) 26:1–17.
30. Jayawardena R, Sooriyaarachchi P, Chourdakis M, Jeewandara C, Ranasinghe P. Mejora de la inmunidad en infecciones virales, con especial énfasis en COVID-19: una revisión. Diabetes Metab Syndr Clin Res Rev. (2020) 14:367–82. doi: 10.1016/j.dsx.2020.04.015
31. Echeverri-Vera JA. La prisionalización, sus efectos psicológicos y su evaluación. *Pensando Psicología.* (2010) 6:157–66. Available online at: https://revistas.ucc.edu.co/index.php/pe/article/view/375
32. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet.* (2020) 395:912–20. doi: 10.1016/S0140-6736(20)30460-8
33. Greenan KA. The Influence of virtual education on classroom culture. *Front Commun.* (2021). 6:1–4. doi: 10.3389/fcomm.2021.641214
34. Solano P, Ustulin M, Pizzorno E, Vichi M, Pompili M, Serafini G, et al. Google-based approach for monitoring suicide risk. *Psychiatry Res.* (2016) 246:581–6. doi: 10.1016/j.psychres.2016.10.030
35. Strotman PK, Novicoff WM, Nelson SJ, Browne JA. Increasing public interest in stem cell injections for osteoarthritis of the hip and knee: a Google trends analysis. *J Arthroplasty.* (2019) 34:1053–7. doi: 10.1016/j.arth.2019.03.002
36. Rossignol L, Pelat C, Lambert B, Flahault A, Chartier-Kastler E, Hanslik T, et al. Method to assess seasonality of urinary tract infections based on medication sales and google trends. *PLoS ONE.* (2013) 8:e76020. doi: 10.1371/journal.pone.0076020
37. Groffman PM, Stylinski C, Nisbet MC, Duarte CM, Jordan R, Burgin A, et al. Restarting the conversation: challenges at the interface between ecology and society. *Front Ecol Environ.* (2010) 8:284–91. doi: 10.1890/090160
38. Hargittai E, Füchslin T, Schäfer MS. How Do Young Adults Engage with Science and Research on Social Media? Some Preliminary Findings and an Agenda for Future Research. *Social Media + Society July.* (2018) 4:1–10. doi: 10.1177/2056305118797720

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 21 December 2021
doi: 10.3389/fnut.2021.785227

# Priority Micronutrient Density of Foods for Complementary Feeding of Young Children (6–23 Months) in South and Southeast Asia

*Flaminia Ortenzi*[1,2*†] *and Ty Beal*[3,4†]

[1] Consultant, Geneva, Switzerland, [2] Global Studies Institute, University of Geneva, Geneva, Switzerland, [3] Global Alliance for Improved Nutrition (GAIN), Washington, DC, United States, [4] Department of Environmental Science and Policy, University of California, Davis, Davis, CA, United States

OPEN ACCESS

**Edited by:**
Alessandra Durazzo,
Council for Agricultural Research and Economics, Italy

**Reviewed by:**
Aly Diana,
Universitas Padjadjaran, Indonesia
Azeb Lelisa Tolossa,
Micronutrient Initiative, Ethiopia
Nadia Shabnam,
National University of Medical Sciences, Pakistan

***Correspondence:**
Flaminia Ortenzi
flamyortenzi@gmail.com

† These authors have contributed equally to this work

**Specialty section:**
This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition

**Received:** 28 September 2021
**Accepted:** 03 December 2021
**Published:** 21 December 2021

**Citation:**
Ortenzi F and Beal T (2021) Priority Micronutrient Density of Foods for Complementary Feeding of Young Children (6–23 Months) in South and Southeast Asia. Front. Nutr. 8:785227. doi: 10.3389/fnut.2021.785227

**Background:** Given their high nutrient requirements and limited gastric capacity, young children during the complementary feeding period (6–23 months) should be fed nutrient-dense foods. However, complementary feeding diets in low- and middle-income countries are often inadequate in one or more essential micronutrients. In South and Southeast Asia infants' and young children's diets are commonly lacking in iron, zinc, vitamin A, folate, vitamin $B_{12}$, and calcium, hereafter referred to as priority micronutrients.

**Objective:** This study aimed to identify the top food sources of priority micronutrients among minimally processed foods for complementary feeding of children (6–23 months) in South and Southeast Asia.

**Methods:** An aggregated regional food composition database for South and Southeast Asia was built, and recommended nutrient intakes (RNIs) from complementary foods were calculated for children aged 6–23 months. An approach was developed to classify foods into one of four levels of priority micronutrient density based on the calories and grams required to provide one-third (for individual micronutrients) or an average of one-third (for the aggregate score) of RNIs from complementary foods.

**Results:** We found that the top food sources of multiple priority micronutrients are organs, bivalves, crustaceans, fresh fish, goat, canned fish with bones, and eggs, closely followed by beef, lamb/mutton, dark green leafy vegetables, cow milk, yogurt, and cheese, and to a lesser extent, canned fish without bones.

**Conclusions:** This analysis provided insights into which foods to prioritize to fill common micronutrient gaps and reduce undernutrition in children aged 6–23 months in South and Southeast Asia.

**Keywords: nutrient density, complementary feeding, South Asia, Southeast Asia, micronutrient deficiencies, 6–23 months, animal-source foods, dark green leafy vegetables**

# INTRODUCTION

The first 2 years of a child's life represent a "critical window" for the achievement of optimal growth and health, for which adequate nutrition is an essential prerequisite. The reversion of stunting becomes difficult after a child reaches 2 years of age, further demonstrating the crucial importance of intervening during this time. This is also the age when growth impairments, deficiencies in certain micronutrients, and common childhood illnesses are most likely to happen (1). In particular, according to the World Health Organization (WHO) and other available literature, the highest occurrence of stunting in low- and middle-income countries is registered during the complementary feeding period, which corresponds to 6–23 months of age (2, 3).

Africa and Asia are the regions where the largest proportion of malnourished children under 5 years of age live. Together, they account for more than nine out of ten of all stunted children and more than nine out of ten of all wasted children globally (4). Within the Asian region, South Asia is home to about a quarter of the world's children under 5 years and has the highest percentages and numbers of stunted (30.7%, $n = \sim54$ million) and wasted (14.1%, $n = 25$ million) children (5), closely followed by Southeast Asia which has the second highest percentages and numbers of stunted (27.4%, $n = \sim15$ million) and wasted (8.2%, $n = 4.6$ million) children (5). Inadequate diets during the complementary feeding period are among the key determinants of the child-stunting and wasting crisis in South and Southeast Asia (2, 6).

The WHO defines complementary feeding as the transition time when other foods and liquids, along with breast milk, are needed to meet a child's nutrient requirements and are gradually introduced in their diet (1, 3). Infants and young children 6–23 months of age have very high nutrient requirements per unit body weight, because of their rapid growth and development rates. Continued breastfeeding can significantly contribute to meeting their nutrient needs, but has to be complemented with a variety of nutritionally adequate and safe family foods (1). Children aged 6–23 months have a limited gastric capacity and can only consume small quantities of food, therefore, complementary foods should have high nutrient density (amount of each nutrient per 100 kcal of food) (1, 2). Unfortunately, in practice this is often not the case, and diets of infants and young children are often not sufficiently diverse and nutrient dense and are lacking in one or more essential micronutrients, especially in low- and middle-income countries, but also in high-income countries [e.g., iron and zinc in the US; (1, 2, 7)].

In particular, in South and Southeast Asia important micronutrient gaps were identified in infants' and young children's diets (7–10). The micronutrients most commonly known to be lacking and those of highest public health significance in the two regions are iron, zinc, vitamin A, folate, vitamin $B_{12}$, and calcium, hereafter referred to as priority micronutrients (7–10). Among the main drivers of micronutrient malnutrition in infants and young children living in South and Southeast Asia are the following: (1) the low micronutrient density and lack of diversity of complementary foods, with most children aged 6–23 months having a primarily cereal-based complementary feeding diet and not consuming the minimum recommended number of food groups each day (6, 7, 10–12); and (2) the inappropriate marketing of nutritionally inadequate, often ultra-processed, complementary foods, and beverages that are promoted as suitable for this age group (11).

Given the crucial importance of adequate nutrition in infancy and early childhood and its impact on children's present and future lives and on societies as a whole, improving the overall quality and diversity of infants' and young children's diets is critical to the achievement of the Sustainable Development Goals (SDGs) (6). The purpose of this study is to identify the top food sources of priority micronutrients, among minimally processed, inherently nutrient-dense foods, to support efforts to reduce micronutrient malnutrition among complementary fed children (6–23 months) in the South and Southeast Asian regions.

# METHODS

## Calculating Recommended Nutrient Intakes

Based on a previously adopted approach to identify affordable nutritious complementary foods (7, 13), Recommended Nutrient Intakes (RNIs) from complementary foods for children aged 6–23 months were calculated from:

- The WHO and the Food and Agriculture Organization (FAO) recommendations (14) for calcium, zinc, and iron;
- The Institute of Medicine (IOM) recommendations (15) for folate, vitamin A, and vitamin $B_{12}$

Average Energy Requirements (AERs) from complementary foods for children 6–23 months of age were calculated from the WHO and the United Nations Children's Fund (UNICEF) recommendations, accounting for average breast milk intake in developing countries (1, 3). The RNI is the intake level sufficient to meet the daily nutrient requirements of almost all individuals (97.5%) in a specific age and gender group (14); while the Average Requirement (AR) is the intake level that is adequate for half (50%) of the individuals in a given population group (16). We decided to use RNIs for the six included micronutrients rather than ARs because this study focuses on achieving micronutrient adequacy for individuals, not on estimating adequacy at the population level [(17); see **Supplementary Material** for a more detailed explanation of how RNIs and AERs from complementary foods were calculated].

## Building a Regional Food Composition Database for South and Southeast Asia

A regional food composition database, representative of the nutritional value of foods in South and Southeast Asia, was built[1], including values for calories, phytate (18), and for the six identified priority micronutrients: vitamin A, folate, vitamin $B_{12}$, calcium, iron, and zinc. Nutrient values were obtained from US Department of Agriculture (USDA) FoodData Central (FDC)

[1]Fortified foods not included.

(19) and from several South and Southeast Asian countries' food composition tables (FCTs) (20): Bangladesh, Indonesia, Laos, Vietnam, and Thailand. Values from FDC were included in the calculations of composite nutrient values, serving as a reference to ensure plausibility of values from the selected national FCTs and, on a few occasions, they were also used to replace missing values from other FCTs[2]. All foods were included in the forms typically consumed, which could be raw (e.g., fruits), cooked (e.g., meat and poultry, pulses), or a mix of both (e.g., vegetables). For foods typically consumed in their cooked form, but with values only available as raw in the included FCTs, weight yields and nutrient retention factors for different cooking methods were applied (21).

Foods presenting low nutrient density variance (e.g., pulses, refined grains) or likely to be targeted as a food group rather than individually in policy and programming (e.g., dark green leafy vegetables, cheese) were aggregated for analysis. For foods analyzed individually, nutrient values were obtained by calculating the medians of country-level composite values from the selected FCTs. Composite values were obtained by averaging nutrient values for different cooking methods (as well as raw foods, if applicable) and for different cuts of the animal in the case of meat and poultry. For aggregated food groups, nutrient values were calculated by averaging regional-level composite values from South and Southeast Asia and from FDC. Composite values were obtained by calculating the medians of nutrient values for several individual foods (e.g., spinach, kale, amaranth leaves, and others) within a given food group (e.g., dark green leafy vegetables), which were derived from multiple included FCTs located in the South and Southeast Asian regions.

The bioavailability of iron and zinc was accounted for in the analysis. Foods included in the regional food composition database were assigned to one of three levels of iron absorption: 20% for ruminant meat, 15% for all other animal-source foods (ASFs), and 10% for all plant-source foods (PSFs), based on the proportion of heme to non-heme iron contained (22). The following heme-iron percentages were assumed: 68% in ruminant meat, including beef (23–25), goat, and lamb/mutton (25, 26); 39% in pork (24, 25, 27–29); 26% in chicken (24, 25, 27–29), fish and seafood (24, 27–30), eggs and dairy (28); and 40% in all other meat, including offal (23, 28, 29). In addition, foods were categorized into four levels of zinc absorption: 44, 35, 30, and 26%, based on the amount of phytate contained in a portion equivalent to one-third of daily mass intake, assuming an energy density of 1.3 kcal/g[3] and considering average energy requirements for a moderately active woman of reproductive age (16). A similar approach was previously used to build a global food composition database [(32); see **Supplementary Material** for a more detailed explanation of the approach adopted to develop the regional food composition database for South and Southeast Asia].

[2]In particular, vitamin $B_{12}$ from the Bangladesh FCT.

[3]Obtained by averaging the energy densities of a minimally processed plant-based, low-fat diet and a minimally processed, animal-based, ketogenic diet (31).

## Aggregate and Individual Priority Micronutrient Density Ratings

Foods were categorized into four levels of priority micronutrient density based on the portion (calories and grams) needed to achieve one-third (for individual ratings) or an average of one-third (for the aggregate rating) of RNIs from complementary foods for the six selected micronutrients (vitamin A, folate, vitamin $B_{12}$, calcium, iron, and zinc). For the aggregate score, the Average Share of Recommended Intakes (*ASRI*) across the six micronutrients (*A*), for a given quantity of grams (*i*), of a given food (*j*), was calculated as:

$$ASRI_{i,j} = \frac{1}{|A|} \sum_{a \in A} \min\{\frac{nutrient\_density_{a,j} * i}{recommended\_intakes_a},\ 1\}$$

A similar approach was previously used to identify nutrient-dense foods for infants and young children and for other population groups (7, 13, 32). As illustrated in the formula, each micronutrient's contribution was capped at 100% of RNIs, meaning that each micronutrient can contribute from 0% up to 50% of the overall score. This choice was made to ensure that foods would only be rated high if they were high in at least two out of the six micronutrients included in the analysis; and to prevent foods with very high densities of individual micronutrients from rating higher for providing amounts well above recommended intakes or even above upper limits. A similar approach was previously applied to identify top food sources of priority micronutrients for other population groups (32) and used to determine micronutrient-dense complementary foods (7, 13).

Foods were ranked according to the following thresholds on Average Requirements (ARs) for energy from complementary foods for children 6–23 months of age and hypothetical ARs for mass, assuming an energy density of 1.3 kcal/g [obtained by averaging the minimum and maximum composite energy densities of the sample complementary feeding diets for breastfed children developed by the World Health Organization (33)]:

- Very high: ≤ one-sixth of ARs for both energy and mass
- High: ≤ one-third of ARs for both energy and mass and < one-sixth of ARs for either energy or mass
- Moderate: ≤ one-third of ARs and > one-sixth of ARs for both energy and mass
- Low: > one-third of ARs for either energy or mass

The above thresholds were chosen by taking into consideration the assumed functional gastric capacity of children aged 6–23 months (30 g/kg body weight/d) and plausible amounts of complementary foods that they could consume at each meal, as well as meal frequency during the complementary feeding period (1, 34).

A slightly different approach was taken for milk (cow and goat), to account for the fact that mass is a less limiting factor for liquids than for solid foods in children 6–23 months of age. The same thresholds on energy and mass as for solid foods were applied, but instead of hypothetical ARs for mass, a maximum daily intake of milk of 400 g was assumed, which was considered plausible for children during the complementary feeding period based on a review of Food-Based Dietary Guidelines (FBDGs) from multiple countries (35, 36).

# RESULTS

## Recommended Nutrient Intakes for Children 6–23 Months of Age

The AR for energy from complementary foods for children aged 6–23 months is 450 kcal/d, accounting for average breast milk intake in developing countries (1). In addition to energy requirements, breast milk contributes to the achievement of RNIs for breastfed infants and young children during the complementary feeding period to different extents depending on the nutrient considered (**Table 1**). For instance, vitamin A requirements are largely covered by breast milk, with only 20% needed from complementary foods. Folate, vitamin $B_{12}$, and calcium requirements are also partially covered by breast milk intake, while for iron, the totality and for zinc, the near totality of RNIs need to be obtained through complementary foods. Moreover, recommended intakes for iron and zinc significantly vary depending on bioavailability: the lower the bioavailability level, the higher the RNIs, as larger amounts of iron and zinc are necessary to meet nutrient requirements.

**TABLE 1** | Recommended Nutrient Intakes (RNIs) for children 6–23 months of age.

| Nutrient | Total recommended nutrient intakes[c] | Proportion required from complementary foods[d] | Recommended nutrient intakes from complementary foods[e] |
|---|---|---|---|
| Vit A (μg RAE) | 367 | 0.2 | 73 |
| Folate (μg DFE) | 127 | 0.6 | 76 |
| Vit $B_{12}$ (μg) | 0.8 | 0.7 | 0.5 |
| Calcium (mg) | 467 | 0.7 | 327 |
| Iron[a] (mg) 20% | 3.5 | 1.0 | 3.5 |
| Iron[a] (mg) 15% | 4.7 | 1.0 | 4.7 |
| Iron[a] (mg) 10% | 7.0 | 1.0 | 7.0 |
| Zinc[b] (mg) R | 2.8 | 0.9 | 2.5 |
| Zinc[b] (mg) SR | 3.5 | 0.9 | 3.1 |
| Zinc[b] (mg) SU | 4.1 | 0.9 | 3.7 |
| Zinc[b] (mg) U | 4.7 | 0.9 | 4.2 |

*Recommended intakes for calcium, iron, and zinc from the World Health Organization and Food and Agriculture Organization of the United Nations (14). Recommended intakes for vitamin A, folate, and vitamin $B_{12}$ from the Institute of Medicine (15).*

*[a]Percentages (20, 15, and 10%) indicate the three levels of iron absorption considered in the analysis.*

*[b]Assuming 300 mg phytate/day and 44% absorption for refined (R) diets, 600 mg phytate/day and 35% absorption for semi-refined (SR) diets, 900 mg phytate/day and 30% absorption for semi-unrefined (SU) diets, and 1,200 mg phytate/day and 26% absorption for unrefined (U) diets.*

*[c]Total recommended intakes to be achieved through the combination of breast milk (or formula) and complementary foods.*

*[d]Assuming that the remaining proportion of recommended intakes would be provided by breast milk (or formula), based on Dewey (37).*

*[e]Recommended intakes to be achieved through complementary foods only, accounting for breast milk intake.*

*RAE, retinol activity equivalent; DFE, dietary folate equivalent; R, refined; SR, semi-refined; SU, semi-unrefined; U, unrefined; Vit, vitamin.*

## Regional Food Composition Database for South and Southeast Asia

A regional food composition database for South and Southeast Asia was built, including a total of 36 individual foods and aggregate food groups, with values for energy, the six priority micronutrients, calculated iron and zinc absorption levels, and phytate (**Table 2**). While most aggregate food groups, such as grains and their products (both whole and refined), "other fruits" and "other vegetables," present low nutrient density variance across included foods and across different countries' FCTs (**Supplementary Table 2**), some, such as dark green leafy vegetables (DGLVs) and fresh fish show greater nutrient density variance across included foods (**Supplementary Tables 2, 4**). For instance, spinach, amaranth leaves, and cassava leaves are more nutrient-dense than lettuce and cabbage. In the case of fresh fish, herring, carp, and mackerel (fatty fish) have higher nutrient values than sea bass and tilapia (lean fish). In addition, for a few foods and food groups, differences in the amounts of some nutrients were identified across the included FCTs. For example, values for folate in pulses and for vitamin A in beef liver are higher in USDA FDC than in South and Southeast Asian countries' FCTs, which may be due to different varieties, soil conditions, types of animal feed, culinary traditions, cooking methods and length, and/or the quality of sampling and analysis, including methods of mass spectrometry, conducted to develop the FCTs.

## Aggregate Priority Micronutrient Density Scores of Foods for Children 6–23 Months of Age

Portion sizes, expressed as calories and grams, required to achieve an average of one-third of RNIs from complementary foods of vitamin A, folate, vitamin $B_{12}$, calcium, iron, and zinc for children aged 6–23 months range from <5 g and kcal for liver (from different sources—beef, goat/lamb, chicken, and pork) to more than 600 g and kcal for refined grain products (**Figure 1**). Foods presenting very high aggregate priority micronutrient density (referred to as "top sources" hereafter) are the following: organs, including liver, spleen, kidney and heart from beef, goat/lamb, chicken, and pork; bivalves (clams, mussels, and oysters); crustaceans; fresh fish, including different species of marine and freshwater fish; goat; canned fish with bones; and eggs. Foods with a high aggregate micronutrient density include beef, lamb/mutton, DGLVs, cow milk, yogurt, and cheese, followed by canned fish without bones which was rated as moderate. All other foods analyzed presented low aggregate priority micronutrient density, including some animal-source foods (goat milk, pork, and chicken) and several plant-source foods.

**TABLE 2 |** Regional food composition database for South and Southeast Asia.

| Food (100 g) | kcal | Vit A (mcg RAE) | Folate (mcg DFE) | Vit $B_{12}$ (mcg) | Calcium (mg) | Iron (mg) | Zinc (mg) | Iron Abs | Zinc Abs | Phytate (mg) |
|---|---|---|---|---|---|---|---|---|---|---|
| Pulses | 140 | 1 | 97 | 0 | 27 | 2.7 | 1.3 | 0.10 | 0.26 | 441 |
| Whole grains[a] | 158 | 0 | 15 | 0 | 13 | 1.7 | 1.1 | 0.10 | 0.26 | 443 |
| Refined grains[a] | 131 | 0 | 5 | 0 | 9 | 0.4 | 0.6 | 0.10 | 0.44 | 42 |
| Whole grain products[a] | 194 | 0 | 84 | 0 | 28 | 1.2 | 1.2 | 0.10 | 0.35 | 77 |
| Refined grain products[a] | 155 | 0 | 6 | 0 | 9 | 0.7 | 0.5 | 0.10 | 0.44 | 49 |
| Millet | 132 | 0 | 23 | 0 | 10 | 2.0 | 1.1 | 0.10 | 0.26 | 200 |
| Nuts | 581 | 1 | 80 | 0 | 76 | 4.3 | 3.1 | 0.10 | 0.26 | 646 |
| Seeds | 561 | 1 | 97 | 0 | 134 | 6.9 | 5.6 | 0.10 | 0.26 | 402 |
| Starchy roots, tubers, and plantains | 101 | 16 | 11 | 0 | 22 | 0.7 | 0.3 | 0.10 | 0.44 | 12 |
| DGLVs | 27 | 286 | 37 | 0 | 102 | 2.0 | 0.3 | 0.10 | 0.44 | 16 |
| Vit A-rich fruits and veg, excl DGLVs | 38 | 126 | 21 | 0 | 20 | 0.4 | 0.2 | 0.10 | 0.44 | 24 |
| Other fruits, excl vit A-rich fruits | 62 | 4 | 17 | 0 | 11 | 0.3 | 0.1 | 0.10 | 0.44 | 10 |
| Other veg, excl DGLVs, and vit A-rich veg | 24 | 15 | 15 | 0 | 19 | 0.5 | 0.3 | 0.10 | 0.44 | 11 |
| Hen egg | 157 | 164 | 45 | 1.1 | 43 | 1.5 | 1.8 | 0.15 | 0.44 | 0 |
| Beef | 273 | 3 | 3 | 1.7 | 10 | 2.5 | 6.3 | 0.20 | 0.44 | 0 |
| Goat | 112 | 0 | 4 | 1.2 | 12 | 2.8 | 3.9 | 0.20 | 0.44 | 0 |
| Lamb/mutton | 210 | 2 | 10 | 2.5 | 13 | 1.7 | 4.1 | 0.20 | 0.44 | 0 |
| Pork | 234 | 1 | 3 | 0.6 | 21 | 1.5 | 2.1 | 0.15 | 0.44 | 0 |
| Chicken | 197 | 17 | 4 | 0.2 | 10 | 0.8 | 1.3 | 0.15 | 0.44 | 0 |
| Fresh cow milk | 67 | 52 | 5 | 0.4 | 120 | 0.1 | 0.4 | 0.15 | 0.44 | 0 |
| Cooked cow milk | 61 | 32 | 5 | 0.5 | 113 | 0.1 | 0.4 | 0.15 | 0.44 | 0 |
| Fresh goat milk | 69 | 45 | 1 | 0.1 | 143 | 0.1 | 0.3 | 0.15 | 0.44 | 0 |
| Yogurt | 61 | 27 | 7 | 0.4 | 121 | 0.1 | 0.6 | 0.15 | 0.44 | 0 |
| Cheese | 358 | 224 | 19 | 0.9 | 691 | 0.4 | 3.3 | 0.15 | 0.44 | 0 |
| Beef liver | 151 | 6,166 | 226 | 60.4 | 10 | 7.7 | 4.3 | 0.15 | 0.44 | 0 |
| Goat/lamb liver | 229 | 7,637 | 237 | 81.1 | 9 | 9.2 | 6.8 | 0.15 | 0.44 | 0 |
| Chicken liver | 147 | 3,492 | 509 | 16.1 | 12 | 9.4 | 3.4 | 0.15 | 0.44 | 0 |
| Pork liver | 139 | 4,924 | 149 | 16.9 | 10 | 15.9 | 5.7 | 0.15 | 0.44 | 0 |
| Heart[b] | 128 | 5 | 15 | 6.1 | 9 | 5.3 | 3.3 | 0.15 | 0.44 | 0 |
| Spleen[b] | 149 | 0 | 4 | 5.0 | 13 | 38.7 | 3.5 | 0.15 | 0.44 | 0 |
| Kidney[b] | 104 | 86 | 53 | 14.7 | 12 | 5.6 | 2.7 | 0.15 | 0.44 | 0 |
| Fresh fish[c] | 120 | 9 | 15 | 4.0 | 30 | 3.1 | 2.1 | 0.15 | 0.44 | 0 |
| Crustaceans | 92 | 3 | 15 | 1.2 | 87 | 1.2 | 3.2 | 0.15 | 0.44 | 0 |
| Bivalves | 106 | 68 | 7 | 17.6 | 68 | 4.0 | 1.9 | 0.15 | 0.44 | 0 |
| Canned fish, without bones | 136 | 20 | 4 | 2.6 | 17 | 1.4 | 0.7 | 0.15 | 0.44 | 0 |
| Canned fish, with bones | 182 | 43 | 10 | 3.9 | 240 | 2.5 | 1.2 | 0.15 | 0.44 | 0 |

*All included foods and aggregate food groups are listed in the forms typically consumed, which could be raw, cooked, or a mix of both, depending on the food/food group considered.*
*[a]The terms “whole grains” and “refined grains” refer to cereals, such as wheat, rice, and barley, while the terms “whole grain products” and “refined grain products” refer to products obtained from cereal flours, such as breads, pasta, noodles, and vermicelli.*
*[b]From different animals, including beef, lamb, pork, and chicken.*
*[c]Includes different species of marine and freshwater fish.*
*Vit, vitamin; RAE, retinol activity equivalent; DFE, dietary folate equivalent; Excl, excluding; Veg, vegetables; Abs, absorption.*

## Individual Priority Micronutrient Density Scores of Foods for Children 6–23 Months of Age

All analyzed foods scored low in at least one of the six priority micronutrients (**Figure 2**), exemplifying the importance of varied diets during the complementary feeding period, with a particular focus on the most nutrient-dense foods, to satisfy all nutrient requirements of infants and young children. For instance, liver and kidney were rated as very high in all included micronutrients except for calcium, for which they scored low. Most animal-source foods (besides chicken and canned fish without bones) and DGLVs scored very high or high in two or more micronutrients. Some plant-source foods presenting a low aggregate score were rated as very high or high in certain micronutrients. For example, vitamin A-rich fruits and vegetables scored very high in vitamin A; pulses and whole

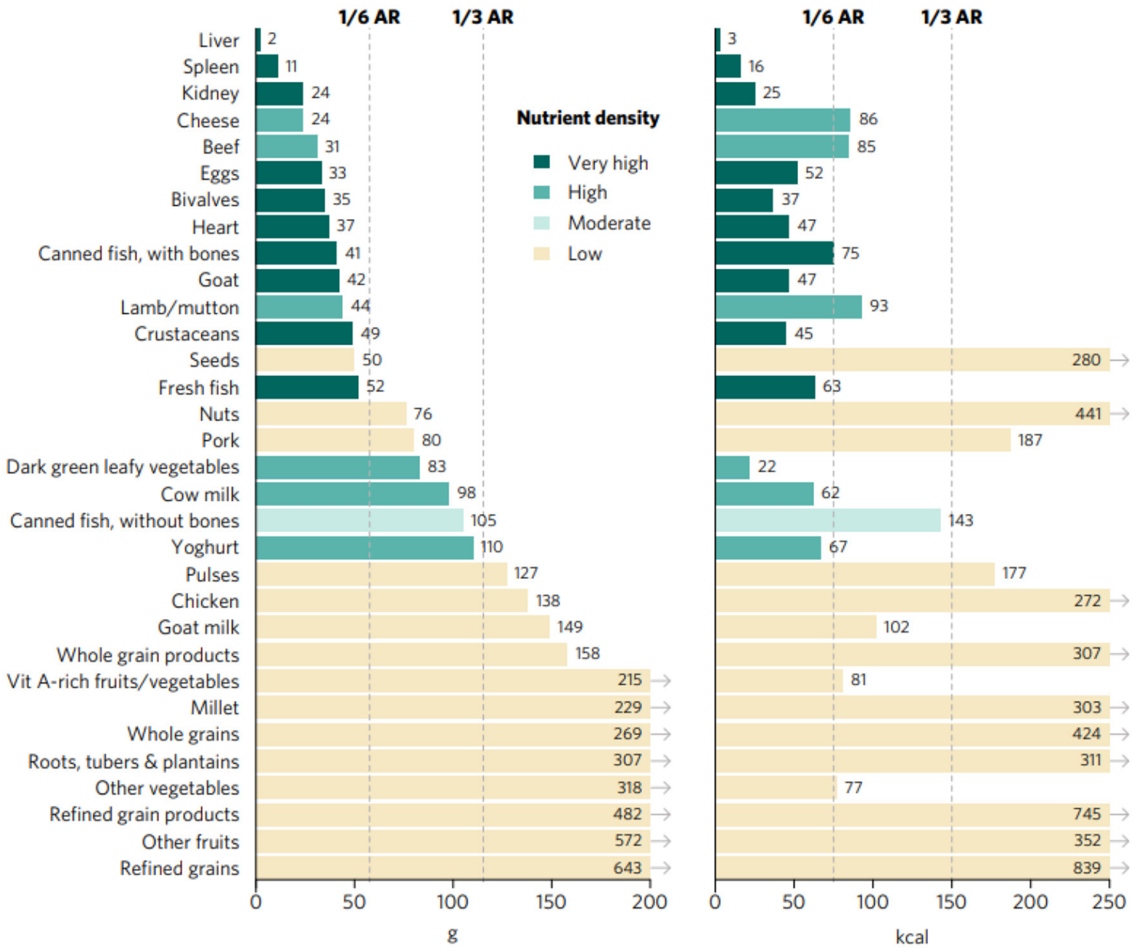


**FIGURE 1 |** Portion sizes (calories and grams) needed to provide an average of one-third of recommended intakes from complementary foods of vitamin A, folate, vitamin B12, calcium, iron, and zinc for children aged 6–23 months. Each micronutrient's contribution to the aggregate score was capped at 100% of recommended intakes. Hypothetical average requirement (ARs) for mass were obtained by assuming an energy density of 1.3 kcal/g.

grain products scored very high in folate; and seeds scored high in zinc and folate. Others, such as starchy roots, tubers and plantains, refined grains and their products, "other vegetables," and "other fruits" were not rated as very high or high in any of the priority micronutrients.

Top sources of iron included organs, fresh fish, bivalves, and goat, closely followed by beef and, to a lower extent, canned fish with bones, lamb/mutton, and pulses (**Figures 2**, **3**). Organs, bivalves, fresh fish, and goat are also top zinc sources, together with crustaceans, beef, lamb/mutton, and eggs, followed by cheese, pork, and seeds. In addition to liver, kidney, bivalves, eggs, goat milk, and cheese, top vitamin A sources included two plant-source foods groups: DGLVs and vitamin A-rich fruits and vegetables, followed closely by cow milk, yogurt, and canned fish with bones. Cheese was identified as the only top source of calcium, closely followed by DGLVs, canned fish with bones, cow milk, goat milk, and yogurt. Top folate sources included liver, kidney, pulses, and whole grain products, followed by eggs, DGLVs, and seeds. Finally, all animal-source foods scored very high in vitamin $B_{12}$, except for chicken and goat milk.

## DISCUSSION

The purpose of this study was to identify top inherent food sources of multiple and individual micronutrients commonly lacking in complementary feeding diets of children aged 6–23 months in South and Southeast Asia. All analyzed animal-source foods, except for canned fish without bones, goat milk, pork and chicken, presented very high (organs, bivalves, crustaceans, goat, eggs, fresh fish, and canned fish with bones) or high (beef, lamb/mutton, cow milk, yogurt, and cheese) aggregate priority micronutrient density; while, DGLVs were the only plant-source food to score high for the aggregate micronutrient density rating. Our findings are in alignment with UNICEF's and WHO's (1, 3, 34, 38) recommendations for complementary feeding, according to which infants and young children should be fed animal-source

| Food | 2+ micronutrients | Iron | Zinc | Vitamin A | Calcium | Folate | Vitamin B12 |
|---|---|---|---|---|---|---|---|
| Liver | Very high | Very high | Very high | Very high | Low | Very high | Very high |
| Spleen | Very high | Very high | Very high | Low | Low | Low | Very high |
| Kidney | Very high | Very high | Very high | Very high | Low | Very high | Very high |
| Bivalves | Very high | Very high | Very high | Very high | Low | Low | Very high |
| Crustaceans | Very high | Low | Very high | Low | Low | Low | Very high |
| Goat | Very high | Very high | Very high | Low | Low | Low | Very high |
| Heart | Very high | Very high | Very high | Low | Low | Low | Very high |
| Hen egg | Very high | Low | Very high | Very high | Low | High | Very high |
| Fresh fish | Very high | Very high | Very high | Low | Low | Low | Very high |
| Canned fish w/bones | Very high | Moderate | Moderate | High | High | Low | Very high |
| DGLVs | High | Low | Low | Very high | High | High | Low |
| Cow milk | High | Low | Low | High | High | Low | Very high |
| Yoghurt | High | Low | Low | High | High | Low | Very high |
| Beef | High | High | Very high | Low | Low | Low | Very high |
| Cheese | High | Low | High | Very high | Very high | Low | Very high |
| Lamb/mutton | High | Moderate | Very high | Low | Low | Low | Very high |
| Canned fish w/o bones | Moderate | Low | Low | Low | Low | Low | Very high |
| Other vegetables | Low | Low | Low | Low | Low | Low | Low |
| Vit A-rich fruits & vegetables | Low | Low | Low | Very high | Low | Low | Low |
| Fresh goat milk | Low | Low | Low | Very high | High | Low | Low |
| Pulses | Low | Moderate | Low | Low | Low | Very high | Low |
| Pork | Low | Low | High | Low | Low | Low | Very high |
| Chicken | Low | Low | Moderate | Low | Low | Low | Moderate |
| Seeds | Low | Low | High | Low | Low | High | Low |
| Millet | Low | Low | Low | Low | Low | Moderate | Low |
| Whole grain products | Low | Low | Low | Low | Low | Very high | Low |
| Roots, tubers & plantains | Low | Low | Low | Low | Low | Low | Low |
| Other fruits | Low | Low | Low | Low | Low | Low | Low |
| Whole grains | Low | Low | Low | Low | Low | Low | Low |
| Nuts | Low | Low | Low | Low | Low | Low | Low |
| Refined grain products | Low | Low | Low | Low | Low | Low | Low |
| Refined grains | Low | Low | Low | Low | Low | Low | Low |

**FIGURE 2 |** Aggregate and individual priority micronutrient density scores for children aged 6–23 months.

foods daily or as often as possible, given that primarily plant-based or vegan diets, even when carefully planned, "*cannot meet nutrient needs at this age unless nutrient supplements or fortified products are used*" (1). Consumption of ASFs remains very low among children aged 6–23 months in South and Southeast Asia (6, 10, 39) and has been found to be strongly associated with infant and young child growth and development (11, 40). Results from the quantitative analysis conducted by UNICEF in its latest Child Nutrition Report (10) show that only 24% of children 6–23 months of age in South Asia consume eggs, fish, and/or meat. In alignment with UNICEF's findings, other recent evidence on South and Southeast Asia highlights that the vast majority of children aged 6–23 months are traditionally fed cereal- and pulse-based complementary foods for the most part (e.g., rice flour and rice porridge), while very few receive meat-, fish-, and/or egg-based complementary foods (2, 6, 9).

Unavailability and unaffordability, as well as insufficient maternal education, were identified as some of the main factors limiting consumption of animal-source foods and nutrient-dense foods in general among infants and young children in the two regions (2, 10, 41, 42). Indeed, while some foods of animal origin, such as liver, eggs and dairy, may be more available and affordable than others (particularly in terms of cost per nutrients provided), efforts to increase access to and knowledge around ASFs in South and Southeast Asia are still needed (43). Although this study does not aim (or claim) to provide solutions to the complex, multifaceted issues of availability, affordability, and desirability of foods, and of complementary feeding practices, it could make a positive contribution by helping policy-makers and program managers to identify the top food sources to prioritize when addressing micronutrient malnutrition in children aged 6–23 months. Policy-makers and program managers could use this knowledge to design and implement policies, measures, and programs toward increasing the availability and affordability of particularly nutrient-dense foods and improving complementary feeding practices among mothers and caregivers. Just to provide a few examples, among others they could (i) incentivize in-country/local production of these foods (e.g., through agricultural subsidies) to decrease the country's reliance on imports; (ii) increase investments in infrastructure services to reduce trade and transportation margins; (iii) provide targeted cash transfers to low-income households with children aged 6–23 months for the purchase of specific nutrient-dense foods; (iv) establish dedicated educational channels within public health systems and set up Social and Behavior Change Communication campaigns (e.g., through mass

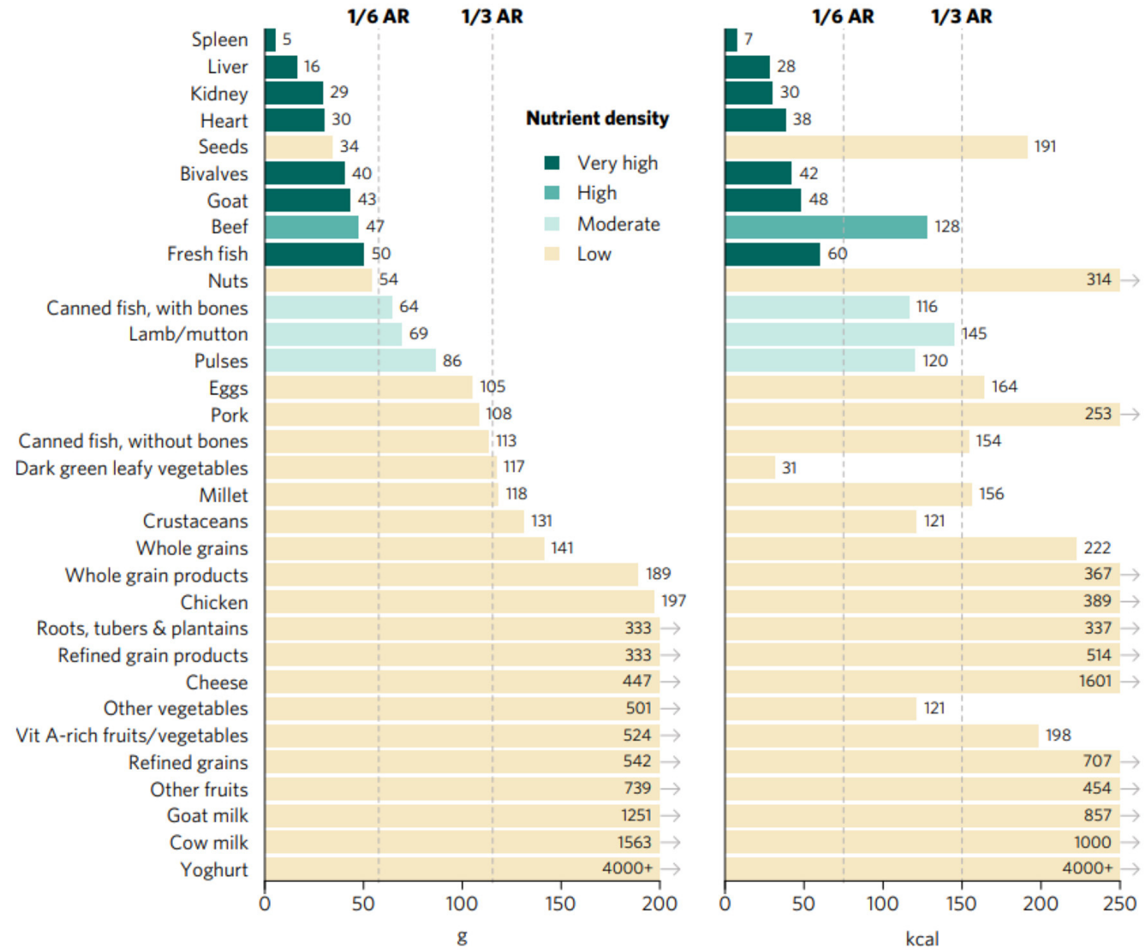


**FIGURE 3 |** Portions sizes (calories and grams) needed to provide one-third of recommended iron intakes from complementary foods for children aged 6–23 months. Hypothetical average requirement (ARs) for mass were obtained by assuming an energy density of 1.3 kcal/g.

media and community mobilization) targeting mothers and caregivers of infants and young children (44).

Despite DGLVs being the only plant-source food with high aggregate priority micronutrient density, other PSFs scored very high or high in individual micronutrients. These PSFs could make an important contribution to the reduction of specific micronutrient gaps in complementary feeding diets, especially when appropriate low-cost, easy-to-implement processing techniques are adopted at the household level to increase absorption of non-heme iron and zinc, by reducing the negative effects of phytate on absorption [e.g., consumption of vitamin C, consumption of animal protein, heating, germination, soaking, and fermentation; (14)]. For instance, vitamin A deficiency in early childhood is widespread in South and Southeast Asia, and can have severe health consequences, which has led to countries' governments carrying out vitamin A supplementation campaigns targeted at infants and young children (6, 8). Consumption of micronutrient-dense fruits and vegetables, particularly DGLVs and vitamin A-rich fruits and vegetables (both scoring very high in vitamin A in our analysis) is very limited during the complementary feeding period in the two regions (2, 6). For instance, in South Asia only one in three (∼33%) children aged 6–23 months receive these foods (2). Therefore, in addition to supplementation and fortification—both key strategies in the fight against micronutrient malnutrition—improving consumption of available, affordable vitamin A food sources in infants and young children would significantly contribute to the reduction of vitamin A deficiency in South and Southeast Asia (6, 43). Other examples of widely available, affordable plant-source foods with very high or high individual micronutrient density include: pulses and whole grain products for folate, and seeds for zinc and folate (in addition to DGLVs) (6, 38, 43).

Among the foods which scored low in all six included micronutrients, some, such as nuts, "other fruits" and "other vegetables" (non-vitamin A-rich and non-DGLVs), are often promoted as nutrient-dense in nutrition policies and programs to improve infants' and young children's diets in South and Southeast Asia and globally (6, 38). However, given the limited gastric capacity of infants and young children (1, 2), these foods could displace more nutrient-dense foods and prevent complementary fed children from obtaining adequate micronutrients necessary for proper growth and development. Though not particularly dense in priority micronutrients, moderate quantities of these foods, which provide energy and other essential nutrients, as well as non-essential beneficial compounds, can contribute to the overall quality and diversity of complementary feeding diets in South and Southeast Asia (2, 7, 38, 39).

This study has several strengths. First, while the importance of nutrient-dense foods, particularly ASFs, for infants and young children has already been extensively acknowledged in previous studies and guidelines (1, 3, 10, 34, 38, 45), this analysis brings added value to the literature by transparently ranking a diverse set of inherent food sources of two or more micronutrients commonly lacking during the complementary feeding period in South and Southeast Asia, and whose deprivation is the cause of significant public health burdens in these regions (6–9). It does so by developing a resource-inexpensive, reproducible approach, which is widely applicable in the two regions considered and could be easily adapted for use in other geographic areas of the world. Second, the approach used for rating foods takes into consideration infants' and young children's limited gastric capacity and plausible amounts of food they could consume at each meal and daily, as well as the adequate meal frequency during the complementary feeding period (1, 34). Third, the micronutrient density analysis was conducted based on an aggregate regional food composition database, which compiles data from multiple countries' FCTs and is reflective of the nutritional value of foods in Southern and Southeastern Asia, unlike other nutrient scoring systems which typically analyze data from a single national FCT, usually USDA FDC (46, 47). Fourth, iron and zinc values were adjusted for bioavailability in different foods. Finally, the methodology adopted is fully transparent, and the analysis is based on publicly available data, as has been recommended for nutrient profiling systems (46, 47). In addition, the results are presented both in written text and figures in a form that is easily interpretable by non-technical audiences, including policy makers and program managers.

The specific focus of this study on priority micronutrients is both a strength and a limitation. Indeed, in addition to the six micronutrients included in the analysis, foods provide energy, protein, essential amino acids, and fatty acids, as well as other essential vitamins and minerals, which are crucial to optimal growth and development in the first 2 years of a child's life and which can also be lacking to some extent in complementary feeding diets (1, 4, 9, 11, 22, 38, 48, 49). Moreover, minimally processed foods of both plant and animal origin contain countless non-essential compounds, including fiber, phytonutrients, and bioactive compounds, with potential beneficial effects on human health (50–53). However, this study focuses on micronutrients that are known to be commonly lacking in the two regions of interest and globally among infants and young children and are hindering optimal growth and development (7–9).

Other limitations should be acknowledged. First, the choice of countries' FCTs used and foods included in the regional food composition database was constrained by limitations in data quality and availability. Indeed, only one country FCT from South Asia (Bangladesh) was considered appropriate for inclusion, while all other selected FCTs are from Southeast Asia. Moreover, soy foods, as well as wild or indigenous fruits, vegetables, nuts, seeds, grains and pulses, many of which are more nutrient-dense than commercial varieties (54), have not been explored. Second, the adopted approach accounts for bioavailability of iron and zinc based on the heme-iron and phytate contents of foods, respectively, which are only two out of numerous factors influencing absorption, particularly the differing micronutrient status, overall diet and genetics of individuals. Third, there can be significant variations in nutrient values of a given food across countries' FCTs, which may be due to different varieties, soil characteristics, climate conditions, quantity and type of fertilizers used, animal feed, production, and processing methods, including local culinary traditions, as well as the quality of sampling, analysis, and reporting processes. However, this study attempts to mitigate such differences and uncertainties by building a regional food composition database with aggregate nutrient values from multiple national FCTs, including the robust USDA FDC. Fourth, as mentioned in the *Results* section, some of the analyzed food groups show high nutrient density variance across included foods, meaning that the overall score of an aggregate food group might not reflect the micronutrient density of all individual foods included. In this regard, certain foods (e.g., fruits and vegetables) may be more likely to be targeted in policies and programming as food groups rather than individually, and consequently they were aggregated despite presenting significant intra-group nutrient density variance.

In conclusion, results from this study clearly show that the introduction of small quantities of priority micronutrient-dense animal—(e.g., organs, fish and shellfish, eggs) and plant—(DGLVs) source foods would significantly contribute to achieving adequacy of micronutrients commonly lacking in complementary feeding diets in South and Southeast Asia. Noticeably, top sources of priority micronutrients should be consumed together with a variety of other nutrient-dense foods, as part of a diverse and balanced diet, able to meet all nutrient requirements of children aged 6–23 months. Our findings could be used to improve current countries' and regional recommendations on complementary feeding, by providing additional insights compared to just common knowledge on the high nutrient-density of ASFs and DGLV. Indeed, this study highlights the nutrient density of specific ASF (e.g., organ meats other than liver—spleen, kidney, heart—, bivalves, canned fish with bones) which are often not included in existing recommendations and whose potential remains largely unexplored. Also, our results show the differences in nutrient density among various ASFs and their relative ranking, enabling policy makers and program managers to prioritize certain ASFs over others for children 6–23 months. For instance, pork and chicken have lower priority micronutrient density compared to organs, ruminant meat, bivalves, eggs, cow milk, and others; therefore, the latter are more ideal types of ASFs to promote for feeding infants and young children on a regular basis. Further analyses are needed to explore ways to integrate these findings into food, agriculture, and nutrition policies and programs aiming to reduce micronutrient malnutrition in the first 2 years of life through the promotion of inherently nutrient-dense foods.

This study focuses specifically on infants and young children living in South and Southeast Asia, however the same approach could be used to analyze the priority micronutrient density of foods for complementary feeding in other regions of the world presenting similar micronutrient gaps in complementary feeding diets, such as Eastern and Southern Africa (55). Future research could build on this analysis, for instance, by expanding the regional food composition database through the inclusion of soy foods and of nutrient-dense wild or indigenous fruits, vegetables, nuts, seeds, grains, pulses (54), and insects (56), the safety and nutritional adequacy of which is currently being studied for potential application in complementary foods (57). In addition, findings from this study could be compared and complemented with affordability and environmental impact metrics, to assess these variables based on priority micronutrient density by expanding on existing approaches (13, 43, 58).

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

FO and TB designed the study and conducted the analyses. Both authors contributed to the article and approved the submitted version.

## FUNDING

This work was funded by contributions from the Ministry of the Foreign Affairs of the Netherlands (grant #4000000622 to GAIN) and the Bill & Melinda Gates Foundation through the Regional Initiatives for Sustained Improvements in Nutrition and Growth (grant INV-008600 to UNICEF). The funder had no role in data collection and analysis, manuscript preparation and revision, or the decision to publish. This study used data from public sources, and all authors had access to the data analyzed as part of this study.

## ACKNOWLEDGMENTS

The authors wish to thank Zivai Murira for reviewing and providing feedback on a draft version of this manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.785227/full#supplementary-material

## REFERENCES

1. Pan American Health Organization. *Guiding Principles for Complementary Feeding of the Breastfed Child*. (2003). Available online at: https://www.who.int/nutrition/publications/guiding_principles_compfeeding_breastfed.pdf (accessed August 16, 2021).
2. Aguayo VM. Complementary feeding practices for infants and young children in South Asia. A review of evidence for action post-2015. *Maternal Child Nutr.* (2017) 13:e12439. doi: 10.1111/mcn.12439
3. WHO Programme of Nutrition. *Complementary Feeding of Young Children in Developing Countries : A Review of Current Scientific Knowledge*. World Health Organization. Report No. WHO/NUT/98.1 (1998). Available online at: https://apps.who.int/iris/handle/10665/65932 (accessed August 16, 2021).
4. Food and Agriculture Organization of the United Nations (FAO). *The State of Food Security and Nutrition in the World (SOFI)*. Rome (2021). Available online at: http://www.fao.org/documents/card/en/c/cb4474en (accessed August 30, 2021).
5. UNICEF, WHO, World Bank. *Joint Child Malnutrition Estimates*. (2021). Available online at: https://data.unicef.org/topic/nutrition/malnutrition/ (accessed August 16, 2021).
6. World Health Organization. Regional Office for South-East Asia. *Regional Nutrition Strategy: Addressing Malnutrition and Micronutrient Deficiencies (2011-2015)*. WHO Regional Office for South-East Asia (2012). Available online at: https://apps.who.int/iris/handle/10665/205804 (accessed August 22, 2021).
7. Beal T, White JM, Arsenault JE, Okronipa H, Hinnouho G-M, Torlesse H, et al. Micronutrient gaps during the complementary feeding period in South Asia: a comprehensive nutrient gap assessment. *Nutr Rev.* (2021) 79(Suppl. 1):26–34. doi: 10.1093/nutrit/nuaa144
8. Anderson VP, Jack S, Monchy D, Hem N, Hok P, Bailey KB, et al. Co-existing micronutrient deficiencies among stunted Cambodian infants and toddlers. *Asia Pac J Clin Nutr.* (2008) 17:72–9. Available online at: http://211.76.170.15/server/APJCN/17/1/72.pdf
9. Laillou A, Pham TV, Tran NT, Le HT, Wieringa F, Rohner F, et al. Micronutrient deficits are still public health issues among women and young children in Vietnam. *PLoS ONE.* (2012) 7:e34906. doi: 10.1371/journal.pone.0034906
10. United Nations Children's Fund (UNICEF). *Fed to Fail? The Crisis of Children's Diets in Early Life.* 2021 Child Nutrition Report. New York, NY: UNICEF (2021). Available online at: https://www.unicef.org/reports/fed-to-fail-child-nutrition (accessed August 30, 2021).
11. Torlesse H, Raju D. *Feeding of Infants and Young Children in South Asia.* Policy Research Working Paper; No. 8655. Washington, D.C.: World Bank (2018). Available online at: https://socialprotection.org/discover/publications/feeding-infants-and-young-children-south-asia (accessed August 30, 2021) doi: 10.1596/1813-9450-8655
12. Kim R, Mejía-Guevara I, Corsi DJ, Aguayo VM, Subramanian SV. Relative importance of 13 correlates of child stunting in South Asia: insights from nationally representative data from Afghanistan, Bangladesh, India, Nepal, and Pakistan. *Soc. Sci. Med.* (2017) 187:144–54. doi: 10.1016/j.socscimed.2017.06.017
13. Ryckman T, Beal T, Nordhagen S, Chimanya K, Matji J. Affordability of nutritious foods for complementary feeding in Eastern and Southern Africa. *Nutr Rev.* (2021) 79:35–51. doi: 10.1093/nutrit/nuaa137
14. World Health Organization, Food and Agriculture Organization of the United Nations. *Vitamin and Mineral Requirements in Human Nutrition*. 2nd ed. (2005). p. 341. Available online at: https://apps.who.int/iris/handle/10665/42716 (accessed August 17, 2021).
15. Institute of Medicine Committee to Review Dietary Reference Intakes for Vitamin D and Calcium. In: Ross AC, Taylor CL, Yaktine AL, Del Valle HB, editors. *Dietary Reference Intakes for Calcium Vitamin D*. Washington,

DC: National Academies Press (The National Academies Collection: Reports funded by National Institutes of Health) (2011).
16. European Food Safety Authority (EFSA). Dietary reference values for nutrients summary report. *EFSA Support Publ.* (2017). 14:e15121E. doi: 10.2903/sp.efsa.2017.e15121
17. Allen LH, Carriquiry AL, Murphy SP. Perspective: proposed harmonized nutrient reference values for populations. *Adv Nutr.* (2020) 11:469–83. doi: 10.1093/advances/nmz096
18. FAO/IZiNCG. *FAO/INFOODS/IZiNCG Global Food Composition Database for Phytate Version 1.0 - PhyFoodComp 1.0*. Rome (2018). Available online at: http://www.fao.org/3/i8542en/I8542EN.pdf (accessed August 30, 2021).
19. U.S. Department of Agriculture (USDA), Agricultural Research Service. *USDA FoodData Central*. Available online at: https://fdc.nal.usda.gov/ (accessed June 18, 2021).
20. FAO. *INFOODS: FAO/INFOODS Databases*. Available online at: http://www.fao.org/infoods/infoods/tables-and-databases/faoinfoods-databases/en/ (accessed June 18, 2021).
21. Bongar A. *Tables on Weight Yield of Food and Retention Factors of Food Constituents for the Calculation of Nutrient Composition of Cooked Foods (Dishes)*. Karlsruhe: Bundesforschungsanstalt für Ernährung (2002). Available online at: http://www.fao.org/uploads/media/bognar_bfe-r-02-03.pdf (accessed August 30, 2021).
22. Beal T, Massiot E, Arsenault JE, Smith MR, Hijmans RJ. Global trends in dietary micronutrient supplies and estimated prevalence of inadequate intakes. *PLoS ONE.* (2017) 12:e0175554. doi: 10.1371/journal.pone.0175554
23. Valenzuela C, López de Romaña D, Olivares M, Morales MS, Pizarro F. Total iron and heme iron content and their distribution in beef meat and viscera. *Biol Trace Elem Res.* (2009) 132:103–11. doi: 10.1007/s12011-009-8400-3
24. Balder HF, Vogel J, Jansen MCJF, Weijenberg MP, van den Brandt PA, Westenbrink S, et al. Heme and chlorophyll intake and risk of colorectal cancer in the Netherlands cohort study. *Cancer Epidemiol Biomarkers Prev.* (2006) 15:717–25. doi: 10.1158/1055-9965.EPI-05-0772
25. Lombardi-Boccia G, Martinez-Dominguez B, Aguzzi A. Total heme and non-heme iron in raw and cooked meats. *J Food Sci.* (2002) 67:1738–41 doi: 10.1111/j.1365-2621.2002.tb08715.x
26. Pourkhalili A, Mirlohi M, Rahimi E. Heme iron content in lamb meat is differentially altered upon boiling, grilling, or frying as assessed by four distinct analytical methods. *Sci World J.* (2013) 2013:e374030. doi: 10.1155/2013/374030
27. Kabat GC, Miller AB, Jain M, Rohan TE. A cohort study of dietary iron and heme iron intake and risk of colorectal cancer in women. *Br J Cancer.* (2007) 97:118–22. doi: 10.1038/sj.bjc.6603837
28. Ronco A, Espinosa E, Calderon J. *A Case-Control Study on Heme/Non-Heme Iron and Breast Cancer Risk 1 MedDocs Publishers of Creative Commons Attribution 4.0 International License Annals of Clinical Nutrition*, MedDocs Publishers LLC (2018).
29. Kongkachuichai R, Napatthalung P, Charoensiri R. Heme and nonheme iron content of animal products commonly consumed in Thailand. *J Food Compos Anal.* (2002) 15:389–98. doi: 10.1006/jfca.2002.1080
30. Taniguchi CN, Dobbs J, Dunn MA. Heme iron, non-heme iron, and mineral content of blood clams (*Anadara* spp.) compared to Manila clams (*V. philippinarum*), Pacific oysters (*C. gigas*), and beef liver (*B. taurus*). *J Food Compos Anal.* (2017) 57:49–55. doi: 10.1016/j.jfca.2016.12.018
31. Hall KD, Guo J, Courville AB, Boring J, Brychta R, Chen KY, et al. Effect of a plant-based, low-fat diet versus an animal-based, ketogenic diet on ad libitum energy intake. *Nat Med.* (2021) 27:344–53. doi: 10.1038/s41591-020-01209-1
32. Beal T, Ortenzi F. *Priority Micronutrient Density in Foods*. (2021). Available online at: https://www.researchsquare.com/article/rs-701840/v1 (accessed August 17, 2021). doi: 10.21203/rs.3.rs-701840/v2
33. World Health Organization. *Complementary Feeding : Family Foods for Breastfed Children*. World Health Organization (2000). p. 52. Available online at: https://apps.who.int/iris/handle/10665/66389 (accessed August 17, 2021).
34. Dewey KG, Brown KH. Update on technical issues concerning complementary feeding of young children in developing countries and implications for intervention programs. *Food Nutr Bull.* (2003) 24:5–28. doi: 10.1177/156482650302400102
35. United Nations Children's Fund (UNICEF). *Food-Based Dietary Guidelines. A Review of National Guidance for Children, Adolescents, and Women*. New York, NY: UNICEF (2020). Available online at: https://www.unicef.org/media/102761/file/2021-Food-based-Dietary-Guidelines-final.pdf (accessed August 30, 2021).
36. Grammatikaki E, Wollgast J, Caldeira S. *Feeding Infants and Young Children. A Compilation of National Food-Based Dietary Guidelines and Specific Products Available in the EU Market*, Ispra, VA: European Commission, Joint Research Centre (2019).
37. Dewey KG. Nutrition, growth, and complementary feeding of the breastfed infant. *Pediatr Clin North Am.* (2001) 48:87–104. doi: 10.1016/S0031-3955(05)70287-X
38. UNICEF. *Improving Young Children's Diets During the Complementary Feeding Period – UNICEF Programming Guidance*. (2020). Available online at: https://www.unicef.org/documents/improving-young-childrens-diets-during-complementary-feeding-period-unicef-programming (accessed August 22, 2021).
39. UNICEF. *From the First Hour of Life: Making the Case for Improved Infant and Young Child Feeding Everywhere*. New York, NY (2016). Available online at: https://data.unicef.org/resources/first-hour-life-new-report-breastfeeding-practices/ (accessed August 22, 2021).
40. *Animal Sourced Foods and Child Stunting: Evidence from 112,887 Children in 46 Countries*. Washington, DC: International Food Policy Research Institute (IFPRI) (2017). Available online at: https://www.ifpri.org/publication/animal-sourced-foods-and-child-stunting (accessed August 30, 2021).
41. Moench-Pfanner R, Bloem MW. ASEAN: insights and considerations toward nutrition programs. *Food Nutr Bull.* (2013) 34(Suppl. 2):S4–7. doi: 10.1177/15648265130342S102
42. Ariff S, Saddiq K, Khalid J, Sikanderali L, Tariq B, Shaheen F, et al. Determinants of infant and young complementary feeding practices among children 6–23 months of age in urban Pakistan: a multicenter longitudinal study. *BMC Nutrition.* (2020) 6:75. doi: 10.1186/s40795-020-00401-3
43. Ryckman T, Beal T, Nordhagen S, Murira Z, Torlesse H. Affordability of nutritious foods for complementary feeding in South Asia. *Nutr Rev.* (2021) 79(Suppl. 1):52–68. doi: 10.1093/nutrit/nuaa139
44. Rawat R, Nguyen PH, Tran LM, Hajeebhoy N, Nguyen HV, Baker J, et al. Social franchising and a nationwide mass media campaign increased the prevalence of adequate complementary feeding in Vietnam: a cluster-randomized program evaluation. *J Nutr.* (2017) 147:670–9. doi: 10.3945/jn.116.243907
45. Osendarp SJM, Broersen B, van Liere MJ, De-Regil LM, Bahirathan L, Klassen E, et al. Complementary feeding diets made of local foods can be optimized, but additional interventions will be needed to meet iron and zinc requirements in 6- to 23-month-old children in low- and middle-income countries. *Food Nutr Bull.* (2016) 37:544–70. doi: 10.1177/0379572116655239
46. Miller GD, Drewnowski A, Fulgoni V, Heaney RP, King J, Kennedy E. It is time for a positive approach to dietary guidance using nutrient density as a basic principle. *J Nutr.* (2009) 139:1198–202. doi: 10.3945/jn.108.100842
47. Drewnowski A, Amanquah D, Gavin-Smith B. Perspective: how to develop nutrient profiling models intended for global use: a manual. *Adv Nutr.* (2021) 12:609–20. doi: 10.1093/advances/nmab018
48. Simopoulos AP. Essential fatty acids in health and chronic disease. *Am J Clin Nutr.* (1999) 70:560–9s. doi: 10.1093/ajcn/70.3.560s
49. Semba RD, Shardell M, Sakr Ashour FA, Moaddel R, Trehan I, Maleta KM, et al. Child stunting is associated with low circulating essential amino acids. *EBioMedicine.* (2016) 6:246–52. doi: 10.1016/j.ebiom.2016.02.030
50. van Vliet S, Kronberg SL, Provenza FD. Plant-based meats, human health, and climate change. *Front Sustain Food Syst.* (2020) 4:128. doi: 10.3389/fsufs.2020.00128
51. Barabási A-L, Menichetti G, Loscalzo J. The unmapped chemical complexity of our diet. *Nat Food.* (2020) 1:33–7. doi: 10.1038/s43016-019-0005-1
52. Jacobs DR, Tapsell LC. Food, not nutrients, is the fundamental unit in nutrition. *Nutr Rev.* (2007) 65:439–50. doi: 10.1111/j.1753-4887.2007.tb00269.x
53. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet.* (2019) 398:1958–72. doi: 10.1016/S0140-6736(19)30041-8
54. Nyirenda D, Musukwa M, Mugode R, Shindano J. *Zambia Food Composition Tables*. 4th ed. Lusaka, Zambia: National Food and Nutrition Commission (2009). Available online at: https://www.nfnc.org.zm/download/zambia-food-composition-tables-4th-edition/ (accessed June 29, 2021).

55. White JM, Beal T, Arsenault JE, Okronipa H, Hinnouho G-M, Chimanya K, et al. Micronutrient gaps during the complementary feeding period in 6 countries in Eastern and Southern Africa: a comprehensive nutrient gap assessment. *Nutr Rev*. (2021) 79:16–25. doi: 10.1093/nutrit/nuaa142
56. Smith MR, Stull VJ, Patz JA, Myers SS. Nutritional and environmental benefits of increasing insect consumption in Africa and Asia. *Environ Res Lett*. (2021) 16:065001. doi: 10.1088/1748-9326/abf06c
57. Mekuria SA, Kinyuru JN, Mokua BK, Tenagashaw MW. Nutritional quality and safety of complementary foods developed from blends of staple grains and honey bee larvae (*Apis mellifera*). *Int J Food Sci*. (2021) 2021e5581585. doi: 10.1155/2021/5581585
58. Poore J, Nemecek T. Reducing food's environmental impacts through producers and consumers. *Science*. (2018) 360:987–92. doi: 10.1126/science.aaq0216

**Author Disclaimer:** The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the funder.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

REVIEW
published: 04 January 2022
doi: 10.3389/fnut.2021.798576

Check for updates

# Branded Foods Databases as a Tool to Support Nutrition Research and Monitoring of the Food Supply: Insights From the Slovenian Composition and Labeling Information System

*Igor Pravst*[1,2,3]*, *Maša Hribar*[1,2], *Katja Žmitek*[1,3], *Bojan Blažica*[4], *Barbara Koroušić Seljak*[4] and *Anita Kušar*[1,3]

[1] *Nutrition Institute, Nutrition and Public Health Research Group, Ljubljana, Slovenia,* [2] *Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia,* [3] *VIST–Faculty of Applied Sciences, Ljubljana, Slovenia,* [4] *Computer Systems Department, Jozef Stefan Institute, Ljubljana, Slovenia*

OPEN ACCESS

**Edited by:**
*Alessandra Durazzo, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Ranga Chimhundu, University of Southern Queensland, Australia*
*Annet J. C. Roodenburg, HAS University of Applied Sciences, Netherlands*

***Correspondence:**
*Igor Pravst*
*igor.pravst@nutris.org*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** *20 October 2021*
**Accepted:** *17 November 2021*
**Published:** *04 January 2022*

**Citation:**
*Pravst I, Hribar M, Žmitek K, Blažica B, Koroušić Seljak B and Kušar A (2022) Branded Foods Databases as a Tool to Support Nutrition Research and Monitoring of the Food Supply: Insights From the Slovenian Composition and Labeling Information System. Front. Nutr. 8:798576. doi: 10.3389/fnut.2021.798576*

Branded foods databases are becoming very valuable not only in nutrition research but also for clinical practice, policymakers, businesses, and general population. In contrast to generic foods, branded foods are marked by rapid changes in the food supply because of reformulations, the introduction of new foods, and the removal of existing ones from the market. Also, different branded foods are available in different countries. This not only complicates the compilation of branded foods datasets but also causes such datasets to become out of date quickly. In this review, we present different approaches to the compilation of branded foods datasets, describe the history and progress of building and updating such datasets in Slovenia, and present data to support nutrition research and monitoring of the food supply. Manufacturers are key sources of information for the compilation of branded foods databases, most commonly through food labels. In Slovenia, the branded food dataset is compiled using standard food monitoring studies conducted at all major retailers. Cross-sectional studies are conducted every few years, in which the food labels of all available branded foods are photographed. Studies are conducted using the Composition and Labeling Information System (CLAS) infrastructure, composed of a smartphone application for data collection and online data extraction and management tool. We reviewed various uses of branded foods datasets. Datasets can be used to assess the nutritional composition of food in the food supply (i.e., salt, sugar content), the use of specific ingredients, for example, food additives, for nutrient profiling, and assessment of marketing techniques on food labels. Such datasets are also valuable for other studies, for example, assessing nutrient intakes in dietary surveys. Additional approaches are also being tested to keep datasets updated between food monitoring studies. A promising approach is the exploitation of crowdsourcing through the mobile application VešKajJeš, which was launched in Slovenia to support consumers in making healthier dietary choices.

**Keywords: food composition database, labeling, pre-packed, nutrition declaration, market, Slovenia, CLAS**

# INTRODUCTION

People's diets are composed of a wide variety of foods and drinks (further referred to as "foods)". In general, we can distinguish between *generic* and *branded foods*. For example, orange juice could be considered a generic food, while on the market, there are a wide variety of branded orange juices, which can have notable differences in taste and nutritional composition. Such products are typically pre-packaged and labeled, and such food labeling information can be a useful resource for the compilation of branded foods composition databases.

It should be mentioned that the composition of foods is sometimes regionally specific, with notable differences between countries. Such differences are more expressed for some nutrients/foods than others. For example, while amino-acid composition in most meats is relatively stable, the content of many other nutrients can vary in foods available worldwide because of different cultivars and agricultural practices, differences in soils and climatic conditions (1) or consumer preferences specific to certain region. Therefore, composition databases for generic food are maintained in different countries and regions. This is also the case for processed branded foods, where between-country differences can be even more notable (2).

## Importance of Food Composition Information

Data about the composition of foods consumed in a population's diet is very important and has a wide variety of uses (**Table 1**). In research, the data can be used in epidemiological dietary studies to investigate nutrient intake and identify populations at risk for deficiencies. In clinical intervention trials, the datasets can be used to account for dietary factors, where diets/foods are co-founding factors in treatments. In food supply studies, the data can be used to investigate changes over time in the composition of foods. This data is also very important in clinical practice (i.e., in dietary counseling and preparation of diets for patients with special dietary needs or medical conditions). For policymakers, the data are used to set targets for reformulation, assess food reformulation programs and make evidence-based policy decisions. Companies use this data to identify opportunities for improving the composition of foods and to provide information for technology (IT) services, which use food composition data to support dietary, lifestyle, and health objectives. These data are also important for consumers to support informed selections of healthier foods and assure food safety, particularly those with special dietary needs, including allergies.

From a public health perspective, monitoring of food composition and labeling of branded foods available in any country provides insights into the market availability and informs the development of reformulation strategies as well as the planning, implementation, and monitoring of different public health interventions (3–5).

## Food Labeling and Composition Data

In the European Union, labeling of branded foods is regulated by Regulation (EU) No 1169/2011 on the provision of food information to consumers (6). In practice, we can distinguish between mandatory, conditionally mandatory, and voluntary food labeling information (examples provided in **Supplementary Table 1**).

A typical example of mandatory food labeling information is the ingredient list, which includes all the food ingredients in descending order of weight, as recorded at the time of their use in the manufacture of the food (6). This information is typically provided in unstructured text format (list), which can also contain embedded complex ingredients, including food additives. In some cases, ingredients in the ingredient list need to be quantified (QUID - Quantitative Ingredient Declaration). For example, if an ingredient appears in the name of the food or is usually associated with that name by the consumer, or if an ingredient is emphasized on the labeling in words, pictures, or graphics.

With a few exceptions, nutrition declaration is also considered mandatory food labeling information (6). In the EU values are typically expressed per 100 g or ml. The mandatory parts of the nutrition declaration include (a) energy value and the amounts of (b) fat, (c) saturates, (d) carbohydrate, (e) sugars, (f) protein, and (g) salt. However, other parts of the nutrition declaration are non-mandatory or conditionally mandatory, including the amounts of (a) mono-unsaturates, (b) polyunsaturates, (c) polyols, (d) starch, (e) fiber, and (f) vitamins or minerals (**Table 2**). Food manufacturers can decide to include this information on food labels using the regulated format (6). However, if such nutrients/constituents are mentioned in nutrition or health claims on the label, their inclusion in the nutrient declaration is conditionally mandatory (7).

## Potential Sources for the Compilation of Branded Foods Databases

Unlike less processed generic foods, the composition of processed pre-packaged 'branded foods' can change considerably in relatively short time periods because food manufacturers adapt formulations to address consumers' expectations and regulatory requirements, foods are removed from the market, and new ones are launched (8).

Datasets with branded food composition and labeling information originate from a variety of sources (**Table 3**).

**TABLE 1 |** Examples of the use of food composition and labeling information.

|  |  |  |  |  |
|---|---|---|---|---|
| **Research** | **Clinical practice** | **Policymakers** | **Businesses** | **Consumers** |
| • Epidemiological dietary studies<br>• Dietary intervention studies<br>• Clinical intervention trials, where diet or foods are considered co-founding factors<br>• Food supply studies<br>• Assessment of exposure to food components | • Nutritional counseling in patients<br>• Preparation of diets for patients with special dietary needs (including allergies) or medical conditions (for example, diabetes)<br>• Identification of dietary risks | • Basis for evidence-based food policy decisions<br>• Setting the targets for reformulation<br>• Assessment of the efficacy of food reformulation programs<br>• Regulatory restrictions related to specific food components (trans fats, additives) | • Identification of opportunities for improving the composition of foods<br>• Comparisons with other foods–use of comparative nutrition claims<br>• Promotion of foods with improved nutritional composition<br>• Providers of IT services, where food composition data is used to support dietary, lifestyle, and health objectives | • Supporting the informed selection of foods<br>• Enabling comparison of different foods<br>• Supporting choices of healthier foods<br>• Assuring food safety, particularly to those with special dietary needs (including allergies) |

**TABLE 2 |** Composition-related information on labels of processed brand foods in the EU (6).

| Food composition information | Note | Mandatory information |
|---|---|---|
| Ingredient list | Ingredients of the food in descending order of weight | ✓ |
| **Nutrition declaration** | | |
| Energy | (kJ/kcal per 100 mg or mL) | ✓ |
| Fat | | ✓ |
| - Saturates | | ✓ |
| - Mono/poly-unsaturates | | *x** |
| Carbohydrate | | ✓ |
| - Sugars | | ✓ |
| - Polyols/starch | | *x* |
| Fiber | | *x** |
| Protein | | ✓ |
| Salt | | ✓ |
| Vitamins & minerals | The units specified in regulation (6) | *x** |

*Nutrition declaration information for specific nutrients/constituents are provided in grams per 100 mg or ml. *Conditionally mandatory (if the constituent is mentioned in nutrition/health claim) (7).*

## Data Provided by Food Manufacturers

Data can be received directly from food manufacturers, but their participation in providing this food composition information to database compilers is voluntary. Food companies rarely decide to share this information to open databases, and if they do, they often only share some of the required information (e.g., information for nutritional labeling of products). There are some good examples of such a collaborative approach for

**TABLE 3 |** Typical sources of branded food composition and labeling information.

| Food manufacturers | Food monitoring | Crowdsourcing |
|---|---|---|
| Sharing food composition information to databases (voluntarily) | Laboratory analyses of available foods (not feasible on a large scale)<br>Data collections from food labels in food stores<br>Web-scraping: data collections from online sources (online shops, web pages of food producers) | Enabling consumers to collect and share data on the composition of foods, i.e., through smartphone or web applications |

creating branded food datasets (9–13), but very few capture the dynamic changes of the food market (13). The progress in the standardization of this area also needs to be noted. Branded foods are usually labeled with EAN/UPC barcodes (14), which are designed for a high-volume scanning environment and are therefore suitable for a retail point-of-sale (POS) system. EAN-13 barcodes are mostly used in Europe. Each barcode is linked with a unique product identifier, usually the GS1 Global Trade Item Number (GTIN) (15). The food name is considered a mandatory attribute in the GTIN Register, while food composition data are not included. It should be mentioned that GS1 also maintains the Global Data Synchronization Network (GDSN), a globally operating, standardized network that enables the exchange of all types of master product data, including food labeling and composition data, between brand owners, manufacturers, suppliers, distributors, and retailers (16). In the GDSN, based on the Global Data Model (17), specific food labeling parameters are considered mandatory information; however, the use of GDSN is voluntary, meaning that not all items in the GTIN Register have a corresponding record in the GDSN. Furthermore, the GDSN record is only available for subscribers and is not open access. Also, some challenges related with EAN barcode numbers also need to be mentioned. While majority of branded foods have assigned standardized GTIN EAN barcodes, some retailers are using their own non-standardized barcoding systems. In such systems similar products (but with different nutritional composition) sometimes have same barcode number. For example, different flavors of same brand/packaging of fruit yogurt can have same barcode number. There are also some challenges if standardized GTIN EAN barcodes are used, because some products can receive new barcode number without any change in the composition, while in some cases manufacturers do not change barcode number even in case or major change in the composition.

## Food Monitoring Studies

Food monitoring studies typically refer to gathering information about foods available in the food supply in any given market (country and region) using a variety of methods. Considering that chemical analysis of thousands of foods is not a feasible option, a typical approach for collecting branded food information is cross-sectional food monitoring studies in food stores, where data is extracted from food labels (18–20). However, to estimate nutritional composition food labeling data is usually calculated based on the ingredients list, rather than gained by chemical analyses; consequently labeling data can be prone to variety of errors such as batch-to-batch variability, incorrect calculations, processing and stability issues (21). Mandatory food composition information on food labels (**Table 2**) is a particularly important resource for compiling food composition databases because such information can be subject to official controls by food authorities (inspection). It should also be noted that tolerances for the control of compliance of nutrient values declared on food labels have been established (22).

A methodological approach for conducting food monitoring studies was established within the GFMG (Global Food Monitoring Group) (19) and INFORMAS (International Network for Food and Obesity/Non-communicable Diseases (NCDs) Research, Monitoring and Action Support) initiatives in 2013 and 2015, respectively (23, 24). These protocols support regular (preferably yearly) regional data collection on food composition and labeling in all major food suppliers, preferably including photographs, across all available food categories. However, this approach is hardly feasible in practice because such data collection is expensive and time-consuming. Therefore, food monitoring studies are commonly conducted with partial data collection focusing on selected food categories and in selected food shops and are not performed very often (2).

While harmonized data collection enables relevant international comparisons of the food supply (2, 25), such cross-sectional food monitoring studies face several challenges, in particular, keeping the database up to date with newly launched and reformulated products (8, 13). Additional methods are therefore important to enable more frequent data collections.

Online food stores are a promising additional resource for the compilation of branded food databases. While most foods are still sold in classic food stores (i.e., supermarkets), online shops are gaining importance. Most recently, the COVID-19 epidemic increased the use of online food purchases (26), which resulted in the further development of online stores and more foods available by delivery. While e-shops sell fewer foods than classic supermarkets, it is expected that market-leading brands are available online, making such an environment interesting for data collection. Recently, an automated big data analysis approach was applied in the UK to exploit the potential of this market. Harrington et al. tested the extraction of food products' data from supermarkets' webpages weekly (27), allowing timely observations of changes in the marketplace.

### Collection of Data Using Crowdsourcing

Although there are several definitions of crowdsourcing, in general, this term is used for outsourcing different tasks to a crowd of people to complete them; the crowdsourcer can be an organization or an individual (28). Crowdsourcing is an innovative approach used for tasks where a vast amount of data needs to be collected and, therefore, can benefit from a larger group of people completing the task. Advantages include reduced costs, speed, flexibility, scalability, diversity, and participation of citizens, while disadvantages are related to accuracy and duplication. While crowdsourcing has rapidly developed in informational sciences, its application in public health and nutrition is also very promising (29, 30). With the rise of the usage of the internet, an innovative way to access and interpret different types of information became widely available also to consumers. Such platforms can commonly serve as an interface to collect crowdsourcing data.

In the context of food monitoring, these emerging new methods and associated technologies create new opportunities to keep up with the rapid changes in the food chain with less effort. A method of data collection using crowdsourcing allows capturing food market changes more regularly and at a considerably reduced cost while simultaneously delivering value for end-users or crowdsources. Examples of web-based crowdsourcing platforms are Open Food Facts (31) and The Open Food Repo (32) initiatives, which enable the collection of food composition data from different countries. The advantage of such an approach is that created datasets can be used to investigate differences in food composition on global markets (33–35).

The increased use of smartphones and their technical capacities also affect the potential of crowdsourcing in this area considerably. The Australian FoodSwitch application for smartphones was developed to collect branded food composition information and has yielded impressive results (36). A particular innovation in the application was incorporating a crowdsourcing function whereby users are able to contribute information on missing or new products. The crowdsourcing tool within the mobile application has enabled a substantial expansion of the underlying database. This shows that an extensive volume of crowdsourced data provides effective real-time, inexpensive tracking of the nutritional composition of foods in the food supply (37) and reveals a unique opportunity for using such an approach in other countries (38).

### Future of Branded Food Datasets

There is no doubt that with developments in this scientific field and progress in information technology, branded foods datasets will gain importance in the future. It should be mentioned that existing datasets and food nutrition security data in general, although widespread, are fragmented and lack critical mass and accessibility. The data are not readily found, accessible, interoperable, or reusable (FAIR). The European project Food Nutrition Security (FNS) Cloud (https://www.fns-cloud.eu/; accessed 11.10.2021), funded by the European commission HORIZON2020 framework program, is addressing this challenge. The FNS-Cloud is developing the first generation 'food cloud' by federating existing and emerging datasets and developing and integrating services to support re-use through the European Open Science Cloud (EOSC). The project's major objectives are to support standardization and demonstrate the usability of such datasets (39).

Food manufacturers are key data sources for the future of branded food composition databases and should be encouraged to share food composition data in open-access databases. Major progress could be achieved if this became a legal requirement. Despite the above-mentioned importance of making food composition data available to research, clinical practice, policymakers, and consumers, such a legal obligation is not expected in the near future. Until then, an approach that combines data provided by food manufacturers where available, and data collected in food monitoring studies, supplemented with data collected from other approaches, such as crowdsourcing and web scraping, is needed. Challenges in connecting different datasets need to be addressed with harmonization in data structure and the development of automatic services.

## COMPILATION OF BRANDED FOODS DATABASE IN SLOVENIA

### History of Monitoring Branded Foods in the Slovenian Food Supply

In Slovenia, data requirements about labeling and composition of branded foods were expressed in 2011, soon after the adoption of Regulation (EC) No 1924/2006 regarding nutrition and health claims made on food products (7). This regulation also provided for monitoring changes in the food supply, particularly the use of nutrition and health claims on foods. It should be mentioned that at that time, harmonized standards for such data collection were not yet available, but the importance of this research topic was highlighted in other countries. For example, Lalor et al. (40) published results of monitoring the Irish food supply in 2010. Their monitoring approach was used to set up food monitoring in Slovenia.

The first food monitoring study in Slovenia was conducted in 2011 within research project V7-1107 'Nutrition and health claims on foods' (41), funded by the Slovenian Research Agency and the Ministry of Agriculture, Forestry, and Food of the Republic of Slovenia. Data collection was described previously (42). In short, researchers visited the grocery stores of three retailers (Mercator, Spar, Hofer) in Ljubljana, which covered the majority of the national food supply. In agreement with the retailers, data collection was done directly in grocery stores; researchers extracted food labeling information into an Structured Query Language (SQL) database, using local notepad computers. A special electronic form was developed, enabling the quick collection of the data into the database. European/International Article Number (EAN) barcodes were used as unique product identifiers. Such an approach avoided duplicates; each product was only collected once, even if it was available on different shelves or in different food stores. The selection of food categories was made according to Lalor et al. (40), with the addition of processed seafood, ready meals, vegetable oils, and plant-based milk/yogurt imitates. The dataset only contained numeric/text information, without any pictures. Altogether, 6,348 unique items were sampled in this study. The following information was collected: date/time stamp, store identification, product EAN/barcode number, food category, food name, manufacturer, use of health and other symbols, use of nutrition/health claims, and nutrition declaration information (energy, fat, saturates, sugar, sodium, fiber). At that time, the labeling of nutritional information on processed branded foods was not yet mandatory; calorie content was available for 65.6% of foods in the dataset, and sodium content for 39.2%.

Confidentiality agreements were signed with major retailers in exchange for their nationwide 12-month sales data. Sales data were provided by each food (EAN barcode number) separately. If the same product (barcode number) was sold at different retailers, the total sale volume was calculated. The barcode number was used to match food products in our dataset with corresponding sales data. Sales data were obtained for 80.4% of foods in our dataset (42).

A major challenge of the 2011 study was that the entire data collection process was conducted within food stores. Researchers had temporary collection points, with laptops connected with the SQL database over the Wi-Fi network. This approach presented a burden for retailers, who had specific requirements to limit the effect of the study on consumers. Another limitation of the monitoring approach was that the dataset produced did not enable further verification of the data or exploitation of parts of the food labeling which were not the subject of the original study. For example, post-data collection studies of marketing approaches to children were not able to be conducted because these were not included in the initial data collection.

Another monitoring approach was used in 2013 within the European FP7 research project CLYMBOL ("Role of health-related CLaims and sYMBOLs in consumer behavior"), funded by the European Commission (43, 44), which was also conducted in Slovenia. While the 2011 food monitoring study collected data for all available foods in the selected food categories, in the CLYMBOL project, foods were sampled with a randomization approach. Food was purchased in stores, and data collection was completed outside the food store (44). Because foods were perishable, all the packages were photographed to enable later processing. The strength of this approach was the ability to verify all recorded information, while limitations included logistic challenges in the randomization, purchasing, and data collection. Also, purchasing foods incurred unnecessary food waste and considerable costs, which were mitigated by donations. Altogether, 2,034 foods were sampled in five countries (about 400 per country) (44).

## Introduction of a Standardized Food Monitoring Approach

In 2015, the national "Nutrition and Public health" research program was initiated in Slovenia with funding from the Slovenian Research Agency. The 2011 branded food composition database had proved very useful for researchers and policymakers, so a major objective of this national research program was also to follow changes in the food supply, with a specific focus on processed branded foods (45), a key contributor in various diet-related non-communicable diseases (46). This time period was also marked by progress in the international harmonization of the methodological approach for conducting food monitoring studies (19, 23, 24), providing guidance for food categorization and data collection, and prioritizing data extraction from photographs of food labels.

A decision was therefore taken to develop an infrastructure to support more efficient food monitoring studies in Slovenia. At that time, the George Institute for Global Health (Australia) 'FoodSwitch' (47) and the University of Toronto (Canada) 'Food Label Information Program (FLIP)' (48) were examples of the most sophisticated infrastructure in this field, and their developers offered important insights, which supported the development of the infrastructure in Slovenia.

In line with this, the Composition and Labeling Information System (CLAS) infrastructure was developed, composed of:

- Data collection via the smartphone application CLAS for use by researchers in food stores; details are specified in **Supplementary Table 1** and **Supplementary Figure 1**. This mobile application works on the Android operating system and directly communicates with the background CLAS SQL database over a Wi-Fi network or through mobile data transfer.
- Data extraction and management via the online CLAS tool; details are specified in **Supplementary Table 2** and **Supplementary Figure 2**. The tool runs on MS Windows Server 2016, while the database runs on MS SQL Server 2016 Standard.

Both tools are interconnected, enabling easier data collection. The smartphone application CLAS can be used for real-time collection by several researchers at the same time without risk of duplication, as barcodes are used as a unique identifier for this purpose, thus ensuring data from a given food is only collected once. From inside food stores, the researcher scans the barcode of a product (using the CLAS application), and if data from this product has not yet been collected in the cross-sectional study, the mobile application requires the researcher to take photos of all sides of the food packaging and to input price information. Timestamp (date, time) and food store identification are saved automatically. All these data and images are directly sent from the mobile application to the online CLAS tool, where they are checked for quality using both automatic controls and a manual check of each product by the researcher (49). If a product with the same barcode was already collected in the same cross-sectional study, only the store identification and time-stamp are saved to the database, and the researcher is notified that product information is already collected, thus removing duplication of work. Data extraction is completed in an online CLAS tool, with the support of Optical Character Recognition (OCR) technology, and supported by manual work and cross-checking (49).

Since 2015, all food monitoring studies in Slovenia have been conducted with this approach, employing data extraction from

**TABLE 4 |** Description of food monitoring studies conducted in Slovenia (2011–2020).

| Year | 2011 | 2015 | 2017 | 2019 | 2020 |
|---|---|---|---|---|---|
| No. of records* | 6,348 | 10,694 | 21,090 | 6,892 | 28,028 |
| Included retailers | Mercator/Spar/Hofer | Mercator/Spar/Hofer | Mercator/Spar/Tuš/Hofer/Lidl | Mercator/Spar/Tuš/Hofer/Lidl | Mercator/Spar/Tuš/Hofer/Lidl/Eurospin |
| Sample | Selected food categories [details in (42)] | Selected food categories [details in (50)] | All food categories, excluding alcoholic drinks | Selected food categories [focus on categories with added sugar] | All food categories (including alcoholic drinks) |
| Nutrition declaration data | ✓ | ✓ | ✓ | ✓ | ✓ |
| Ingredient list data | *x* | ✓ | ✓ | ✓ | ✓ |
| Price | *x* | ✓ | ✓ | ✓ | ✓ |
| Notes | Manual collection of data from food labels in food stores. | Collection of data using CLAS infrastructure from pictures of food labels taken in food stores. All pictures are archived. | | | |

**Number of records before removing ineligible foods for specific studies/analyses (i.e., removal of items which combined toys, items with different types of foods in the same package).*

food labeling photographs (**Table 4**). In 2015, food monitoring studies were conducted in the same three retailers as in 2011 (Mercator, Spar, Hofer), while retailers Tuš and Lidl were added in 2017 and Eurospin in 2020. Food monitoring in 2015 built on experiences gained in 2011 (42) and still included only selected food categories. A cross-sectional study in 2017 investigated almost all available pre-packed foods (excluding alcoholic beverages); however, alcoholic beverages were also included in 2020. A cross-sectional study in 2019 was conducted as a partial study, focusing only on food categories with higher sugar content, specifically addressing objectives of the national research project "Sugars in human nutrition: availability in foods, dietary intakes, and health effects" (51).

As in the 2011 food monitoring study, arrangements were made with major retailers, who provided their nationwide 12-month sales data, and barcode numbers were used to match food products in the dataset with corresponding sales data.

In line with the increasing number of retailers in the food monitoring studies and the inclusion of additional food categories, the number of sampled foods has increased considerably since 2011; over 28,000 unique products were sampled in the last data collection (2020). With consideration that each regular data collection (2011, 2015, 2017, 2020) included the same/additional retailers/food categories, the compiled dataset can be used to provide interesting insights about the market-life of foods available in the food supply (**Figure 1**). For example, out of 6,348 products sampled in 2011, only 2,285 (−65%) were still found in 2015, and 1,526 (−80%) were found in 2020. Similarly, only about half of products from the complete food monitoring in 2017 were still on the market in 2020. These trends indicate that while a small proportion of branded foods in the food supply has a long market-life, many foods are quickly removed from the market and substituted with new products. Food supply studies are challenging because branded food databases become outdated very quickly. For this reason, monitoring is essential, and enforcing the obligation of sharing data openly for producers would be beneficial to the field.

## Introduction of Crowdsourcing

In 2019 a smartphone application called "VešKajJeš" (#VKJ; English translation: You know what you eat) was launched in Slovenia as part of the collaborative "IRIO" ('Innovative solutions for informed choices') project. The application is available for Android and iOS platforms. The project was launched by the Nutrition Institute, 'Jožef Stefan' Institute, and the Slovenian Consumer Organization and supported by the Slovenian Ministry of Health. The app enables users (consumers) to scan the barcode (EAN) of selected food and receive feedback information on the product's nutritional composition. It also interprets the nutritional information based on the nutrient profile by using the food traffic light labeling system (**Figure 2**), supporting healthier food choices.

The crowdsourcing function of the application is activated when a user scans a food barcode that is either not yet included in the branded food database or when there is a difference between the nutritional composition of the scanned food and the information presented in the mobile application VešKajJeš (**Figure 3**).

Crowdsourcing information is stored and processed in a special web application (bazil.si), which was developed by the 'Jožef Stefan' Institute within the IRIO project (specified in **Supplementary Table 3**). Bazil.si enables researchers to view data sent by users and transcribe the data of interest from the labels that can be seen on the images. Products are processed based on search volume, so the products that are most scanned by users are added to the database as soon as possible to ensure a high rate of positive answers by the VešKajJeš application. Extracted parameters include EAN number, product name, category, and data from mandatory nutrition declarations. Since its release in mid-2019, the application users have contributed 11,482 unique items, 9,348 of which were processed. The application currently

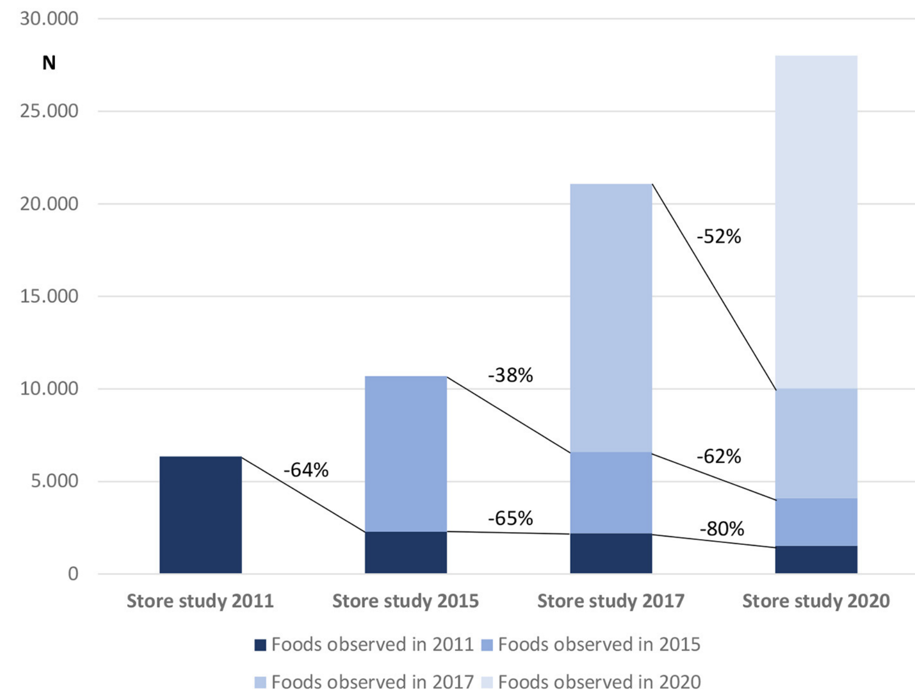


**FIGURE 1 |** Number of foods collected in regular food monitoring studies in Slovenia.

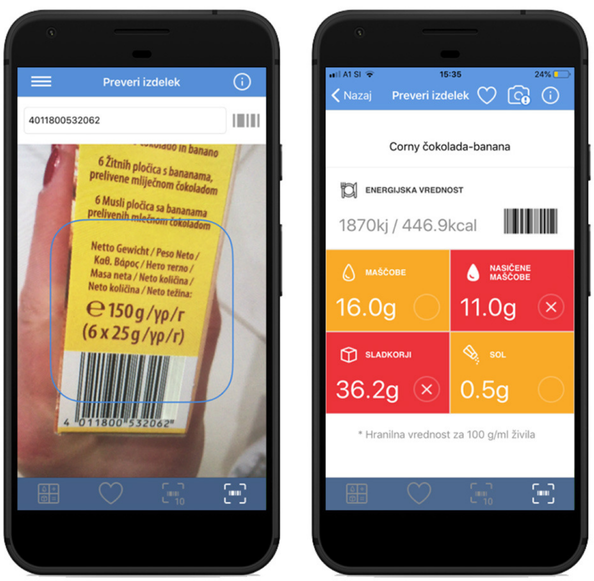


**FIGURE 2 |** Branded food composition database can support consumers in interpreting food composition information. Example of the Slovenian smartphone application VešKajJeš (iOS/Android).

has 24,000 active users (over 1% of the whole population of Slovenia) who make approximately 250,000 inquiries per year. In general, the application is well-received by the general public, and was also highlighted on the Information Society Multiconference with 'the information strawberry' award for best IT achievement in year 2019 (52). It is worth mentioning that for the success of the application, a well-planned and executed media campaign was and still is crucial.

A dataset compiled from crowdsourcing is currently used as a complementary data resource for the mobile app #VKJ. This enables the composition of newly identified foods to be also available to other app users. This dataset is also used within the previously mentioned FNS Cloud project (39), which is investigating if crowdsourcing can be used as an efficient, effective, and low-cost method for generating reliable food composition datasets. Both CLAS and Bazil.si are inter-connected with Application Programming Interface (API) to support continuous data transfers (**Figure 4**).

## THE EXPLOITATION OF USE OF BRANDED FOODS DATABASE IN NUTRITION RESEARCH

The first Slovenian food monitoring study in 2011 identified a major challenge: many foods were not labeled with full nutrition declaration data. Considering that nutritional profiling of foods can be only be conducted if sufficient details on nutrition composition are available, a relevant question was whether such missing information could be supplemented from other sources. Eržen et al. (53) conducted a comparative evaluation of the use of generic food composition databases for nutrient profiling. The study was conducted with two nutrient profiling models

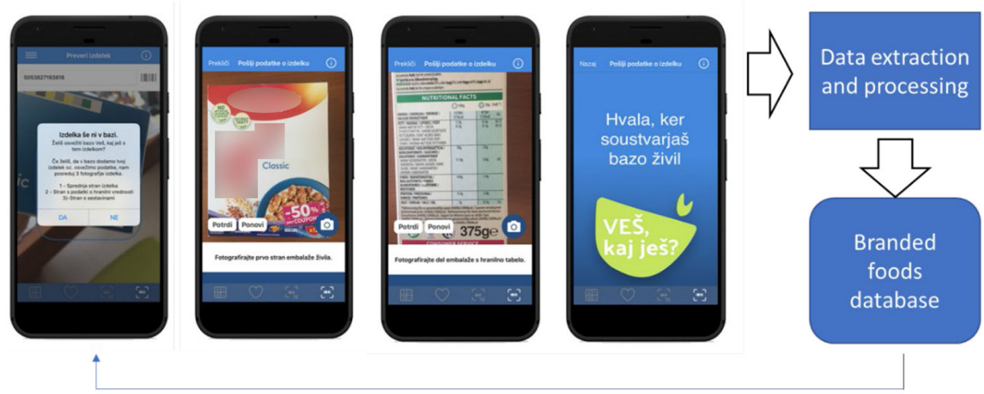


**FIGURE 3 |** Crowdsourcing approach used in Slovenia with VešKajJeš smartphone application.

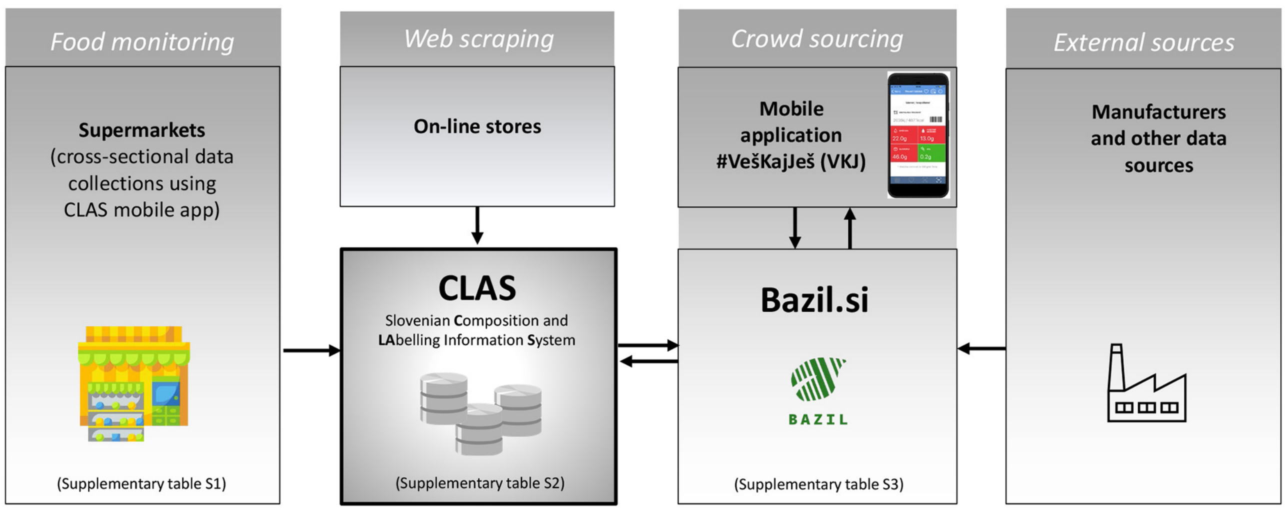


**FIGURE 4 |** Schematic presentation of data-pathways in the Slovenian Composition and Labeling Information System (CLAS).

(UK Ofcom, Australian FSANZ). Moderate/good agreement was observed for the dataset compiled from food labels, compared with the dataset compiled using the food matching technique (54). This was particularly important in the 2011 dataset, as, at that time, labeling of nutrition declaration was not yet mandatory in the European Union. However, the situation changed considerably after the implementation of harmonized EU food labeling Regulation (EU) No 1169/2011 (6) in December 2016, which made this information part of mandatory food labeling. Nevertheless, some types of information which are commonly used in nutrition research (i.e., content of dietary fiber) are still not part of mandatory food labeling, meaning that the food-matching approach is still relevant.

Datasets collected in Slovenian food monitoring studies were either directly or indirectly used in numerous studies in public health and nutrition research. We have reviewed all types of data exploitation, which resulted in peer-reviewed scientific publications. We identified $N = 18$ studies, focused on analyses of branded foods datasets (**Table 5**), and $N = 11$ studies in which branded foods datasets were used indirectly as a source for estimating the nutritional composition of branded foods collected in other studies (**Supplementary Table 4**).

## Studies Directly Exploiting Branded Food Databases

### Assessments of Nutritional Composition of Food in the Food Supply

Several studies have focused on assessing the nutritional composition of processed foods available in the food supply using nutrition declaration information. Korošec et al. (42) used the 2011 dataset to investigate the importance of sales data in assessing the food supply. A similar study was also conducted on the 2015 dataset (49) to investigate time-changes in the sodium content in branded foods as a result of voluntary food

**TABLE 5 |** Examples of studies directly exploiting Slovenian branded food databases (2011-2020; details provided in **Supplementary Table 4**).

| Type of use | Examples |
|---|---|
| Assessments of nutritional composition of food in the food supply | Salt/sodium (42, 49); Free/total sugar (50, 55) |
| Assessments of the use of specific ingredients | Partially hydrogenated vegetable oils/fats as sources of trans fatty acids (54); food fortification (56) |
| Assessments of the use of food additives | Sweeteners (51, 57), titanium dioxide (58) |
| Assessments using nutrient profiling approach | Methodological/assessment of the use of external data for missing values (53); international comparisons (2); nutrient profiling of the food supply (20, 59, 60) |
| Assessment of food marketing techniques on food labels | Nutrition and health claims on foods (41, 61), children marketing (62), gluten-free claim (63) |

*List only include studies where Slovenian branded food databases were exploited directly. Food monitoring studies datasets were used for several other studies, for example, to support in the sampling of foods in the food supply for assessment of the content of trans fatty acids in foods (64, 65), for assessment of salt iodization (66), for assessment of dietary intakes (67–71), food marketing (25, 72–74) and offers in vending machines (75, 76). Studies exploiting additional Slovenian branded foods composition datasets (44, 66, 77–79) are also not included.*

reformulation activities. The study identified a trend of reduced sodium content in cheese and a neutral trend in bread and meat products, while higher sodium content was observed in ready meals, highlighting that additional efforts are needed for sodium reduction (49).

A similar approach was used by Zupanič et al. (50) on the 2015 dataset, but with a focus on total/free sugar content in branded foods. A particular challenge in this study was that in the EU, food labels only provide information on total sugar content, but not for added and/or free sugar. Therefore, an algorithm for estimating free sugar was established based on the protocol developed by Bernstein et al. (80), where labeled food ingredient information is used together with nutrition information data. Altogether, more than half of the products in the dataset contained free sugar; categories with the highest content were honey/syrups, jellies, chocolate and sweets, jam/spreads, and cereal bars. Using sales data, chocolate/sweets and soft drinks were identified as major contributors to free sugar in the food supply. In 2015, the Slovenian food industry accepted a pledge to lower sugar content in beverages. To investigate the efficacy of industry self-regulation on trends in free sugar content in branded foods, a similar study was conducted using the 2017 dataset (55). The study identified soft drinks as the most important source of free sugar, responsible for 28% of all free sugar sold with branded foods. Few positive trends were observed, and the study also highlighted that in the categories with the highest share of free sugar, brands with higher market share were commonly sweeter than category averages.

## Assessments of the Use of Specific Ingredients, Including Additives

Food composition datasets were also explored with a specific focus on particular ingredients reported in labeled food ingredient lists. One particularly interesting food ingredient is partially hydrogenated oils (PHO), a source of trans fatty acids (TFAs), which is a key risk factor for cardiovascular disease (81). After a surprising report by Stender et al. (82), which identified that PHO-containing biscuits in Slovenia are still a source of TFAs, Zupanič et al. (54) investigated Slovenian branded foods datasets compiled in 2015 and 2017 to identify food categories where PHOs were most commonly used. In this study, biscuits were identified as a food category with the most frequent usage of PHO, but the study also showed a notably lower usage of this ingredient in 2017. Same approach has been used for the exploitation of food fortification practices with vitamin D (56).

Similarly, Hafner et al. investigated trends in the use of different artificial sweeteners in drinks in the Slovenian food supply in the years 2017, 2019 (51) and 2020 (57), while Blaznik et al. (58) exploited usage of titanium dioxide (food additive E 171) in the food supply. The latter case is an excellent example of the exploitation of branded food composition databases for rapid public health use, supporting food policers and food control authorities. Titanium dioxide has been used as a very efficient white food pigment for decades (83), but later research showed that a notable proportion of this additive are nanoparticles (84, 85) and highlighted genotoxicity concerns (86, 87). After a re-evaluation of the safety of this additive, the EFSA published an official opinion in May 2021 (88), concluding that because a concern for genotoxicity could not be excluded, this additive could not be considered safe anymore. As a rapid response, the last Slovenian branded foods dataset (CLAS 2020) was investigated for the use of $TiO_2$ and compared with the 2017 dataset. Preprints of the study were published 1 month after the publication of the EFSA opinion in June 2021 (89), and a peer-reviewed publication was published in August 2021 (58). Chocolate and sweets, and chewing gums were identified as categories in which $TiO_2$ is most commonly used; its use significantly decreased (by 50%) between 2017 and 2020.

## Assessments Using Nutrient Profiling Approach

While previous examples investigated the food supply with a focus on specific nutrients or food ingredients, another approach employs nutrient profiling, in which several nutrients/constituents are considered at the same time. 'Nutrient profiling' is described as the science of classifying foods according to their nutritional composition for reasons related to preventing disease and promoting health (90, 91). We should note that there is no consensus on the superiority of a specific nutrient profiling system. Models are used for very different purposes (adjusting the food supply, food marketing, supporting healthy dietary choices, etc.) and also differ considerably in the profiling algorithms, i.e., in the number/selection of food categories, the involved nutrients and food constituents, and regarding involving of scoring (92–94). Two common general approaches are the use of category-based cut-off values, as in case the WHO Office for Europe nutrient profile (WHOE)

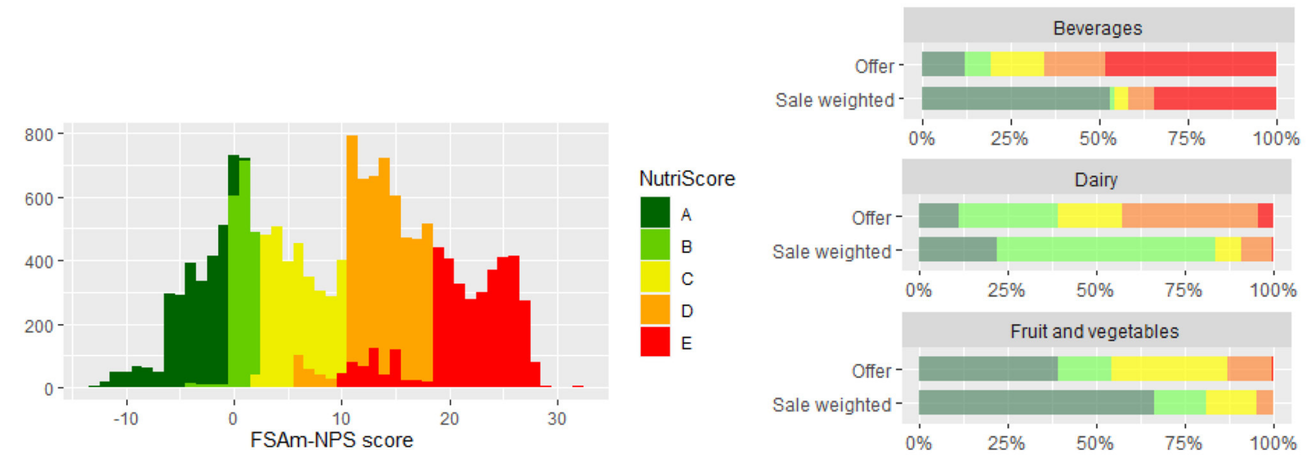


**FIGURE 5 | (Left)**: Distribution of Food Standards Agency Nutrient Profile (FSAm-NPS) scores and corresponding Nutri-Score grades for foods and drinks in the Slovenian food supply. **(Right)**: Distribution of Nutri-Score grades across selected categories using sale-weighting. Reproduced from Hafner et al. (60) with approval of authors.

(95), or a scoring approach with consideration of different nutrients/constituents, as in the case of the United Kingdom Food Standards Agency nutrient profiling system (NPS) (96), Australian Nutrient Profiling Scoring Criterion (NPSC) and Health Star Rating (HSR) (97), and French Nutriscore (NS) (98, 99).

Dunford et al. (2) used the 2015 version of the Slovenian branded foods dataset in an international comparison of the healthiness of packaged foods/beverages from 12 countries (Australia, Canada, Chile, China, India, Hong Kong, Mexico, New Zealand, Slovenia, South Africa, the UK, and the USA) using the HSR nutrient profiling system. Specific food categories were highlighted as particularly problematic in the Slovenian food supply, for example, beverages, snacks, and meat products. Altogether, Slovenia ranked in the middle between the countries with the highest overall nutrient profile (UK, USA, Australia, and Canada) and the countries with the lowest ranks (India, Hong Kong, China, and Chile) (2). The 2015 dataset was also used to compare different front-of-package labeling schemes (20) and investigate the nutritional quality of the foods labeled with health-related claims (59). The latter study reported that about 68% and 33% of the foods labeled with health-related claims passed NPSC and WHOE criteria, respectively, highlighting the need for stricter regulations to use such claims on foods.

Nutrient profiles are a very timely topic because of the current discussions in the European Union on the possibility of implementing a mandatory front-of-pack nutritional labeling scheme. Nutri-Score (NS) is an example of such a scheme, which was originally developed in France (98, 99) but is currently voluntarily used in several other European countries (100). NS presents an upgrade of the NPS; it is a scoring system based on the content of selected nutrients/constituents per 100 g of food. This nutrient profile model grades the nutritional quality of products with a 5-color/letter scale from dark green (A; healthy) to dark orange (E; least healthy) (101). Very recently, Hafner et al. used the sale-weighting approach to evaluate the power of NS for discriminating the nutritional quality of branded foods using the 2017 branded foods dataset (60). The study indicated the very good ability of this nutrient profile model to compare between foods (**Figure 5**, left). This was confirmed also within food (sub)categories. Furthermore, it was shown that the availability of foods does not always reflect their sales; notable differences between available and sold foods were reported, particularly in beverages, dairy products, fruits, and vegetables (**Figure 5**, right).

### Assessment of Food Marketing Techniques on Food Labels

In addition to the previously mentioned evaluation of the nutritional quality of food labeled with nutrition and health claims (59), several other studies also exploited branded food databases to assess food marketing techniques on food labels. The first such study employed the 2011 dataset to investigate consumers' exposure to nutrition/health claims on branded foods (41); 37 and 15% of foods were labeled with nutritional or health claims, respectively (this was 45 and 11%, after correction for market-shares). The majority of health claims were general non-specific and/or function claims. Later on, researchers exploited the 2015 dataset to investigate the use of gluten-free claims (63), heart images (61), and marketing to children (62). Studies identified several challenges of such poorly regulated marketing, which often promotes foods with questionable overall nutritional composition. For example, the use of cartoon characters and other children's marketing techniques on food labels was almost exclusively linked to unhealthy foods (62), indicating the need for regulation in this area.

## Indirect Use of Branded Food Databases in Nutrition Research

In addition to the above-mentioned direct exploitation of branded food databases, datasets have also been used indirectly to support nutrition research (**Supplementary Table 3**). A relevant example is supported in the sampling of foods in the food supply. Food labeling only includes some types of information, while others must be investigated using laboratory investigations. However, sampling in such studies is very challenging because the tens of thousands of foods that are available on the market cannot be easily analyzed. In such cases, branded foods datasets can be used to support the selection of foods for laboratory analyses. For example, Kušar et al. (64) and Mencin et al. (65) used the 2017 dataset to select branded foods for determining the number of trans fatty acids (TFA) analytically. Studies have identified specific food categories with higher content of TFAs, supporting the Slovenian government in implementing a regulatory ban on such foods (102).

We should also mention that branded foods datasets are a useful resource in conducting dietary surveys. In line with guidance on EU Menu Methodology (103), nationally representative food consumption studies should employ at least two 24-h recalls, meaning that the collected datasets contain thousands of different food items, for which nutritional composition needs to be estimated. While in some cases, this can be done using generic food databases, estimation of the composition of branded foods is much more reliable if branded food datasets are also employed. The Slovenian national dietary survey SI.Menu was conducted in 2017/2018 (104), and the 2017 edition of the branded foods dataset was used to estimate the nutritional composition of reported foods. This approach has already supported several epidemiological studies, which investigated diet-related-risk factors and population intakes of total/free sugars (69), dietary fiber (70), TFAs (68), vitamin D (67) and folate (71).

Studies investigating food marketing are another example of using branded food datasets for estimating the nutritional composition of processed foods. Such studies typically result in very large collections of advertisements (i.e., from television, magazines), which are linked to specific foods. Considering that majority of food marketing is for branded foods (2), such products can be linked with branded food databases while missing information is searched in other sources. Such an approach was used in analyses of food marketing to children in magazines (73) and on television (25, 72). Considering that television marketing was identified as particularly problematic, the Slovenian government introduced regulatory restrictions for marketing within programs targeting children, and this intervention was also evaluated (74) using a branded food database.

Similarly, the nutritional quality of foods available in vending machines was also investigated. Rozman et al. reported the poor nutritional quality of snacks (75) and beverages (76) available in vending machines in health and social care institutions in Slovenia.

## CHALLENGES AND CONCLUSIONS

Branded foods databases are becoming valuable not only in nutrition research but also for clinical practice, policymakers, businesses, and end-users. In contrast to generic foods, branded foods are marked by rapid changes in the food supply because of reformulations, the introduction of new foods, and the removal of existing ones from the market. Also, different branded foods are available in different countries. This not only complicates the compilation of branded foods datasets but also makes such datasets out-of-date quite quickly.

In an ideal situation, the composition of branded foods would be provided in an open-access format automatically by food manufacturers, like in the Netherlands. There is no indication that such data sharing will become mandatory in the near future; therefore, branded food databases also need to employ other approaches to assure that data are up-to-date. A standard approach is conducting food monitoring studies in supermarkets to collect the information presented on food labels. Such an approach has several challenges: (1) studies can be only conducted periodically in selected stores, (2) sophisticated infrastructure is needed to conduct such studies efficiently, (3) depending on the extent of the data collected, such studies are related to notable man-power and costs, (4) a long time period is needed to complete such studies: when data-collection and analyses are completed, the situation in the food supply could already be different, (5) the accuracy of the data is limited to the accuracy of the data provided on food labels. Therefore, complementary approaches are used to supplement such cross-sectional datasets, for example, more regular data collections of the (much more limited) offer of branded foods in online stores and crowdsourcing to collect data about the most relevant foods in real-time.

Branded food databases can be used directly for assessments of nutritional composition of food in the food supply (i.e., salt, sugar content), use of specific ingredients, for example, food additives, for nutrient profiling, and the assessment of marketing techniques on food labels. Such datasets are also valuable for other studies, for example, for nutrient intake assessments in dietary surveys.

## AUTHOR CONTRIBUTIONS

IP: conceptualization and first draft. IP, MH, KŽ, BB, BK, and AK: manuscript writing—review and editing. All authors have read and agreed to the published version of the manuscript.

## FUNDING

In Slovenia Composition and Labeling Information System (CLAS) has been conducted by the Nutrition Institute (Ljubljana, Slovenia) within the national research programme "Nutrition and Public Health" (P3-0395) and "Infrastructure programme for monitoring of the composition and labelling of foods" (IO-0054), funded by the Slovenian Research Agency. Development of the infrastructure and data collections were further supported by several national research projects (L7-1849, V3-1901, L4-9305, L3-9290, L4-7552, L3-7538, V3-1501, V7-1107, P2-0098), funded by Slovenian Research Agency, the Ministry of Health of Republic of Slovenia, and Ministry of Agriculture, Forestry, and

Food of the Republic of Slovenia; by national health promotion programmes (IRIO/mobile application VeškajJeš/Do you know what you eat? and Do you know what you drink?), funded by the Ministry of Health of Republic of Slovenia, and by the Food Nutrition Security Cloud project (FNS-Cloud), which received funding from the European Union's Horizon 2020 Research and Innovation programme (H2020-EU.3.2.2.3.—A sustainable and competitive agri-food industry) under grant agreement no. 863059. Information and views in this report do not necessarily reflect the official opinion or position of the European Union. Neither European Union institutions and bodies, nor any person acting on their behalf, may be held responsible for the use that may be made of the information contained herein.

## ACKNOWLEDGMENTS

The authors would like to thank the retailers for granting access to their stores to collect data for the study. We acknowledge the support of researchers at the Nutrition Institute (including Sanja Krušič and Živa Lavriša) and students from the Biotechnical Faculty (University of Ljubljana) and BIC (Ljubljana) in data collections. Further we acknowledge support of information technology experts, collaboration in the development of CLAS (Miha Plesnik) and BAZIL (Urban Škovrc, Matevž Ogrinc), and the Slovenian Consumer Organization–our partner in national health promotion programmes and owner of the VešKajJeš mobile application. We also acknowledge support of FNS-Cloud (see Funding for details) beneficiaries, particularly Paul Finglas (Quadram Institute, Norwich, UK), Eileen Gibney (University College Dublin, Ireland), Kurt Gedrich (Technical University of Munich, Mark Roe (EUROFIR; Brussels, Belgium), Germany), Susanne Westenbrink (RIVM; Netherland), Karl Presser (Premotec, Switzerland) and Mateja Podlogar (GS1 Slovenia, Ljubljana, Slovenia).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.798576/full#supplementary-material

## REFERENCES

1. Greenfield H, Southgate DAT. *Food Composition Data: Production, Management, and Use*. Food & Agriculture Org. (2003).
2. Dunford EK, Mhurchu CN, Huang L, Vandevijvere S, Swinburn B, Pravst I, et al. A comparison of the healthiness of packaged foods and beverages from 12 countries using the Health Star Rating nutrient profiling system, 2013-2018. *Obes Rev*. (2019) 20:107–15. doi: 10.1111/obr.12879
3. Beaudry M, Delisle H. Public('s) nutrition. *Public Health Nutr*. (2005) 8(6a):743–8. doi: 10.1079/PHN2005777
4. Walls HL, Peeters A, Loff B, Crammond BR. Why education and choice won't solve the obesity problem. *Am J Public Health*. (2009) 99:590–2. doi: 10.2105/AJPH.2008.156232
5. World Health O. *General Meeting of the WHO Global Coordination Mechanism on the Prevention and Control of Noncommunicable Diseases: Meeting Report: International Conference Centre, Geneva, Switzerland, 5-6 November (2018)*. Geneva: World Health Organization (2019).
6. EC. Regulation (EU) no 1169/2011 of the European parliament and of the council of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004 (2011). Available online at: http://data.europa.eu/eli/reg/2011/1169/2018-01-01 (accessed September 27, 2021).
7. EC. Regulation (EC) No 1924/2006 of the European Parliament and of the Council on nutrition and health claims made on foods (2006). Available online at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02006R1924-20141213 (accessed March 29, 2020).
8. Kapsokefalou M, Roe M, Turrini A, Costa HS, Martinez-Victoria E, Marletta L, et al. Food composition at present: new challenges. *Nutrients*. (2019) 11:1714. doi: 10.3390/nu11081714
9. Kretser A, Murphy D, Starke-Reed P. A partnership for public health: USDA branded food products database. *J Food Compos Anal*. (2017) 64:10–2. doi: 10.1016/j.jfca.2017.07.019
10. Perrin C, Battisti C, Chambefort A, Digaud O, Duplessis B, Volatier J-L, et al. Range of processed foods available in France and nutrition labelling according to the type of brand. *J Food Compos Anal*. (2017) 64:156–62. doi: 10.1016/j.jfca.2017.08.009
11. Seeuws C. Belgian branded food products database: inform consumers on a healthy lifestyle in a public-private partnership. *J Food Compos Anal*. (2017) 64:39–42. doi: 10.1016/j.jfca.2017.07.008
12. Ocké MC, Westenbrink S, van Rossum CTM, Temme EHM, van der Vossen-Wijmenga W, Verkaik-Kloosterman J. The essential role of food composition databases for public health nutrition – Experiences from the Netherlands. *J Food Comp Anal*. (2021) 101:103967. doi: 10.1016/j.jfca.2021.103967
13. Westenbrink S, van der Vossen-Wijmenga W, Toxopeus I, Milder I, Ocké M. LEDA, the branded food database in the Netherlands: Data challenges and opportunities. *J Food Comp Anal*. (2021) 102:104044. doi: 10.1016/j.jfca.2021.104044
14. GS1. EAN/UPC barcodes (2021). Available online at: https://www.gs1.org/standards/barcodes/ean-upc (accessed September 28, 2021).
15. GS1. Global Trade Item Number (GTIN) (2021). Available online at: https://www.gs1.org/docs/idkeys/GS1_GTIN_Executive_Summary.pdf (accessed September 28, 2021).
16. GS1. GS1 Global Data Synchronisation Network (GS1 GDSN) (2021). Available online at: https://www.gs1.org/services/gdsn (accessed September 28, 2021).
17. GS1. GS1 Global Data Model (2021). Available online at: https://www.gs1.org/standards/gs1-global-data-model (accessed September 28, 2021).
18. The Food Monitoring G. Progress with a global branded food composition database. *Food Chemistry*. (2013) 140:451–7. doi: 10.1016/j.foodchem.2012.10.065
19. Dunford E, Webster J, Metzler AB, Czernichow S, Ni Mhurchu C, Wolmarans P, et al. International collaborative project to compare and monitor the nutritional composition of processed foods. *Eur J Prev Cardiol*. (2012) 19:1326–32. doi: 10.1177/1741826711425777
20. Pivk Kupirovič U, Hristov H, Hribar M, Lavriša Ž, Pravst I. Facilitating consumers choice of healthier foods: a comparison of different front-of-package labelling schemes using slovenian food supply database. *Foods*. (2020) 9:399. doi: 10.3390/foods9040399
21. Albuquerque TG, Nunes MA, Oliveira M, Costa HS. Compliance of declared vs. analysed values with EU tolerance limits for mandatory nutrients in prepacked foods. *Food Chem*. (2020) 302:125330. doi: 10.1016/j.foodchem.2019.125330
22. EC. Guidance document for competent authorities, tolerances for the control of compliance of nutrient values declared on a label with EU legislation (2012). [15.2.2021]. Available online at: https://ec.europa.eu/food/sites/food/files/safety/docs/labelling_nutrition-vitamins_minerals-guidance_tolerances_1212_en.pdf (accessed February 2, 2021).

23. Neal B, Sacks G, Swinburn B, Vandevijvere S, Dunford E, Snowdon W, et al. Monitoring the levels of important nutrients in the food supply. *Obesity Reviews.* (2013) 14(S1):49–58. doi: 10.1111/obr.12075
24. Rayner M, Vandevijvere S. *INFORMAS Protocol: Food Labelling Module*. Oxford: University of Oxford (2017).
25. Kelly B, Vandevijvere S, Ng S, Adams J, Allemandi L, Bahena-Espina L, et al. Global benchmarking of children's exposure to television advertising of unhealthy foods and beverages across 22 countries. *Obes Rev.* (2019) 20(Suppl 2):116–28.doi: 10.1111/obr.12840
26. Janssen M, Chang BPI, Hristov H, Pravst I, Profeta A, Millard J. Changes in food consumption during the COVID-19 pandemic: analysis of consumer survey data from the first lockdown period in Denmark, Germany, and Slovenia. *Front Nutr.* (2021) 8:635859. doi: 10.3389/fnut.2021.635859
27. Harrington RA, Adhikari V, Rayner M, Scarborough P. Nutrient composition databases in the age of big data: food DB, a comprehensive, real-time database infrastructure. *BMJ open.* (2019) 9:e026652. doi: 10.1136/bmjopen-2018-026652
28. Estellés-Arolas E, González-Ladrón-de-Guevara F. Towards an integrated crowdsourcing definition. *Journal of Information Science.* (2012) 38:189–200. doi: 10.1177/0165551512437638
29. Wazny K. Applications of crowdsourcing in health: an overview. *J Glob Health.* (2018) 8:010502. doi: 10.7189/jogh.08.010502
30. Soon JM, Saguy IS. Crowdsourcing: A new conceptual view for food safety and quality. *Trends in Food Science & Technology.* (2017) 66:63–72. doi: 10.1016/j.tifs.2017.05.013
31. OFF. Open Food Facts: the free food products database (2021). [25.8.2021]. Available online at: https://world.openfoodfacts.org/ (accessed August 2, 2021).
32. OFR. The Open Food Repo: A community-driven open database for barcoded food products (2021). Available online at: https://www.foodrepo.org/ (accessed September 27, 2021).
33. Szabo de Edelenyi F, Egnell M, Galan P, Hercberg S, Julia C. Ability of the front-of-pack nutrition label Nutri-Score to discriminate nutritional quality of food products in 7 European countries (Spain, Switzerland, Belgium, Italy, UK, Netherlands and Sweden) and consistency with nutritional recommendations. Fecha de acceso: 20 de agosto de (2019).
34. Szabo de Edelenyi F, Egnell M, Galan P, Druesne-Pecollo N, Hercberg S, Julia C. Ability of the Nutri-Score front-of-pack nutrition label to discriminate the nutritional quality of foods in the German food market and consistency with nutritional recommendations. *Archives of public health = Archives belges de sante publique.* (2019) 77:28. doi: 10.1186/s13690-019-0357-x
35. Dréano-Trécant L, Egnell M, Hercberg S, Galan P, Soudon J, Fialon M, et al. Performance of the front-of-pack nutrition label nutri-score to discriminate the nutritional quality of foods products: a comparative study across 8 European countries. *Nutrients.* (2020) 12:1303. doi: 10.3390/nu12051303
36. Dunford E, Trevena H, Goodsell C, Ng HK, Webster J, Millis A, et al. FoodSwitch: A Mobile Phone App to Enable Consumers to Make Healthier Food Choices and Crowdsourcing of National Food Composition Data. *JMIR mHealth uHealth.* (2014) 2:e37. doi: 10.2196/mhealth.3230
37. Crino M, Sacks G, Dunford E, Trieu K, Webster J, Vandevijvere S, et al. Measuring the healthiness of the packaged food supply in Australia. *Nutrients.* (2018) 10:702. doi: 10.3390/nu10060702
38. Dunford EK, Neal B. FoodSwitch and use of crowdsourcing to inform nutrient databases. *Journal of Food Composition and Analysis.* (2017) 64:13–17. doi: 10.1016/j.jfca.2017.07.022
39. FNS-Cloud. Food Nutrition Security (FNS) Cloud: Project overview (2021). Available online at: https://www.fns-cloud.eu/overview/ (accessed October 11, 2021).
40. Lalor F, Kennedy J, Flynn MA, Wall PG. A study of nutrition and health claims–a snapshot of what's on the Irish market. *Public Health Nutr.* (2010) 13:704–11. doi: 10.1017/S1368980009991613
41. Pravst I, Kušar A. Consumers' exposure to nutrition and health claims on pre-packed foods: Use of sales weighting for assessing the food supply in Slovenia. *Nutrients.* (2015) 7:9353–68. doi: 10.3390/nu7115474
42. Korosec Ž, Pravst I. Assessing the average sodium content of prepacked foods with nutrition declarations: the importance of sales data. *Nutrients.* (2014) 6:3501–15. doi: 10.3390/nu6093501
43. Hieke S, Kuljanic N, Wills JM, Pravst I, Kaur A, Raats MM, et al. The role of health-related claims and health-related symbols in consumer behaviour: design and conceptual framework of the CLYMBOL project and initial results. *Nutr Bull.* (2015) 40:66–72. doi: 10.1111/nbu.12128
44. Hieke S, Kuljanic N, Pravst I, Miklavec K, Kaur A, Brown KA, et al. Prevalence of nutrition and health-related claims on pre-packaged foods: A five-country study in Europe. *Nutrients.* (2016) 8:1–16. doi: 10.3390/nu8030137
45. NUTRIS. Nutrition and Public Health (P3-0395) Research Programme (2021). Available online at: https://nutris.org/en/research-programe-nutrition-and-public-health (accessed October 4, 2021).
46. Hodge A. Ultra-processed food and health. *Public Health Nutr.* (2021) 24:3177–8. doi: 10.1017/S1368980021002585
47. GIGH. *The George Institute for Global Health: FoodSwitch*. (2021). Available online at: https://www.georgeinstitute.org/projects/foodswitch (accessed October 4, 2021).
48. UT. *University of Toronto: FLIP - Food Label Information Program (FLIP)*. (2021). Available online at: https://labbelab.utoronto.ca/projects/the-canadian-food-supply/ (accessed October 4, 2021).
49. Pravst I, Lavriša Ž, Kušar A, Miklavec K, Žmitek K. Changes in average sodium content of prepacked foods in Slovenia during 2011–2015. *Nutrients.* (2017) 9:952. doi: 10.3390/nu9090952
50. Zupanič N, Miklavec K, Kušar A, Žmitek K, Fidler Mis N, Pravst I. Total and free sugar content of pre-packaged foods and non-alcoholic beverages in Slovenia. *Nutrients.* (2018) 10:151. doi: 10.3390/nu10020151
51. Hafner E, Hribar M, Hristov H, Kušar A, Žmitek K, Roe M, et al. Trends in the use of low and no-calorie sweeteners in non-alcoholic beverages in Slovenia. *Foods.* (2021) 10:387. doi: 10.3390/foods10020387
52. IJS. *Proceedings of the 22nd International Multiconference Information Society - Volume A*. (2019). Available online at: http://library.ijs.si/Stacks/Proceedings/InformationSociety/2019/IS2019_Complete.pdf (accessed October 20, 2021).
53. Eržen N, Rayner M, Pravst I. A comparative evaluation of the use of a food composition database and nutrition declarations for nutrient profiling. *Journal of Food and Nutrition Research.* (2015) 54:93–100.
54. Zupanič N, Hribar M, Pivk Kupirovič U, Kušar A, Žmitek K, Pravst I. Limiting trans fats in foods: use of partially hydrogenated vegetable oils in prepacked foods in Slovenia. *Nutrients.* (2018) 10:355. doi: 10.3390/nu10030355
55. Zupanic N, Hribar M, Mis NF, Pravst I. Free sugar content in pre-packaged products: does voluntary product reformulation work in practice? *Nutrients.* (2019) 11:2577. doi: 10.3390/nu11112577
56. Krušič S, Hribar M, Havfner E, Žmitek K, Pravst I. Use of branded food composition databases for the exploitation of food fortification practices: a case study on vitamin D in the Slovenian food supply. *Front. Nutr.* (2021) 8:775163. doi: 10.3389/fnut.2021.775163
57. Hafner E, Pravst I. The sharp rise in the use of low- and no-calorie sweeteners in non-alcoholic beverages in Slovenia: an update based on 2020 data. *Frontiers In Nutrition.* (2021) 8:9. doi: 10.3389/fnut.2021.778178
58. Blaznik U, Krušič S, Hribar M, Kušar A, Žmitek K, Pravst I. Use of Food Additive Titanium Dioxide (E171) before the Introduction of Regulatory Restrictions Due to Concern for Genotoxicity. *Foods.* (2021) 10:1910. doi: 10.3390/foods10081910
59. Pivk Kupirovič U, Miklavec K, Hribar M, Kušar A, Žmitek K, Pravst I. Nutrient profiling is needed to improve the nutritional quality of the foods labelled with health-related claims. *Nutrients.* (2019) 11:287. doi: 10.3390/nu11020287
60. Hafner E, Pravst I. Evaluation of the ability of nutri-score to discriminate the nutritional quality of prepacked foods using a sale-weighting approach. *Foods.* (2021) 10:1689. doi: 10.3390/foods10081689
61. Miklavec K, Hribar M, Kusar A, Pravst I. Heart images on food labels: a health claim or not? *Foods.* (2021) 10. doi: 10.3390/foods10030643
62. Lavriša Ž, Pravst I. Marketing of foods to children through food packaging is almost exclusively linked to

unhealthy foods. *Nutrients.* (2019) 11:1128. doi: 10.3390/nu11051128
63. Lavriša Ž, Hribar M, Kušar A, Žmitek K, Pravst I. Nutritional composition of gluten-free labelled foods in the slovenian food supply. *Int J Env Res Public Health.* (2020) 17:8239. doi: 10.3390/ijerph17218239
64. Kusar A, Hribar M, Lavrisa Z, Zupanic N, Kupirovic UP, Hristov H, et al. Assessment of trans-fatty acid content in a sample of foods from the Slovenian food supply using a sales-weighting approach. *Public Health Nutr.* (2021) 24:12–21. doi: 10.1017/S1368980020001949
65. Mencin M, Abramovic H, Zlatic E, Demsar L, Piskernik S, Schreiner M, et al. Content of trans-fatty acid isomers in bakery products on the Slovenian market. *Lwt-Food Science and Technology.* (2021) 143. doi: 10.1016/j.lwt.2021.111095
66. Zmitek K, Pravst I. Iodisation of salt in Slovenia: increased availability of non-iodised salt in the food supply. *Nutrients.* (2016) 8. doi: 10.3390/nu8070434
67. Hribar M, Hristov H, Lavriša Ž, Koroušić Seljak B, Gregorič M, Blaznik U, et al. Vitamin D intake in Slovenian adolescents, adults, and the elderly population. *Nutrients.* (2021) 13:1–19. doi: 10.3390/nu13103528
68. Zupanič N, Hribar M, Hristov H, Lavriša Ž, Kušar A, Gregorič M, et al. Dietary intake of trans fatty acids in the Slovenian population. *Nutrients.* (2021) 13:207. doi: 10.3390/nu13010207
69. Zupanič N, Hristov H, Gregorič M, Blaznik U, Delfar N, Koroušić Seljak B, et al. Total and free sugars consumption in a Slovenian population representative sample. *Nutrients.* (2020) 12:1729. doi: 10.3390/nu12061729
70. Koroušić Seljak B, Valenčič E, Hristov H, Hribar M, Lavriša Ž, Kušar A, et al. Inadequate intake of dietary fibre in adolescents, adults, and elderlies: results of Slovenian representative SI. *Menu Study Nutrients.* (2021) 13:3826. doi: 10.3390/nu13113826
71. Pravst I, Lavriša Ž, Hribar M, Hristov H, Kvarantan N, Seljak BK, et al. Dietary intake of folate and assessment of the folate deficiency prevalence in Slovenia using serum biomarkers. *Nutrients.* (2021) 13:3860. doi: 10.3390/nu13113860
72. Korosec Z, Pravst I. Television food advertising to children in Slovenia: analyses using a large 12-month advertising dataset. *Int J Public Health.* (2016) 61:1049–57. doi: 10.1007/s00038-016-0896-7
73. Lavrisa Z, Erjavec K, Pravst I. Trends in marketing foods to children in Slovenian magazines: a content analysis. *Public Health Nutr.* (2018) 21:3344–53. doi: 10.1017/S1368980018002513
74. Lavrisa Z, Hristov H, Kelly B, Pravst I. Regulating children's exposure to food marketing on television: are the restrictions during children's programmes enough? *Appetite.* (2020) 154. doi: 10.1016/j.appet.2020.104752
75. Rozman U, Pravst I, Kupirovič UP, Blaznik U, Kocbek P, Turk SŠ. Sweet, fat and salty: snacks in vending machines in health and social care institutions in Slovenia. *Int J Env Res Public Health.* (2020) 17:7059. doi: 10.3390/ijerph17197059
76. Rozman U, Mis NF, Kupirovic UP, Pravst I, Kocbek P, Strauss M, et al. Nutritional quality of beverages available in vending machines in health and social care institutions: do we really want such offers? *Journal of Health Population and Nutrition.* (2021) 40. doi: 10.1186/s41043-021-00250-1
77. Kaur A, Scarborough P, Hieke S, Kusar A, Pravst I, Raats M, et al. The nutritional quality of foods carrying health-related claims in Germany, The Netherlands, Spain, Slovenia and the United Kingdom. *Eur J Clin Nutr.* (2016) 70:1388–95. doi: 10.1038/ejcn.2016.114
78. Pravst I, Kusar A, Miklavec K, Hieke S, Raats M, Rayner M. Use of nutrition and health claims and symbols on prepacked foods in Europe: from consumer exposure to public health implications. In: Astley S, editor. *Health Claims and Food Labelling. Food Chemistry, Function and Analysis*. p. 79–93.
79. Debeljak K, Pravst I, Košmelj K, Kač M. "Healthier" and "less heathy" classifications according to three nutrient profiling systems relative to nutrition and health claims on food labels. *Acta alimentaria*. (2015). 44:561. doi: 10.1556/066.2015.44.0028
80. Bernstein J, Schermel A, Mills C, L'Abbe M. Total and Free Sugar Content of Canadian Prepackaged Foods and Beverages. *Nutrients*. (2016) 8:582. doi: 10.3390/nu8090582
81. Stender S. *Trans Fatty Acids and Heart Health*. Reference Module in Food Science: Elsevier (2019).
82. Stender S, Astrup A, Dyerberg J. Tracing artificial trans fat in popular foods in Europe: a market basket investigation. *Bmj Open.* (2014) 4. doi: 10.1136/bmjopen-2014-005218
83. European Commission. Commission (EC) No 1333/2008 of the European Parliament and of the Council of 16 December 2008 on food additives. *Official Journal of the European Union*. (2011) 354:16–18. Available online at: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32008R1333 (accessed November, 2021).
84. Dorier M, Béal D, Marie-Desvergne C, Dubosson M, Barreau F, Houdeau E, et al. Continuous *in vitro* exposure of intestinal epithelial cells to E171 food additive causes oxidative stress, inducing oxidation of DNA bases but no endoplasmic reticulum stress. *Nanotoxicology.* (2017) 11:751–61. doi: 10.1080/17435390.2017.1349203
85. Proquin H, Rodríguez-Ibarra C, Moonen CG, Urrutia Ortega IM, Briedé JJ, de Kok TM, et al. Titanium dioxide food additive (E171) induces ROS formation and genotoxicity: contribution of micro and nano-sized fractions. *Mutagenesis.* (2017) 32:139–49. doi: 10.1093/mutage/gew051
86. French Agency for Food Environmental and Occupational Health & Safety. AVIS de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail relatif aux risques liés à l'ingestion de l'additif alimentaire E171. Saisine n° 2019-SA-0036. 2019:44.
87. Netherlands Food and Consumer Product Safety Authority (NVWA). Opinion of BuRO on possible health effects of the food additive titanium dioxide (E171). In: Research OfRA, editor: *Ministry of Agriculture Nature and Food Quality*. (2019). p. 30.
88. EFSA Panel on Food Additives Flavourings, Younes M, Aquilina G, Castle L, Engel K-H, Fowler P, et al. Safety assessment of titanium dioxide (E171) as a food additive. *EFSA Journal.* (2021) 19:e06585. doi: 10.2903/j.efsa.2021.6585
89. Blaznik U, Krušič S, Hribar M, Kušar A, Žmitek K, Pravst I. Use of food additive titanium dioxide (E171) before the introduction of regulatory restrictions due to concern for genotoxicity. *Preprints.* (2021) 10:1910. doi: 10.20944/preprints202106.0483.v1
90. WHO. *Nutrient Profiling: Report of a WHO/IASO Technical Meeting*. London, UK: World Health Organisation (2011).
91. Foltran F, Verduci E, Ghidina M, Campoy C, Any KD, Widhalm K, et al. Nutritional profiles in a public health perspective: a critical review. *J Int Med Res.* (2010) 38:318–85. doi: 10.1177/147323001003800202
92. Drewnowski A, Amanquah D, Gavin-Smith B. Perspective: How to Develop Nutrient Profiling Models Intended for Global Use: A Manual. *Advances in Nutrition.* (2021) 12:609–20. doi: 10.1093/advances/nmab018
93. Garsetti M, de Vries J, Smith M, Amosse A, Rolf-Pedersen N. Nutrient profiling schemes: overview and comparative analysis. *Eur J Nutr.* (2007) 46 Suppl 2:15–28. doi: 10.1007/s00394-007-2002-7
94. Labonte ME, Poon T, Gladanac B, Ahmed M, Franco-Arellano B, Rayner M, et al. Nutrient Profile Models with Applications in Government-Led Nutrition Policies Aimed at Health Promotion and Noncommunicable Disease Prevention: A Systematic Review. *Advances in Nutrition.* (2018) 9:741–88. doi: 10.1093/advances/nmy045
95. WHO. WHO Regional Office for Europe nutrient profile model (2015). Available online at: http://www.euro.who.int/__data/assets/pdf_file/0005/270716/Nutrient-children_web-new.pdf?ua=1 (accessed February 26, 2018).
96. Rayner M, Scarborough P, Lobstein T. The UK Ofcom Nutrient Profiling Model - Defining 'healthy' and 'unhealthy' foods and drinks for TV advertising to children In: Nuffield Department of Population Health, editor. University of Oxford (2009).
97. Dunford EK, Huang L, Peters SAE, Crino M, Neal BC, Ni Mhurchu C. Evaluation of alignment between the health claims nutrient profiling scoring criterion (NPSC) and the health star rating (HSR) nutrient profiling models. *Nutrients.* (2018) 10. doi: 10.3390/nu10081065
98. Julia C, Kesse-Guyot E, Touvier M, Méjean C, Fezeu L, Hercberg S. Application of the British Food Standards Agency nutrient profiling system in a French food composition database. *British*

*Journal of Nutrition*. (2014) 112:1699–705. doi: 10.1017/S0007114514002761

99. Julia C, Fialon M, Galan P, Deschasaux M, Andreeva VA, Kesse-Guyot E, et al. Are foods 'healthy' or 'healthier'? front of pack labelling and the concept of healthiness applied to foods. *British Journal of Nutrition*. (2021) 1–17. doi: 10.1017/S0007114521001458
100. Delhomme V. Front-of-pack nutrition labelling in the European Union: a behavioural, legal and political analysis. *European Journal of Risk Regulation*. (2021) 1–24. doi: 10.1017/err.2021.5
101. Julia C, Hercberg S, World Health Organization. Regional Office for E. Development of a new front-of-pack nutrition label in France: the five-colour Nutri-Score. *Public health panorama*. (2017) 03:712–25. Available online at: https://apps.who.int/iris/handle/10665/325207 (accessed November, 2021).
102. RS. Decree laying down maximum permitted levels of trans fatty acids in foodstuffs. In Slovenian: Pravilnik o največji dovoljeni vsebnosti transmaščobnih kislin v živilih (2018). Available online at: https://www.uradni-list.si/glasilo-uradni-list-rs/vsebina/2018-01-0801/ (accessed August 30, 2019).
103. Guidance on the EU Menu methodology. *EFSA Journal*. (2014) 12. doi: 10.2903/j.efsa.2014.3944
104. Gregorič M, Blaznik U, Delfar N, Zaletel M, Lavtar D, Koroušić Seljak B, et al. Slovenian national food consumption survey in adolescents, adults and elderly. *EFSA Supporting Publications*. (2019) 16. doi: 10.2903/sp.efsa.2019.EN-1729

**Conflict of Interest:** IP has led and participated in various other research projects in the area of nutrition, public health, and food technology, which were (co) funded by the Slovenian Research Agency, Ministry of Health of the Republic of Slovenia, the Ministry of Agriculture, Forestry, and Food of the Republic of Slovenia, and in the case of specific applied research projects, also by food businesses.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

BRIEF RESEARCH REPORT
published: 04 January 2022
doi: 10.3389/fnut.2021.775163

# Use of Branded Food Composition Databases for the Exploitation of Food Fortification Practices: A Case Study on Vitamin D in the Slovenian Food Supply

*Sanja Krušič*[1], *Maša Hribar*[1,2], *Edvina Hafner*[1,2], *Katja Žmitek*[1,3] *and Igor Pravst*[1,2,3]*

[1] Nutrition Institute, Nutrition and Public Health Research Group, Ljubljana, Slovenia, [2] Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia, [3] VIST—Faculty of Applied Sciences, Ljubljana, Slovenia

OPEN ACCESS

**Edited by:**
*Alessandra Durazzo, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Ewa Sicinska, Warsaw University of Life Sciences, Poland*
*Yefim Roth, University of Haifa, Israel*
*Áine Hennessy, University College Cork, Ireland*

***Correspondence:**
*Igor Pravst*
*igor.pravst@nutris.org*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** *13 September 2021*
**Accepted:** *12 November 2021*
**Published:** *04 January 2022*

**Citation:**
*Krušič S, Hribar M, Hafner E, Žmitek K and Pravst I (2022) Use of Branded Food Composition Databases for the Exploitation of Food Fortification Practices: A Case Study on Vitamin D in the Slovenian Food Supply. Front. Nutr. 8:775163. doi: 10.3389/fnut.2021.775163*

Vitamin D deficiency is a worldwide public health concern, which can be addressed with voluntary or mandatory food fortification. The aim of this study was to determine if branded food composition databases can be used to investigate voluntary fortification practices. A case study was conducted using two nationally representative cross-sectional datasets of branded foods in Slovenia, collected in 2017 and 2020, and yearly sales data. Using food labeling data we investigated prevalence of fortification and average vitamin D content, while nutrient profiling was used to investigate overall nutritional quality of the foods. In both datasets, the highest prevalence of vitamin D fortification was observed in meal replacements (78% in 2017; 100% in 2020) and in margarine, corresponding to high market share. Other food categories commonly fortified with vitamin D are breakfast cereals (5% in 2017; 6% in 2020), yogurts and their imitates (5% in 2017; 4% in 2020), and baby foods (18% in both years). The highest declared average content of vitamin D was observed in margarine and foods for specific dietary use (7–8 μg/100g), followed by breakfast cereals (4 μg/100g), while the average content in other foods was below 2 μg/100g. Only minor differences were observed between 2017 and 2020. Major food-category differences were also observed in comparison of the overall nutritional quality of the fortified foods; higher overall nutritional quality was only observed in fortified margarine. Our study showed that branded food composition databases are extremely useful resources for the investigation and monitoring of fortification practices, particularly if sales data can also be used. In the absence of mandatory or recommended fortification in Slovenia, very few manufacturers decide to add vitamin D, and even when this is the case, such products are commonly niche foods with lower market shares. We observed exceptions in imported foods, which can be subject to fortification policies introduced in other countries.

**Keywords: vitamin D, food fortification, fortification, food supply, Europe, Slovenia**

## INTRODUCTION

Vitamin D (VitD) is a fat-soluble vitamin family usually encompassing ergocalciferol (D2) and cholecalciferol (D3) (1, 2). It is a pro-hormone with a well-established role in musculoskeletal health and other functions (3–5). Discussion about the importance of sufficient VitD status during the COVID-19 epidemic should also be mentioned (6–9), but consensus has not yet been reached on this topic.

The chemical structure of VitD was described in the early years of the last century. After small quantities were first found in butterfat and cod liver oil (10), it was later observed that it can be biosynthesized after sun exposure (11). It is now well established that the exposure of skin to sufficiently intensive ultraviolet B (UVB) sunlight is a key source of VitD for humans (12); UVB photons enter the skin and photolyze 7-dehydrocholesterol into previtamin D3, which is then transformed into vitamin D3 (13, 14). However, the cutaneous biosynthesis is affected by various personal and environmental factors (15–19), and when it is insufficient to ensure adequate VitD status, dietary intake of VitD becomes of major importance. Unfortunately, both sources combined are commonly not enough, making VitD deficiency one of the most frequent micronutrient deficiencies globally (20). The situation in Slovenia, a mid-latitude European country (45° and 46°N), is also far from perfect; considerable seasonal variations in VitD status were reported; with about 80% of the adult population having insufficient VitD status, and about 40% VitD deficiency during winter time (21).

The poor VitD status can be improved with increased dietary intake of VitD. We can distinguish three dietary sources of this vitamin: (a) VitD naturally present in foods, such as cod liver oil and oily fish, such as sardines, mackerel, and salmon (22, 23) (however, very few foods are good VitD sources); (b) medicines or dietary supplements that contain VitD; (c) foods that are enriched or fortified with VitD. While some researchers carefully distinguish between food enrichment and fortification, based on the food matrix and/or purpose (24), herein we only use the term "fortification". According to the World Health Organization (WHO), food fortification is a "practice of deliberately increasing the content of an essential micronutrient (i.e., vitamins) in a food to improve the nutritional quality of the food supply and provide a public health benefit with minimal health risk" (25). In general, VitD content in foods can be increased with different strategies. Most commonly used fortification approach is simply adding VitD to processed foods. On one hand the nutritional quality of food crops can be improved through plant breeding, agronomic practices, or modern biotechnology (26–30). Examples of these approaches are feeding hens with VitD to increase its content in eggs, likewise with livestock animals in relation to meat, and UV exposure of mushrooms or yeast (29, 31). In Europe, UV-treated yeast, which is a good source of vitamin D2, was approved as a novel food in various food categories (32, 33).

In the European Union (EU), the addition of micronutrients to foods was harmonized in 2006 with the adoption of Regulation 1925/2006 on the addition of vitamins, minerals, and other certain substances to foods (34, 35). The legislation enables the adding of micronutrients to foods in cases of deficiency in the population, possibly improving the nutritional status in the population or certain groups of the population, or if their use is scientifically supported. Fortification is not allowed for unprocessed foods (fruit, vegetables, meat, fish, etc.) and alcoholic beverages. The legislation also defines the chemical substances that can be used for fortification. In the case of VitD, both cholecalciferol and ergocalciferol can be used (36).

To efficiently improve VitD intake in the general population with the fortification of foods, several issues should be considered. Fortified foods must contain a sufficient amount of VitD, they should be consumed by the majority of the population, and they must meet specific standards for bioavailability, storage stability, and cooking conditions. They should also not exceed the amount of VitD that could have adverse effects (27, 37). The United States of America (USA) and some other industrialized countries, such as Great Britain, introduced VitD fortification back in the 1930s and 1940s (26). Because early efforts were mainly focused on the prevention of rickets in children, cow's milk was first selected as a fortification matrix. This was followed by other foods, including dairy products, margarine, hot dogs, peanut butter, and others (38). In the Great Britain in the 1960s and in the USA in the 1980s, uncontrolled or accidental high intakes of VitD through fortified products occurred. Poor monitoring and large excesses of fortification of some dairy products led to hypercalcemia in some cases, and as a result, fortification of dairy products with VitD was then banned in some countries (39–41).

Although voluntarily addition of vitamins to foods is harmonized across the EU (36), fortification practices vary between countries. In most countries, including Slovenia, there are no requirements for mandatory fortification and no formal recommendations on this topic. However, some countries adapted their national policies or guidelines. In Finland for example, the voluntary fortification of foods was recommended in 2003, when the government encouraged the addition of VitD to margarine/fat spreads (10 μg/100 g of food) and fluid milk products (0.5 μg/100ml of milk) (35, 42, 43). The recommended fortification content doubled in 2010 (44). Although this was a voluntary option, it was followed by most manufacturers, and a notable reduction in VitD deficiency was observed in the population (42, 45).

Foods fortified with vitamins are commonly labeled with various nutrition and health claims (46), which can be very attractive for consumers (47). Previous research highlighted issues related with the overall nutritional quality of such foods, which can be high in energy, fat, sugar or salt content (48–50), because this area is still not regulated in the EU (51). Similar might apply for the foods fortified with specific vitamins, but this area has not yet been investigated.

Considering the high prevalence of VitD deficiency in Slovenia (52), the government is searching for the most feasible policy approaches to address this public health problem. While the COVID-19 epidemic has resulted in notable changes in dietary behaviors (53), and also increased the supplementation of VitD in the general population (54), a feasible long-term solution is needed to ensure an optimal VitD status in the general population. A possible route forward is the introduction of national VitD fortification guidelines, but data about existing fortification practices is needed before the introduction of an evidence-based policy decision.

With thousands of different foods in the food supply, the investigation of voluntary fortification practices is very challenging. Branded food composition databases have been shown to be excellent resource for investigating nutritional quality and the content of certain nutrients and additives in the food supply (55–57). The objective of this case study was, therefore, to explore the possibility of the use of the nationally representative branded food composition database to investigate VitD supplementation practices. Our goals were to identify food categories with added VitD and to determine the prevalence of fortification and the typical amount of added VitD in those categories. Nutrient profiling was used to investigate the overall nutritional quality (healthiness) of the fortified foods, in comparison with non-fortified foods. Study utilized two cross-sectional datasets (from the years 2017 and 2020), compiled within the national research program "Nutrition and Public Health" and the EC-funded "Food Nutrition Security Cloud" project (FNS-Cloud; www.fns-cloud.eu). Market-share differences were addressed with the use of nation-wide 12-month sales data, provided by major food retailers. Fortification practices were investigated using both datasets, while nutrient profiling was conducted on the 2017 dataset.

# MATERIALS AND METHODS

## Data Collection and Categorization

The source of branded food composition data was the Slovenian Composition and Labeling Information System (CLAS). CLAS is an online tool for monitoring the supply of the prepacked foods in Slovenia, maintained by the Nutrition Institute (Ljubljana, Slovenia) (58). The tool was developed within the national research program "Nutrition and Public Health", funded by a Slovenian research agency. In Slovenia, this tool was first tested for a food supply study in 2015 on specific food categories (55, 59), while the first complete food supply study (on all categories of prepacked foods) was conducted in year 2017 (56, 60), and repeated in year 2020. Within this monitoring program, food labels of prepacked foods in the Slovenian food supply are photographed. In the CLAS tool, the data about the nutritional composition and food ingredients are extracted. To ensure representativeness, the data was collected from shops of all the major retailers with a nationwide market. In 2017, five retailers were included (Mercator, Spar, Tuš, Hofer, and Lidl), while in 2020, we additionally included the retailer Eurospin. The data collections were conducted in Ljubljana (Slovenia).

For this case study, we utilized food composition datasets compiled in years 2017 and 2020. EAN (EAN; European Article Number) barcode numbers were used as unique food identifiers. This approach enabled us to avoid duplicate entries of foods, which are available in shops of different retailers. The data collection approach was previously described in detail (56, 60). In short, the collected information included the product's name, the list of the ingredients, nutritional values, packaging volume, price, and barcode number. Each product was assigned to one of the 16 parent categories and 57 categories using a classification system developed within the Global Food Monitoring Initiative (GFMI) (61). The food categorization system was nationally adapted with additional sub-categories on the third level (57).

While the CLAS datasets include all the prepacked foods available in the selected grocery stores at the time of sampling, this case study only included food categories in which we found products fortified with VitD. Altogether, the CLAS contained 21,090 products in 2017, and 28,028 products in year 2020. VitD fortification was identified with the food ingredients lists as provided by the manufacturers on the food labels. All items in the dataset in which the ingredient list wording contained relevant VitD-related terms ("vitamin D," "cholecalciferol," "calciferol," "ergocalciferol," "D2," or "D3") were manually checked by a researcher to identify products fortified with VitD. After exclusion of food supplements, the following food categories included products with added VitD in either the 2017 or 2020 dataset: beverages; bread and bakery products; cereal and cereal products; confectionery; dairy and imitates; edible oils and oil emulsions; foods for special dietary use; sauces and spreads. All further analyses were done on these food categories. Number of foods in the datasets containing these food categories were 13,393 in 2017, and 16,064 in year 2020. The content of VitD was taken from the nutrition declaration on the label and corresponded to the composition of food as sold (without further preparation). We did not conduct laboratory analyses of foods.

Reformulation practices were investigated on a sub-sample of products, which were found in both the 2017 and 2020 datasets. Matching was conducted using barcode numbers. If the 2020 dataset contained a similar product with a different barcode number, it was not included into comparison. In the original datasets with all food categories, $N = 10{,}034$ foods (47.6% of the 2017 sample; 35.8% of the 2020 sample) were found in both years. In the dataset with only the above-mentioned food categories with VitD added to the foods, $N = 6{,}534$ foods were found in both years (48.8% of the 2017 sample and 40.7% of the 2020 sample).

Previously described sale-weighting approach (57) was used to account for different market-shares of different foods. This approach was applied to all items, for which 12-month national sales data was available. The 12-month volume sales data were provided by the major retailers, covering a majority of the national food retail market. These sales data were connected with the datasets using barcode numbers. In the above-mentioned selected food categories, the sales data were available for 9,258 (69.1%) products in 2017, and for 10,923 (68.0%) in 2020 dataset.

Evaluation of the overall nutritional quality of foods was investigated using two nutrient profiling models, namely the WHO Europe (WHOE) profile (62), and the Nutri-Score (NS) (63). WHOE was developed to restrict the advertising of unhealthy foods to children, while NS is a scheme for front-of-package nutrition labeling, which assigns foods into five grades A-B-C-D-E; dark green grade A is assigned to products with highest overall nutrition quality, and dark orange grade E to those with the lowest quality. This scheme is implemented for voluntarily use in some European countries (France, Germany, Belgium, Luxembourg, Netherlands, Switzerland and Spain) (64). We compared nutrient profiles of VitD fortified foods with

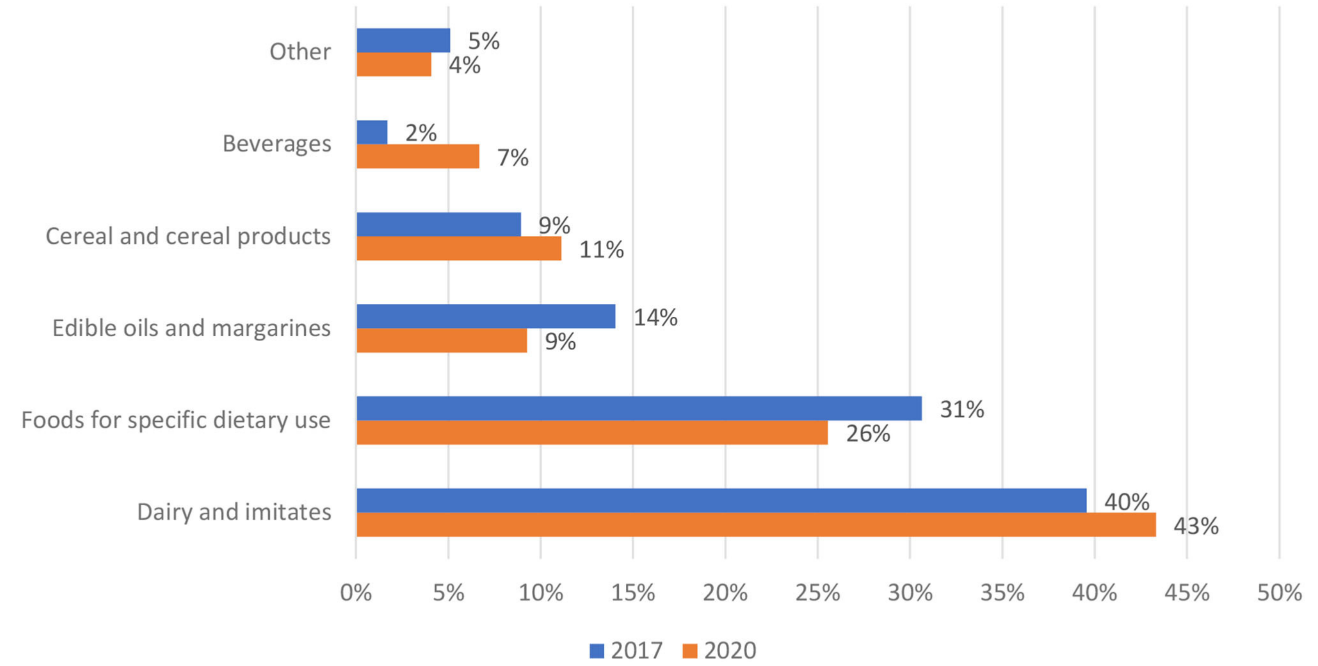


**FIGURE 1 |** Distribution of prepacked foods fortified with vitamin D per food (sub)categories in years 2017 ($N = 235$) and 2020 ($N = 270$) (Slovenia).

those that were not fortified. Nutrient profiling methodology is presented in details elsewhere (65). In short, nutrient profile was determined with consideration of the nutritional composition of food, standardized to 100 g. In addition to the content of energy, total sugars, saturated fats, salt, dietary fiber and proteins, the percentage of fruits/vegetables/pulses/nuts/specific oils was also considered. Information provided on food labels were used for calculations (65, 66). Foods passing WHOE criteria (permitted for marketing) were assigned as having higher nutritional quality. Considering that both used nutrient profiling models are category specific, comparisons were done within (sub)categories. Analyses was done using 2017 dataset; out of the 13,393 products in the original dataset, $N = 2,327$ foods were excluded: 828 due to missing food composition information that is needed for profiling, and 1,499 for other reasons, i.e., food category not covered in the nutrient profiling algorithms for example foods for specific dietary use, which are excluded form algorithms because of specific nutritional needs of children) (67).

## Data Processing and Statistical Analyses

The food composition data were processed using Microsoft Analysis Services Client Tools 13.0, Microsoft SQL Server Management Studio 13.0, Microsoft Excel 2019 (Microsoft, Redmond, Washington, DC, USA), Microsoft Data Access Components (MDAC) 10.0, and the CLAS (Nutrition Institute, Ljubljana, Slovenia). Data processing was performed using Microsoft Excel 2019 (Microsoft, Redmond, Washington, DC, USA).

For statistical evaluation, we calculated sale-weighted and non-weighted proportions of foods fortified with VitD in different food categories, and the average VitD contents.

For the proportions of food fortified with VitD, we used a descriptive analysis and Wilson score interval (68). A 95% confidence interval (95% CI) is provided. A two-tailed $z$-test was used to identify increases in proportions in VitD fortification in 2020, compared to year 2017; negative value present lower proportion in 2017. Same statistical test was also used to compare within-category proportions of VitD fortified foods that pass WHOE nutrient profile in the 2017 dataset, in comparison with non-fortified foods. The level of significance was set at $p < 0.05$.

# RESULTS AND DISCUSSION

Among 21,090 foods in the year 2017, VitD was added to 235 foods. In the 2020 dataset, VitD fortification was found in 270 out of 28,028 foods. All further results are reported for reduced datasets of food categories in which VitD fortification was observed ($N = 13,393$ and 16,064, for years 2017 and 2020, respectively).

The distribution of prepacked foods fortified with VitD in 2017 and 2020 is presented in **Figure 1**. In both years, the highest proportion of fortified foods was within dairy and imitates, which represent about two-fifths of the sample (40% in 2017; 42% in 2020), followed by foods for specific dietary use (for example, meal replacements, weaning foods, and infant formulas)-—31% in 2017 and 26% in 2020. In 2017, additional food subcategories with a notable proportion of fortification were edible oils and margarines (14%), cereal and cereal products (9%) and beverages

**TABLE 1 |** (Sub)category proportions of foods fortified with vitamin D in the food supply for 2017 and 2020 (Slovenia).

| Food category | 2017 | | | | | 2020 | | | | | Two sample *z*-test for proportions |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *N* | VitD Cont. *N* | % (95% CI) | Sale-weighted proportion (%) | Average VitD Cont. (µg per 100 g or ml) | *N* | VitD cont. *N* | % (95% CI) | Sale-weighted proportion (%) | Average VitD cont. (µg per 100 g or ml) | *p*-value |
| Beverages | 2,454 | 4 | 0.2 (0.0; 0.3) | 0.03 | 1.6 | 3,385 | 18 | 0.5 (0.3; 0.8) | 0.3 | 1.4 | 0.02 |
| Bread and bakery products | 2,105 | 8 | 0.4 (0.2; 0.8) | 0.1 | 2.9 | 2,389 | 5 | 0.2 (0.0; 0.4) | / | * | ns |
| Cereal and cereal products | 1,854 | 21 | 1.1 (0.7; 1.7) | 1.0 | 4.3 | 2,196 | 30 | 1.4 (0.9; 1.9) | 0.8 | 4.2 | ns |
| – Breakfast cereals | 375 | 18 | 4.8 (2.6; 7.0) | 11.8 | 4.1 | 487 | 29 | 6.0 (3.9; 8.1) | 9.9 | 4.3 | ns |
| Confectionery | 2,213 | 2 | 0.1 (-0.0; 0.2) | 0.01 | * | 2,470 | 3 | 0.1 (0.0; 0.3) | / | 4.2 | ns |
| Dairy and imitates | 2,843 | 93 | 3.3 (2.6; 4.0) | 1.5 | 2.3 | 3,459 | 117 | 3.4 (2.8; 4.0) | 2.2 | 1.9 | ns |
| – Yogurt products | 806 | 39 | 4.8 (3.4; 6.3) | 4.4 | 1.1 | 911 | 39 | 4.3 (3.0; 5.6) | 5.4 | 1.3 | ns |
| – Flavored yogurt | 419 | 10 | 2.4 (0.9; 3.8) | 1.8 | 1.2 | 386 | 8 | 2.1 (0.7; 3.5) | 4.5 | 1.3 | ns |
| – Flavored y. drinks | 119 | 22 | 18.5 (11.5; 25.5) | 23.4 | 0.9 | 199 | 22 | 11.2 (6.7; 15.4) | 21.0 | 1.0 | 0.03 |
| – Plain yogurt | 235 | 0 | / | / | / | 284 | 1 | 0.7 (−0.3; 1.6) | / | 0.8 | ns |
| – Yogurt imitates | 33 | 7 | 21.2 (7.3; 35.2) | 31.2 | 1.5 | 42 | 8 | 19.1 (7.2; 30.9) | 33.3 | 1.1 | ns |
| – Milk | 114 | 4 | 3.5 (0.1; 6.9) | 0.4 | 0.8 | 155 | 7 | 4.5 (1.3; 7.8) | 0.3 | 0.9 | ns |
| – Milk imitates | 150 | 28 | 18.7 (12.4; 24.9) | 10.6 | 1.0 | 185 | 42 | 22.7 (16.7; 28.7) | 31.4 | 0.9 | ns |
| Edible oils; oil emulsions | 550 | 33 | 6.0 (4.2; 8.0) | 8.6 | 7.1 | 617 | 26 | 4.2 (2.6; 5.8) | 6.2 | 8.0 | ns |
| – Margarines | 62 | 33 | 53.2 (40.8; 65.7) | 70.5 | 7.1 | 53 | 25 | 47.2 (33.7; 60.6) | 65.8 | 8.1 | ns |
| Foods for spec. dietary use | 281 | 72 | 25.6 (16.6; 26.1) | 9.8 | 6.1 | 251 | 69 | 27.5 (22.0; 33.0) | 10.9 | 6.8 | ns |
| – Baby foods | 244 | 43 | 17.6 (12.8; 23.0) | 8.8 | 6.7 | 222 | 40 | 18.0(13.0; 23.1) | 10.5 | 7.7 | ns |
| – Infant formula | 15 | 15 | 100 | 100 | 9.3 | 20 | 20 | 100 | 100 | 13.3 | ns |
| – Weaning foods | 193 | 28 | 14.5 (9.5; 19.5) | 4.4 | 6.4 | 180 | 20 | 11.1 (6.5; 15.7) | 4.3 | 5.7 | ns |
| – Meal replacements | 37 | 29 | 78.4 (65.1; 91.6) | 58.0 | 5.5 | 29 | 29 | 100 | 100 | 5.9 | <0.01 |
| Sauces and spreads | 1,093 | 2 | 0.2 (-0.1; 0.4) | 0.03 | 4.5 | 1,298 | 3 | 0.2 (0.0; 0.8) | 0.5 | 1.5 | ns |

*VitD, Vitamin D; VitD cont. N, number of vitamin D-containing food products; Data presented for food categories with at least one product with VitD in either the 2017 or 2020 dataset. 95% CI-95% confidence interval; N, number of all products; ns, not significant; na, not applicable; *data for VitD content was not available (VitD was declared only in the ingredients); / data not available.*

(2%). In 2020, a somewhat higher proportion was observed in cereal and cereal products (11%) and beverages (7%), and lower in edible oils and margarines (9%). The remaining categories, each with less than a 3% share, are presented as combined (**Figure 1**, "other"); bread and bakery products (3.4% in 2017; 1.9% in 2020), sauces and spreads (0.9% in 2017; 1.1 in 2020) and confectionery (0.9% in 2017; 1.1 in 2020).

While each dataset presents cross-sectional data for a specific time period, a comparison of the same foods in both datasets provided very interesting insights into food reformulation practices. For this purpose, we matched products using the EAN barcode number as the product identifier. Focusing on the food categories in which VitD fortifications were observed, 6,534 foods were found in the datasets for both observation years, and 118 of those were fortified with VitD. In 109 (87.9%) of those products, we did not observe any changes in the VitD fortification practice in 2020, meaning that the content of added VitD did not change in these products in the observation period. However, in eight products (one from cereals, five from cakes, muffins, and pastry, and two from the chocolate and sweets category), the producers discontinued VitD fortification by 2020. We only found one product that did not have VitD fortification in 2017, but the manufacturer started adding VitD by 2020. These results do not indicate reformulation of foods by adding VitD to the existing formulations. Our data revealed that a higher number of VitD-fortified foods in 2020 corresponded with additional foods in the 2020 sample, that were not present in the 2017 dataset.

The penetration of VitD fortification in specific food (sub)categories is presented in **Table 1**. In both years, the largest share of VitD-fortified foods per category was represented by meal replacements (78% in 2017; 100% in 2020) and margarine, comprising about half of the sample (53% in 2017; 47% in 2020), followed by milk imitates (19% in 2017; 23% in 2020), yogurt imitates (21% in 2017; 19% in 2020), baby foods (18% in both years), flavored yogurt drinks (19% in 2017; 11% in 2020) and weaning foods (15% in 2017;11% in 2020). Except for breakfast cereals (5% in 2017; 6% in 2020), the share for other categories was <5%. It should be noted that the difference between both observation years was significant only in three food (sub)categories. Flavored yogurt drinks had lower proportions of VitD-fortified products in 2020, while the opposite was observed in beverages and meal replacements. In the latter case, beverages with added VitD were mostly soft drinks and energy drinks.

Considering that different prepacked foods have very different market shares, we calculate the per-category sale-weighted proportions of VitD fortification (**Table 1**). In this analyses we used 12-months sales data, provided by largest food retailers in Slovenia, covering the majority of the national market. It should be noted that the sales data were available for the majority of foods in both datasets (69% in 2017, and 68% in 2020), and for even more foods fortified with VitD (81% in 2017 and 82% in 2020). A general observation was that in most food categories with low proportions of VitD-fortified foods (<5%), the sale-weighted proportion of VitD fortification is even lower, while the opposite was observed in categories with higher proportions

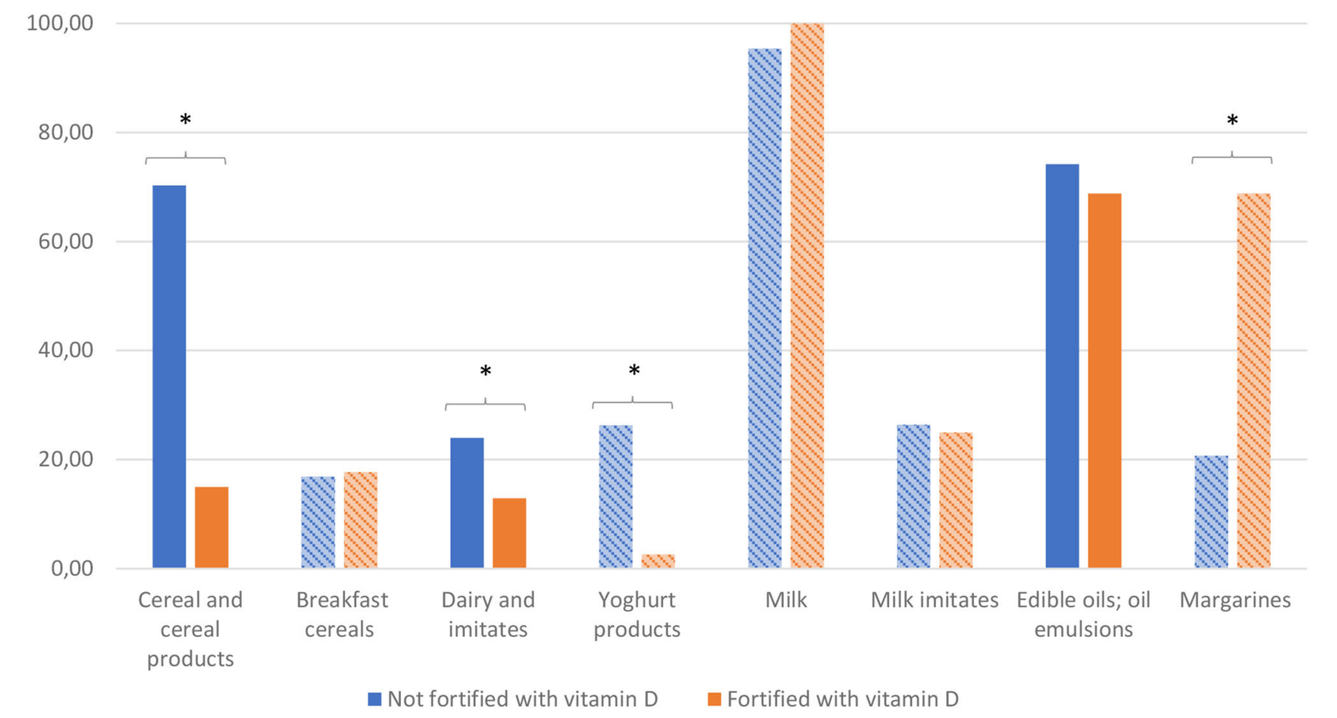


**FIGURE 2 |** (Sub)category proportions (%) of (healthier) foods passing WHOE nutrient profile model for samples of foods (1) which are not fortified, or (2) are fortified with vitamin D. *Significant differences. Subcategories are shown in bars using diagonal lines pattern. Profiling using the WHO Europe (WHOE) nutrient profile (62).

of fortification. The exceptions are foods for specific dietary use (where 25.6% were fortified in 2017 and 27.5% in 2020, corresponding to 9.8 and 10.9% of the market share, respectively), and milk imitates (in 2017, 18.7% of available milk imitates were fortified with VitD, but these only presented 10.6% of the market share). Interestingly, the situation changed notably in 2020, when the market share of VitD-fortified milk imitates increased to 31.4%. We should also mention some food (sub)categories, where sale-weighting resulted in very interesting observations. In both observation years, the market share of fortified breakfast cereals was almost double (11.8 and 9.9% in 2017 and 2020, respectively), in comparison with the proportion of such food on the market. Notable sale-weighing effects were also observed in fortified margarine (which presented 70.5 and 65.8% of the market share, respectively), yogurt imitates (31.3 and 33.3%, respectively), and flavored yogurt drinks (23.4 and 21.0%, respectively).

We also investigated the VitD content in fortified foods (**Table 1**). On average, more VitD was added in the categories of margarines and edible oils (7–8 µg/100g), foods for specific dietary use (6–7 µg/100g)—especially in infant formula (9–13 µg/100g), and breakfast cereals (4 µg/100g). It should be noted that some of these foods are typically consumed in smaller portions, <100 g per day. In foods that are typically consumed in portions above 100 g, lower levels of VitD were observed. For example, around 1 µg/100g in yogurts and imitates, and milk and imitates. The average VitD content in fortified beverages was 1.6 µg/100g in 2017, and 1.4 µg/100g in 2020. These results are quite comparable with the French study, which reported the VitD fortification of dairy products, breakfast cereals, edible oils, and margarines; the average VitD content varied from 0.8 µg/100g in dairy products to 8 µg/100g in margarines (69). And with a study covering several European countries, where the content of VitD in milk and dairy was 0.1–1.2 µg/100g, and 7–8 µg/100g in margarines (70).

Further we compared overall nutritional quality of the foods fortified with VitD with those, which were not fortified. Analyses was done using the 2017 dataset. First, we employed WHOE nutrient profiling model, which enable distinguishing foods of lower (less healthy) and higher (healthier) overall nutritional quality. **Figure 2** provide results of analysis for the selected food (sub)categories, in which at least 2% of foods were fortified with VitD. To provide further insights, Nutri-Score (NS) nutrient profile model was also used, grading foods into five different grades of the overall nutritional quality (with dark green grade A appointed to those with highest nutritional quality, and grade E to those with lowest quality). Because such a comparison is relevant only in food subcategories with sufficient number of

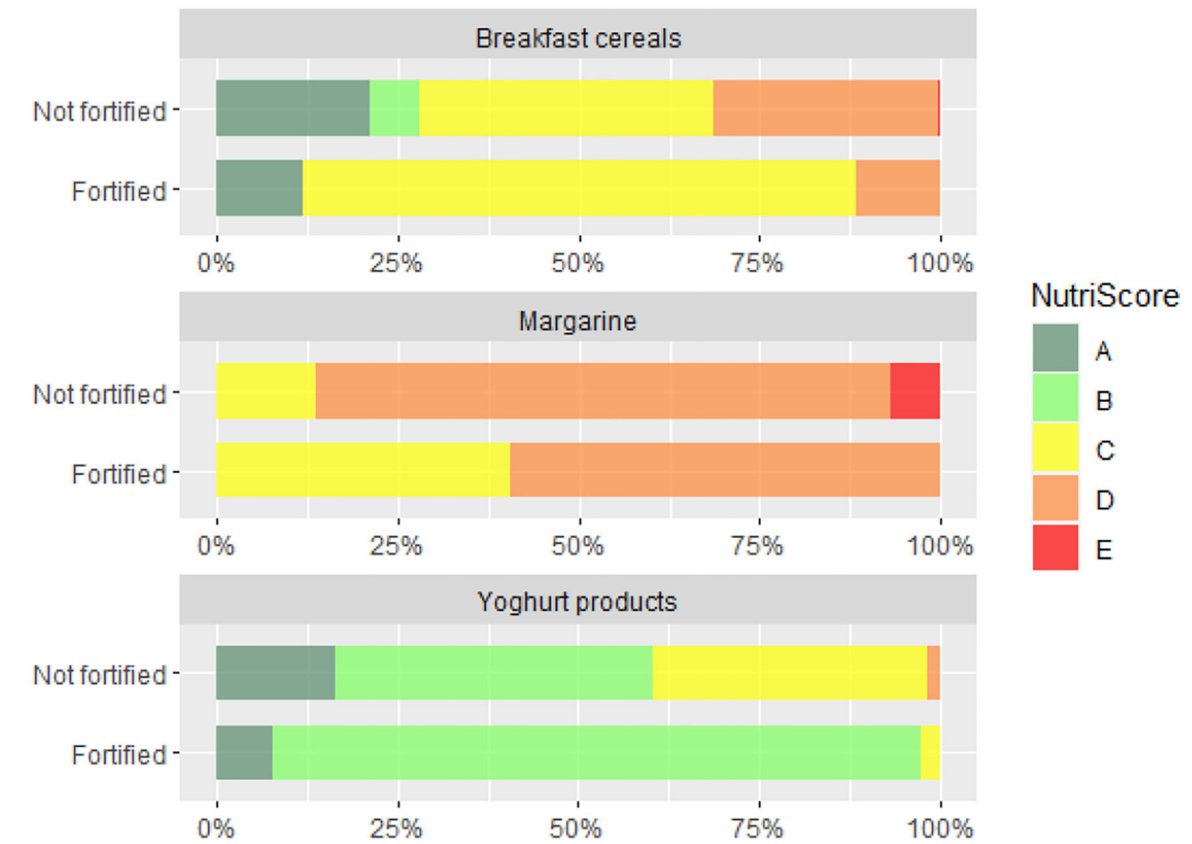


**FIGURE 3 |** Distribution of Nutri-Score grades (A-B-C-D-E) in breakfast cereals, margarine and yogurt products which are (1) not fortified and (2) fortified with vitamin D.

fortified products, **Figure 3** presents comparison for breakfast cereals, margarine and yogurt products.

Profiling with WHOE showed significant lower nutritional quality of VitD fortified cereal and cereal products. When we take a closer look into a subcategory of breakfast cereals, which are typically ultra-processed foods, no significant differences were observed (**Figure 2**). However, the use of Nutri-Score revealed interesting observation, that vast majority of VitD fortified breakfast cereals is in the middle quality grade C, while in non-fortified foods both higher and lower grades are much more evenly distributed (**Figure 3**). Among dairy products and imitates, the WHOE also highlighted lower nutritional quality in subcategory of yogurts. This was due to very strict sugar content cut-off in the WHOE model for this category (10 g sugar/100 g food), which cannot be compensated with other parameters, i.e., lower (saturated) fat content. On the other hand, Nutri-Score—where final score is calculated with consideration of positive and negative nutrients, graded both fortified and non-fortified yogurts notably better. Interestingly, margarine was the only food category, where nutritional quality of fortified products was rated better than for non-fortified foods, by both WHOE and NS. However, promotion of excess consumption of such high fat foods not appropriate, which makes margarine a limited source of vitamin D for the general population. In other European countries, fat spreads, breakfast cereals, milk, and certain baby foods are most often fortified with VitD (71, 72). Our results also show that in both years, margarine was the most common food category fortified with VitD, and also contained the highest amount of this vitamin per 100 g. This observation is similar to the situation in the Netherlands, where margarine was also the most frequently consumed fortified food product (73). It should be noted that margarine with the highest market share in our study were from major European manufacturers, which are also sold in other EU markets. In Europe, the voluntary fortification of margarine has been practiced since 1925, but with the advent of the second world war, when margarine became a major replacement for butter (74), some European countries also required mandatory margarine fortification with VitD (75–77). In the UK, the mandatory fortification of margarine was revoked in 2014 (78), but most margarine are still voluntarily fortified.

Interestingly, while milk is also commonly fortified with VitD in some European countries (particularly in northern Europe), as well as in Canada and the USA—either voluntarily or mandatorily (79, 80)—this is not the case in Slovenia. Only about 4% of milk was fortified with this vitamin, and the market share of such products was much lower (<0.5%). This can be explained by the fact that in contrast to margarine, the majority of milk in Slovenia originates from local suppliers.

As in Europe and elsewhere in the world, challenges with VitD status have been observed among various population groups (81–88). The data from the 4-year (2013–2017) food-based solutions for optimal VitD nutrition and health through the life cycle (ODIN) project revealed that 1 in 8 (13%) of the EU population are VitD deficient, and 40% insufficient (85, 89). The situation can be much worse in seasons in which sun exposure is not sufficient for the biosynthesis of VitD in human skin. For example, during the winter in Slovenia, about 40% of the population was VitD deficient, while insufficiency was observed in 80% of the adult population (21). With the absence of VitD biosynthesis, sufficient dietary intake of VitD is needed, but typical intakes in most populations are rather low. For example, in Slovenia the estimated daily intake of VitD in adults is only 2.9 μg, much below the recommended intake of 20 μg/day (90–93). Also in Canada, the majority of the population consumed very low amounts of VitD, and mandatory fortification of milk was, therefore, implemented (94). However, except in one age group (1–50 years) VitD intake is still below the estimated average requirement (95), which is consistent with reports of low serum 25-OH-Vitamin D concentrations in different population groups (96–98). European food consumption surveys also show low VitD intakes in Europe (44, 99–103), with a very limited contribution of fortified foods (104). However, improvements in serum concentrations of 25-OH-Vitamin D and a lower prevalence of VitD deficiency were reported in a long-term study after 11 years of mandatory fortification in Finland (45). Similar observations were also reported in another Finnish study, 2 years after implementing the mandatory fortification of milk and margarine (42). The available data suggest that improving VitD status through food fortification (when implemented at the population level) is cost-effective (105).

All the findings point to the need to increase the intake of vitD through food in Slovenia, as it is difficult to expect people to change their diet in order to increase the intake of VitD with regular foods, rich in this vitamin, such as fish. On the other hand, increased consumption of other dietary sources of vitamin D (i.e., liver, fortified margarine, eggs) is also not in line with the dietary guidelines for healthy eating. While Slovenia has not implemented mandatory fortification of foods with VitD, a very recent Slovenian study has estimated changes in dietary VitD intake in a hypothetical scenario of mandatory milk fortification with 2 μg of VitD per 100 ml (93). Study results showed a notable increase in the predicted VitD intake, but the expected intake would be still much below the recommended intake. However, we should mention very positive experiences with food fortification in Slovenia. Mandatory iodisation of salt was implemented very successfully back in 1953 (106), addressing iodine deficiency (107). WHO also highlighted the importance of mandatory micronutrient fortification in the case of a high prevalence of deficiencies in certain populations (25).

Evaluation of the overall nutritional quality of the foods fortified with VitD showed differences between food categories; in many cases fortified foods had low overall nutritional quality, which can be explained with the fact that currently fortification is more commonly practiced among ultra-processed foods, which are typically having lower nutritional quality (108). Fortification enables food manufacturers to use nutrition and health claims, which can be very attractive to consumers and could mask overall poor nutritional quality of the food (109). According to the EU nutrition and health claims regulation, such practices should be prohibited with the introduction of nutrient profiles—back in year 2009 (110). While this part of the regulation has not yet been implemented (111), our study highlighted that this is very relevant. It should pointed out that our observations reflect the situation of fully voluntarily market-driven addition of VitD to foods. In case of mandatory fortification, VitD would be more commonly added to less processed foods with higher overall nutritional quality.

The strength of this study is in the use of the nationally representative cross-sectional branded food composition datasets at two time points (2017 and 2020), and in the use of sales data to address market share differences. Such an approach has been used in the past to investigate various public health risks [e.g., the content of salt (55), sugar (56), and additives (57)]. Some study limitations also need to be mentioned. While our datasets were representative of the national food supply market, the data collections were conducted in major retailers with shops distributed nationwide. This means that some foods that are only available in smaller local shops are not presented. However, such products have low market shares, and we do not expect that they would notably affect the study conclusions. We also did not have access to sales data for all products in our datasets, but it should be mentioned that the sales data were provided by major retailers, covering the majority of the national market. We should mention that majority of the VitD fortified foods were international brands, sold across the European Union, making study results very relevant for larger region. For example, less than 9% of the fortified foods in our dataset ($N = 19$ and $N = 24$, in the 2017 and 2020 dataset, respectively) were produced in Slovenia, while over 90% were imported. A limitation of our study is that the VitD content in foods was not determined in the laboratory, but taken from food labels. While we did not have the capacity for the analyses of the foods, further studies should also investigate this topic, particularly if mandatory fortification is introduced. We should note that previous studies highlighted that, due to the instability of VitD in certain food matrices, the amount of this vitamin in fortified foods can be lower than declared (112). We should also note that the addition of VitD to food does not necessarily mean that such product has a better nutritional composition, and nutrient profiling analyses was therefore applied. However, a limitation is, that nutrient profiling analyses was only done on the 2017 sample, because the 2020 dataset has not yet been updated with estimates for missing data on the content of certain nutrients/constituents, which are used in the nutrient profiling algorithms (i.e., dietary fiber content).

While previous research highlighted notable brand differences in the nutritional quality of foods (113–115), our dataset of VitD fortified foods was not large enough for such an analysis. However, this topic is very interesting for further research, which need to focus into larger datasets, for example into foods with added other vitamins and/or also minerals.

## CONCLUSIONS

We showed that branded food composition databases are extremely useful resources for the investigation and monitoring of food fortification practices, particularly if sales data can also be used. In Slovenia, the fortification of foods with VitD is fully voluntary, without any formal recommendations. The comparison of nationally representative branded food datasets compiled in cross-sectional studies in 2017 and 2020 showed minor differences between both years. Overall, the highest prevalence of fortification was observed in margarine, where VitD was added to about half of the products. Other food categories that are more commonly fortified with VitD are breakfast cereals, yogurts and their imitates, and foods for specific dietary use. The highest average content of VitD was observed in margarine. Major food-category differences were also observed in comparison of the overall nutritional quality of the fortified foods; higher overall nutritional quality was only observed in fortified margarine. In the absence of a mandatory or recommended fortification, very few manufacturers decide to add VitD, and even when this is the case, such products are commonly niche foods with lower market shares. We observed exceptions in imported foods, which can be subject to fortification policies introduced in other countries.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

IP: conceptualization. MH and SK: data collection. IP and KŽ: methodology. EH: nutrient profiling. SK: formal analysis and writing—original draft preparation. MH, KŽ, and IP: manuscript review. All authors: manuscript writing—review and editing. All authors have read and approved the final version of the manuscript.

## FUNDING

The study was supported by the national research programme "Nutrition and Public Health" (P3-0395) and "Infrastructure programme for monitoring of the composition and labelling of foods" (IO-0054), funded by the Slovenian Research Agency; and the research project "Challenges in achieving adequate vitamin D status in the adult population" (L7-1849), funded by the Slovenian Research Agency and the Ministry of Health of the Republic of Slovenia. We also acknowledge support of the Nutrition Institute in the FNS-Cloud project, which received funding from the European Union's Horizon 2020 Research and Innovation programme (H2020-EU.3.2.2.3—A sustainable and competitive agri-food industry under Grant Agreement No. 863059).

## ACKNOWLEDGMENTS

We are grateful to all the retailers for granting access to their stores to collect data for the study. We also acknowledge collaborating researchers at the Nutrition Institute and students from the Biotechnical Faculty (University of Ljubljana) and BIC (Ljubljana) for their help in the data collection. We also acknowledge the support of Hristo Hristov (Nutrition Institute, Ljubljana, Slovenia) for his support in the data analyses.

## REFERENCES


1. Institute of Medicine. *Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride*. (1997). Washington, DC: The National Academies Press. p. 448.
2. Holick MF. *McCollum Award Lecture, 1994: vitamin D—new horizons for the 21st century*. *Am J Clin Nutr*. (1994) 60:619–30. doi: 10.1093/ajcn/60.4.619
3. Autier PM, Pizot BC, Mullie P. Vitamin D status and ill health: a systematic review. *Lancet Diabet Endocrinol*. (2014) 2:76–89. doi: 10.1016/S2213-8587(13)70165-7
4. Zittermann A. Vitamin D in preventive medicine: are we ignoring the evidence? *Br J Nutr*. (2003) 89:552–72. doi: 10.1079/BJN2003837
5. Zittermann A, Pilz S, Hoffmann H, Marz W. Vitamin D and airway infections: a European perspective. *Eur J Med Res*. (2016) 21:10. doi: 10.1186/s40001-016-0208-y
6. Barazzoni R, Bischoff SC, Breda J, Wickramasinghe K, Krznaric Z, Nitzan D, et al., ESPEN expert statements and practical guidance for nutritional management of individuals with SARS-CoV-2 infection. *Clin Nutr*. (2020) 39:1631–8. doi: 10.1016/j.clnu.2020.03.022
7. Meltzer DO, Best TJ, Zhang H, Vokes T, Arora V, Solway J. Association of vitamin D status and other clinical characteristics with COVID-19 test results. *JAMA Netw Open*. (2020) 3:e2019722. doi: 10.1001/jamanetworkopen.2020.19722
8. Merzon ED, Gorohovski TA, Vinker S, Cohen AG, Green I, et al. Low plasma 25(OH) vitamin D level is associated with increased risk of COVID-19 infection: an Israeli population-based study. *FEBS J*. (2020) 287:3693–702. doi: 10.1111/febs.15495
9. Derbyshire EJ, Calder PC. Respiratory tract infections and antibiotic resistance: a protective role for vitamin D? *Front Nutr*. (2021) 8:652469. doi: 10.3389/fnut.2021.652469
10. McCollum EV, Davis M. Observations on the isolation of the substance in butter fat which exerts a stimulating effect on growth. *J Biol Chem*. (1914) 19:245–50. doi: 10.1016/S0021-9258(18)88306-5
11. Hess AF, Unger LL. The cure of infantile rickets by sunlight. *JAMA*. (1921) 19:39–41. doi: 10.1001/jama.1921.02630270037013
12. Holick MF. Evolution and function of vitamin D. *Recent Results Cancer Res*. (2003) 164:3–28. doi: 10.1007/978-3-642-55580-0_1
13. Holick MF. *Ultraviolet B radiation: the vitamin D connection*. *Adv Exp Med Biol*. (2017) 996:137–54. doi: 10.1007/978-3-319-56017-5_12
14. Holick MF, MacLaughlin JA, Doppelt SH. Regulation of cutaneous previtamin D3 photosynthesis in man: skin pigment is not an essential regulator. *Science*. (1981) 211:590–3. doi: 10.1126/science.6256855
15. Holick MF. Resurrection of vitamin D deficiency and rickets. *J Clin Invest*. (2006) 116:2062–72. doi: 10.1172/JCI29449
16. Holick MF. Vitamin D and sunlight: strategies for cancer prevention and other health benefits. *Clin J Am Soc Nephrol*. (2008) 3:1548–54. doi: 10.2215/CJN.01350308
17. Cuppari L, Garcia Lopes MG, Kamimura MA. Vitamin D biology: from the discovery to its significance in chronic kidney disease. *J Ren Nutr*. (2011) 21:113–6. doi: 10.1053/j.jrn.2010.10.007
18. Wacker M, Holick MF. Sunlight and vitamin D: a global perspective for health. *Dermatoendocrinol*. (2013) 5:51–108. doi: 10.4161/derm.24494
19. Bouillon R. Comparative analysis of nutritional guidelines for vitamin D. *Nat Rev Endocrinol*. (2017) 13:466–79. doi: 10.1038/nrendo.2017.31
20. Adams JS, Hewison M. Update in vitamin D. *J Clin Endocrinol Metab*. (2010) 95:471–8. doi: 10.1210/jc.2009-1773
21. Hribar M, Hristov H, Gregorič M, Blaznik U, Zaletel K, Oblak A, et al. Nutrihealth study: seasonal variation in vitamin D status among the slovenian adult and elderly population. *Nutrients*. (2020) 12:1838. doi: 10.3390/nu12061838
22. Holick MF. Vitamin D: importance in the prevention of cancers, type 1 diabetes, heart disease, and osteoporosis. *Am J Clin Nutr*. (2004) 79:362–71. doi: 10.1093/ajcn/79.3.362
23. Holick MF. Vitamin D deficiency. *N Engl J Med*. (2007) 357:266–81. doi: 10.1056/NEJMra070553
24. Dary O, Mora JO. Food fortification: technological aspects. In: Caballero B, Editor. *Encyclopedia of Human Nutrition (Third Edition)*. (2013). Academic Press: Waltham. p. 306–14
25. WHO/FAO, Department of Nutrition for Health and Development, Editor. *Guidelines on Food Fortification with Micronutrients*. (2006). p. 341.
26. Calvo MS, Whiting SJ. Survey of current vitamin D food fortification practices in the United States and Canada. *J Steroid Biochem Mol Biol*. (2013). 136:211–3. doi: 10.1016/j.jsbmb.2012.09.034
27. Whiting S, Calvo M. *Vitamin D Fortification and Supplementation Policies to Correct Vitamin D Insufficiency/Deficiency Globally*. London: Academic Press (2018). p. 91–108.
28. World Health Organization (WHO). *Biofortification of Staple Crops*. (2016). Available from: https://www.who.int/elena/titles/biofortification/en/ (accessed November 23, 2021).
29. Cashman KD, Kiely M. Chapter 63—vitamin D and food fortification. In: Feldman D, Editor. *Vitamin D (Fourth Edition)*. (2018). Cambridge: Academic Press. p. 109–27
30. Hayes A, Cashman KD. Food-based solutions for vitamin D deficiency: putting policy into practice and the key role for research. *Proc Nutr Soc*. (2017). 76:54–63. doi: 10.1017/S0029665116000756
31. Pilz S, März W, Cashman KD, Kiely ME, Whiting SJ, Holick MF, et al. Rationale and plan for vitamin D food fortification: a review and guidance paper. *Front Endocrinol*. (2018) 9:373. doi: 10.3389/fendo.2018.00373
32. Efsa Panel on Dietetic Products Nutrition Allergies, Scientific Opinion on the safety of vitamin D-enriched UV-treated baker's yeast. *EFSA J*. (2014) 12:3520. doi: 10.2903/j.efsa.2014.3520
33. D. Turck, J. Castenmiller, S. De Henauw, K.I. Hirsch-Ernst, J. Kearney, et al. EFSA panel on nutrition novel foods food allergens, safety of extended uses of UV-treated baker's yeast as a novel food pursuant to Regulation (EU) 2015/2283. *EFSA J*. (2021) 19:e06602. doi: 10.2903/j.efsa.2021.6602
34. European Parliament Council of the European Union, Regulation (EC) No 1925/2006 of the European Parliament and of the Council of 20 December 2006 on the addition of vitamins and minerals and of certain other substances to foods. *Official J Eur Union*. (2006) L 404/26, 1–13.
35. Hennessy Á, Walton J, Flynn A. The impact of voluntary food fortification on micronutrient intakes and status in European countries: a review. *Proc Nutr Soc*. (2013) 72:433–40. doi: 10.1017/S002966511300339X
36. European Commission (EC). *Community Register on the Addition of Vitamins and Minerals and of Certain Other Substances to Foods. Regulation (EC) No 1925/2006 on the Addition of Vitamins and Minerals and of Certain Other Substances to Foods*. 01. 01. 2021. Available from: https://ec.europa.eu/food/system/files/2021-01/labelling_nutrition-vitamins_minerals-comm_reg_en.pdf (accessed September 09, 2021).
37. Flynn A, Moreiras O, Stehle P, Fletcher RJ, Müller DJ, Rolland V. Vitamins and minerals: a model for safe addition to foods. *Eur J Nutr*. (2003) 42:118–30. doi: 10.1007/s00394-003-0391-9
38. Mertz W. Food fortification in the United States. *Nutr Rev*. (1997) 55:44–9. doi: 10.1111/j.1753-4887.1997.tb01594.x
39. Seelig MS. Vitamin D and cardiovascular, renal, and brain damage in infancy and childhood. *Ann N Y Acad Sci*. (1969) 147:539–82. doi: 10.1111/j.1749-6632.1967.tb41272.x
40. Blank, S., Scanlon KS, Sinks TH, Lett S, Falk H. An outbreak of hypervitaminosis D associated with the overfortification of milk from a home-delivery dairy. *Am J Public Health*. (1995) 85:656–9. doi: 10.2105/AJPH.85.5.656
41. Moulas AN, Vaiou M. Vitamin D fortification of foods and prospective health outcomes. *J Biotechnol*. (2018) 285:91–101. doi: 10.1016/j.jbiotec.2018.08.010
42. Piirainen T, Laitinen K, Isolauri E. Impact of national fortification of fluid milks and margarines with vitamin D on dietary intake and serum 25-hydroxyvitamin D concentration in 4-year-old children. *Eur J Clin Nutr*. (2007) 61:123–8. doi: 10.1038/sj.ejcn.1602506
43. Scientific Advisory Committee on Nutrition. *Update on Vitamin D*. (2007). p. 77.
44. Hirvonen T, Sinkko H, Valsta L, Hannila ML, Pietinen P. Development of a model for optimal food fortification: vitamin D among adults in Finland. *Eur J Nutr*. (2007) 46:264–70. doi: 10.1007/s00394-007-0660-0
45. Jääskeläinen T, Itkonen ST, Lundqvist A, Erkkola M, Koskela T, Lakkala K, et al. The positive impact of general vitamin D food fortification policy on vitamin D status in a representative adult Finnish population: evidence from an 11-y follow-up based on standardized 25-hydroxyvitamin D data. *Am J Clin Nutr*. (2017) 105:1512–20. doi: 10.3945/ajcn.116.151415
46. Pravst I, Kušar A. Consumers' exposure to nutrition and health claims on pre-packed foods: Use of sales weighting for assessing the food supply in Slovenia. *Nutrients*. (2015) 7:9353–68. doi: 10.3390/nu7115474

47. Hodgkins CE, Egan B, Peacock M, Klepacz N, Miklavec K, Pravst I, et al., Understanding how consumers categorise health related claims on foods: a consumer-derived typology of health-related claims. *Nutrients*. (2019) 11:539. doi: 10.3390/nu11030539
48. Kaur A, Scarborough P, Hieke S, Kusar A, Pravst I, Raats M, et al. The nutritional quality of foods carrying health-related claims in Germany, The Netherlands, Spain, Slovenia and the United Kingdom. *Eur J Clin Nutr*. (2016) 70:1388–95. doi: 10.1038/ejcn.2016.114
49. Kaur A, Scarborough P, Matthews A, Payne S, Mizdrak A, Rayner M. How many foods in the UK carry health and nutrition claims, and are they healthier than those that do not? *Publ Health Nutr*. (2016) 19:988–97. doi: 10.1017/S1368980015002104
50. Pivk Kupirovič U, Miklavec K, Hribar M, Kušar A, Žmitek K, Pravst I. Nutrient profiling is needed to improve the nutritional quality of the foods labelled with health-related claims. *Nutrients*. (2019) 11:287. doi: 10.3390/nu11020287
51. Pravst I, Kusar A, Miklavec K, Hieke S, Raats M, Rayner M. Use of nutrition and health claims and symbols on prepacked foods in europe: from consumer exposure to public health implications. In: Astley S, Editor. *Health Claims and Food Labelling*. (2020). p. 79–93.
52. Del Valle HB, Yaktine AL, Taylor CL, Ross AC. *Dietary reference intakes for calcium and vitamin D*. (2011). Washington (DC): National Academies Press.
53. Janssen M, Chang B, Hristov H, Pravst I, Profeta A, Millard J. Changes in food consumption during the COVID-19 pandemic: analysis of consumer survey data from the first lockdown period in Denmark, Germany and Slovenia. *Front Nutr*. (2021). In press. doi: 10.3389/fnut.2021.635859
54. Zmitek K, Hribar M, Lavrisa Z, Hristov H, Kusar A, Pravst I. Socio-demographic and knowledge-related determinants of vitamin D supplementation in the context of the COVID-19 pandemic: assessment of an educational intervention. *Front Nutr*. (2021) 8:290. doi: 10.3389/fnut.2021.648450
55. Pravst I, Lavriša Ž, Kušar A, Miklavec K, Žmitek K. Changes in average sodium content of prepacked foods in Slovenia during 2011–2015. *Nutrients*. (2017) 9:952. doi: 10.3390/nu9090952
56. Zupanič N, Hribar M, Fidler Mis N, Pravst I. Free sugar content in pre-packaged products: does voluntary product reformulation work in practice? *Nutrients*. (2019) 11:2577. doi: 10.3390/nu11112577
57. Blaznik U, Krušič S, Hribar M, Kušar A, Žmitek K, Pravst I. Use of food additive titanium dioxide (E171) before the introduction of regulatory restrictions due to concern for genotoxicity. *Foods*. (2021) 10:1910. doi: 10.3390/foods10081910
58. Pravst I, Hribar M, Žmitek K, Blažica B, Korousic Seljak B, Kušar A. Branded foods databases as a tool to support nutrition research and monitoring of the food supply: Insights from the Slovenian Composition and Labelling Information System (CLAS). *Front Nutr*. In press. doi: 10.3389/fnut.2021.798576
59. Zupanič N, Miklavec K, Kušar A, Žmitek K, Fidler Mis N, Pravst I. Total and free sugar content of pre-packaged foods and non-alcoholic beverages in Slovenia *Nutrients*. (2018) 10:151. doi: 10.3390/nu10020151
60. Zupanič, N, Hribar M, Pivk Kupirovič U, Kušar A, Žmitek K, Pravst I. Limiting trans fats in foods: use of partially hydrogenated vegetable oils in prepacked foods in Slovenia. *Nutrients*. (2018) 10:355. doi: 10.3390/nu10030355
61. Dunford E, Webster J, Metzler AB, Czernichow S, Ni Mhurchu C, Wolmarans P, et al. International collaborative project to compare and monitor the nutritional composition of processed foods. *Eur J Prev Cardiol*. (2012) 19:1326–32. doi: 10.1177/1741826711425777
62. WHO. *WHO Regional Office for Europe Nutrient Profile Model*. (2015). Available from: https://www.euro.who.int/__data/assets/pdf_file/0005/270716/Nutrient-children_web-new.pdf (accessed May 19, 2021).
63. Julia C, Hercberg S, World Health Organization Regional Office for Europe. Development of a new front-of-pack nutrition label in France: the five-colour nutri-score. *Public Health Panorama*. (2017) 03:712–25.
64. Delhomme V. Front-of-pack nutrition labelling in the European Union: a behavioural, legal and political analysis. *Eur J Risk Regul*. (2021) 12:1–24. doi: 10.1017/err.2021.5
65. Hafner E, Pravst I. Evaluation of the ability of nutri-score to discriminate the nutritional quality of prepacked foods using a sale-weighting approach. *Foods*. (2021) 10:1689. doi: 10.3390/foods10081689
66. Eržen N, Rayner M, Pravst I. A comparative evaluation of the use of a food composition database and nutrition declarations for nutrient profiling. *J Food Nutr Res*. (2015) 54:1–8. doi: 10.13140/2.1.2096.0000
67. Santé publique France. *Nutri-Score Frequently Asked Questions—Scientific & Technical*. (2021). p. 1–31. www.santepubliquefrance.fr
68. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. *Am Stat*. (1998) 52:119–26. doi: 10.1080/00031305.1998.10480550
69. Dhaussy A. Vitamin D recommendations, fortification in France, and communication. *OCL*. (2014) 21:D305. doi: 10.1051/ocl/2013063
70. Milešević J, Samaniego L, Kiely M, Glibetić M, Roe M, Finglas P. Specialized food composition dataset for vitamin D content in foods based on European standards: Application to dietary intake assessment. *Food Chem*. (2018) 240:544–9. doi: 10.1016/j.foodchem.2017.07.135
71. European Union. *Commission Directive 2006/141/EC on Infant Formulae and Follow-on Formulae and Amending Directive 1999/21/EC*. (2006). Available from: https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32006L0141 (accessed September 9, 2021).
72. Cashman KD, Kiely M. Tackling inadequate vitamin D intakes within the population: fortification of dairy products with vitamin D may not be enough. *Endocrine*. (2016) 51:38–46. doi: 10.1007/s12020-015-0711-x
73. van Rossum CTM, Fransen HP, Verkaik-Kloosterman J, Buurma-Rethans EJM, Ocke MC. *Dutch National Food Consumption Survey 2007–2010 : Diet of Children and Adults Aged 7 to 69 Years*. (2011). Available from: https://www.rivm.nl/publicaties/dutch-national-food-consumption-survey-2007-2010-diet-of-children-and-adults-aged-7-to (accessed September 1, 2021).
74. European Comission. Council regulation (EC) No 2991/94 of 5 December 1994 laying down standards for spreadable fats. *EUR-Lex*. (1994) L 316/2, 1–6.
75. Sioen I. Fortified margarine and fat spreads. In: Preedy VR, Srirajaskanthan R, Patel VB, editors. *Handbook of Food Fortification and Health: From Concepts to Public Health Applications Volume 1*. (2013). New York, NY: Springer. p. 159–71.
76. DH (Department of Health). *Dietary Reference Values for food energy and nutrients for the United Kingdom. Report of the Panel on Dietary Reference Values of the Committee on Medical Aspects of Food Policy*. (1991). Available from: https://www.gov.uk/government/publications/coma-reports (accessed September 1 2021).
77. Laaksi IT, Ruohola JP, Ylikomi TJ, Auvinen A, Haataja RI, Pihlajamäki HK, et al. Vitamin D fortification as public health policy: significant improvement in vitamin D status in young Finnish men. *Eur J Clin Nutr*. (2006) 60:1035–8. doi: 10.1038/sj.ejcn.1602414
78. Department for Environment Food and Rural Affairs. *Consultation on Revoking Regulation 4 of the Spreadable Fats (Marketing Standards) and Milk and Milk Products (Protection of Designations) (England) Regulations 2008—Margarine Fortification*. (2014). Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/287673/margarine-fortification-sum-resp-201403.pdf (accessed September 3 2021).
79. Department of Health Regulations Amending Certain Regulations Made Under the Food and Drugs Act. *Nutrition Symbols, Other Labelling Provisions Partially Hydrogenated Oils and Vitamin D*. (2018). Available from: https://canadagazette.gc.ca/rp-pr/p1/2018/2018-02-10/pdf/g1-15206.pdf (accessed September 9 2021).
80. Itkonen ST, Erkkola M, Lamberg-Allardt CJE. Vitamin D fortification of fluid milk products and their contribution to vitamin D intake and vitamin D status in observational studies—a review. *Nutrients*. (2018) 10:1054. doi: 10.3390/nu10081054
81. Spiro A, Buttriss JL. Vitamin D: An overview of vitamin D status and intake in Europe. *Nutr Bull*. (2014) 39:322–50. doi: 10.1111/nbu.12108
82. Cashman KD, Dowling KG, Škrabáková Z, Gonzalez-Gross M, Valtueña J, De Henauw S, et al. Vitamin D deficiency in Europe: pandemic? *Am J Clin Nutr*. (2016) 103:1033–44. doi: 10.3945/ajcn.115.120873
83. Palacios C, Gonzalez L. Is vitamin D deficiency a major global public health problem? *J Steroid Biochem Mol Biol*. (2014) 144:138–45. doi: 10.1016/j.jsbmb.2013.11.003
84. Holick MF. The vitamin D deficiency pandemic: Approaches for diagnosis, treatment and prevention. *Rev Endocr Metab Disord*. (2017) 18:153–65. doi: 10.1007/s11154-017-9424-1

85. Kiely M, Cashman KD. Summary outcomes of the ODIN project on food fortification for vitamin D deficiency prevention. *Int J Environ Res Public Health*. (2018) 15:2342. doi: 10.3390/ijerph15112342
86. Lips P. Vitamin D status and nutrition in Europe and Asia. *J Steroid Biochem Mol Biol*. (2007) 103:620–5. doi: 10.1016/j.jsbmb.2006.12.076
87. Pludowski P, Grant WB, Bhattoa HP, Bayer M, Povoroznyuk V, Rudenka E, et al. Vitamin D status in Central Europe. *Int J Endocrinol*. (2014) 2014:589587. doi: 10.1155/2014/589587
88. Manios Y, Moschonis G, Lambrinou CP, Tsoutsoulopoulou K, Binou P, Karachaliou A, et al. A systematic review of vitamin D status in southern European countries *Eur J Nutr*. (2018) 57:2001–36. doi: 10.1007/s00394-017-1564-2
89. European Comission. *Final Report Summary—ODIN (Food-based solutions for Optimal vitamin D Nutrition and health through the life cycle)* (2018). Available from: https://cordis.europa.eu/project/id/613977/reporting (accessed November 23, 2021).
90. World Health Organization. *Vitamin and Mineral Requirements in Human Nutrition*. (2004). USA: World Health Organization.
91. Bresson J-L, Burlingame B, Dean T, Fairweather-Tait S, Heinonen M, Hirsch-Ernst K-I, et al. *Dietary Reference Values for Vitamin D* (2016). doi: 10.2903/j.efsa.2016.4547
92. German Nutrition Society, New reference values for vitamin D. *Ann Nutr Metab*. (2012) 60:241–6. doi: 10.1159/000337547
93. Hribar M, Hristov H, Lavriša Ž, Koroušić Seljak B, Gregorič M, Blaznik U, et al. Vitamin D intake in slovenian adolescents, adults, and the elderly population. *Nutrients*. (2021) 13:3528. doi: 10.3390/nu13103528
94. Government of Canada. *Food and Drug Regulations C.R.C., c. 870*. (2021). Available from: https://laws-lois.justice.gc.ca/eng/regulations/c.r.c.,_c._870/FullText.html (accessed September 9, 2021).
95. Vatanparast H, Calvo MS, Green TJ, Whiting SJ. Despite mandatory fortification of staple foods, vitamin D intakes of Canadian children and adults are inadequate. *J Steroid Biochem Mol Biol*. (2010) 121:301–3. doi: 10.1016/j.jsbmb.2010.03.079
96. Rucker D, Allan JA, Fick GH, Hanley DA. Vitamin D insufficiency in a population of healthy western Canadians. CMAJ. (2002) 166:1517–24.
97. Vieth R, Cole DE, Hawker GA, Trang HM, Rubin LA. Wintertime vitamin D insufficiency is common in young Canadian women, and their vitamin D intake does not prevent it. *Eur J Clin Nutr*. (2001) 55:1091–7. doi: 10.1038/sj.ejcn.1601275
98. Roth DE, *Martz P, Yeo R, Prosser C, Bell M, Jones AB.* Are national vitamin D guidelines sufficient to maintain adequate blood levels in children? *Can J Public Health*. (2005) 96:443–9. doi: 10.1007/BF03405185
99. Department of Health. *National Diet and Nutrition Survey: Headline results from Years 1 and 2 (combined) of the rolling programme 2008-9–2009-10*. (2011). Available from: https://www.gov.uk/government/publications/national-diet-and-nutrition-survey-headline-results-from-years-1-and-2-combined-of-the-rolling-programme-2008-9-2009-10 (accessed September 3 2021).
100. Whitton C, Nicholson SK, Roberts C, Prynne CJ, Pot GK, Olson A, et al. National Diet and Nutrition Survey: UK food consumption and nutrient intakes from the first year of the rolling programme and comparisons with previous surveys. *Br J Nutr*. (2011) 106:1899–914. doi: 10.1017/S0007114511002340
101. Hill TR, O'Brien M, Cashman KD, Flynn A, Kiely M. Vitamin D intakes in 18-64-y-old Irish adults. *Eur J Clin Nutr*. (2004) 58:1509–17. doi: 10.1038/sj.ejcn.1602001
102. Tetens I, Biltoft-Jensen A, Spagner C, Christensen T, Gille M-B, Bügel S, et al. Intake of micronutrients among Danish adult users and non-users of dietary supplements. *Food Nutr Res*. (2011) 55:1–8. doi: 10.3402/fnr.v55i0.7153
103. Dufour A, Wetzler S, Touvier M, Lioret S, Gioda J, Lafay L, et al. Comparison of different maximum safe levels in fortified foods and supplements using a probabilistic risk assessment approach. *Br J Nutr*. (2010) 104:1848–57. doi: 10.1017/S0007114510002862
104. Flynn A, Hirvonen T, Mensink GBM, Ocké MC, Serra-Majem L, Stos K, et al. Intake of selected nutrients from foods, from fortification and from supplements in various European countries. *Food Nutr Res*. (2017) 53:1–51. doi: 10.3402/fnr.v53i0.2038
105. Aguiar M, Andronis L, Pallan M, Högler W, Frew E. Preventing vitamin D deficiency (VDD): a systematic review of economic evaluations. *Eur J Public Health*. (2017) 27:292–301. doi: 10.1093/eurpub/ckw270
106. Zmitek K, Pravst I. Iodisation of salt in Slovenia: increased availability of non-iodised salt in the food supply. *Nutrients*. (2016) 8:434. doi: 10.3390/nu8070434
107. Ministry of Health Slovenia. *Rules on salt iodisation (Uradni list RS, št. 47/18)* (2018). Available from: http://www.pisrs.si/Pis.web/pregledPredpisa?id=PRAV12884 (accessed November 23, 2021).
108. Dunford EK, Ni Mhurchu C, Huang L, Vandevijvere S, Swinburn B, Pravst I, et al. The comparative healthiness of 394,815 packaged food and beverages from 12 countries. *Obesity Rev*. (2019) 20:107–15. doi: 10.1111/obr.12879
109. Miklavec K, Pravst I, Grunert KG, Klopčič M, Pohar J. The influence of health claims and nutritional composition on consumers' yoghurt preferences. *Food Qual Pref*. (2015) 43:26–33. doi: 10.1016/j.foodqual.2015.02.006
110. EC. *Regulation (EC) No 1924/2006 of the European Parliament and of the Council on Nutrition and Health Claims Made on Foods*. (2006). Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02006R1924-20141213 (accessed March 29, 2020).
111. Kusar A, Zmitek K, Lahteenmaki L, Raats MM, Pravst I. Comparison of requirements for using health claims on foods in the European Union, the USA, Canada, and Australia/New Zealand. *Compr Rev Food Sci Food Saf*. (2021) 20:1307–32. doi: 10.1111/1541-4337.12716
112. Termova Rakuša Ž. Development of stability-indicating methods for stability evaluation of hydrophilic and lipophilic vitamins and design of their stabilization. In: *Faculty of Pharmacy*. (2021). University of Ljubljana. p. 421.
113. Gerstner E. Do higher prices signal higher quality? *J Market Res*. (1985) 22:209–15. doi: 10.1177/002224378502200210
114. Roth Y. Do brands serve as reliable signals of nutritional quality? The case of breakfast cereals. *J Food Prod Market*. (2016) 23:1–23. doi: 10.1080/10454446.2017.1244787
115. Angelino D, Rosi A, Dall'Asta M, Pellegrini N, Martini D. Evaluation of the nutritional quality of breakfast cereals sold on the Italian market: The Food Labelling of Italian Products (FLIP) Study. *Nutrients*. (2019) 11:2827. doi: 10.3390/nu11112827


**Author Disclaimer:** The information and views in this report do not necessarily reflect the official opinion or position of the European Union. Neither the European Union Institutions and bodies nor any person acting on their behalf may be held responsible for the use of the information contained herein.

**Conflict of Interest:** We acknowledge that IP has led and participated in various other research projects in the areas of nutrition, public health, and food technology, which were (co)funded by the Slovenian Research Agency, Ministry of Health of the Republic of Slovenia, the Ministry of Agriculture, Forestry and Food of the Republic of Slovenia and, in cases of specific applied research projects, also by food businesses. IP and KŽ are members of a National Workgroup responsible for the development of recommendations for assuring adequate vitamin D status among the Slovenian population.

The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 13 January 2022
doi: 10.3389/fnut.2021.762814

# Sensory-Related Industrial Additives in the US Packaged Food Supply

Marilyn Tseng[1]*, Camille J. Grigsby[1], Abigail Austin[1], Samir Amin[2] and Aydin Nazmi[2]

[1] Department of Kinesiology and Public Health, California Polytechnic State University, San Luis Obispo, CA, United States,
[2] Department of Food Science and Nutrition, California Polytechnic State University, San Luis Obispo, CA, United States

**Background:** Increasing evidence suggests that ultra-processed foods (UPFs) lead to elevated risk of obesity-related conditions, but UPF measurement has been criticized for its subjectivity and lack of clarity on biological mechanism. Sensory-related industrial additives (SRIAs) are a defining feature of UPFs and may encourage overconsumption by enhancing the sensory quality of foods. However, practical challenges have prevented systematic incorporation of SRIAs into UPF measurement.

**Objective:** The objectives of this work were to describe a new, open-source ingredient list search method and to apply this method to describe the presence of SRIAs in US packaged foods.

**Methods:** We developed computer coding to search for 64 common SRIAs related to sweetness, flavor, appearance, and texture in 241,688 foods in the US Branded Food Products Database (BFPD). The BFPD includes manufacturer-provided ingredient lists for ∼300,000 branded and private label food items. We determined the total number of SRIAs (0–64) and the number of different types of SRIAs (sweetness, flavor, appearance, texture, 0–4) in each food, then calculated the percent of all foods with SRIAs. This was done for all foods, and by food group for 224,098 items with food group data.

**Results:** Most (64.9%) foods in the BFPD contained at least one SRIA, and more than a third had at least three. Sweets (89.5%), beverages (84.9%), and ready-to-eat (RTE) foods (82.0%) were the most likely to contain SRIAs. With respect to SRIA types, 25.7% of all food items had at least three of the four types of SRIAs examined, with texture-related additives being the most common. Among sweets, 20% had all four types of SRIAs.

**Discussion:** This work confirms the high prevalence of SRIAs in US packaged foods. They are ubiquitous in sweets, beverages, and RTE foods, but also present in substantial proportions of other food groups. Quantifying the presence of SRIAs in ingredient lists offers a novel way to identify UPFs for research; to distinguish more vs. less ultra-processed foods; and to test whether UPFs increase risk for obesity-related conditions through additives that enhance the product's sensory qualities.

**Keywords: artificial food colors, emulsifiers, industrial additives, ingredient list, sweeteners, ultra-processed food, USDA branded food products database**

OPEN ACCESS

**Edited by:**
*Alessandra Durazzo, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Elizabeth Dunford, University of New South Wales, Australia*
*Marcela Reyes, University of Chile, Chile*

***Correspondence:**
*Marilyn Tseng*
*mtseng@calpoly.edu*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** 22 August 2021
**Accepted:** 06 December 2021
**Published:** 13 January 2022

**Citation:**
*Tseng M, Grigsby CJ, Austin A, Amin S and Nazmi A (2022) Sensory-Related Industrial Additives in the US Packaged Food Supply. Front. Nutr. 8:762814. doi: 10.3389/fnut.2021.762814*

# INTRODUCTION

A steadily growing body of evidence links highly or "ultra-processed foods" (UPFs) to cardiometabolic conditions and cancer (1–6). Existing frameworks for identifying foods as "ultra-processed" variously consider extent of modification from a food's original form; its ingredients (e.g., the number and types of additives); the use of industrial processing methods, as opposed to methods used in home or culinary preparations; and the purpose of processing (e.g., food safety, convenience, palatability) (7, 8).

However, processed food classification systems have been challenged for their subjectivity, inconsistency, ambiguity, and lack of clarity on biological mechanisms specific to UPFs (7, 9). Such criticisms raise doubts about the validity of findings linking UPF consumption to disease risk, and about implications for policy decisions relating to UPFs. Current systems also treat UPFs as a binary characteristic, such that a flavored yogurt might fall into the same category as a food made entirely of extracted substances. A less subjective way to classify foods that also allows for distinguishing levels of ultra-processing would be a useful complement to current classification frameworks, allow for more quantitative explorations of associations between UPFs and disease, and potentially serve as a useful basis to guide consumers and policy-makers.

The presence of industrial additives as ingredients figures prominently in most major frameworks to classify UPFs (7, 8, 10) and offers a potentially more objective indicator of one component of ultra-processing (11). A food additive is defined broadly as any substance added to food to perform a specific function (12, 13)—for example, to improve safety, slow spoilage, improve or maintain nutritional value, as well as "to improve taste, texture and appearance" (12). Despite their recognized importance to the safety, shelf-life, and nutritional value of foods (12, 14), additives have also been implicated in a variety of health outcomes (15–17). However, additives that enhance the sensory qualities of foods are of particular relevance if UPFs are thought to increase risk by encouraging overconsumption and contributing to excess adiposity (18, 19). Thus, we focus on *sensory-related* industrial additives (SRIAs) as "classes of additives whose function is to [...] give the final product sensory properties especially attractive to see, taste, smell and/or touch" (11). Quantifying SRIAs in foods provides a means to evaluate one mechanism by which UPFs have been suggested to increase risk for obesity-related conditions.

SRIAs are a defining feature of UPFs but have not been systematically incorporated into measurement of UPFs in research because of practical challenges in accessing and searching ingredient lists for variably expressed or easily misspelled ingredients. The relatively recent release of a publicly accessible database with ingredient lists for packaged foods presents the opportunity to explore a new method of identifying and describing the presence of SRIAs in packaged foods in the US. The objectives of this paper are: (1) to describe a new method that searches ingredient lists for common SRIAs affecting four aspects of food: sweetness, flavor (other than sweetness), texture, and appearance; and (2) to apply this method to describe the presence of SRIAs in US packaged foods.

# MATERIALS AND METHODS

## Identification of SRIAs

A graphical scheme illustrating our approach is shown in **Figure 1**. We compiled an initial list of 42 common SRIAs from the US Food and Drug Administration (12) and International Food Information Council (14), including only ingredients and additives specified as sweeteners (nutritive and non-nutritive), color additives, flavors or flavor enhancers, fat replacers, emulsifiers, stabilizers, thickeners, binders, texturizers, and that were also deemed not to be commonly used in domestic or culinary settings (e.g., vanilla extract, pectin). We used the primary function assigned by these websites to categorize each additive as relating to: texture (including emulsifiers, thickeners, and fat replacers), flavor (flavoring agents or flavor enhancers, other than sweetness), sweetness, or appearance (colors, dyes, and glazing agents) (Note that for the remainder of the paper, our use of the term "flavor" excludes sweetness as a flavor). We identified 13 additional different names for sugars in ingredient lists (20) and consulted with an expert in culinary science and food technology (SA) to confirm that the list included sweeteners used primarily in industrial and not in domestic or culinary settings.

We also added to our list seven SRIAs that occurred in at least 0.5% of the 126,556 food products in a recent analysis of French food and beverage products (21). Subsequent iterative verifications of the completeness of our search (described below) resulted in the addition of two more SRIAs to our search. For these last nine SRIAs, we assigned primary function based on online resources (22, 23). Our final list of 64 SRIAs included 21 related to texture, six related to flavor (excluding sweetness), 26 related to sweetness (including 20 nutritive and six non-nutritive), and eleven related to appearance (all colors and dyes except carnauba wax) (**Table 1**).

## Branded Food Products Database

We searched for each SRIA within ingredient lists available in the Branded Food Products Database (BFPD). The BFPD is a publicly available database that includes nutrient information and ingredient lists voluntarily provided by food companies for packaged food items (24). In 2019, Baldridge et al. (25) reported that the database "represents >80% of all food and beverage products sold in the US over the past three years." Manufacturers and retailers submit data to the database using a portal provided by Label Insight, a company specializing in consumer product label data aggregation and analysis; or by directly synchronizing their data through the GS1 Global Data Synchronization Network, if participating in the network. About 95% of food item data in the BFPD are derived from Label Insight.

We used the December 2019 version, which contained 331,271 items. We excluded 37,033 items with duplicate UPC codes, 50,020 items with duplicate brand and food names, 2,529 items

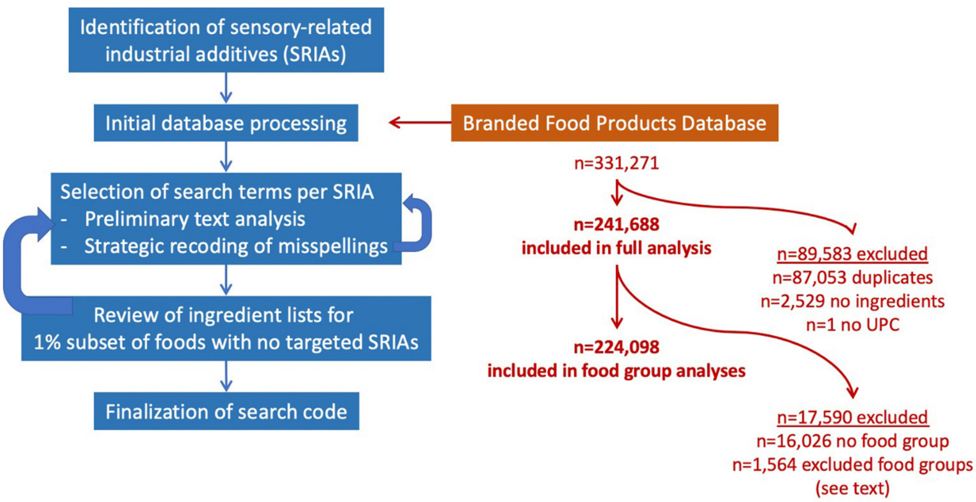


**FIGURE 1 |** Graphical scheme illustrating study approach and database used. Initial identification of sensory-related industrial additives (SRIAs) was compiled from US Food and Drug Administration (12), International Food Information Council (14), and SugarScience (20) websites. Following preliminary editing of the Branded Food Products Database, an iterative process was used to identify search terms for each SRIA, which included a preliminary text analysis and selective recoding of misspellings. Ingredient lists for a 1% subset of foods with no SRIAs were reviewed for errors or omissions, and search terms were revised. The process was repeated until no further modifications were deemed necessary. Also shown are numbers of BFPD food items included in analyses.

with no ingredients, and one item with no UPC, leaving 241,688 foods in our analysis.

Of 241,688 foods, 225,662 had a field representing 227 food categories. We combined food categories into 79 food groups (representing seven broad food categories) developed by the Economic Research Service to correspond with the 2015 Dietary Guidelines for Americans and to capture information on convenience and processing (26). Because the BFPD often did not provide more specific information to categorize foods more specifically, we collapsed the 79 food groups further; for example, all vegetables were combined into one group that included starchy vegetables, tomatoes, dark green vegetables, other red and orange vegetables, beans, lentils and peas or legumes, and other/mixed vegetables. Four food groups (egg and egg substitutes; tofu and meat substitutes; vitamins and meal supplements; alcohol) had fewer than 700 items, suggesting that foods in those groups were being captured in other categories and that the included foods were not representative of all foods in that group. After excluding those four groups, we ended with 224,098 food items in fifteen food groups (**Table 2**).

## Search Strategy and Iteration

We edited the unstructured text in the ingredients field of the BFPD—for example, converting all text to lower case, replacing or standardizing symbols, and removing multiple blanks. Preliminary examinations of ingredient lists revealed substantial variability in the spelling and expression of different additives (for examples of this variability, see **Supplementary Table 1**). We used the Text Explorer platform in JMP, which breaks unstructured text data (such as text in ingredient lists) into terms ("tokens"), to identify misspellings or alternative spellings for each word for the additives of interest.

After identifying the range of misspellings and alternative spellings, we selected the smallest character string that would identify a given additive, in an iterative process in both JMP and SAS to confirm that searching for that string would not identify unwanted items (i.e., false positives). We recoded other misspellings as necessary using the TRANWORD function in SAS. We then searched for the character string using the INDEX function, which searches for specific character strings in a specified field. Examples of this approach are provided in **Supplementary Table 1**.

To verify the completeness of our search, we reviewed the ingredient lists of a randomly selected subset of 1% (850–900 foods) of foods identified as not having any of our targeted additives, specifically to identify SRIAs missed by our initial searches, as well as SRIAs of potential interest not on our original list. We revised our SAS code to address omissions and repeated the process of reviewing ingredient lists of randomly selected subsets of foods until we noted no further necessary revisions to our search. The final version of SAS code used to produce the results presented here are available in **Supplementary File 1**.

## Comparison With Independent Database

To evaluate the completeness and accuracy of our search, we compared our search results to data from Open Food Facts (27). Open Food Facts (OFF) is a publicly available international database of >650,000 food products, including almost 350,000 US items. Data on nutrient content and ingredients are submitted by manufacturers and consumers. Included in OFF is a data field indicating the specific additives present in each food, identified by E number, a number assigned to identify food additives approved for use in the European Union. According to a March 2021 email from S. Gigandet (stephane@openfoodfacts.org),

**TABLE 1 |** List of sensory-related industrial additives (SRIA) sought in ingredient lists.

| Function | Additive | E number[a] |
|---|---|---|
| **Texture** | | |
| | 1. Lecithin | E322 |
| | 2. Mono- and diglycerides | E471 |
| | 3. Polysorbates | E432-E436 |
| | 4. Sorbitan monostearate | E491 |
| | 5. Di-, tri-, poly-, and pyro-phosphates | E450-E452 |
| | 6. Polyglycerol polyricinoleate | E476 |
| | 7. Sodium stearoyl-2-lactylate | E481 |
| | 8. Ammonium phosphatides | E442 |
| | 9. Sodium caseinate | |
| | 10. Guar gum | E412 |
| | 11. Carrageenan | E407 |
| | 12. Xanthan gum | E415 |
| | 13. Locust (carob) bean gum | E410 |
| | 14. Gum arabic (acacia gum) | E414 |
| | 15. Hydrogenated oils | |
| | 16. Cellulose, cellulose gum, cellulose gel, crystalline cellulose, carboxymethylcellulose | E460, E466 |
| | 17. Modified food starch | E1401 |
| | 18. Alginic acid and alginates | E400-E405 |
| | 19. Olestra | |
| | 20. Polydextrose | E1200 |
| | 21. Whey protein concentrate | |
| **Flavor (excluding sweetness)** | | |
| | 1. Artificial flavor | |
| | 2. Natural flavor | |
| | 3. Disodium guanylate | E627 |
| | 4. Disodium inosinate | E631 |
| | 5. Autolyzed yeast extract | |
| | 6. Hydrolyzed vegetable proteins | |
| **Sweetness** | | |
| | 1. High fructose corn syrup | |
| | 2. Sorbitol | E420 |
| | 3. Mannitol | E421 |
| | 4. Xylitol | E967 |
| | 5. Erythritol | E968 |
| | 6. Maltitol | E965 |
| | 7. Polyglycitol syrup | E964 |
| | 8. Cane juice | |
| | 9. Corn syrup solids | |
| | 10. Dextrin | E1400 |
| | 11. Dextrose | |
| | 12. Fructose | |
| | 13. Fruit juice concentrate | |
| | 14. Glucose | |
| | 15. Invert sugar | |
| | 16. Maltodextrin | |
| | 17. Maltol | E636 |
| | 18. Refiner's syrup | |
| | 19. Rice syrup | |
| | 20. Maltose | |
| | 21. Acesulfame potassium | E950 |
| | 22. Aspartame | E951 |
| | 23. Neotame | E961 |
| | 24. Saccharin | E954 |
| | 25. Stevia | E960 |
| | 26. Sucralose | E955 |
| **Appearance** | 1. FD&C blue no. 1 | E133 |
| | 2. FD&C blue no. 2 | E132 |
| | 3. FD&C green no. 3 | E143 |
| | 4. FD&C red no. 3 | E127 |
| | 5. FD&C red no. 40 | E129 |
| | 6. FD&C yellow no. 5 | E102 |
| | 7. FD&C yellow no. 6 | E110 |
| | 8. Artificial color | |
| | 9. Titanium dioxide | E171 |
| | 10. Caramel color | E150 |
| | 11. Carnauba wax | E903 |

(Continued)

TABLE 1 | Continued

[a]Not all additives that we included in our search have E numbers.

OFF contributors have developed a multilingual ingredients and additives taxonomy that contains different names and synonyms for additives in many languages. Analysts then use a text parser to analyze ingredient lists and search for specific additives and classes of additives that would identify the food as "ultra-processed," based on published literature describing the NOVA classification system.

We compared results among 125,502 barcode-matched foods for 43 of our 64 selected SRIAs that had E numbers. For each SRIA, we determined the number of foods identified as having that additive by our search vs. by OFF. We then calculated the proportion of OFF-identified foods that our search missed, and the proportion of foods we identified that OFF missed.

In an initial comparison across the 43 SRIAs, the median percent of foods identified by our search as having a given SRIA missed by OFF was 37.9%. In contrast, the median percent of foods that OFF identified as having a given SRIA missed in our search was 1.9%. For example, of 6,285 foods we identified as containing the colorant FD&C yellow #6, OFF missed 2912 (46.3%). In contrast, our search missed 387 (10.3%) of the 3,769 foods OFF identified as containing FD&C yellow #6. OFF missed at least half of all the foods we identified as containing a given additive for 15 of the 43 SRIAs, and at least 5% of foods for all SRIAs except one. Our search missed at least half of OFF-identified foods for two of the 43 SRIAs, and at least 5% for eleven of the SRIAs.

We looked more closely at eleven SRIAs for which we missed >5% of foods captured by OFF. For each of these eleven SRIAs,

**TABLE 2 |** Description and distribution of 15 food groups (representing seven food categories) used in analysis of Branded Food Products Database food items ($N = 224{,}098$).

| Food group | Description | *N* (%) |
|---|---|---|
| **Grains** | • Breads and buns<br>• Pastas, noodles, rice<br>• Flours, doughs, crusts<br>• Cereals and cereal products<br>• Stuffing | 21,801 (9.7) |
| **Vegetables** | • Prepared/processed/pre-packaged vegetables<br>• Canned or frozen vegetables or beans | 12,782 (5.7) |
| **Fruits and juices** | • Prepared/processed fruit<br>• Canned or frozen fruit<br>• Fruit juices, concentrates, nectars | 8,873 (4.0) |
| **Milk products** | | |
| Cheese | – | 10,290 (4.6) |
| Milk, cream, yogurt | – | 6,613 (3.0) |
| **Meat and protein** | | |
| Meats | • Prepared/processed beef, pork, or poultry<br>• Frozen meat, poultry, patties, burgers, sausages<br>• Canned meat<br>• Sausages, hot dogs, cold cuts | 9,904 (4.4) |
| Fish and seafood | • Prepared/processed fish<br>• Canned or frozen fish or seafood | 4,260 (1.9) |
| Nuts | • Nuts<br>• Nut and seed butters | 1,482 (0.7) |
| Ready-to-eat foods | • Prepared/packaged/ready-made meals, dishes, sandwiches, salads, and pizzas<br>• Frozen dinners, entrees, sides, breakfast foods<br>• Canned soups and stews | 18,005 (8.0) |
| **Other foods** | | |
| Dressings/condiments | • Dressings, mayonnaise, ketchup, mustard, sauces, dips<br>• Oils, butter, spreads<br>• Pickles<br>• Herbs, spices, seasonings | 35,253 (15.7) |
| **Beverages** | | |
| Coffee and tea | – | 2,294 (1.0) |
| Water | – | 2,724 (1.2) |
| All other non-alcoholic | • Soda, soft drinks<br>• Energy, protein, sports drinks | 9,696 (4.3) |
| Sweets | • Cookies, cakes, pastries, pies, puddings<br>• Frozen desserts<br>• Confections, chocolate, candy<br>• Fruit spreads<br>• Dessert sauces and toppings<br>• Sugar | 49,039 (21.9) |
| Salty snacks | • Chips, pretzels, crackers, popcorn, snack bars | 31,082 (13.9) |

we examined a subset of up to 50 foods identified by OFF as having the SRIA but not identified by our search, specifically examining their ingredient lists to determine how our search might have failed to see the SRIA. This led us to revise our search for some additives—for example, expanding our search for steviol glycoside to include rebaudioside A, and our search for polyglycerol polyricinoleate to include "pgpr." After these modifications, the median percent of foods missed by our search decreased to 1.7%. A review of four remaining SRIAs for which we missed >5% of OFF-identified foods showed that our search correctly excluded the food >96% of the time. For example, a review of all instances showed that our search was correct in 242 (98%) of 247 instances for blue 1, and 372 (96%) of 387 instances for yellow 6. Examples of when our search incorrectly excluded the food were typographical errors ("blue !" instead of blue 1, or "yellow 5 and 36" instead of yellow 5 and 6), or instances in which an E number was listed without the letter "E" immediately preceding it.

## Analysis

For each food, we determined the total number of SRIAs (0–64) and the number of different types of SRIAs (0–4, i.e., additives used for texture, flavor, sweetness, and appearance). We also calculated proportions with at least one SRIA, each different type of SRIA, at least three SRIAs, and at least three of the four types of SRIAs, overall and by food group. The last two (3+ SRIAs, 3+ types of SRIAs) were arbitrarily selected as indicators of higher degree of ultra-processing.

# RESULTS

The percentages of foods with SRIAs are shown in **Table 3** and **Figure 2**. Of 241,688 foods in the BFPD, almost two-thirds (64.9%) contained at least one SRIA from our list of 64 (**Figure 2**). Of these, over half (58.9%) had three or more SRIAs (**Figure 2**). Overall, the most prevalent SRIAs were for texture, which were present in 45.1% of all foods, followed by flavor (42.2%), sweetness (38.4%), then appearance (19.8%) (**Table 3**). Over a quarter (26.5%) of all foods contained at least three of the four SRIA types investigated (**Table 3**).

Among foods with information on food group, sweets (89.5%), non-alcoholic beverages (84.9%), and ready-to-eat (RTE) foods (82.0%) were the most likely to contain SRIAs (**Table 4**). Over half of the items in these categories contained three or more SRIAs. In contrast, 17.9% of items in the vegetable category contained SRIAs, and 8.2% contained three or more. Sweets (50.4%), non-alcoholic beverages (45.7%), and RTE foods (37.2%) were also the most likely to contain at least three of the four SRIA types, compared with <10% of fruits/juices, nuts, cheese, and vegetables.

Prevalence of each of the four types of SRIAs and the specific SRIAs occurring in at least 10% of foods are shown in **Supplementary Tables 2–5**. Sweets were the most likely to contain texture related SRIAs, with 72.3% containing at least one (**Table 4**). The most common texture related SRIA used in sweets was lecithin, present in almost half of all products (**Supplementary Table 2**). Over half of RTE and milk products

**TABLE 3 |** Prevalence of sensory-related industrial additives (SRIA) and types of SRIA in food items in the Branded Food Products Database (*n* = 241,688).

| | *N* | % |
|---|---|---|
| Any SRIAs | 156,738 | 64.9 |
| **Number of SRIAs** | | |
| 1 | 37,060 | 15.3 |
| 2 | 27,430 | 11.4 |
| 3–4 | 38,305 | 15.8 |
| 5–7 | 30,637 | 12.7 |
| ≥8 | 23,306 | 9.6 |
| **Any SRIA related to** | | |
| Texture | 108,907 | 45.1 |
| Flavor[a] | 101,969 | 42.2 |
| Sweetness | 92,887 | 38.4 |
| Appearance | 47,941 | 19.8 |
| **Number of SRIA types** | | |
| 1 | 48,473 | 20.1 |
| 2 | 44,134 | 18.3 |
| 3 | 41,561 | 17.2 |
| 4 | 22,570 | 9.3 |

[a]*Excludes sweetness.*

contained texture related SRIAs, with modified starches being the most common in RTE products, and carrageenan and modified starches the most common in milk products. The prevalence of texture related SRIAs was <10% for coffee and tea, water, and fruits and juices.

Flavor-related SRIAs were most prevalent in non-alcoholic beverages (71.6%), water products (67.1%), and sweets (66.2%). Natural flavors were by far the most common flavor related SRIA across food categories (**Supplementary Table 3**). Sweeteners were also the most prevalent in non-alcoholic beverages (63.6%), with high fructose corn syrup being the most common (**Supplementary Table 4**). Appearance-related SRIAs were present in <10% of foods for nine of the 15 food categories (**Supplementary Table 5**). However, they occurred in 42.2% of sweets/desserts, with the most common SRIAs being FD&C dyes.

# DISCUSSION

## Summary and Discussion of Possible Mechanisms

Primary findings from this analysis were that most foods in the BFPD contained at least one SRIA, and over a third of foods had at least three. More than 80% of sweets, non-alcoholic beverages, and RTEs contained SRIAs. Additionally, more than a quarter of all food items in the BFPD had at least three of the four different categories of SRIAs examined, with texture related additives being the most common. Among sweets, 20% had all four categories.

Our analysis also points to specific SRIAs that appear the most frequently across food groups. For example, SRIAs that occurred the most frequently in sweets included lecithin, natural and artificial flavors, modified starches, mono/diglycerides, and dextrose. In RTE foods, frequently occurring SRIAs were natural flavors, modified food starches, xanthan gum, dextrose, and maltodextrin.

Industrial additives are frequently mentioned as a concern in consuming UPFs. A primary mechanism linking UPFs to disease risk is through enhanced sensory qualities that lead to overconsumption (18, 19). In a study conducted by Hall et al. (18), 20 adults were randomized to receive either ultra-processed or unprocessed diets for two weeks, followed by the alternate diet for two weeks. The two diets were matched for energy density and energy, macronutrient, fiber, sugar, and sodium content, and participants were instructed to consume as much or as little as desired. Hall et al. found that while participants lost weight during the unprocessed diet, they had greater energy intake and gained weight during the ultra-processed diet; this occurred despite no significant difference in reported palatability of the meals.

Small and DiFeliceantonio (19) suggest that additives in UPFs might co-opt existing pathways in which metabolic signals after consuming a food are conveyed from the gastrointestinal system to the brain. There, a rise in dopamine, part of the reward circuit response, reinforces the value of consuming that food, producing a reward response disproportionate to the food's caloric or nutrient content and thereby encouraging overconsumption. Notably, neural processing of these reinforcing signals appears to be independent of conscious perceptions about food, such as ratings of food liking or sensory pleasure. As a second possible mechanism, Fazzino et al. (28) suggest that foods with "multiple palatability-inducing ingredients" might weaken the sensory-specific satiety response, resulting in delayed signals for eating cessation and leading again to overconsumption (29). The extent to which these mechanisms are due to the presence of SRIAs, potentially combined with their relatively higher sodium and sugar content (30), has also yet to be explored, with implications for identification of specific additives of potential concern and product reformulation.

Worth noting is the possibility of product reformulation to *reduce* overconsumption. A recent pooled analysis found that energy intake rates are higher with consumption of UPF compared with less processed foods (31). In addition to suggesting a mechanism by which UPF might increase consumption independent of their SRIA content, the finding also suggests the possibility of reformulating UPFs to reduce intake rates—for example, by altering their form or texture (31, 32), taking advantage of technologies made possible with "ultra-processing".

## Comparison With Related Studies

Baldridge et al. (25) classified >230,000 food and beverage products in Label Insight's Open Data database into 59 food categories, mapping each to NOVA categories representing levels of processing. Overall, they found that 70.9% of the 230,156 food and beverage products were ultra-processed. Ultra-processed food products comprised over 90% of convenience foods, sauces/dressings, bread/bakery products, and 100% of snack foods and confectionery. The discrepancy in results between the Baldridge et al.'s and our study is most likely because Baldridge

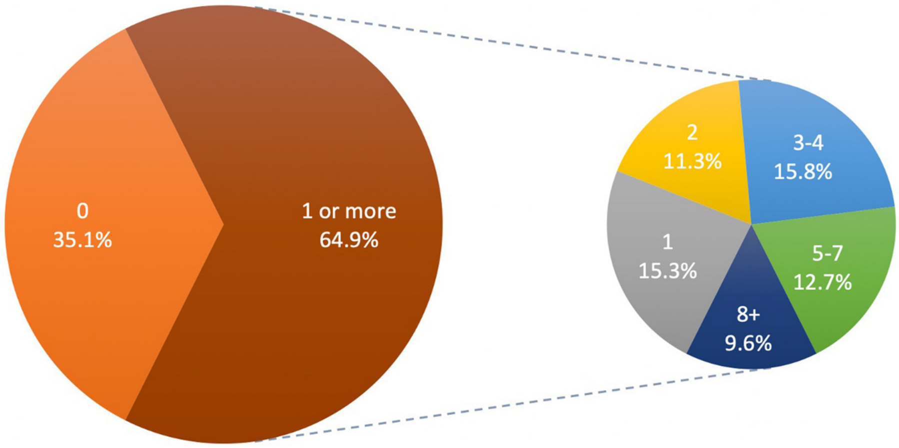


**FIGURE 2 |** Proportions of foods in the Branded Food Products Database ($n = 241,688$) with one or more sensory-related industrial additives (SRIAs).

**TABLE 4 |** Prevalence (%) of sensory-related industrial additives (SRIA) in food items in the Branded Food Products Database, by food group ($n = 224,098$).

| | ≥1 SRIA[a] | ≥3 SRIA[a] | Texture | Flavor | Sweetness | Appearance | ≥3 SRIA types[b] |
|---|---|---|---|---|---|---|---|
| Sweets | **89.5** | **68.1** | **72.3** | **66.2** | **53.5** | 42.2 | **50.4** |
| Beverages | **84.9** | **57.7** | 41.1 | **71.6** | **63.6** | 36.4 | 45.7 |
| Ready-to-eat foods | **82.0** | **53.3** | **64.6** | **52.8** | **54.0** | 20.8 | 37.2 |
| Water | **70.0** | 26.0 | 9.4 | 67.1 | 36.8 | 9.6 | 12.0 |
| Meats | **66.0** | 19.0 | 25.0 | 34.2 | 46.3 | 7.5 | 11.5 |
| Milk/cream/yogurt | **65.1** | 43.2 | **56.0** | 51.2 | 34.1 | 8.7 | 29.0 |
| Dressings/condiments | **54.8** | 25.4 | 38.0 | 32.6 | 26.7 | 15.6 | 18.2 |
| Salty snacks | **56.7** | 29.6 | 34.3 | 36.0 | 36.8 | 13.4 | 20.5 |
| Fruits and juices | **55.2** | 13.9 | 7.6 | 31.5 | 45.3 | 9.2 | 8.0 |
| Coffee and tea | **53.7** | 16.0 | 9.7 | 43.4 | 28.8 | 12.3 | 10.1 |
| Grains | **52.0** | 28.0 | 40.7 | 22.7 | 24.8 | 9.6 | 12.6 |
| Fish | 45.5 | 16.3 | 39.8 | 15.5 | 16.2 | 4.3 | 10.1 |
| Nuts | 42.4 | 5.4 | 36.0 | 7.8 | 13.7 | 0.3 | 1.3 |
| Cheese | 40.5 | 15.7 | 37.8 | 12.0 | 8.2 | 4.8 | 6.7 |
| Vegetables | 17.9 | 8.2 | 11.2 | 12.7 | 8.3 | 3.5 | 5.9 |

*Boldface indicates prevalence ≥50%.*
[a]*Number of individual SRIAs, with a possible maximum of 64.*
[b]*Number of types of SRIAs, with a possible maximum of four (texture, flavor, sweetness, appearance).*

et al. assigned entire categories of food to NOVA categories. In contrast, our estimates are based on actual ingredients listed. It is worth noting, however, that our findings are broadly similar: a significant majority of packaged food and beverage products in the US are "ultra-processed" according to the NOVA framework, including most sweets, beverages, convenience foods, meat, dairy, sauces/dressings, and snacks.

Dunford et al. (33) used a similar approach to our study—compiling a list of terms to search for in ingredient lists—to determine the proportion of branded food and beverage products containing non-nutritive sweeteners, using data from Label Insight's Open Data initiative. They found that the percent of products containing non-nutritive sweeteners in the US was 4.37%, close to our estimate of 4.29% (not shown). Dunford et al. also found that ∼30–40% of soda, sports, and water drinks contained non-nutritive sweeteners, comparable with our estimates for beverages and waters. These findings indicate that a strategy of searching for additives of interest in ingredient lists can produce similar results, even when exact search terms vary. Of note, applying the same approach to Nielsen Homescan Consumer Panel data, Dunford et al. (34) found substantial changes in the prevalence of households purchasing non-nutritive sweeteners over time and differences by race/ethnicity, suggesting the additional insight gained from examining purchase data.

Chazelas et al. (21) examined additives in over 126,000 food and beverage products in the French market available in OFF. As described above in our Methods section, OFF searches parsed text in ingredient lists for specific additives (e.g., whey, invert sugar) or classes of additives (e.g., sequestrants, glazing agents) that are indicative of ultra-processed foods. They found that 53.8% of foods in the French packaged food supply contained at least one additive, with the highest proportions occurring in beverages, sweets, and convenience foods. Although Chazelas et al. did not limit their search to sensory related additives but included other categories such as preservatives, their estimate is lower than our estimate of 63.8%. The lower prevalence may be due to differences in the packaged food supply between France and the US, or differences in search procedures. Consistent with our findings, however, texture-related additives, such as lecithins, modified starches, and xanthan gum, were among the most frequently occurring additives in ingredient lists.

Batada et al. (35) estimated the prevalence of artificial food colors in 810 food and beverage products marketed toward children in one major supermarket in North Carolina. They found that 43.2% of products contained artificial food colors, with the most common being red 40 (29.8%), blue 1 (24.2%), yellow 5 (20.5), and yellow 6 (19.5%). Product categories with the highest percentages of artificial food colors were candies (96.3%), fruit-flavored snacks (94.7%), drink mixes/powders (89.7%), frozen breakfasts (85.7%), and toaster pastries (66.7%). The substantially higher percentages found by Batada et al. indicate the importance of appearance related SRIAs in marketing to children. Batada et al. (35) selected products only if they displayed a cartoon character or "bright and bubbly, child-friendly lettering" on the front of the package; advertised a child-oriented prize or incentive; and/or were thought to be a traditional children's item (e.g., fruit-flavored snacks).

## Strengths and Limitations

A limitation of this study is that it did not include all possible SRIAs due to the labor-intensive nature of searching ingredient lists. While we focused on only 64 SRIAs, E numbers exist for 41 colors, 19 sweeteners, 63 emulsifiers, stabilizers, thickeners, and gelling agents, and over 150 other additives with potentially relevant functions related to texture, flavor, and appearance (36). We chose to focus on the additives that appeared most indicative of industrial food production, and not typically used in domestic or culinary settings. However, our method is expandable; our shared program may be revised to include other SRIAs as deemed appropriate, and to be applied to other databases, including more updated versions of the BFPD, or even purchase data if linked to ingredient lists.

The prevalence of SRIAs may be underestimated because of incomplete ingredient lists, as well as variability in how the same SRIA can be written into ingredient lists, despite specific guidance to industry for food labeling (37). In addition, the BFPD does not represent all food products in the US marketplace. However, as the largest such analysis with over 240,000 items, our study offers a meaningful characterization of the US retail packaged food supply. Our results are relatively consistent with findings using other methods to identify additives or ultra-processed foods. Taken together, these studies collectively demonstrate the ubiquity of SRIAs in packaged food products, and most particularly in sweets, beverages, and RTE foods.

Although individual SRIAs have been linked to adverse health effects (17, 38, 39), estimating dosage or amounts of specific additives ingested by consumers was outside the scope of this study. However, our work is intended to build on previous studies on ultra-processed foods that treat ultra-processing as a binary characteristic. A binary classification system for processed foods is a potentially simple basis for developing guidelines for consumers and policymakers. But consumers may find some utility in distinguishing between UPFs containing one industrial sweetener from UPFs containing SRIAs that alter texture, flavor, and appearance. In research, distinguishing among foods that are more vs. less ultra-processed based on the number of SRIAs in a food offers a strategy for examining gradients in risk with degree of this aspect of ultra-processing. Considering the purpose of the additive is also potentially important. For example, Drewnowski (40) noted that >90% of plant-based alternatives to dairy milk contain industrial ingredients, using a similar search strategy as the one used here, but also including vitamins, minerals, and preservatives. Whether risk differs according to the number and types of additives in a food is a question that has not been addressed in previous research. Our approach offers a relatively straightforward way to quantify degree of one aspect of ultra-processing, using the presence of SRIAs as an indicator of the sensory qualities of a food.

Strengths of this study include its use of a large, publicly available database, and the development of open-source code meant to serve as a basis for other work. In addition, given the impracticality of recommending reduced intake of as broad and heterogeneous a class of foods as UPFs, quantifying the presence of SRIAs is a novel concept with implications for research exploring more specific mechanisms by which UPFs might affect health (41), and for developing actionable recommendations with respect to both consumer choice and industry product reformulation.

## Future Research

Our findings suggest three directions for future research. First, comparing this method with existing frameworks to classify UPFs will be useful both to quantify level of agreement across methods and to assess their relative usefulness in predicting disease risk. Second, a detailed description of correlations between SRIA occurrence and sodium, sugar, and fat content will help address whether and how SRIA and nutrient effects on disease risk might be disentangled. Third, examining the presence of SRIAs in purchase and intake data will provide necessary information on the level at which the US population is buying and consuming foods with SRIAs.

# CONCLUSION

In summary, our examination of the largest available database of packaged foods in the US confirms the ubiquity of SRIAs, especially in sweets, beverages, and RTE foods, but in all food groups to some extent. Quantifying the presence of SRIAs in foods offers a new approach to examine the aggregate effects of SRIAs on health outcomes, and to examine the assumption that UPFs increase disease risk through the presence of SRIAs and through effects on sensory qualities.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: Branded Food Products Database https://fdc.nal.usda.gov; Open Food Facts https://us.openfoodfacts.org.

## AUTHOR CONTRIBUTIONS

MT conceptualized the project, led the data analyses, contributed to data interpretation, and drafted the manuscript. CG assisted in selection of additives to include, developed the initial coding to search for text, investigated and edited ingredient text data prior to searching, conducted the comparisons with Open Food Facts, and contributed to data analyses and interpretation. AA assisted in selection of additives to include, developed the final coding to correct and search for text, reviewed output for errors and oversights, and contributed to data analyses and interpretation. SA provided input on sweeteners to include in searches. AN contributed to conceptualization and drafting the manuscript. All authors contributed to the article and approved the submitted version.


## FUNDING

This research was generously supported in part by the William and Linda Frost Fund in the Cal Poly College of Science and Mathematics.

## ACKNOWLEDGMENTS

The authors thank Miriam Silliman, Grace Vandervort, and Geth Wu for their assistance in preliminary background research and analyses. The authors also thank Stéphane Gigandet for his generosity in sharing his expertise, including the search strategies and code used for Open Food Facts.


## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.762814/full#supplementary-material

## REFERENCES


1. Moradi S, Entezari M, Mohammadi H, Jayedi A, Lazaridi A, Kermani M, et al. Ultra-processed food consumption and adult obesity risk: a systematic review and dose-response meta-analysis. *Crit Rev Food Sci Nutr*. (2021). doi: 10.1080/10408398.2021.1946005. [Epub ahead of print].
2. Juul F, Vaidean G, Lin Y, Deierlein ANP. Ultra-processed foods and incident cardiovascular disease in the framingham offspring study. *J Am Coll Cardiol.* (2021) 77:1520–31. doi: 10.1016/j.jacc.2021.01.047
3. Zhong G, Gu H, Peng Y, Wang K, Wu Y, Hu T, et al. Association of ultra-processed food consumption with cardiovascular mortality in the US population: long-term results from a large prospective multicenter study. *Int J Behav Nutr Phys Act.* (2021) 18:14. doi: 10.1186/s12966-021-01081-3
4. Chen X, Zhang Z, Yang H, Qiu P, Wang H, Wang F, et al. Consumption of ultra-processed foods and health outcomes: a systematic review of epidemiological studies. *Nutr J.* (2020) 19:10. doi: 10.1186/s12937-020-00604-1
5. Elizabeth L, Machado P, Zinöcker M, Baker P, Lawrence M. Ultra-processed foods and health outcomes: a narrative review. *Nutrients.* (2020) 12:33. doi: 10.3390/nu12071955
6. Fiolet T, Srour B, Sellem L, Kesse-Guyot E, Allès B, Méjean C, et al. Consumption of ultra-processed foods and cancer risk: results from NutriNet-Santé prospective cohort. *BMJ.* (2018) 360:11. doi: 10.1136/bmj.k322
7. Sadler CR, Grassby T, Hart K, Raats M, Solović M, Timotijevic L. Processed food classification: conceptualisation and challenges. *Trends Food Sci Technol*. (2021) 112:149–62. doi: 10.1016/j.tifs.2021.02.059
8. Crino M, Barakat T, Trevena H, Neal B. Systematic review and comparison of classification frameworks describing the degree of food processing. *Nutr Food Technol.* (2017) 3:12. doi: 10.16966/2470-6086.138
9. Gibney MJ, Forde CG, Mullally D, Gibney ER. Ultra-processed foods in human health: a critical appraisal. *Am J Clin Nutr.* (2017) 106:717–24. doi: 10.3945/ajcn.117.160440
10. Bleiweiss-Sande R, Chui K, Evans W, Goldberg J, Amin S, Sacheck J. Robustness of food processing classification systems. *Nutrients.* (2019) 11:21. doi: 10.3390/nu11061344
11. Monteiro CA, Cannon G, Levy RB, Moubarac JC, Louzada ML, Rauber F, et al. Ultra-processed foods: what they are and how to identify them. *Public Health Nutr*. (2019) 22:936–41. doi: 10.1017/S1368980018003762
12. US Food and Drug Administration (FDA). *International Food Information Council (IFIC), Overview of Food Ingredients, Additives & Colors. U.S. Food & Drug Administration*. (2010). Available online at: https://www.fda.gov/food/food-ingredients-packaging/overview-food-ingredients-additives-colors#foodadd (accessed June 10, 2019).
13. European Food Safety Authority. *Food Additives: Introduction*. Available online at: https://www.efsa.europa.eu/en/topics/topic/food-additives (accessed June 10, 2019).
14. International Food Information Council Foundation. *What's in Our Food: Understanding Common Food Ingredients*. Available online at: https://foodinsight.org/wp-content/uploads/2014/08/FINAL-WIF-8-22-12.pdf (accessed June 19, 2019).
15. Kemp A. Food additives and hyperactivity. *BMJ.* (2008) 336:1144. doi: 10.1136/bmj.39582.375336.BE
16. Marion-Letellier R, Amamou A, Savoye G, Ghosh S. Inflammatory bowel diseases and food additives: to add fuel on the flames!. *Nutrients*. (2019) 11:1111. doi: 10.3390/nu11051111
17. Chassaing B, Koren O, Goodrich J, Poole A, Srinivasan S, Ley R, et al. Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. *Nature.* (2015) 519:92–6. doi: 10.1038/nature14232
18. Hall KD, Ayuketah A, Brychta R, Cai H, Cassimatis T, Chen KY, et al. Ultra-processed diets cause excess calorie intake and weight gain: an inpatient randomized controlled trial of ad libitum food intake. *Cell Metab.* (2019) 30:67–77.e63. doi: 10.1016/j.cmet.2019.05.020
19. Small DM, DiFeliceantonio AG. Processed foods and food reward. *Science.* (2019) 363:346–7. doi: 10.1126/science.aav0556
20. SugarScience. *Hidden in Plain Sight: 61 Names for Sugar*. Available onine at: https://sugarscience.ucsf.edu/hidden-in-plain-sight/#.YNuMsy9h3OR (accessed June 19, 2019).
21. Chazelas E, Deschasaux M, Srour B, Kesse-Guyot E, Julia C, Alles B, et al. Food additives: distribution and co-occurrence in 126,000 food products of the French market. *Sci Rep.* (2020) 10:3980. doi: 10.1038/s41598-020-60948-w
22. Stichting Food-Info. *Food-Info*. Available online at: http://food-info.net (accessed March 1, 2021).
23. UK Food Standards Agency. *Approved Additives and E Numbers*. Available online at: https://www.food.gov.uk/business-guidance/approved-additives-and-e-numbers#emulsifiers-stabilisers-thickeners-and-gelling-agents (accessed March 1, 2021).
24. Kretser A, Murphy D, Starke-Reed P. A partnership for public health: USDA branded food products database. *J Food Comp Analysis.* (2017) 64:10–12. doi: 10.1016/j.jfca.2017.07.019

25. Baldridge AS, Huffman MD, Taylor F, Xavier D, Bright B, Van Horn LV, et al. The healthfulness of the US packaged food and beverage supply: a cross-sectional study. *Nutrients.* (2019) 11:15. doi: 10.3390/nu11081704
26. U.S. Department of Agriculture ERS. *National Household Food Acquisition and Purchase Survey (FoodAPS): Nutrient Coding Overview*. Washington, DC: U.S. Department of Agriculture (2016).
27. *Open Food Facts [database on the Internet]*. Available online at: https://world.openfoodfacts.org (accessed April 18, 2020).
28. Fazzino TL, Rohde K, Sullivan DK. Hyper-palatable foods: development of a quantitative definition and application to the US food system database. *Obesity.* (2019) 27:1761–68. doi: 10.1002/oby.22639
29. de Macedo IC, de Freitas JS, da Silva Torres IL. The influence of palatable diets in reward system activation. *A Mini Review*. *Adv Pharmacol Sci*. (2016) 2016:7238679. doi: 10.1155/2016/7238679
30. Tseng M, Neill DB, Teaford SF, Nazmi A. Alternative myplate menus: effects of ultra-processed foods on saturated fat, sugar, and sodium content. *J Nutr Educ Behav.* (2017) 50:258-266.e251. doi: 10.1016/j.jneb.2017.10.009
31. Forde CG, Mars M, de Graaf K. Ultra-processinng or oral processing? A role for energy density and eating rate in moderating energy intake from processed foods. *Curr Dev Nutr*. (2020) 4:7. doi: 10.1093/cdn/nzaa019
32. Rolls BJ, Cunningham PM, Diktas HE. Properties of ultraprocessed foods that can drive excess intake. *Nutr Today*. (2020) 55:109–15. doi: 10.1097/NT.0000000000000410
33. Dunford EK, Taillie LS, Miles DR, Eyles H, Tolentino-Mayo L, Ng SW. Non-nutritive sweeteners in the packaged food supply-an assessment across 4 countries. *Nutrients*. (2018) 10:13. doi: 10.3390/nu10020257
34. Dunford E, Miles DR, Ng SW, Popkin BM. Types and amounts of non-nutritive sweeteners purchased by US households: a comparison of 2002 and 2018 nielsen homescan purchases. *J Acad Nutr Diet.* (2020) 120:1662–71. doi: 10.1016/j.jand.2020.04.022
35. Batada A, Jacobson MF. Prevalence of artificial food colors in grocery store products marketed to children. *Clin Pediatr.* (2016) 55:1113–9. doi: 10.1177/0009922816651621
36. U.K. Food Standards Agency. *Approved Additives and E Numbers*. Available online at: https://www.food.gov.uk/business-guidance/approved-additives-and-e-numbers (accessed July 1, 2021).
37. US Department of Health Human Services, Food Drug Administration, Nutrition CfFSaA. *A Food Labeling Guide: Guidance for Industry*. College Park, MD: Office of Nutritiion, Labeling, and Dietary Supplements (2013).
38. Neto H, Ausina P, Gomez L, Leandro J, Zancan P, Mauro Sola-Penna M. Effects of food additives on immune cells as contributors to body weight gain and immune-mediated metabolic dysregulation. *Front Immunol.* (2017) 8:11. doi: 10.3389/fimmu.2017.01478
39. Zhu Y, Olsen S, Pauline Mendola P, Halldorsson T, Rawal S, Hinkle S, et al. Maternal consumption of artificially sweetened beverages during pregnancy, and offspring growth through 7 years of age: a prospective cohort study. *Int J Epidemiol.* (2017) 46:1499–508. doi: 10.1093/ije/dyx095
40. Drewnowski A. Perspective: identifying ultra-processed plant-based milk alternatives in the USDA branded food products database. *Adv Nutr*. (2021) 12:2068–75. doi: 10.1093/advances/nmab089
41. Tobias DK, Hall KD. Eliminate or reformulate ultra-processed foods? Biological mechanisms matter. *Cell Metab*. (2021) 33:2314–5. doi: 10.1016/j.cmet.2021.10.005


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

DATA REPORT
published: 19 January 2022
doi: 10.3389/fnut.2021.800317

# The Malaysian Food Barometer Open Database: An Invitation to Study the Modernization of Malaysian Food Patterns and Its Economic and Health Consequences

*Jean-Pierre Poulain*[1,2,3,4]*, *Laurence Tibère*[1,2,3,4], *Elise Mognard*[2,3,4], *Cyrille Laporte*[1,2,3,4], *Tristan Fournier*[5], *Ismail Mohd Noor*[4], *Anindita Dasgupta*[4,6], *Yasmine Alem*[1,2,3,4], *Kremlasen Naidoo*[1,2,3,4], *Anne Dupuy*[1,2,3], *Amandine Rochedy*[1,2,3], *Pradeep Kumar Nair*[4] and *Neethianhantan Ari Ragavan*[1,4,5]

[1] Centre d'Études et de Recherche: Travail, Organisation, Pouvoir (CERTOP) UMR CNRS 5044, Université de Toulouse, Toulouse, France, [2] Chair ≪ Food Studies: Food, Cultures & Health ≫, Université de Toulouse, Toulouse, France, [3] Faculty of Social Sciences and Leisure Management, University Malaysia, Subang Jaya, Malaysia, [4] Center for Asian Modernisation Studies (CAMS), Taylor's University, Subang Jaya, Malaysia, [5] Institut de Recherche Interdisciplinaire sur les Enjeux Sociaux (IRIS), Centre National de la Recherche Scientifique (CNRS), Paris, France, [6] Faculty of Social Sciences and Leisure Management, Taylor's University, Subang Jaya, Malaysia

OPEN ACCESS

**Edited by:**
*Alessandra Durazzo, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Mohamed Merdji, Audencia Nantes School of Management, France*
*Bhanu Prakash Bylapudi, HCG Cancer Hospital, India*

***Correspondence:**
*Jean-Pierre Poulain*
*poulain@univ-tlse2.fr*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** 22 October 2021
**Accepted:** 06 December 2021
**Published:** 19 January 2022

**Citation:**
*Poulain J-P, Tibère L, Mognard E, Laporte C, Fournier T, Noor IM, Dasgupta A, Alem Y, Naidoo K, Dupuy A, Rochedy A, Nair PK and Ragavan NA (2022) The Malaysian Food Barometer Open Database: An Invitation to Study the Modernization of Malaysian Food Patterns and Its Economic and Health Consequences. Front. Nutr. 8:800317. doi: 10.3389/fnut.2021.800317*

**Keywords: open science, sociology of food, health, nutrition, meals norms, food practices, food transition, compacted modernization**

## INTRODUCTION

The study was conducted within the framework of the chair of "Food Studies: Food, Cultures and Health" created jointly by Taylor's University (Malaysia) and the University of Toulouse Jean Jaurès (France). The Malaysian Food Barometer (MFB) is a recurrent survey focusing on the transformations of food patterns of the Malaysian population. MFB works with national representative samples from a socio-anthropological theoretical standpoint.

Since the 1990s, two major nutrition surveys namely the first Malaysian Adult Nutrition Survey MANS 2003 (1), which was repeated in 2014. However, these studies focus on individual food consumption, assessed in terms of nutritional composition of the diet, including some basic sociodemographical characteristics. MFB is a complementary survey to MANS (2003 and 2014) that explore in depth the sociological and ethnological dimensions of Malaysian food habits. The MFB collects face-to-face data on food-related practices including social norms, attitudes, cultural representations, and routines and their supposed sociocultural and demographic determinants. The focus includes the repartition of food practices at home and away from home, sources of food, food socialization of the meals and food intakes, food temporality, and perception of risks in food. A series of indexes were assigned to individuals within the phenomena of transition (demographic, nutrition, food, and protein) and of "compacted modernization" experienced by some Asian countries (2–5).

The theoretical objective of the project consists of studying the social, ethnic, and cultural diversification of food patterns of Malaysians and its changes in the context of modernization. The applied objective of the research is to analyze: first the impact of these transformations on the technical and economic organization of the food industry and service sector and second the consequences of changes in food consumption in terms of public health, especially on non-communicable diseases.

## MALAYSIAN SOCIETY

The study of Malaysian eating patterns is a scientific challenge due to the multicultural character of the society coupled with the fact that it is facing a very rapid modernization. This phenomenon that we can call "compacted" modernization (6, 7) that impacts the ways to eat and the values system attached to food in a given population (8). Compacted modernity is characterized by some sociodemographic transformations, linked to the demographic transition like rapid urbanization, decrease of fertility, and reduction in the size of the household including some socioeconomic changes like increase in purchasing power and emergence of a middle class (9, 10). The epidemiological transition can be referred to as changes and the causes of mortality from epidemic diseases to cardiovascular diseases, cancers, and degenerative diseases (11–13). All these structural transformations have affected the lifestyles and the food habits of the various ethnic groups, which make the Malaysian population (14). Many countries, including Malaysia, are facing obesity epidemic that raises concerns about the negative effects on health that has stimulated researchers to focus on the food cultures and lifestyle transformations leading to the development of this issue.

Although most of the above characteristics can be found in modern countries, Malaysian social context has in addition two distinct characteristics. The first one is its multi-ethnicity character. Officially Malaysian society consists of three main ethnic groups, i.e., Malay, Chinese, and Indian (and a few minorities groups). Each group has its own food culture with its emblematic dishes, its taboos, and restrictions, its eating rituals, its meal structures, and its symbolic dimensions of food. However, "ethnic" categories are not homogeneous in Malaysia. For example, Indians may belong to different religions: Hindus, Muslims, Sikhs, Buddhists, Christians, or members of New Religious Movements. They can also speak different languages like Bahasa Malaysia, English, and Indian languages like Tamil, Hindi, Urdu, Malayalam, etc. In addition, they may be identified by a caste, by region of origin, etc. and may live in Malaysia for several generations or just arrived. Similarly, for the "Chinese" people, they may be Buddhist, Taoist, Christian, Muslim, or free thinker. They may speak Hakka, Hokkien, Cantonese, Teochew, and Mandarin. They may have several ethnic belonging, Hakka, Cantonese, and Wu. Furthermore, there are Malaysians in the official "Others" categories, such as the non-Malay Bumiputra, Dusun, Iban, and Kadazan. We must add foreigners living in Malaysia to this heterogeneous group, including expats (executives and domestic helpers), and international retirees on the "second home" program. But boundaries between the three main groups are not totally hermetic. There is even a certain "porosity" resulting from overlapping religious affiliation and language competence and interpersonal relationships across ethnic boundaries and through friendship. This "porosity" results also from the usage behind the primary "race" identity, religious conversion, and "crossbreeding" (multi-ethnic people) from historical institutionalized mixed marriages (e.g., in the Baba-Nyonya community).

The second characteristic is the high frequency of food consumed outside home, which is one among the highest rates in world. Historically, eating out-of-home, especially street food, is an important tradition in Asian countries. With the increase of urbanization, the opportunities for Malaysians to eat out have increased greatly, and the costs are sometimes cheaper than the home-made meals.

The MFB is a tool to identify and study the sociocultural determinants of the Malaysian food habits. It focuses simultaneously on the practices and on the representations of food cultures. The aim of MFB is to help understand the food life styles and the different food contexts of the various Malaysian ethnic groups and "middle class" in order to elucidate their process of making food decisions.

## THEORETICAL FRAMEWORK

The theoretical framework of this database (see **Figure 1**) breaks with the dominant reading grid in nutrition, based on theories of rational choice or programmed action. The adopted theoretical stand postulates that a significant part of eating behavior is "unthoughtful," and that dietary decisions are embodied in behavioral scenario predefined by societies and cultures. It is what we use to call according to a long anthropological tradition "pattern." Thus, social influences can come from: (1) social positions of eaters (ethnicity, income, education, job, marital status, etc.), (2) food patterns defined by cultures (meals structures, routines, and eating contexts), and (3) social interactions with people who prepare the food and with whom the meals are taken and shared. The database, therefore, explores blind spots of traditional nutritional surveys. It opens a dialog with the nutritional sciences. "Food patterns" are "used" by eaters, and by the household members, who shop, cook, and prepare meals. They are used also by all those who directly or indirectly participate in the production, processing, and distribution of food products and services. They constitute the cognitive infrastructure, which allows these actors to coordinate and contribute to the functioning of the food system, its social "orchestration" (15).

On the question of decision, there are rooms for decision-making, but they are embedded in the structures and categories of food systems. Food models are therefore the infrastructure on which decisions are made. Eaters use these patterns like if they were "natural," and therefore they have little awareness of their existence (16, 17).

Thus, this theoretical framework stands at a distance from the classic nutritional approach, explicitly included in the theories of rational choice. Therefore, decisions of the eaters are nested in culturally defined action scenarios. There are some decisions, but they take place in a socially determined room of freedom. Then the theoretical framework considers: (1) the multicultural characteristics of Malaysian society, (2) the transformations of this society under the influence of rapid modernization ("compacted modernization"), (3) the consequences of this modernization on "food social space,"

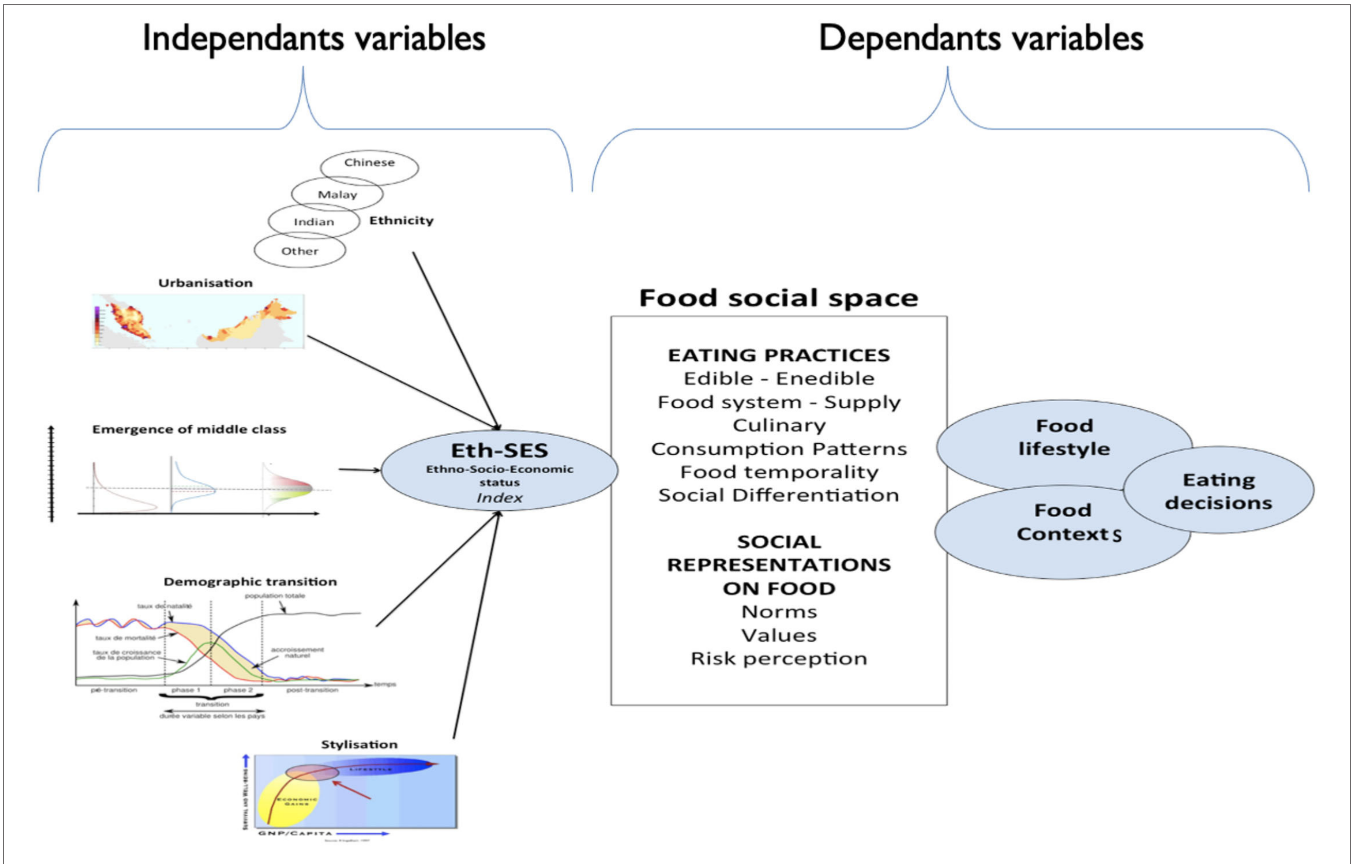


**FIGURE 1 |** Malaysian Food Barometer (MFB) conceptual framework.

both on food representations and on behavior patterns, and (4) contextualized decisions.

## METHODOLOGY

The MFB is a recurrent survey that studies the socioeconomic, demographic, and cultural determinants of food consumption. It focuses on the transformations of food practices and representations. It studies the mutations of food habits and tries to identify their possible influences on health, especially on non-transmissible diseases in which food is involved. It complements traditional nutritional surveys and aims to participate in the development of prevention programs (5). MFB is based on a mixed methodology developed at the national level. The qualitative approach precedes and prepares basis for the quantitative survey. The method is based on the distinction between social norms and practices. This approach has been developed over the last three decades in France and used in several national surveys, including the Barometer "Santé Nutrition" of the National Institute for Prevention and Health Education (INPES).

The data on eating habits are based on a "24-h recall" approach. The interviewees were asked to list all foods and drinks intake of the day before, both during and between main meals. It is quite a common method used to assess individuals food intake mainly used in nutrition science surveys, but we have adapted and modified it taking into account the sociocultural dimensions we wish to investigate. In the early 1980s, sociologists and economists were engaged in intense research activity focusing on food consumption surveys. The results of their work generated some theoretical debates, which gave rise to methodological advances that subsequently benefitted our approach to the current investigation. These include awareness of the need to consider the status of the variables and the data collection techniques.

For the quantitative research on eating practices, we encountered an obstacle related to the use of declarative methods to uncover, or at least get data as close as possible to the actual behavior of individuals. For example, when we ask individuals to describe the meals they ate the day before and if they have not eaten "as usual," or if they ate differently from their usual pattern, "what they think they should have done," they feel uncomfortable. Indeed, what should their answer be? What they actually did, or what they usually do?

The problem, methodologically, is that all individuals do not solve this dilemma in the same way. Some of them, respecting the instructions of the interviewer, faithfully describe the food intake of the day before, whereas others, eager to report their usual way of eating, are tempted to change their statement from the actual to the usual, to reduce the cognitive dissonance they feel. All seek to translate what they think is the reality of their food practices. In the second case, the data collected can be said to be more related to their perception of "social norms," which are a mix of social and nutritional requirements than to their actual practices. Thus, the data obtained have a fairly weak empirical value because it represents neither a complete picture of the real behavior of individuals, nor of the social representations (norms and values) relating to food in the social group being studied. In an attempt to resolve this ambiguity, some studies have developed and accommodated the development of a collection method, which facilitates the distinction between practices and norms, using a questionnaire administered during a face-to-face interview (18, 19). This is done by first inviting people to say what they consider to be a "proper meal," a "proper breakfast," a "proper lunch," etc. This is presented to them as taking place in an ideal setting when nothing has disturbed the material organization of the preparation and consumption of these meals. This method is an extension of the work of Mary Douglas (20) on "deciphering a meal." Through this process, the social norms are collected for the meals under consideration. In the second step, when the interviewee is "liberated" from the normative pressures by his or her statement, another series of questions is proposed to help the individual to rebuild his or her food day. The interviewer begins by specifying that what now interests the research team is what really happened, what has really been eaten. The enumerator explains that, working at the level of the total population, it is not a problem if the meals eaten by the interviewee differ from what has been said in the first part of the questionnaire, when the informant tells what she or he thinks should be done, or what she or he usually does.

The first type of data corresponds to social norms, that is to say, provides an aggregate of guidelines that are rooted in cultural, social, and family traditions. They result from the specific socialization of an individual. But these norms are also impacted by the prevailing discourses of public health in relation to diet, or by the pressure of prevailing models of desirable body shape. The second type of data always retains the status of declarative data but is much closer to the actual practices of individuals. With such a method, the data collected gained precision, and it is possible to distinguish norms and practices and their relationships with each other, particularly for the exploration of various forms of change. The improvement of data collection methods is an important issue in this research. The ability to distinguish between norms and practices allows a deeper understanding of the transformation of eating habits. The distinction between norms and practices is a solution devised to improve the empirical quality of the data, but it has a cost because it greatly increases the length of the questionnaire and almost always requires face-to-face data collection.

## Questionnaire and Variables

The MFB studies the social, ethnic, and cultural diversification of food habits in Malaysia. It studies the evolution of food consumption, both at home and outside the home and identifies the consequences in terms of market factors and public health. It was based on a national representative sample of 2,000 respondents of 15 years old and above. The sampling methodology is a semi-randomized approach, based on the regions within Malaysia and their degree of urbanization (21) than a quota system based on age and ethnicity was also applied.

The questionnaire (see **Annex 1**) has 6 main parts: sociodemographics and ethnicity indicators, food norms, food intake in the last 24 h (recall), cooking practices, social representations of food, health and risk issues. It comprises 66 items and more than 1,400 variables and 58 closed and multiple-choice questions, consisting of standard questions used in sociology to describe the sociodemographics of a population (22), and questions that have been used in prior studies (2, 18, 23–25). The questionnaire was translated in three languages (English, Malay, and Chinese), and retro-translated to ensure the right meaning of the questions.

A complete presentation of the variables is available in the report (2). Here, we would like to highlight three essential points: the variables dedicated to ethnicity, the variables relating to eating habits, and some indexes of social position.

1. For ethnicity, data are not just what "is written on identity cards," but include:
   - Assigned ethnicity.
   - Self-declaration.
   - Declaration of interviewee for the ascendants and spouse and spouse ascendants.
   - Hierarchy by the individual of ethnicity, religion, and citizenship.
2. Indexes of social position:
   - The "Modernization index" is a weighted combination of level of urbanization, size of household, level of income, evolution of income, level of education. Modernization index = (Urb × 3) + (Size household × 2) + (Income × 3) + (Evolution income × 1) + (Education × 2).
   - The "Ethnicity index" is a weighted combination of assigned group (Malay, Chinese, Indian, and others), self-declaration, auto-definition, self-declaration of the family (spouse, parents and grandparents of interviewee and spouse), religion, and intensity of religion practice.
   - EthSocPos categories combine "Assigned ethnicity" and "Income available per person" 12 categories are created:
   - 4 ethnicities (non-Malay Bumiputra, Malay, Indian, and Chinese).
   - 3 income categories (20% lowest, 60% middle, and 20% highest).

   Finally, it allows to a multicriteria approach of ethnic affiliations and social positions to analyze and sort out the influences of determinations of ethno-cultural origin and those linked to social positions.
3. Variables relating to the levels of the "act of eating." The collection of food consumption data considers:

- Sociocultural representations relating to food and eating practices.
- Social norms defining meals.
- Food practices of the day before the survey reconstructed using a 24-h reminder (adapted to avoid injunctions from Western categories, etc.).
- Social contexts of consumption (places and people who are sharing the intake).
- Social interactions.
- Distribution of food intakes between home and out of home for the days before and the week (7 days) preceding the survey.

## DATA COLLECTION

The data collection was done face to face by enumerator trained by the research team. The objective was to help them to understand the key points of the methodology and the use of the questionnaire. The training of enumerators included role plays during which they alternated from the position of interviewer to that of interviewees. These sessions were followed by debriefings. Enumerators were all fluent in English plus at least one of the languages in which the questionnaires had been translated.

Particular attention was paid to the use of the food day reconstruction tool: the interviewee's 24-h food intake. The data collection were pitched in the middle ground between a qualitative approach and a quantitative approach. The questions were intended to be a sort of guide to help the interviewee to remember the food intake composition and structure, the time of consumption, the conditions of the acquisition, and the socio-technical contexts of consumption, including the ethnicity of the persons with whom the food was shared. The form and sequence of questions have been developed to reflect the Malaysian context, including the high frequency of consuming food away from home, or as takeaway food, and the large variety in eating places (see **Figure 2**) .

The data collection was conducted on "normal" periods of the year (out of the religious or national feasts) between January and May 2013.

## ALREADY PUBLISHED RESULTS AND WORK PERSPECTIVES

Some results have been produced and published (2–5, 8, 26, 27). But the interest of the MFB database is far from exhausted. It allows a deep description of food patterns and an analysis of their social determinants like:

- Describe the food habits, practices, social norms, social representations, and beliefs, pertaining to food, food patterns, and temporalities.
- Measure frequency of eating out.
- Identify food lifestyles with a focus on ethnicity.

**Thanks you very much. Now we will proceed to the same questions, but for the second food or drink intake that your remember.**

**2. Your second meal, food or drink intake of yesterday**

| Where did you have your second food or drink intake of yesterday? | | Where did the food came from? | How do you call this intake? | Could you describe the content(s) of your meal, food or drink intake? | Could you describe the social context of this intake? | |
|---|---|---|---|---|---|---|
| ☐ At home | ☐ your place | ☐ cooked by you<br>☐ cooked by friend/familly | Name : | Drink : | ☐ Alone | ☐ working<br>☐ watching tv |
| | ☐ friend's place | ☐ delivery* :<br>☐ bring from outside* : | At what time ? | Individual food items : | | ☐ on a computer<br>☐ other distraction : |
| ☐ Outside | ☐ in the office | ☐ hawkers, street food (1)<br>☐ conveniance store, supermarket, pasar mini (2) | | | | How many adults :<br>How many childs : |
| | ☐ in a restaurant | ☐ fast-food (3)<br>☐ food court (4) | Lenght : | Shared food items : | ☐ With someone | Guest ethnicity : |
| | ☐ on the run | ☐ mamak (5)<br>☐ restaurant (6) | | | | |

*** coding the source of food among the 6 propositions bellow**

**FIGURE 2 |** Adapted 24-h recall, food intake recording method.

- Analyze the correlation between lifestyles and above-mentioned characteristics.
- Explore the correlation between lifestyles and BMI.
- Analyze risk perceptions linked to food.

## CONCLUSION

The MFB database is made available in an "open science" philosophy. Let us remember the spirit of Open Science. It is a movement which consists in opening the research process to all types of actors concerned by the topic (scientific partners, economical actors, public institutions, citizens, NGOs, etc.). It uses the opportunity represented by the "digital mutation" to make research data easily accessible and reusable (28). In our case, the MFB1 report has already been freely accessible since 2015, and the MFB2 will be available shortly. The provision of the database itself is a new step. Sharing research results and data, not only between scientists but also with all stakeholders in society, is a way of disseminating and a way of developing partnership relations with all the different actors concerned by the subject of the research. But it is also a way of advancing knowledge, a way to co-construct the knowledge. And finally, the objective of open science is to accelerate the process of scientific production itself (29, 30).

It is possible to carry out comparative analysis between different time periods in Malaysia and with other countries. It can be used by researchers, public and private decision-makers, educators, and students. Researchers to carry out original work based on secondary data processing, or to make comparisons with other data of the same nature. As part of educational activities, academics and students can do secondary analysis or compare their data collected on subsets of the national population (region, group ethnic or social group) with the database. Public and private decision-makers of the public health sector or food industry and food service sector will find data to support strategical decisions.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Taylor's University Ethical Committee. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

J-PP and NR obtained the funding. J-PP, LT, CL, EM, and IN designed the study, involved in data collection, and supervision. J-PP, LT, CL, EM, TF, ADa, ADu, AR, YA, and KN contributed to the data analysis. J-PP drafted the manuscript. All authors critically revised the manuscript.


## FUNDING

The project was made possible with the help of Malaysian and international public and private support. Academics partners: Taylor's University (TUC Chair 2012-3 & TRC2013) 31%, Toulouse University (ISTHIA-TTUC2012-3) 31%, Ministry of Higher Education Malaysia, Long Research Grant Scheme (LRGS) "Social National Cohesion", (Prof. Shamsul, LRGS/BU/2011/UKM/CMN) 6%, CNRS-France, LIA "Food, Cultures and Health" (LIA-CNRS France-Malaisie2016) 2%. Industrial partners: Observatory of Food Habits of French Dairy Industry (CNIEL-MFB 2012-4) 23%, Coca-Cola-Malaysia 5%, and Nestlé-Malaysia 2%. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.


## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.800317/full#supplementary-material

## REFERENCES


1. Ministry of Health Malaysia. *Malaysian Adult Nutrition Survey (2003)*. Putrajaya: Ministry of Health (2008).
2. Poulain JP, Tibère L, Laporte C, Mongard E. *Malaysian Food Barometer*. Kuala Lumpur: Taylor's University Press (2014).
3. Poulain JP, Laporte C, Tibére L, Mognard E, Ragavan N. Malaysian Food Barometer (MFB): a study of the impact of compressed modernization on food habits. *Malays J Nutr*. (2020) 26: 1–17. doi: 10.31246/mjn-2019-0042
4. Fournier T, Tibère L, Laporte C, Mognard E, Ismail MN, Sharif SP, et al. Eating patterns and prevalence of obesity: lessons learned from the Malaysian Food Barometer. *Appetite*. (2016) 107:362–71. doi: 10.1016/j.appet.2016.08.009
5. Poulain JP, Smith W, Laporte C, Tibère L, Ismail MN, Mognard E, et al. *Studying the Consequences of Modernization on Ethnic Food Patterns: Development of the Malaysian Food Barometer (MFB)*. Anthropology of food (2016). Available online at: http://aof.revues.org/7735 (accesed December 22, 2021).
6. Chang KS. Compressed modernity in South Korea: constitutive dimensions, manifesting units, and historical conditions. In: *Routledge Handbook of Korean Culture and Society*. London; New York, NY: Routledge (2016), p. 45–61.
7. Rouleau-Berger L, Neng L. Compressed modernity and Chinese temporalities. *Temporalités*. (2017) 26. doi: 10.4000/temporalites.3810
8. Poulain JP, Laporte C, Tibère L, Mognard E, Neethia Ragavan NA. Le *Malaysian Food* Barometer. *L'impact de la modernisation compressée sur l'alimentation*. Sciences de la société [En ligne] (2019). p. 104. Available online at: http://journals.openedition.org/sds/10621 (accesed December 22, 2021).

9. Shamsul AB. From Orang Kaya Baru to Malayu Baru: cultural construction of the Malay 'new rich. In: Pinches M, editor. *Culture and Privilege in Capitalist Asia*. London; New York, NY: Routledge (1999). p. 104–28.
10. Embong AR. New middle classes in Asia. In: Ritzer G, editor. *The Blackwell Encyclopedia of Sociology*. Oxford: Blackwell Publishing (2007). p. 3198–201). doi: 10.1002/9781405165518.wbeosn019
11. Ismail MN. The nutrition and health transition in Malaysia. *Public Health Nutr.* (2002) 5:191–5. doi: 10.1079/PHN2001293
12. Hirschman C, Guest P. The emerging demographic transition of Southeast Asia. *Popul Dev Rev.* (1990) 16:121–52. doi: 10.2307/1972532
13. Poulain JP. Food in transition: the place of food in the theories of transition. *Sociol Rev.* (2021) 69:702–24. doi: 10.1177/00380261211009092
14. Tee ES. *Selected Annotated Bibliography of Research Public Actions and Resource Materials (1980-1993)*. Kuala Lumpur: Food, nutrition and health promotion in Malaysia (1994).
15. Warde A. Consumption and theories of practice. *J Consumer Cult.* (2005) 5:131–53. doi: 10.1177/1469540505053090
16. Poulain JP. *Towards a Sociological Theory of Eating: A Review of Alan Warde's The Practice of Eating*. Anthropology of food (2020). Available online at: URL: http://journals.openedition.org/aof/10866 (accesed December 22, 2021).
17. Poulain JP. *Dictionnaire des Cultures Alimentaires*. Paris: PUF. (2018).
18. Poulain JP. The contemporary diet in France: "De-structuration" or from commensalism to "vagabond feeding." *Appetite.* (2002) 8:1–13. doi: 10.1006/appe.2001.0461
19. Poulain JP, Guignard R, Michaud C, Escalon H. Les repas, distribution journalière, structure, lieux et convivialité. In: Escalon H, Bossard C, Beck F, editors. *Baromètre Santé Nutrition 2008*. Paris: INPES (2010). p. 186–211.
20. Douglas M. Deciphering a meal. In: Geertz C, editor. *Myth, Symbol, and Culture*. New York, NY: W W Norton & Company (1971).
21. Census. *Malaysian Government*. (2017). Available online at: https://www.dosm.gov.my/v1/index.php?r=colum/ctwoByCat&parent_id=115&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09 (accesed December 22, 2021).
22. Desrosieres A. *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press (1998).
23. Rozin P, Fischler C, Shields C, Masson E. Attitudes towards large numbers of choices in the food domain: A cross-cultural study of five countries in Europe and the USA. *Appetite.* (2006) 46:304–8.
24. Dupuy A. *Plaisirs alimentaires. Socialisation des enfants et des adolescents, Sociologie, Collection Tables des hommes, PUR–En coédition avec les Presses universitaires François-Rabelais* Tours (2013).
25. Poulain JP. *The Sociology of Food: Eating and the Place of Food in Society*. London: Bloomsbury. (2017).
26. Tibère L, Mognard E, Aloysius M, Laporte C, Poulain JP. Without Nasi Lemak, no life in Malaysia!. *APJIHT.* (2018) 8:51–66.
27. Drewnowski A, Mognard E, Gupta S, Mohd Noor, I.smail, Karim N, Tibère L, Laporte C, et al. Socio-cultural and economic drivers of plant and animal protein consumption in Malaysia: the SCRiPT study. *Nutrients.* (2020) 12:1530. doi: 10.3390/nu12051530
28. CNRS. *Roadmap for Open Science*. (2019). Available online at: https://www.science-ouverte.cnrs.fr/wp-content/uploads/2019/11/CNRS_Roadmap_Open_Science_18nov2019.pdf (accesed December 22, 2021).
29. Nielsen M. *Reinventing Discovery: The New Era of Networked Science*. Princeton, NJ: Princeton University Press (2012). doi: 10.1515/9781400839452
30. Mirowski P. The future (s) of open science. *Soc Stud Sci.* (2018) 48:171–203. doi: 10.1177/0306312718772086


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 24 January 2022
doi: 10.3389/fnut.2021.749596

Check for updates

# "How to Select a Representative Product Set From Market Inventory?" A Multicriteria Approach as a Base for Future Reformulation of Cookies

*Carole Liechti[1], Julien Delarue[1,2], Isabelle Souchon[3], Véronique Bosc[1] and Anne Saint-Eve[1]**

[1] University of Paris Saclay, UMR SayFood, AgroParisTech, INRAE, Palaiseau, France, [2] Department of Food Science and Technology, University of Davis, Davis, CA, United States, [3] Avignon University, UMR SQPOV, INRAE, Avignon, France

OPEN ACCESS

**Edited by:**
Massimo Lucarini,
Council for Agricultural Research and Economics, Italy

**Reviewed by:**
Antonis Vlassopoulos,
Agricultural University of Athens, Greece
Quoc Cuong Nguyen,
Ho Chi Minh City University of Technology, Vietnam
Sercan Karav,
Çanakkale Onsekiz Mart University, Turkey

***Correspondence:**
Anne Saint-Eve
anne.saint-eve@inrae.fr

**Specialty section:**
This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition

**Received:** 29 July 2021
**Accepted:** 13 December 2021
**Published:** 24 January 2022

**Citation:**
Liechti C, Delarue J, Souchon I, Bosc V and Saint-Eve A (2022) "How to Select a Representative Product Set From Market Inventory?" A Multicriteria Approach as a Base for Future Reformulation of Cookies. Front. Nutr. 8:749596. doi: 10.3389/fnut.2021.749596

Consuming too much fat, sugar, and salt is associated with adverse health outcomes. Food reformulation is one possible strategy to enhance the food environment by improving the nutritional quality of commercial products. However, food reformulation faces many hindrances. One way to alleviate some of these hindrances is to embrace a multicriteria approach that is based on a market inventory. In this objective, additional sensory screening and water content analyses allow going beyond nutrition and composition information on the packaging. However, due to feasibility reasons for later in-depth analyses, it is necessary to work with several reduced and manageable products. To the best of the authors' knowledge, in the literature, there is no sample selection approach taking into account multiple criteria as a base for future food reformulation. The overall aim of this paper is to propose a method to select the best representative products from the market base, for future reformulation by going beyond nutrition and composition information on the packaging. This approach considered therefore nutrition, composition, economic, water content, and sensory information with the example of the cookies market. The first step is the creation of an extensive cookie database including sensory and water content information. In total 178 cookies among the French market were identified, then focus was placed on 62 chocolate chip cookies only. Sensory screening and water content analyses of all 62 products were conducted. The second step is to make an informed subset selection, thanks to a cluster analysis based on 11 nutrition, composition, and water content variables. A representative subset of 18 cookies could be derived from the obtained clusters. The representativity was evaluated with statistical uni- and multivariate analyses. Results showed a broad variety of chocolate chips cookies with a large nutritional, compositional, water content, and sensory differences. These results highlighted the first paths for future reformulation in this product category and showed the importance to include physical product information beyond the information on the packaging. This complete database on the selected cookies constituted a solid base for identifying future reformulation levers, in order to improve the nutritional quality and health.

**Keywords: market inventory, multiple criteria, sample selection, food reformulation, nutrition, health**

## INTRODUCTION

Food reformulation is an interesting lever to reduce over-consumed nutrients (such as sugar, fat, and salt) and to enhance our diet and health (1, 2). A successful reformulation makes it possible to move toward a healthier food offer without largely changing consumers eating habits. It was shown that a reduction of undesirable nutrients such as trans fatty acids, saturated fat, sodium, and sugar of more than 17,000 foods and beverages would lead to a reduced intake of nutrients which are to limit and improve public health (3).

However, a recent study across 20 European countries showed that many packaged foods and drinks still have too high fat, sugar, and salt, as well as too low fiber content (4). These results of concern are surprising as the link between overconsumed adverse nutrients and negative health outcomes is well known (5). For example in the UK, most of the companies between 2015 and 2018 did not reach an overall sugar reduction of 5% among the top five product categories which contribute the most to the high sugar intake (biscuits, cereal bars, breakfast cereal, chocolate, and sugar confectionery, yogurts) among the UK population (6).

These findings underline the fact that food reformulation still has a lot of unused potentials, very likely due to many technological and sensory barriers. Indeed, reformulating biscuits is challenging. For instance, decreasing sugar and fat content is difficult because of their multiple functional properties in the food matrix, and in particular in sweet bakery products (7–9). Moreover, sugar and fat are strong drivers of preferences. Any modification might have huge consequences on liking, pleasure experiences, and food choices (10, 11). This may contribute to the hesitant willingness for voluntary food reformulation on the part of the food industry as cost and time effectiveness are not immediately granted.

To overcome these barriers and to encourage industries to reformulate healthier versions of their products, we argue that is necessary to see food reformulation in a comprehensive way because of the multifactorial nature of the determinants of food preferences and the multiple interactions between food components in the food matrix (10–14). Focusing only on nutritional and compositional changes would thus inevitably lead to dead-ends or missed opportunities. Improving the nutrition quality by food reformulation is complex and needs to integrate different dimensions such as food composition, physicochemical properties, and sensory perception. To address this challenge, it would be useful to create a solid base relying on multiple criteria in order to anticipate possible interactions and to achieve a successful reformulation in the long term.

Biscuits structure is highly dependent on processing conditions (temperature, moisture, time) and formulation (presence of sugars and fats) (15). Water content is thus a key property for biscuits' structure, texture, and fracture properties. For example, biscuits with a lower water content tend to show a higher measured fracture than biscuits with higher water content. In other words, biscuits with a higher water content tend to be softer than biscuits with lower water content (16, 17). As a result water content may indirectly drive biscuit preferences (15). In addition, water content is very important for products stability and shelf-life (18).

Besides the above-mentioned relationship between cookies' water content, structure, and texture, it is further very important to better understand consumers' perception of cookies' texture.

A sensory analysis (temporal dominance of sensation) with a trained panel showed that "hardness" was the first dominant attribute resulting from biscuits formulations (19). Interestingly, fat and/or sugar reduction leads to an increased hardness of the biscuits (20, 21). Therefore, we assume that the sensory variable "hardness" is an important attribute when it comes to food reformulation among a complex (high sugar and fat at the same time) food matrix, such as sweet bakery products.

Moreover, biscuits can usually be divided into two subcategories: soft or crunchy. In some cases, this information is available on the front packaging. However, some brands could have an intermediate texture, therefore a texture somewhere between hard and soft. Due to his missing texture information, it is thus impossible for the consumers to know in advance what type of cookie texture they buy until they actually eat them (or at least open the package and manipulate the cookies).

This article proposes a guide on how to create such a base for future reformulation. Cookies were chosen as a case study of prime relevance. The first aim was to create a database with the help of comprehensive market inventory, taking into account multiple criteria by going beyond nutrition and composition information on the packaging. This multicriteria approach considers easily available information from the packaging (nutrition, composition, and price) and adds complementary information obtained thanks to simple analyses such as water content measurement and sensory screening.

However, conducting any advanced measurement with such a high number of products would be very time-intensive. For example, developing a protocol for any physicochemical analysis for all the products with their large product diversity is complex, especially when conducting the measurements in triplicates. The same goes for sensory panelists. Testing such a high number of products would be too demanding and might be physically impossible, because of sensory fatigue, or increased number of evaluation sessions.

This approach aims to provide a holistic view of the market which could benefit the industries by evaluating their competitors' products and ultimately gaining a better understanding of their own product positioning. But the food market is complex to analyze, with many different recipes from many different manufacturers. An investigation of the product category under consideration is thus a first necessary step to identify the diversity of existing products. A sensory screening and a water content analysis of all included cookies were thus necessary to get the first-level view of product texture. Then, further in-depth analyses are usually needed to gain greater knowledge of the existing recipes and to guide reformulation. To make these analyses realistic and compatible with experimental constraints, a second step in the proposed approach is to define a subset of products that would be representative of the market and yet be of manageable size. In this objective, we suggest making an informed selection based on a multicriteria analysis. To the best of the authors' knowledge, there is a lack of a common subset selection method adapted to food reformulation and that takes into account multiple variables.

# MATERIALS AND METHODS

This section describes the four steps used for this multicriteria approach, from the analysis of the cookie market to the subset selection and the representativity checks (**Figure 1**, steps 1–4). First of all, to identify the potential for reformulation and to explore the diversity of recipes of the "commercial cookies" product category, online analysis of the French market was conducted (step 1). The focus was set on a uniform cookie variety. In order to identify possible levers for food reformulation in a later step, it is important to first have a broad view of products' characteristics. For this, we needed to select a representative subset of products while maintaining a good vision of the market diversity. Therefore, sensory screening and water content analyses were first conducted on all chocolate chip cookies (step 2). Then, the cookies were clustered based on available nutrition, composition, and water content data. A subset of products was then proposed based on additional 11 compositions, and sensory and economic criteria ranked for their importance in the selection (step 3). Finally, a check of the subset representativity was performed with uni- and multi-dimensional statistical analyses (step 4).

## Cookie Market Inventory (Step 1)

All the cookie brands available in 2019 on the online French market with the specific mention "Cookie" on the packaging were included in this study (**Figure 1**, step 1). All the different varieties were initially considered, including cookies with inclusions, fillings, and coatings and specialites such as gluten allergen and sugar free products. A database with a total of 178 cookies was created, using information from the packaging and websites. Available nutrition, composition, and price information from the packaging and online websites were collected. In addition to that, the Nutri-Score (letters A-E) was calculated based on the Rayner computation for each product (22).

### Focus on Uniform Cookie Category: Chocolate Chip Cookies

We aimed to focus on a uniform cookie variety in order to work on a homogeneous product set and to make future reformulation work consistent. Cookies with chocolate inclusions showed the broadest range and greatest diversity within the market offer. Therefore, the decision was made to consider cookies containing chocolate inclusions only. Besides, it is worth noting that chocolate chips are important carriers of sugar and fat which makes them a potentially interesting lever for reformulation.

We excluded all cookies without chocolate inclusions, those with fillings and coatings, and specialties cookies such as gluten-free, sugar-free, and allergen-free cookies. Finally, a total of 62 cookies with chocolate inclusions were included in this study (**Figure 1**, step 2 excluded cookies).

### Variables and Criteria Among the Chocolate Chip Cookie Database

Besides the cookie information from the packaging and online websites, the chocolate chip cookie database was further completed with water content analysis (section Water Content) and sensory screenings (visual perception shape chocolate inclusion, and cookie surface) (sections Screening of the Texture in Hand – Screening of the Cookie Surface and the Shape of the Chocolate Inclusion) of the 62 cookies. All quantitative variables were included for the cluster analyses [section Subset Selection (Step 3) and Cookie Clustering (Step 3)] and representativity checks [sections Statistical Representativity Check (Step 4) and Validation With Multi-Dimensional Statistical Analysis (Step 4)], whereas additional ranked criteria were considered for the subset selection among clusters (section Additional Ranking Criteria). All variables and criteria used for the chocolate chip cookies database are presented in **Supplementary Tables 1, 2**.

#### *Quantitative Variables*

In this method, 14 quantitative nutrition, composition, and water content variables were used to constitute the database: kcal, fat, saturated fat, carbohydrates, sugar, protein, fiber and salt content per 100 g, number of technological additives (important for baking and conservation properties such as baking powder, emulsifier, thickening agents, antioxidants, and humectant), number sensory additives (important for sensory properties such as colorings and artificial sweeteners), number additives (technological and sensory additives), number ingredients, water content, and the calculated Rayner score.

Furthermore, 11 variables (fat, saturated fat, carbohydrates, sugar, protein, fiber, and salt content per 100 g, number technological additives, number sensory additives, number ingredients, and water content) were included for the clustering [section Subset Selection (Step 3)], while 12 variables [kcal, fat, saturated fat, carbohydrates, sugar, protein, fiber and salt content per 100 g, number additives (technological and sensory additives), number ingredients, water content, and the calculated Rayner score] were included for the representativity check [section Statistical Representativity Check (Step 4)].

#### *Additional Ranking Criteria*

Additional ranked compositions, sensory and economic criteria with their 40 subgroups were also included in the database (**Figure 2**). Cookies' availability was not included as a criterion but as a constraint, as for obvious practical reasons products must be available for further analyses.

All 11 criteria were ranked according to their possible impact on the food structure, sensory perception, and liking. In addition to these criteria, product availability at the time of the study was a strong constraint and was taken into account for the final selection.

Obviously, composition criteria such as major cookies ingredients and chocolate ingredients were considered as most impactful on the food matrix due to their high quantity in the recipe.

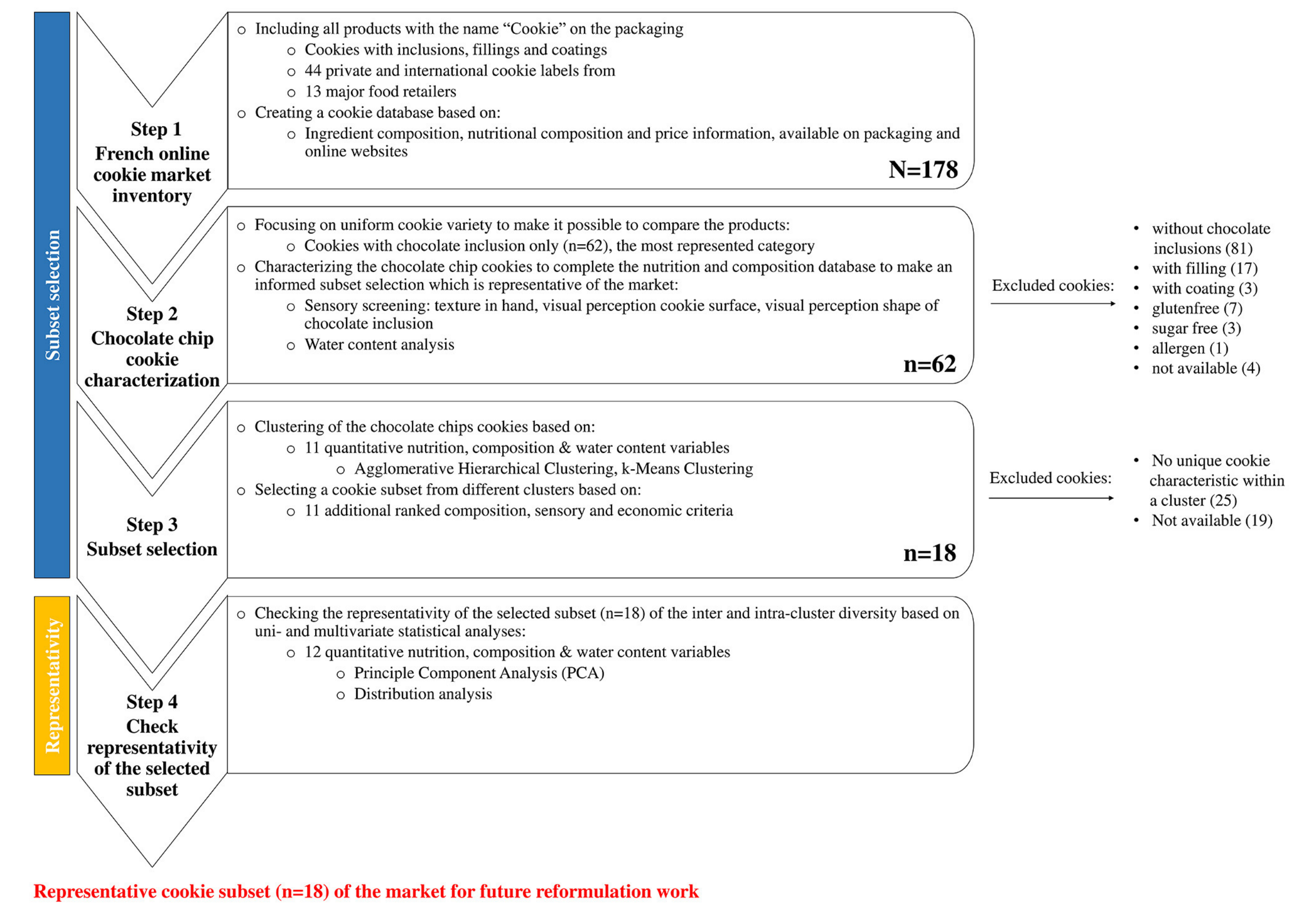


**FIGURE 1 |** Overview of the four steps from the online cookie market inventory to the selected subset and its check of the representativity.

Cookies' weight and price were included as well. These two criteria lesser impact reformulation, sensory perception, and liking. However, the cookie weight might play an important role in portion size and kcal intake. Besides, as healthier reformulated products should be available for all, we also included the price that is an important determinant of purchasing behavior.

Moreover, six criteria contained qualitative information (type of flour, type of sugar, type of fat, type of chocolate inclusion, shape of chocolate inclusion, and surface cookie), whereas five criteria contained quantitative information (amount of chocolate and cacao powder, amount of chocolate inclusion, texture cookie, weight of cookie and price per kg cookie).

All 11 qualitative and quantitative criteria were considered as categorial criteria with subgroups. For the 5 quantitative criteria, we calculated their quintile rank in order to obtain five subgroups of equal frequencies such as very low, low, middle, high, and very high. For the criterion texture in hand, we calculated the tertile rank for the score in order to distinguish between three main textures (soft, intermediate, and hard) (**Table 1**).

## Chocolate Chip Cookie Characterization (Step 2): Sensory Screening and Water Content Analysis

Further analyses were conducted in order to go beyond composition and nutritional values from the packaging and to better understand the major sensory and water content characteristics of the products (**Figure 1**, step 2). Three sensory screenings (perceived texture in hand; visual perception cookie surface and shape of chocolate inclusion) were performed on all 62 chocolate chip cookies. The goal of the sensory screening was to categorize the cookies according to their most striking visual and texture characteristics. This type of evaluation is relatively easy and differences between the cookies were expected to be quite obvious. Under such conditions, it is thus possible to reach sufficient power, even with a very small number of panelists. Besides the sensory screening, the water content was measured for all 62 chocolate chip cookies. Measuring the water content among baked bakery products will provide information about cookies structure and texture characteristics.

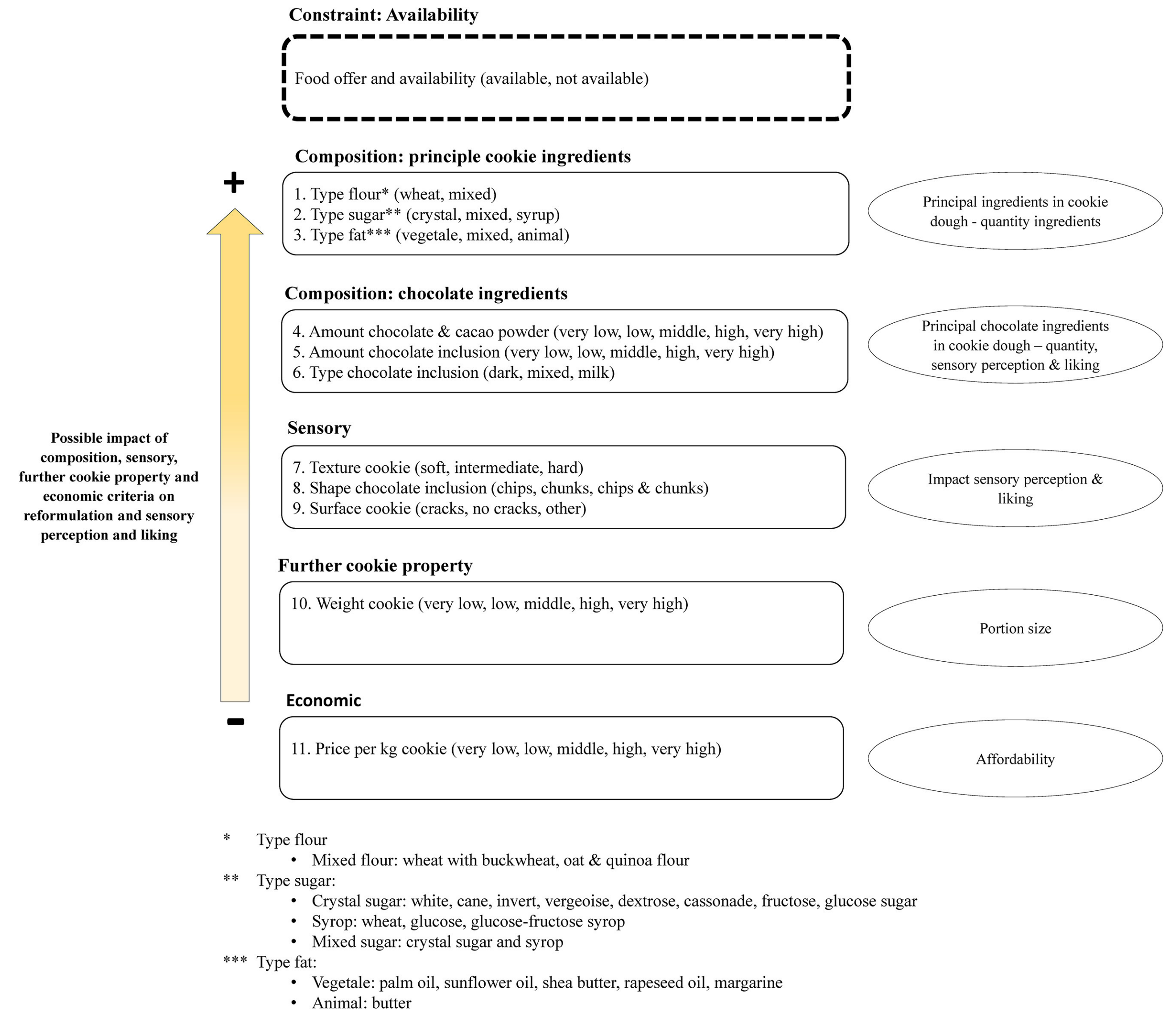


**FIGURE 2 |** Overview of the 11 additional ranked criteria with their 40 subgroups (in brackets) and the availability as constraint.

## Screening of the Texture in Hand

In order to gain insights into product texture, three trained subjects evaluated the hardness of all chocolate chip cookies. They were told to break the cookies in two halves by hand and to report their perceived hardness on an unstructured scale from 0 to 10, where 0 indicated soft and 10 hard. The evaluation took place over six sessions and was conducted in sensory booths in a sequential monadic way, following a balanced order over the panel and sessions. Samples were coded with random three-digit numbers and presented in blind.

## Screening of the Cookie Surface and the Shape of the Chocolate Inclusion

Cookies' surface aspect and the shape of the chocolate inclusions were evaluated by a cookie expert who was trained over 1 year with all 62 chocolate chip cookies. The visually perceived cookie surface was grouped into the qualitative subgroups "cracks," "no cracks," and "other" which means neither cracks nor no cracks. The visually perceived shape of the chocolate inclusion was grouped into three qualitative subgroups "chips," "chunks," and "chips and chunks."

## Water Content

The water content was determined by the oven drying method and adapted from (23). First, all chocolate chip cookies were crushed and grinded with a mortar for 15 s. After grinding, 3 g of ground cookies were weighed in a round aluminum dish. It was further put in the oven (EM10, Chopin, France) for 18 h at 103°C. The time was set by 18 h as the weight after drying did not change anymore. The sample was then placed in a desiccator

**TABLE 1 |** Rank calculation of the quintiles and tertile of the subgroups.

| Criteria (quintiles) | Very low (%) | Low (%) | Middle (%) | High (%) | Very high (%) |
|---|---|---|---|---|---|
| Amount chocolate and cocoa powder in % (per 100 g) | 1.3–2.3 | 2.4–2.6 | 2.7–3.1 | 3.2–5.7 | 5.8–15.3 |
| Amount chocolate inclusion in % (per 100 g) | 5.5–22.6 | 22.7–28 | 28.1–30.5 | 30.6–37 | 37.1–40 |
| **Criteria (quintiles)** | **Very low (g)** | **Low (g)** | **Middle (g)** | **High (g)** | **Very high (g)** |
| Weight cookie in g | 3.12–16.5 | 16.6–16.6 | 16.7–23 | 23.1–25 | 25.1–70 |
| **Criteria (quintiles)** | **Very low (€)** | **Low (€)** | **Middle (€)** | **High (€)** | **Very high (€)** |
| Price per kg cookie | 2.9–4.5 | 4.6–7.5 | 7.6–10.4 | 10.5–16.7 | 16.8–29.3 |
| **Criterion (tertile)** | **Soft** | **Intermediate** | **Hard** | | |
| Texture in hand (score 0–10) | 0–5 | 5.1–7.2 | 7.3–9.4 | | |

for 1 h before weighing. All measurements were performed in triplicates among three different cookies and the results averaged.

The mass loss was determined by weighing the sample before and after drying to constant weight:

$$water\ content\ in\ \% = \frac{weight\ (g)\ cookie\ before\ oven\ -\ weight\ (g)\ cookie\ after\ oven}{weight\ (g)\ cookie\ before\ oven} * \mathbf{100}$$

## Subset Selection (Step 3)

In order to best represent the market diversity and to make an informed subset selection, cookies were clustered based on 11 nutrition, composition, and water content variables defined in section Quantitative Variables (**Figure 1**, step 3). This allowed the selection of cookies from each cluster with different cookie characteristics.

In sensory science, Agglomerative Hierarchical Clustering (AHC) and K-means clustering are possible methods to group product characteristics or consumers in the same clusters based on their similarities and is, therefore, a suitable tool to contribute to decisions for product development (24, 25).

The subset selection from each cluster was done with the help of 11 additional ranked criteria defined in section Additional Ranking Criteria. As well other authors have used several ranked criteria for product selection (26, 27).

Within a cluster, those cookies with unique characteristics (subgroup) were selected. Please find more detailed information in **Supplementary Table 3**. Cookies that were marked as "not available" were excluded. The cookie numbers per cluster are shown in **Supplementary Table 4**.

## Statistical Representativity Check (Step 4)

In order to evaluate the representativity of the selected subset, we applied multi-dimensional analyses with 12 quantitative nutrition, composition, and water content variables (section Quantitative Variables). To check the subset based on their intra- and inter-cluster diversity, a principal component analysis (PCA) was performed on the 12 variables. To validate the representativity of the subset, we compared the distribution of the 62 chocolate chip cookies and the 18 selected cookies based on the 12 variables.

# DATA ANALYSIS

All statistical analyses (AHC, K-means ANOVA clustering, PCA, linear regression Histogram, and Linear Regression) were conducted with XLSTAT version 2018.1.1 (Addinsoft, New York, USA), where alpha = 0.05 was considered as the needed significance level.

## Cookie Clustering (Step 3)

To characterize and cluster cookies according to their nutrition, composition, and water content characteristics, we first ran an AHC with Euclidean distance, Wards' Method, and centered and reduced data. This analysis allows to visually define the optimal numbers of clusters. Data from the AHC were organized in a table with 11 columns (nutrition, composition, and water content analyses) and 62 rows (commercial chocolate chips).

Defining the "good" number of cookies clusters is a matter of balance between precision (the higher the number of clusters or products to be selected, the higher the precision of the selection) and feasibility (the higher the number of products, the more difficult to run additional analyses).

Once the number of clusters was selected, K-means clustering was applied (Trace (W) criterion on reduced and centered data the differences among the seven clusters was evaluated with an ANOVA). Conducting first an AHC followed by K-means is a common method in sensory science to obtain robust clusters.

## Validation With Multi-Dimensional Statistical Analysis (Step 4)

To visualize the seven clusters from the K-means clustering, the 10 active and two supplementary quantitative variables were plotted on a 2-dimensional map by using PCA. This allowed us to evaluate and check the multivariate nutrition, composition, and water content variables based on their intra- and inter-cluster diversity. For the PCA, we used Pearson correlation with a significance level alpha = 0.05 and standardized (n) data. Missing data were replaced by mean or mode. For the validation axes, F1 and F2 were considered. To evaluate the relationship between the texture in hand and cookies' water content, a linear regression was applied. The PCA is a convenient tool to visualize and plot obtained clustering data to detect class diversity (28).

To check the representativity of the selected subset, we visually compared the distributions of all 62 chocolate chip cookies and the 18 selected cookies based on the 12 quantitative variables thanks to histogram plots.

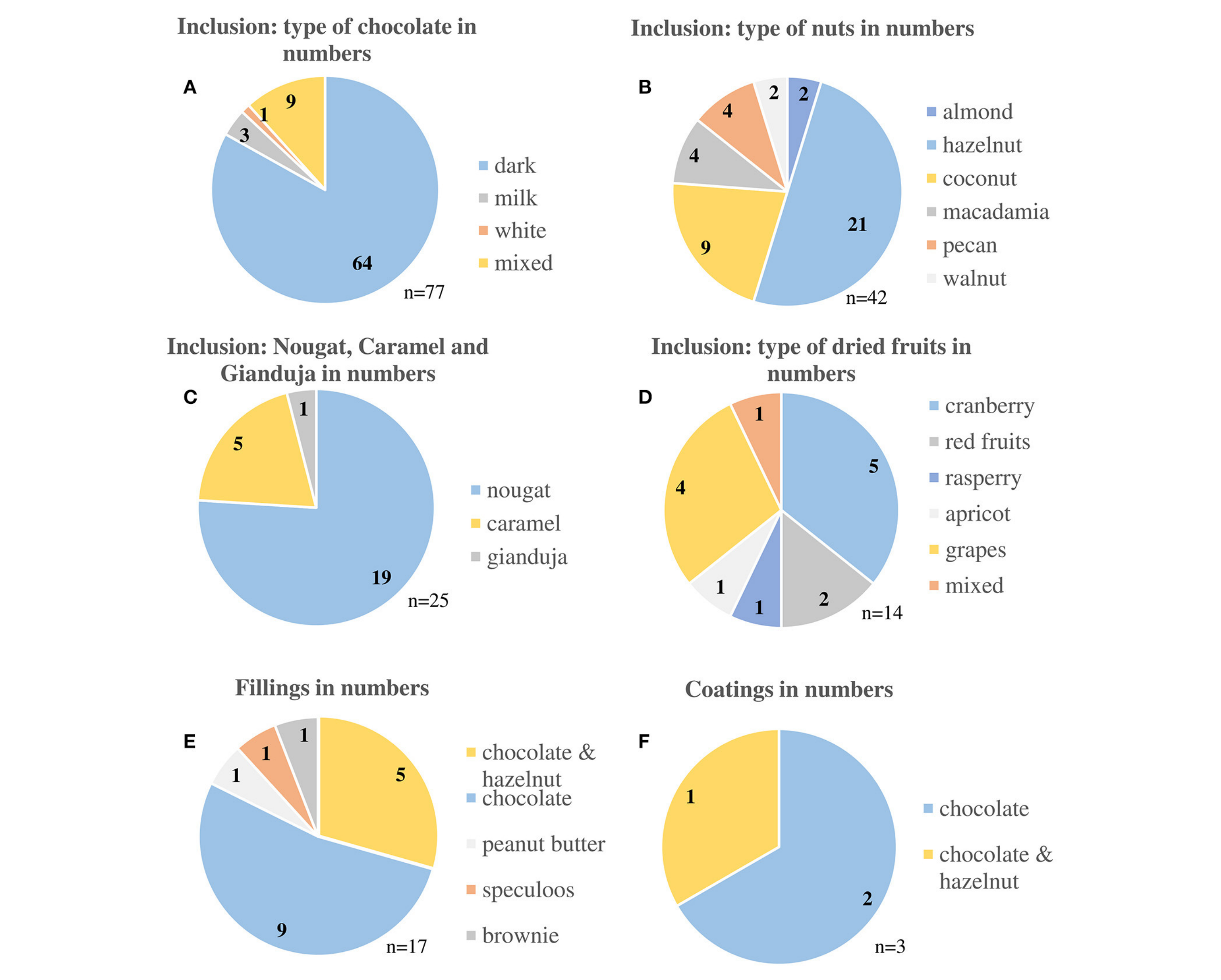


**FIGURE 3 |** Cookie varieties from the online French market inventory in 2019. Cookies **(A–D)** are cookies with different inclusions and cookies **(E)** and **(F)** are cookies with fillings and coatings.

# RESULTS

## Cookie Varieties and Compositional Diversity Among the Entire Cookie Database With Private and International Labels

The results showed that three main cookie varieties were identified from the whole database with 178 cookies: 158 (88.8%) cookies with inclusions, 17 (9.5%) cookies with fillings, and 3 (1.7%) cookies with coatings. Among the cookies with inclusions, products with chocolate chips 77 (48.7%), nuts 42 (26.6%), nougat, caramel, and gianduja 25 (15.8%), and dried fruits 14 (8.9%) were the most represented cookies on the market. Around 20 (11.2%) cookies do have fillings and coatings, while more cookies have fillings than coatings. The vast majority of cookies on the French market thus have chocolate inclusions with dark chocolate or nuts inclusions with almonds. **Figure 3** presents detailed information about the different inclusions, fillings, and coatings of commercial French cookies.

As shown in **Table 2**(A), the ranges (from the minimum to maximum values among the 178 cookies) for nutritional values (per 100 g) were: 183 kcal, 27 g sugar, 22.8 g carbohydrates, 16.4 g fat, and 16.4 g saturated fat. The calculated Rayner score showed that 170 (95.5%) cookies had a Nutri-Score of E, while 6 (3.3%) cookies had a Nutri-Score of D and 2 (1.2%) cookies had a Nutri-Score of B.

## Nutrition, Water Content, and Sensory Diversity Among the Chocolate Chip Cookie Database

In this study, we set the focus on 62 cookies with chocolate inclusions, representing 27 different private labels and international brands gathered from 12 retailers. As shown in **Table 2**(B), the broadest ranges for nutritional values per 100 g

**TABLE 2 |** Nutritional values of the entire- (A) and chocolate chip cookie database (B).

| | Databases A: 178 cookies B: 62 cookies | Min. value | Max. value | Difference Min. and Max. value | Mean ± SD |
|---|---|---|---|---|---|
| Kcal | A | 389 | 572 | 183 | 499.5 ± 25.5 |
| | B | 433 | 518 | 85 | 495.4 ± 15.2 |
| Fat (g) | A | 16 | 32.4 | 16.4 | 25.4 ± 3.1 |
| | B | 17.1 | 28 | 10.9 | 24.3 ± 2.0 |
| Saturated fat (g) | A | 3.6 | 20 | 16.4 | 12 ± 3.3 |
| | B | 5.9 | 18 | 12.1 | 12.6 ± 2.7 |
| Carbohydrates (g) | A | 48 | 70.8 | 22.8 | 59.6 ± 3.8 |
| | B | 57 | 70.8 | 13.8 | 61.6 ± 2.5 |
| Sugar (g) | A | 0* | 43 | 27 | 32.1 ± 6.0 |
| | B | 27 | 41.8 | 14.8 | 34.5 ± 3.0 |
| Protein (g) | A | 3.5 | 13 | 9.5 | 6.2 ± 1.1 |
| | B | 4.5 | 7.6 | 3.1 | 6 ± 0.8 |
| Fiber (g) | A | 0.8 | 10 | 9.2 | 3.9 ± 1.5 |
| | B | 1.8 | 5.7 | 3.9 | 3.6 ± 0.9 |
| Salt (g) | A | 0.03 | 2 | 1.97 | 0.7 ± 0.4 |
| * for sugar free cookies | B | 0.2 | 1.5 | 1.3 | 0.8 ±0.3 |

*0g of sugar per 100g was for the special cookies "sugar free" with artificial sweeteners.*

were found for kcal (85), sugar (14.8 g), carbs (13.8 g), saturated fat (12.1 g), and fat (10.9 g). Additionally, cookies with chocolate inclusion (B) demonstrated a slightly higher mean content of saturated fat, carbohydrates, sugar, and salt content, whereas we found a slightly lower mean content of protein and fiber compared to the whole cookie database (A). For cookies with chocolate inclusions, 61 cookies (98.4%) have a Nutri-Score E and only one (1.6%) cookie a Nutri-Score D. The Rayner score ranged between 14 (Nutri-Score D) and 28 (Nutri-Score E).

The mean value of the measured water content among the 62 cookies was 3.9 ± 1.7%, whereas the values ranged from min. 2.1% to max. 9.3%. For the perceived texture in hand, the mean value was 5.6 ± 2.5, with a range from min. 0 and max. 9.4 on a scale from 0 to 10 (0: soft; 10: hard). (**Supplementary Figure 1**). As expected, a significant negative correlation ($p < 0.0001$, $r = -0.773$, $r^2 = 0.579$) was observed between the variable texture in hand and the water content.

The numbers of cookies for the additional ranking criteria "texture in hand," "shape chocolate inclusion" and "surface cookie" are shown in **Figure 4** with their subgroups. Among the 62 chocolate chip cookies, most products had chips as a chocolate shape and most of the cookies' surfaces were characterized with cracks. All types of cookie textures are well distributed, in soft, intermediate, and hard. All criteria with their subgroups are shown in **Supplementary Table 2**.

On the PCA score map in **Figure 5A**, the cookies were grouped and represented with different colors based on seven clusters, obtained by K-means clustering (further variable information in **Supplementary Table 1**). The broadest variability was explained by axes F1 and F2 with a variance of 49.12%. Results of a conducted ANOVA demonstrated significant differences between almost all clusters ($p < 0.05$), except for the smallest ingredient components fiber and the salt content.

The loadings in **Figure 5B** show the relationships between 10 active and two supplementary variables. The strongest positive correlations were found between the fat and kcal content ($r = 0.828$), between the Rayner score and the number additives ($r = 0.538$), and between the Rayner score and the salt content ($r = 0.447$). On the other hand, the strongest negative correlations were found between the kcal and water content ($r = -0.793$), the protein content and the number of additives ($r = -0.659$), and between the fat and the water content ($r = -0.603$).

**Figure 5A** shows broad nutrition, composition, and water content diversity among the 62 chocolate chip cookies. Some cookies on the right side of the plot (axis F1) are characterized by a high carbohydrate, sugar, and water content and a higher number of total additives (technological and sensory) and ingredients. On the left side (axis F1), some cookies are characterized by their high protein and kcal content. Furthermore, cookies on the bottom (axis F2) had higher fat and saturated fat content with a higher Rayner score. Cookies on the top (axis F2) tend to have a higher fiber content. Moreover, and as expected, the Rayner score showed a positive correlation with fat content ($r = 0.288$), saturated fat ($r = 0.538$) and salt content (0.447), and a negative correlation with the fiber content ($r = -0.281$). Cookies of clusters 1 and 3 had a high Rayner score with high fat, saturated fat, and lower fiber content. On the other hand, most of the cookies belonging to clusters 5, 6, and 7 presented a lower Rayner score with higher fiber content, and lower fat and saturated fat content than other cookies. Cookies from cluster 5 showed the highest kcal content, while those of cluster 7 were characterized with high sugar content.

**Figure 6** presents all 62 commercial chocolate chips cookies. Although this study concentrated on a single product category only, we can observe a broad visual cookie variety. Notably, they differ in their size, their dough color, the shape of the chocolate inclusion, the quantity of the chocolate inclusion appearing on the surface, or the cracks on the cookie surface.

## The Selected Subset and Its Representativity

To best represent the diversity of all chocolate chip cookies, a subset of 18 cookies was also proposed. This accounts for almost a third of the initial 62 cookies. In total, 3 cookies were selected in clusters 1, 2, 3, and 5, while 2 cookies were selected in clusters 4, 6, and 7 (**Figure 5** and **Supplementary Table 4**).

### Labels and Retailers

The subset of cookies included 13 different private and international labels from 7 retailers. In comparison to the entire cookie database with 178 products, this is almost on third (29.5%) of all labels and more than half (53.8%) of all retailers. Considering the chocolate chip cookie database, the subset included about half (48.1%) of all labels and more than half (58.3%) of all retailers (**Figure 7**).

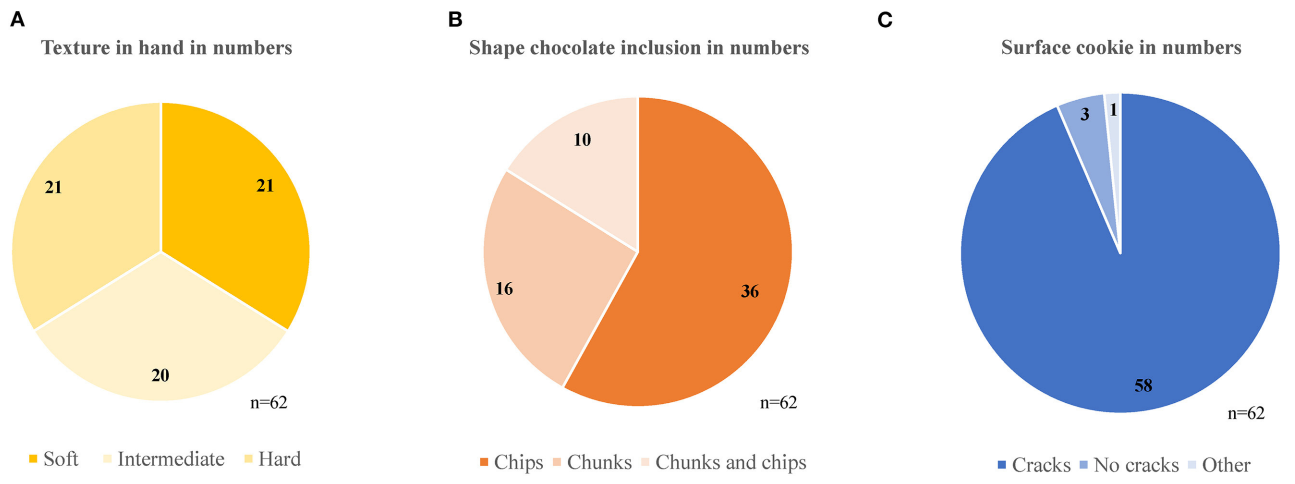


**FIGURE 4 |** Results of the three sensory screenings on all 62 chocolate chip cookies with the additional ranked criteria **(A)** texture in hand, **(B)** shape of the chocolate inclusion and, **(C)** the surface of the cookie. Those three criteria are part of the 11 additional ranked criteria in order to be able to make an informed subset selection which is representative of the market.

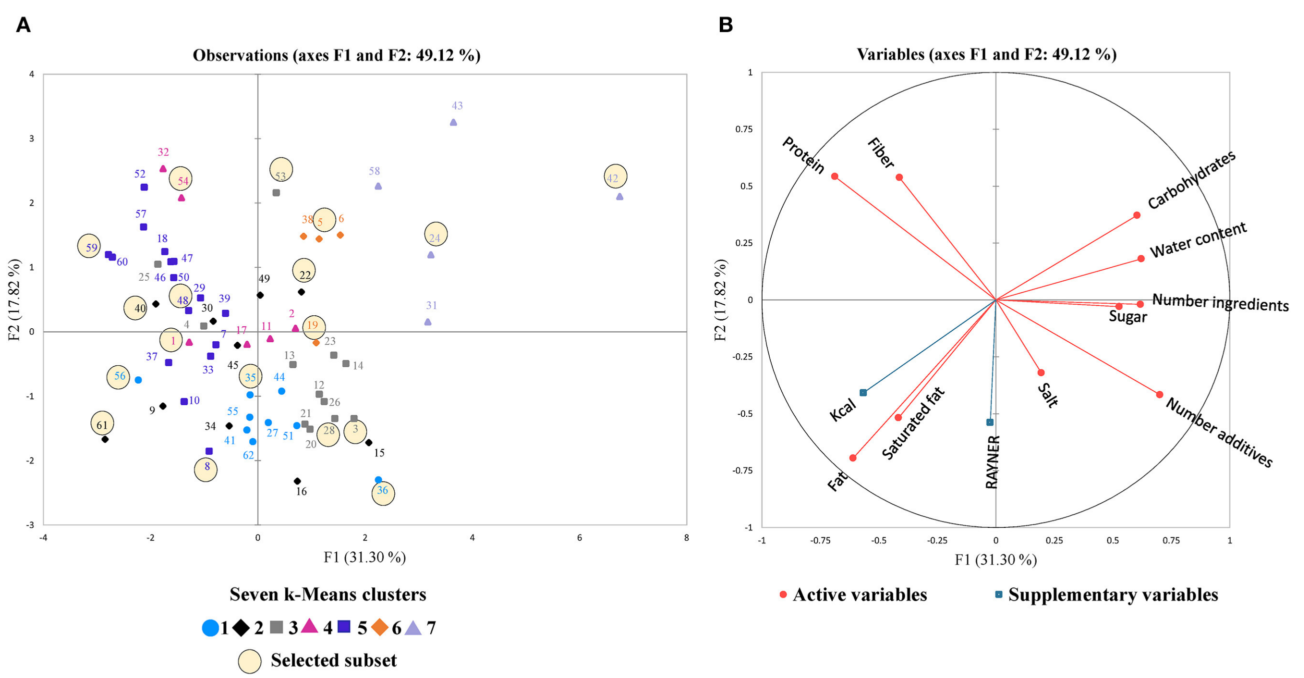


**FIGURE 5 |** Principal component analysis (PCA) with axes F1 and F2. **(A)** observation plot including 62 cookies with chocolate inclusion colored based on seven different K-means clusters and the 18 selected cookies for the subset in black circles, representative of the market. **(B)** PCA correlation matrix with 10 active nutrition, composition, and water content variables (Fat, saturated fat, carbohydrates, sugar, protein, fiber and salt in g per 100 g, water content in %, number ingredients and number additives), and two supplementary variables (Rayner score, Kcal).

## Quantitative Nutrition, Composition, and Water Content Variables

The 18 selected cookies are highlighted on the PCA of the 62 chocolate chip cookies in **Figure 5A**. The subset shows broad nutrition, composition, and water content diversity based on 12 plotted variables. Moreover, the selected cookies of each cluster showed a balanced distribution of extreme cookies within and between clusters based on axes F1 and F2.

As can be seen in **Figure 8** we observed a similar distribution between the chocolate chips cookies and the selected subset

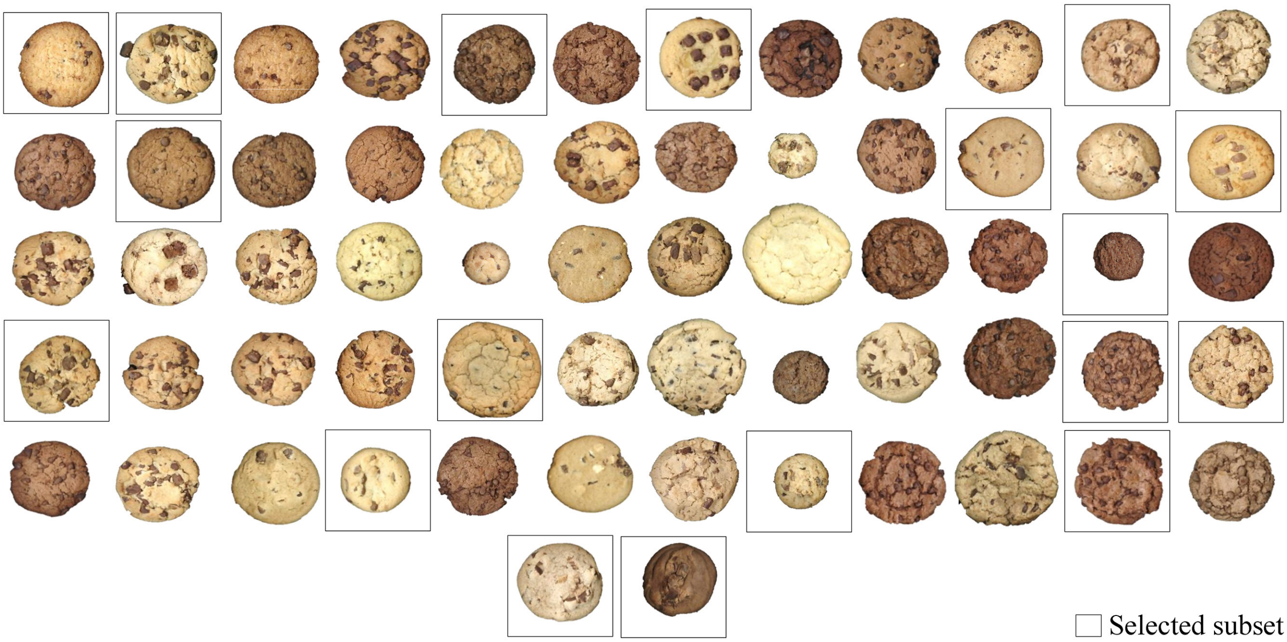


**FIGURE 6 |** This figure shows 62 pictures of the commercial chocolate chip cookies, including the selected subset of 18 cookies.

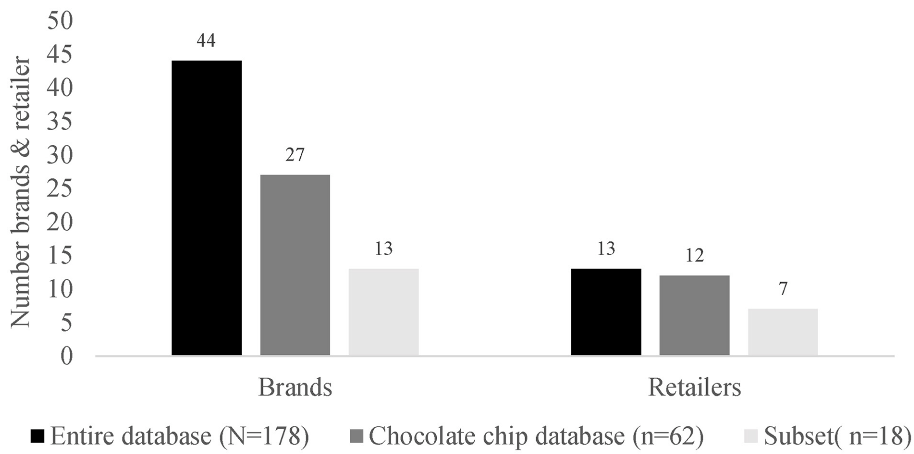


**FIGURE 7 |** Representativity of the brands and retailers of the selected subset.

for almost all variables. The subset led to a slightly different distribution for the kcal, fiber, protein, salt, and water content. The subset also accounted for minimal and maximum values for most of the variables.

## Additional Ranking Criteria

As shown in **Supplementary Table 5**, the selected subset considered 37 out of the 40 subgroups. Three subgroups (criterion type sugar: syrup, criterion amount chocolate, and cacao powder: middle and criterion weight cookie: high) remained unrepresented. Although we did not seek strict representativeness in terms of group sizes, it can be noted that larger subgroups in the chocolate chip cookie database also had more products in the selected subset. The selected subset contained cookies as well with smaller subgroups in the chocolate chip cookie database, such as the subgroup "type of flour, mixed," several subgroups from the criteria "amount chocolate and cacao powder" and "surface cookie" and as well the subgroup "type of chocolate, milk." In addition, each additional ranked criterion was considered when selecting the cookies from the seven clusters.

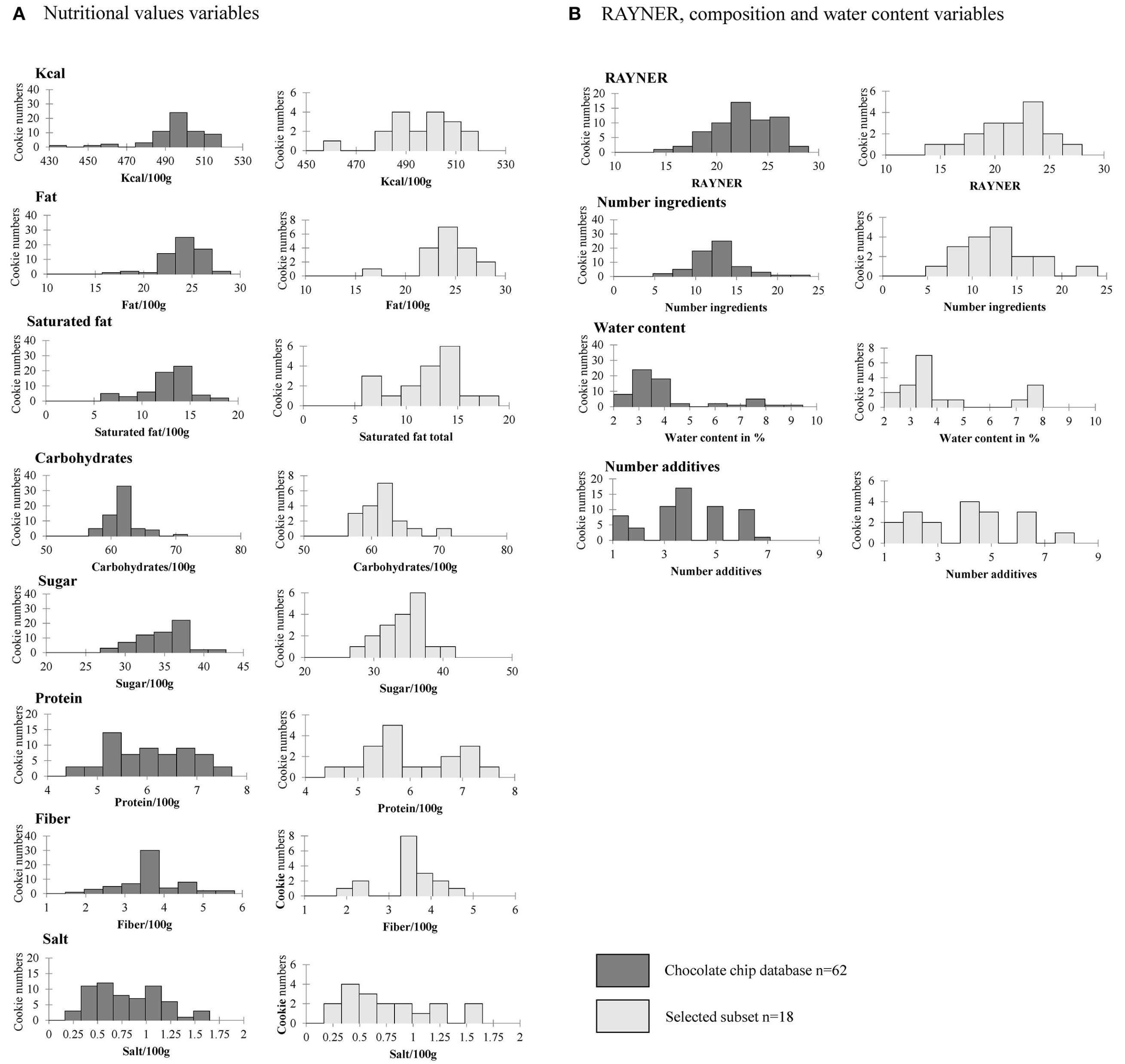


**FIGURE 8 |** Comparison of 12 nutrition, composition and water content variables among chocolate chip cookies and the selected subset based on distributions with **(A)** nutritional values variables and **(B)** Rayner score, composition, and water content variables.

# DISCUSSION

## Focus on a Single Product Category (Cookie) and Specific Cookie Variety (Chocolate Chip Cookies)

Inventorying one product category only is not so frequent in the field of food reformulation. Further, sensory studies often usually deal with a small number of products without considering the full diversity of the market. However, this study showed that working with a single product category has several advantages.

First of all, focusing on a single product category allows increasing the product number in a database, which will lead to an increased product diversity what is important for the later reformulation. As well conducting analyses implies working on physically available products from the market, as that information is not available on the packaging.

Moreover, it is possible to compare food products' nutritional, sensory, and water content characteristics. Although this study focused on chocolate chip cookies only, it was shown that this product category provided large nutrition, composition, water content, and sensory diversity. Those various characteristics are necessary to better understand the links between products' composition, perception, and liking in a later step. Although focusing on cookies only, the measured water content ranged from 2 to almost 10%, with significant consequences on the perceived texture in hand. Further, it was observed that cookies with an increased kcal content had a lower water content and were perceived as harder in hand. A possible explanation might be due to baking parameters, with a higher water loss during baking leading to an increased kcal content and more hard texture. On the opposite, cookies with a higher water content had either a lower baking time or temperature, which lead to a lower kcal content with a softer texture. Besides the baking parameters, an increased sugar content for clusters 7 might as well lead to a softer texture as sugar is bringing moisture to the cookie dough. Therefore, we suggest that the water content provides important information when it comes to food reformulation among baked cereal products.

A past study showed difficulties in comparing the kcal content and setting cut-offs for nutritional values among different product categories, as the products' composition and the portion sizes were too different (3). Comparing only one product category would make it possible to compare all nutritional values and might help therefore setting pertinent cut-off levels for reformulation.

Focusing on cookies with chocolate chips and excluding the other types of cookies, presents the advantage to have a realistic overview of the range of products, but also led to a slightly decreased range for minimum and maximum nutritional values in this study. The used approach might be more inclusive compared with other studies that usually focus on a specific product or a specific ingredient category (29, 30). Moreover, other studies prioritized leading brands, high price segments, or premium and private label brands for their dataset (31, 32). Considering only market leader brands might have relieved cookies availability in our study. However, we suggest that especially recipes from different brands, including the leader and niche brands, may provide a more diverse product variety and represents the current market for one specific product category.

## The Selected Subset and Its Representativity

This study illustrated how a representative subset of products can be selected from detailed market inventory, including multiple criteria by going beyond the nutrition and composition information on the packaging. The selected subset of 18 cookies in this study is representative of 23 variables and criteria. Including different retailers, private labels and international brands might contribute to a broader range of recipes and product diversity in the subset as when only focusing on selected retailers and international or private labels.

This PCA permits to visually assess that the selected subset was representative of the market, with the presence of "extreme" cookies within a single cluster but as well between the seven clusters. However, only two cookies were selected from cluster 7, although this cluster showed the highest variance. Due to the unavailability of some of the corresponding brands on the market, it was not possible to select more cookies from cluster 7. To avoid this situation, we could have excluded cookies that are difficult to purchase from the start. However, this would have limited our knowledge of the market. In addition to this multifactorial analysis, the unidimensional distributions confirmed the good representativity of the subset for almost all 13 variables.

The results in this study showed that almost all subgroups were represented in the selected subset. Several subgroups remained unselected, although they were higher ranked than the remaining criteria. However, these subgroups had already a low representativity in the chocolate chip cookie database, which increases the difficulty to be selected. One possible approach to increase the chance for a selection even for weakly represented subgroups would be to create subgroups based on extreme values, rather than the calculation of balanced quintiles ranks. Moreover, each criterion was considered for selecting cookies among the clusters. However, the selected cookie numbers were not higher among criteria that were higher ranked. Instead, more cookies were selected on lower-ranked criteria. A possible way to solve this disbalance would be to set higher numbers of cookies to be selected at higher-ranked criteria and lower numbers of cookies to be selected at lower-ranked criteria.

The representativity step is a critical step as the main purpose is primarily to help industrials and decision-makers to anticipate the levers of reformulation. Therefore, working with a reduced subset that is representative of the market is required to conduct further in-depth sensory, physicochemical, and liking analyses.

## Potential for Reformulation Among the Product Category Commercial "Chocolate Chip Cookies" and Prospects for the Future Reformulation Work

We found a large heterogeneity among the chocolate chip cookies in terms of nutritional, compositional, water content, and sensory aspects. Large ranges for minimum and maximum values for fat and sugar indicate a potential for certain commercial cookies to reduce their fat and sugar content. Likewise, we identified the potential to increase the fiber content among certain cookies. Moreover, most of the cookies were graded at the highest Nutri-Score (E), which is associated with poor nutritional quality food. However, this study identified commercial cookies with a reduced Rayner score (but only one cookie with a reduced Nutri-Score D), which is associated with lower fat, saturated fat, and higher fiber content. This rise the question of whether the Rayner respectively the Nutri-Score is an optimal tool to reformulate among a product category which is known to have very high sugar and fat contents, as the effect of the sugar and fat reduction on the score itself might be low. On the other hand, a lump of sugar and fat reduction in too large steps implies many constraints. Therefore, besides nutritional values, further parameters such as the level of processing and the number of additives should be considered for reformulation.

Sweet biscuits are important contributors to children's unhealthy diet, composed of high caloric, highly processed, and palatable foods (6, 33–35). Our data confirm the interest and the possibilities of reformulation among the commercial "chocolate chip cookies" product category. There clearly is scope for improvement of some cookies' nutritional profile, especially by reducing fat and sugar content and increasing fiber content.

Besides the nutritional diversity, we identified three types of cookie textures: soft, intermediate, and hard. Texture properties play a key role in sensory perception and as a driver of preferences for food choices (36, 37). Moreover, besides nutrient manipulation, studies have shown that texture modification might be successful to decrease the "obesogenic" eating style. Indeed, food oral processing, impacted by texture, could influence satiety and satiation of individuals with faster eating rates and shorter oral exposure time (38). This confirms the high interest to expand nutritional and compositional databases with water content variables and sensory criteria.

Eventually, we will use the selected subset to identify future reformulation levers, while maintaining sensory perception and liking. This will imply further in-depth sensory and physicochemical analyses, in order to better understand the link between products composition, sensory and physicochemical properties. This understanding is crucial for a successful reformulation, as this multicriteria approach anticipates potential limitations among the reformulation process.

## CONCLUSION

The creation of a database based on a comprehensive market inventory with a focus on a single product category and that considers multiple variables allows to describe and compare the diversity of products. It also sets the basis for future reformulation. Besides nutrition and composition information collected from packaging, simple additional characterizations are useful to better assess product diversity. This allowed making an informed and representative subset selection whose representativity was checked by uni- and multi-variate analyses.

Thanks to this subset selection, it will be possible to conduct in-depth sensory, physicochemical, and hedonic investigations that are required to successfully reformulate the products as one possible answer to the improvement of the food offer. Besides nutrient modification, texture modification with possible impact on the oral process, satiation, and satiety is a further promising reformulation lever.

It is also necessary to create a multicriteria database to ideally identify some healthier formulation solutions which minimize changes in sensory perception, liking, and cost-effectiveness. Increasing the potential for voluntary reformulation among industries on one hand and providing a tool to drive public policies, on the other hand, might strengthen the reformulation as an even more impactful lever to enhance our food environment.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

CL: conceptualization, data curation, formal analysis, investigation, visualization, and writing–original draft. JD and AS-E: conceptualization, funding acquisition, supervision, and writing–review and editing. VB and IS: conceptualization, investigation, supervision, writing–review, and editing. All authors contributed to the article and approved the submitted version.

## FUNDING

This project was supported by the international STOP project (Science and Technology in childhood Obesity Policy) which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 774548.

## ACKNOWLEDGMENTS

A special thank to the students from AgroParisTech for contributing to the sensory screening and water content analysis.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2021.749596/full#supplementary-material

## REFERENCES

1. Gressier M, Swinburn B, Frost G, Segal AB, Sassi F. What is the impact of food reformulation on individuals' behaviour, nutrient intakes and health status? A systematic review of empirical evidence. *Obes Rev*. (2021) 22:e13139. doi: 10.1111/obr.13139
2. Spiteri M, Soler L-G. Food reformulation and nutritional quality of food consumption: an analysis based on households panel data in France. *Eur J Clin Nutr*. (2018) 72:228–35. doi: 10.1038/s41430-017-0044-3
3. Nijman CAJ, Zijp IM, Sierksma A, Roodenburg AJC, Leenen R, van den Kerkhoff C, et al. A method to improve the nutritional quality of foods and beverages based on dietary recommendations. *Eur J Clin Nutr*. (2007) 61:461–71. doi: 10.1038/sj.ejcn.1602548
4. Storcksdieck Genannt Bonsmann S, Robinson M, Wollgast J, Caldeira S. The ineligibility of food products from across the EU for marketing to children according to two EU-level nutrient profile models. *PLoS ONE*. (2019) 14:e0213512. doi: 10.1371/journal.pone.0213512
5. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. (2019) 393:1958–72. doi: 10.1016/S0140-6736(19)30041-8

6. Bandy LK, Scarborough P, Harrington RA, Rayner M, Jebb SA. The sugar content of foods in the UK by category and company: A repeated cross-sectional study, 2015-2018. *PLoS Med*. 2021 18:e1003647. doi: 10.1371/journal.pmed.1003647
7. Ghotra BS, Dyal SD, Narine SS. Lipid shortenings: a review. Food Research International [Internet]. 2002 Jan 1 [cited 2021 Mar 29];35(10):1015–48. Available from: https://www.sciencedirect.com/science/article/pii/S0963996902001631
8. Miller RA, Dann OE, Oakley AR, Angermayer ME, Brackebusch KH. Sucrose replacement in high ratio white layer cakes. *J Sci Food Agric*. (2017) 97(10):3228–32. doi: 10.1002/jsfa.8170
9. Pareyt B, Talhaoui F, Kerckhofs G, Brijs K, Goesaert H, Wevers M, et al. The role of sugar and fat in sugar-snap cookies: Structural and textural properties. *J Food Eng*. (2009) 90:400–8. doi: 10.1016/j.jfoodeng.2008.07.010
10. Marty L, Nicklaus S, Miguet M, Chambaron S, Monnery-Patris S. When do healthiness and liking drive children's food choices? The influence of social context and weight status. *Appetite*. (2018) 125:466–73. doi: 10.1016/j.appet.2018.03.003
11. Nguyen SP, Girgis H, Robinson J. Predictors of children's food selection: The role of children's perceptions of the health and taste of foods. *Food Qual Prefer*. (2015) 40:106–9. doi: 10.1016/j.foodqual.2014.09.009
12. Batista LF, Marques CS, Pires AC dos S, Minim LA, Soares N de FF, Vidigal MCTR. Artificial neural networks modeling of non-fat yogurt texture properties: effect of process conditions and food composition. *Food Bioprod Process*. (2021) 126:164–74. doi: 10.1016/j.fbp.2021.01.002
13. Monnet A-F, Laleg K, Michon C, Micard V. Legume enriched cereal products: A generic approach derived from material science to predict their structuring by the process and their final properties. *Trends Food Sci Technol*. (2019) 86:131–43. doi: 10.1016/j.tifs.2019.02.027
14. Vasanthakumari P, Jaganmohan R. Process development and formulation of multi-cereal and legume cookies. *J Food Process Preserv*. (2018) 42:e13824. doi: 10.1111/jfpp.13824
15. Hough G, Buera MDP, Chirife J, Moro O. Sensory Texture of Commercial Biscuits as a Function of Water Activity. *J Texture Stud*. (2001) 32:57–74. doi: 10.1111/j.1745-4603.2001.tb01034.x
16. Baltsavias A, Jurgens A, van Vliet T. Properties of Short-Dough Biscuits in Relation to Structure. *J Cereal Sci*. (1999) 29:245–55. doi: 10.1006/jcrs.1999.0250
17. Baltsavias A, Jurgens A, Vliet TV. Factors affecting fracture properties of short-dough biscuits. *J Texture Stud*. (1997) 28:205–19. doi: 10.1111/j.1745-4603.1997.tb00111.x
18. Faridi H. *The Science of Cookie and Cracker Production*. Chapman and Hall (1994). Available online at: https://scholar.google.com/scholar_lookup?title=The$+$Science$+$of$+$cookie$+$and$+$cracker$+$production&author$=$Faridi%2C$+$H.&publication_year$=$1994 (accessed December 6, 2021)
19. Laguna L, Varela P, Salvador A, Fiszman S. A new sensory tool to analyse the oral trajectory of biscuits with different fat and fibre contents. *Int Food Res J*. (2013) 51:544–53. doi: 10.1016/j.foodres.2013.01.003
20. Laguna L, Varela P, Salvador A, Sanz T, Fiszman SM. Balancing Texture and Other Sensory Features in Reduced Fat Short-Dough Biscuits. *J Texture Stud*. (2012) 43:235–45. doi: 10.1111/j.1745-4603.2011.00333.x
21. Manohar RS, Rao PH. Effect of sugars on the rheological characteristics of biscuit dough and quality of biscuits. *J Sci Food Agric*. (1997) 75:383–90.
22. Rayner M. Nutrient profiling for regulatory purposes. *Proc Nutr Soc*. (2017) 76:230–6. doi: 10.1017/S0029665117000362
23. Upadhyay S, Khan S, Tiwari R, Kumar S, Rautela I, Muktawat P, et al. Nutritional and sensory evaluation of herbal cookies. *Int J Food Sci*. (2017) 2:156–60.
24. Pagliuca MM, Scarpato D. The olive oil sector: a comparison between consumers and "experts" choices by the sensory analysis. *Procedia Econ*. (2014) 17:221–30. doi: 10.1016/S2212-5671(14)00897-1
25. Alves WAL, Araújo SAD, Pessota JH, Santos RABOD. A methodology for sensory evaluation of food products using self-organizing maps and K-means algorithm. *Appl Mech Mater*. (2013) 263–266:2191–4. doi: 10.4028/www.scientific.net/AMM.263-266.2191
26. Mora M, Urdaneta E, Chaya C. Emotional response to wine: Sensory properties, age and gender as drivers of consumers' preferences. *Food Qual Prefer*. (2018) 66:19–28. doi: 10.1016/j.foodqual.2017.12.015
27. Thompson JL, Drake MA, Lopetcharat K, Yates MD. Preference Mapping of Commercial Chocolate Milks. *J Food Sci*. (2004) 69:S406–13. doi: 10.1111/j.1365-2621.2004.tb09958.x
28. Baxter MJ. Standardization and transformation in principal component analysis, with applications to archaeometry. *J R Stat Soc*. (1995) 44:513–27. doi: 10.2307/2986142
29. Pearson W, Schmidtke L, Francis IL, Blackman JW. An investigation of the Pivot© Profile sensory analysis method using wine experts: Comparison with descriptive analysis and results from two expert panels. *Food Qual Prefer*. (2020) 83:103858. doi: 10.1016/j.foodqual.2019.103858
30. Sune F, Lacroix P, Huon de Kermadec F. A comparison of sensory attribute use by children and experts to evaluate chocolate. *Food Qual Prefer*. (2002) 13:545–53. doi: 10.1016/S0950-3293(02)00057-5
31. Martinez C, Santa Cruz MJ, Hough G, Vega MJ. Preference mapping of cracker type biscuits. *Food Qual Prefer*. (2002) 13:535–44. doi: 10.1016/S0950-3293(02)00087-3
32. Schouteten JJ, Verwaeren J, Lagast S, Gellynck X, De Steur H. Emoji as a tool for measuring children's emotions when tasting food. *Food Qual Prefer*. (2018) 68:322–31. doi: 10.1016/j.foodqual.2018.03.005
33. Costa CS, Del-Ponte B, Assunção MCF, Santos IS. Consumption of ultra-processed foods and body fat during childhood and adolescence: a systematic review. *Public Health Nutr*. (2018) 21:148–59. doi: 10.1017/S1368980017001331
34. Forde CG, Mars M, de Graaf K. Ultra-Processing or oral processing? a role for energy density and eating rate in moderating energy intake from processed foods. *Curr Dev Nutr*. (2020) 4. doi: 10.1093/cdn/nzaa019
35. Hall KD, Ayuketah A, Brychta R, Cai H, Cassimatis T, Chen KY, et al. Ultra-Processed Diets Cause Excess Calorie Intake and Weight Gain: An Inpatient Randomized Controlled Trial of Ad Libitum Food Intake. *Cell Metab*. (2019) 30:67–77.e3. doi: 10.1016/j.cmet.2019.05.008
36. Brown WE, Braxton D. Dynamics of food breakdown during eating in relation to perceptions of texture and preference: a study on biscuits11This paper was presented at the third rose Marie Pangborn memorial symposium (Ålesund, Norway, 9–13 August 1998). *Food Qual Prefer*. (2000) 11:259–67. doi: 10.1016/S0950-3293(99)00014-2
37. Jeltema M, Beckley J, Vahalik J. Model for understanding consumer textural food choice. *Food Sci Nutr*. (2015) 3:202–12. doi: 10.1002/fsn3.205
38. Fogel A, Goh AT, Fries LR, Sadananthan SA, Velan SS, Michael N, et al. A description of an 'obesogenic' eating style that promotes higher energy intake and is associated with greater adiposity in 4.5year-old children: Results from the GUSTO cohort. *Physiol Behav*. (2017) 176:107–16. doi: 10.1016/j.physbeh.2017.02.013


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fnut.2021.796082

Check for updates

# Multi-Dimensional Dataset of Open Data and Satellite Images for Characterization of Food Security and Nutrition

*David S. Restrepo*[1*], *Luis E. Pérez*[1], *Diego M. López*[1], *Rubiel Vargas-Cañas*[2] and *Juan Sebastian Osorio-Valencia*[3]

[1] Telematics Department, Telematics Engineering Research Group, Universidad del Cauca, Popayán, Colombia, [2] Physics Department, Dynamic Systems, Instrumentation and Control Research Group, Universidad del Cauca, Popayán, Colombia, [3] Department of Global Health, University of Washington, Seattle, WA, United States

OPEN ACCESS

**Edited by:**
Alessandra Durazzo,
Council for Agricultural Research and Economics, Italy

**Reviewed by:**
Esperanza Nieto,
Universidad Nacional Abierta y a Distancia UNAD, Colombia
Paulo Lopez-Meyer,
Intel, United States

***Correspondence:**
David S. Restrepo
dsrestrepo@unicauca.edu.co

**Specialty section:**
This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition

**Received:** 15 October 2021
**Accepted:** 10 December 2021
**Published:** 27 January 2022

**Citation:**
Restrepo DS, Pérez LE, López DM, Vargas-Cañas R and Osorio-Valencia JS (2022) Multi-Dimensional Dataset of Open Data and Satellite Images for Characterization of Food Security and Nutrition. Front. Nutr. 8:796082. doi: 10.3389/fnut.2021.796082

**Background:** Nutrition is one of the main factors affecting the development and quality of life of a person. From a public health perspective, food security is an essential social determinant for promoting healthy nutrition. Food security embraces four dimensions: physical availability of food, economic and physical access to food, food utilization, and the sustainability of the dimensions above. Integrally addressing the four dimensions is vital. Surprisingly most of the works focused on a single dimension of food security: the physical availability of food.

**Objective:** The paper proposes a multi-dimensional dataset of open data and satellite images to characterize food security in the department of Cauca, Colombia.

**Methods:** The food security dataset integrates multiple open data sources; therefore, the Cross-Industry Standard Process for Data Mining methodology was used to guide the construction of the dataset. It includes sources such as population and agricultural census, nutrition surveys, and satellite images.

**Results:** An open multidimensional dataset for the Department of Cauca with 926 attributes and 9 rows (each row representing a Municipality) from multiple sources in Colombia, is configured. Then, machine learning models were used to characterize food security and nutrition in the Cauca Department. As a result, The Food security index calculated for Cauca using a linear regression model (Mean Absolute Error of 0.391) is 57.444 in a range between 0 and 100, with 100 the best score. Also, an approach for extracting four features (Agriculture, Habitation, Road, Water) of satellite images were tested with the ResNet50 model trained from scratch, having the best performance with a macro-accuracy, macro-precision, macro-recall, and macro-F1-score of 91.7, 86.2, 66.91, and 74.92%, respectively.

**Conclusion:** It shows how the CRISP-DM methodology can be used to create an open public health data repository. Furthermore, this methodology could be generalized to other types of problems requiring the creation of a dataset. In addition, the use of satellite images presents an alternative for places where data collection is challenging. The model and methodology proposed based on open data become a low-cost and effective solution that could be used by decision-makers, especially in developing countries, to support food security planning.

**Keywords: data mining, food security, machine learning, remote sensing, satellite imagery, dataset**

## INTRODUCTION

In 2015, United Nations member states adopted the Sustainable Development Goals (SDG), a commitment to accomplishing before 2030 a set of universal actions to end poverty, protect the planet, and promote peace and prosperity in the world (1). They placed a high priority on food security and nutrition (SDG). The target for 2030 included, among others, ending hunger and all forms of malnutrition; ensuring access to safe, nutritious, and sufficient food all year round; doubling the agricultural productivity and incomes of small-scale food producers and ensuring sustainable food production systems; and implementing resilient agricultural practices that increase productivity and production. Considering the above, establishing food security actions and measuring their impact is critical for decision-making in a country or specific region like a department or municipality.

According to the Food and Agriculture of the United Nations (FAO), food security encompasses four dimensions: physical availability of food economic and physical access to food, food utilization, and the sustainability of the dimensionsmentioned above (2). These dimensions must be taken into account to calculate a comprehensive food security index that allows characterizing and comparing the food security levels of a country or geographic region. Some works have been proposed around creating a food security index in which the four dimensions proposed by FAO are integrated. Among them, the one that stands out the most is the Global Food Security Index (GFSI), introduced by the Economist Intelligence Unit (EIU) (3). While this index meets the dimensions and has proven to be accurate (4), it is an index that is difficult to calculate mainly due to the large amount of data required and the costs involved in collecting this data. Therefore, in most cases, the GFSI is applied only at the country level and not at the regional level.

There are other alternatives for calculating a food security index without requiring a large amount of data. Some examples are the Global Hunger Index (GHI) (5) or The Food Insecurity Experience Scale (FIES) from FAO (6). However, these indices do not meet all the dimensions described by GFSI. They focus on nutritional and physical access issues, which gives us an incomplete perspective on the problem. The challenge then for governments is to have comprehensive and complete sources of information to characterize the food security situation at the national, regional, or local level to support decision-making and achieve the SDGs.

According to the information described in the Colombian National Survey of Nutritional Situation document (7), food insecurity in Colombia per household was 54.2%, which means that at least one out of every two households does not guarantee food security. In households with food insecurity, 13.8% corresponds to moderate food insecurity, and 8.5% corresponds to severe food insecurity. This situation is aggravated in the rural area, where it represents 64.1%. According to GFSI data, Colombia's food security index corresponds to 63.2 points (this value indicates how favorable the food security environment is in the country). This result is the most complete that can be found at the country level. However, Colombia is a large country with multiple landscapes and economies, with food security affected by them. Therefore, for decision-making at the departmental or municipal level, it is necessary to calculate an adapted and specific food security index. However, no data is available at this level.

For the Department of Cauca, one of the more vulnerable and underserved regions in Colombia; collecting the data necessary to calculate a complete food security index is a highly complex and expensive process due to the difficulties in accessing some locations. Given this problem raises the possibility of using satellite data to cover remote areas through images (8), in addition to other meteorological data and combine these with different data obtained from other sources such as nutrition or agriculture surveys, censuses, or health reports, so that a completely open data multidimensional dataset that allows calculating the GFSI for this department can be obtained.

Machine Learning models (ML) can be potentially used for the calculation of food security. Nevertheless, current models are focused on analyzing the first FAO's dimension of physical availability of food through the prediction of crop yield (9–24), and are not used to calculate an index itself. On the other hand, other techniques such as GHI and FIES used for calculating a food security index at a regional level mainly focus on nutritional or economic and physical access aspects (25). As a result, it leaves out the four dimensions described by FAO.

This paper aims to build a multi-dimensional dataset of open data and satellite images to characterize food security in the Department of Cauca, Colombia. The dataset is structured according to the four-dimensional model of food security proposed by FAO. An ML-based model that calculates the GFSI for the Department of Cauca is trained and tested with the GFSI data to validate the dataset.

## MATERIALS AND METHODS

There are many methodologies and indicators to measure food security. They can broadly be classified based on the analysis of (a) primary data sources (expert opinion and community perspectives) and (b) secondary data (26). In building a comprehensive dataset, using open and secondary data sources with varied nature and format, and large volumes of information, it was decided to follow the Cross-Industry Standard Process for Data Mining methodology (CRISP-DM) (27). CRISP-DM is the most widely used methodology in data mining projects analyzing large datasets. This methodology consists of 6 phases (Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment). In this work, the first five phases of the methodology were addressed. The last phase, deployment, is outside the scope of this paper.

Following the methodology we expected to obtain a dataset that addresses the four dimensions described by FAO. Additionally, the quality of the database is validated using an ML model to calculate the GFSI in some municipalities of Cauca. The phases are explained below.

## Business Understanding

In this phase, we reviewed the existing literature to determine the causes of food security and nutritional problems, as well as possible consequences and existing measures. The results were categorized into the four dimensions described by FAO:

- Physical availability of food. This dimension includes the quantity of food offered, which comprises food production, stocks, and net trade.
- Economic and physical access to food. In addition to having physical availability, it must also be guaranteed that the population can acquire these foods. To obtain food, they must have economic resources and easy access (i.e., roads, avenues, routes).
- Food utilization. Having access to food is not always enough. Aspects such as a healthy diet, optimal food preparation, a good distribution of nutrients, and quality food must also be considered.
- The sustainability of the dimensions mentioned above is one factor that makes measuring food security so complex. In addition, environmental, political, or economic variables are considered in this dimension and may also influence food security indices over time.

The GFSI is the selected food security index as it is the only one that covers the four dimensions described by FAO. The challenge would be in the data collection required for its calculation. Due to its complexity, it is an index calculated at a country level but not for specific regions like departments or municipalities. Then the use of a methodology for data mining is necessary to create a dataset and the respective ML model. It confirms the need to use CRISP-DM as the methodology for this research.

## Data Understanding

Once we understood the problem, we started with the data collection. The data was structured according to the fields provided by the GFSI and the FAO dimensions For example, in the case of Cauca, the data collected were:

- Satellite Imagery: the Google Earth Engine (GEE) platform (28) offers a dataset of Landsat-8 satellite imagery from April 2013 to the present. This satellite provides an image with an approximate frequency of once every 16 days and has different bands where the B4, B3, and B2 bands are the Red, Green, and Blue bands, respectively. These bands can be used to build an RGB image with a resolution of 30 m/px, and the B8 is the Panchromatic band with a resolution of 15 m/px and can be used to increase the resolution of the RGB image. These data were collected due to the potential for obtaining information on crop and water sources (physical availability), roads (economic and physical access), and temporal scale (sustainability of the dimensions over time).
- Nutrition data: Colombia's nutrition data are obtained from the National Survey of Nutritional Situation (ENSIN) (29). This is a survey carried out every 5 years by a public entity of the Colombian government to classify the nutritional situation by departments and municipalities. It has information on economic income and sociodemographic data belonging to the economic and physical access dimension, and the food consumed and the vitamins of a person (food utilization). This is the most comprehensive survey in the country related to nutrition and food security. The available data correspond to the year 2015.
- Agricultural Census: the agricultural census data come from the National Statistics Office (DANE) (30), and correspond to the 2014 census. In this data, farmers are asked questions about the conditions of some crops, access to financing, infrastructure, and planted area. These data are part of the dimension of the physical availability of food. These surveys are carried out at the municipality level by department.
- Health Records: health data is usually the most complex data to obtain due to the resources and tools necessary to get it and the limitations of information privacy. However, it is possible to access public records that are part of the National Public Health Surveillance System (SIVIGILA) (31). From this system, information about the weekly admissions to hospitals can be obtained. In this dataset, data on acute malnutrition in children under 5 years of age can be obtained from January 2016 to the present. Also, data on mortality due to malnutrition from January 2013 to the present. These data would allow us to analyze sustainability.
- Meteorological Data: within the category of meteorological data, many types of data from different sources are relevant to this study, such as temperature, soil type, precipitation, soil maps, humidity, etc. However, due to the complexity of processing these data, we only use the temperature and precipitation, that are part of the physical availability of food dimension. Therefore, two primary data sources were considered: WorldClim (32) and Google Earth Engine–GEE. The first provides pre-processed information on monthly temperature and precipitation data from 1960 to 2018, while in the second, this temporal resolution depends on the data source.

## Data Preparation

In this phase, data preprocessing was done to have it ready for developing the model. This phase includes data selection, cleaning, changing formats, creating new variables, and merging. In this case, the data preparation was divided into two sections according to the two existing data modalities, satellite images, and metadata.

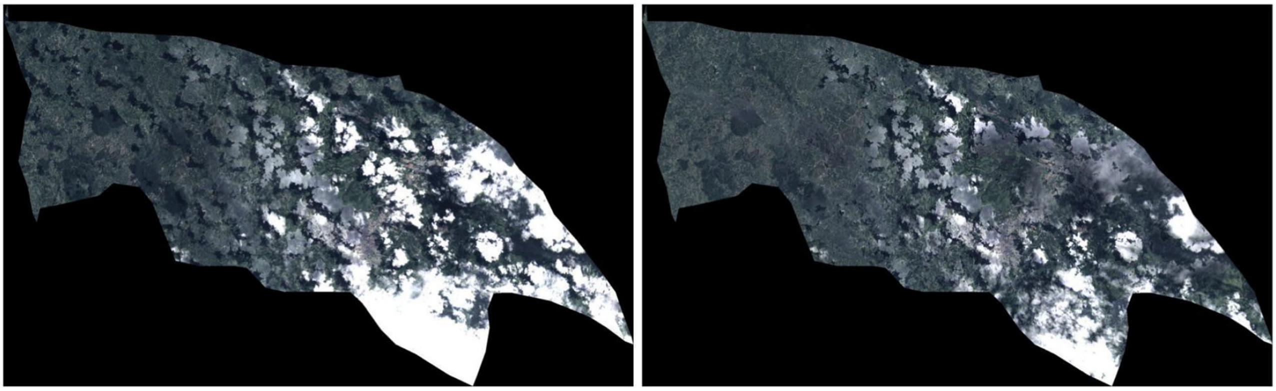


**FIGURE 1 |** Resulting image choosing different percentages in the composite. **(A)** Landsat satellite composite image of the municipality of Popayán using images collected during 2 months. **(B)** Landsat satellite composite image of the municipality of Popayán using images collected during 4 months.

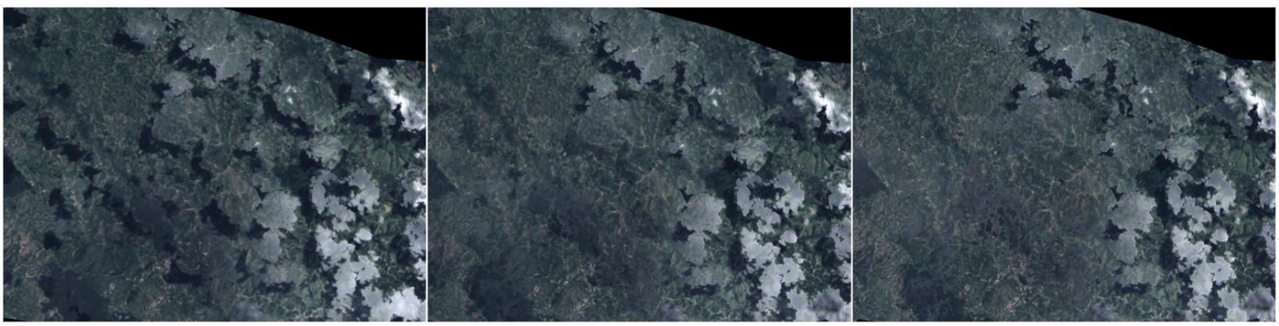


**FIGURE 2 |** Resulting image choosing different percentages in the composite. **(A)** Composite of landsat images using the pixel value in the percentile 20th. **(B)** Composite of landsat images using the pixel value in the percentile 50th. **(C)** Composite of landsat images using the pixel value in the percentile 75th.

## Data Preparación for Satellite Imagery

### *Image Selection*

Because the proposal of this section is only an alternative to the lack of data and in this case it is not the data used for the calculation of the GFSI, images were used only for the capital of the department of Cauca (Popayán). Then, the GEE platform was used to crop the images of the selected satellite. For this case, Landsat-8 was used, and the limits of each image were defined with the shapefile of the municipality.

### *Cloud and Shadow Removal*

Monthly images of the satellite were created trying to eliminate as many clouds and shadows as possible. To create the image, a composition was made of a series of images captured in an interval of 3 months in the future and 1 month in the past in the area to be studied in a given month. Such a long time interval was used to get more image samples and cover more areas without clouds or shadows. A smaller quantity of images resulted in an image of poorer quality, as seen in **Figure 1**. The method used here calculated the percentage of clouds in each pixel and selected the pixels with the lowest rate of clouds. Among those pixels with the lowest ratio of clouds, the value was taken at the 75th percentile. The reason for using the 75th percentile value was that pixels with shadows correspond to the darkest values (lowest percentiles), as shown in **Figure 2**. Therefore, the pixels with the highest cloud index (the lightest) were eliminated.

### *Resolution Increase*

The resolution of the images was increased in the bands B4, B3, and B2, corresponding to the colors red, green, and blue. To do this, a transformation of the image was made from RGB to HSV, and was operated with band 8 (Panchromatic) at 15 m/px. Finally, a reconversion from HSV to RGB is done again with the values scaled to 15 m/px, as shown in **Figure 3**.

### *Features Extraction*

To use satellite images in conjunction with metadata from other sources, there are several techniques. We selected the extraction of features from the image due to its simplicity and interpretability. The characteristics extracted were agriculture, inhabitants, roads, and water. For this, deep learning models are trained to detect the presence of the four characteristics using

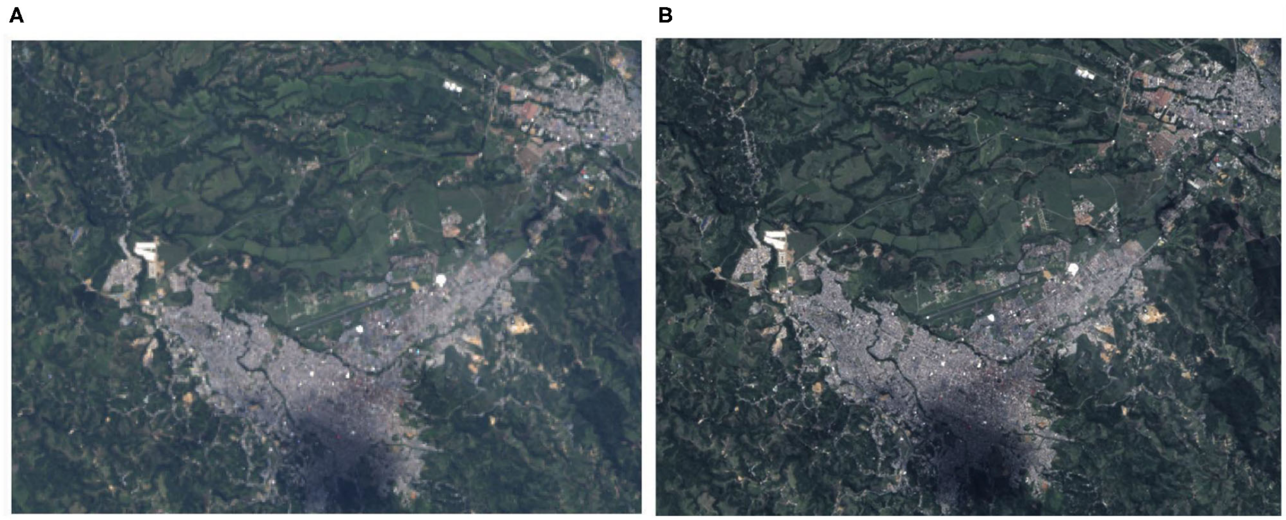


**FIGURE 3 |** Increased resolution. **(A)** Landsat satellite image of the municipality of Popayán with the original RGB resolution of 30 m/px. **(B)** Landsat satellite image of the municipality of Popayán with the increased RGB resolution of 15 m/px.

the dataset "Planet: Understanding the Amazon from Space" available at Kaggle (33). The "Planet" dataset has 40,479 Images of 256 × 256 pixels, each image tagged with the features present in it. Once the images were selected, the dataset was randomly divided into two parts, 80% used in training and the test data with the remaining 20%. The training dataset was subdivided randomly into 90% for training and 10% for validation. The task was a multi-label classification in which models such as ResNet 50 (34), VGG 16 (35), and Visual Transformers (ViT) (36). The VGG 16 model was also trained using the pre-trained weights from ImageNet. To evaluate each model, precision, recall, and the f1-score were calculated to predict each class in the 20% of test data. The model selected as the best was the model with the highest f1-score (**Figure 8**) because it is the most balanced metric from those calculated.

The best model was a ResNet 50 model trained from scratch. The original architecture of ResNet 50 was modified to have as input a tensor of shape 256, 256, 3 because of the shape of the images in Planet dataset. The output was also modified to have a number of neurons as the number of features to predict, each neuron with a sigmoid activation function. The model was trained using batches of 32 images as input, each image with a random augmentation (flipped, shift between 0 and 15%, zoom between 0 and 40%, and rotation between 0 and 30%). The output was a matrix with batch size as the number of rows and each feature as a column indicating if there is presence of a given feature in the image (1) or not (0). To train the model the loss function used was binary cross entropy in conjunction with an Adam optimizer with 0.001 as learning rate, 0.9 as beta 1, 0.999 as beta 2, and 1e-07 as epsilon. The model was trained using early stop comparing the validation losses to avoid overfitting. If the validation losses increased during 5 consecutive epochs by a value >1e-3, the training would stop and return to the best weights.

Once the model was selected, we made predictions on the dataset of images extracted from GEE. As an image of a municipality is usually much larger than the images in which the model was trained, this image must be divided into 256 × 256 pixel patches, as seen in **Figure 4**. In **Figure 4**, it is also noticed that the mask (black region) covers all or most of the patch. For this reason, only patches with relevant information were selected. To select the patches with relevant information, a threshold of 30% was defined for the maximum number of clouds (totally white pixels) and mask (totally black pixels) admitted in each patch. If a patch does not contain at least 70% of the terrain information, it was removed.

With the 256 × 256 patches, the prediction was made using the ResNet 50 model on each patch. This told us whether there is a presence of the classes in each image or not. With the results of these predictions, an estimate was calculated for the number of water sources, crops, populated areas, or roads in a specific department or municipality.

## Data Preparation for Metadata

- Nutrition data and Agricultural Census: the data from these sources contains many variables, therefore we just selected those directly related to the GFSI, ensuring that the data was numeric data or binary fields (yes/no) to facilitate its processing. In the selected variables the "no" values were assigned to 0 and "yes" values to 1. Also, within the data cleaning process, the empty data was replaced with 0 so that it did not have any weight in the final count when we were counting the proportion of yes values over the total. The rows with missing data with other numerical values were dropped, except in the case of the vitamins table, there were many variables with empty data. In this case, we decided to impute the missing data with the average of the corresponding variable. Once the data was cleaned, a grouping of each table

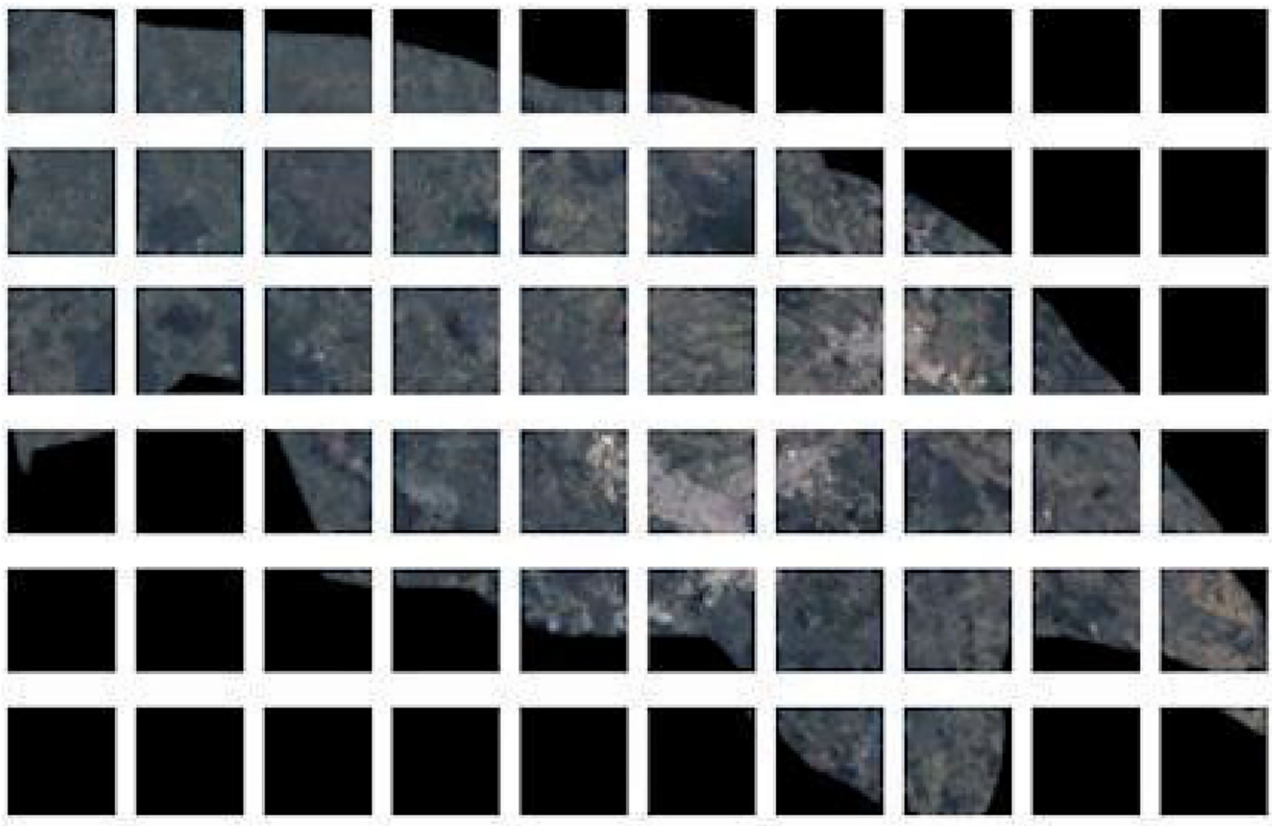

**FIGURE 4 |** Patch generation. Satellite image of municipality of Popayán divided into patches of images of 256 × 256 pixels used to generate the predictions of features found in each patch.

was made concerning the key variable that relates all the tables. Next, all the tables were grouped by the municipality to have a single value for each column per municipality adding the number of ones (yes) and dividing over the total of answers to have the proportion of yes.

- Health Records: health data provided information on hospital admissions. Therefore, they have data irrelevant for the GFSI. For this reason, the fields that had to do only with nutrition were filtered, and a dataset was generated where the rows are the municipalities, and the columns are the number of reports in a specific week in chronological order.
- Meteorological Data: this data was extracted from raster files containing the pixel value (temperature or precipitation in a specific month) and the position (latitude and longitude). Therefore, having the geographic coordinates of a location, these data was extracted at that point to build a dataset.

Finally, all the nutrition, agriculture, health, and meteorological data were integrated into a single dataset, joining them by a municipality code. As a result, the dataset obtained contains 926 columns, representing different multidimensional characteristics, and 9 rows, each row representing a municipality of Department of Cauca where data could be obtained (Popayán, Balboa, Bolívar, Paez, Patía, San Sebastián, Santander de Quilichao, Timbío, Timbiquí). Although not all columns were used to calculate the GFSI, having so many characteristics allows this dataset to be used for further research.

A dataset with all 42 municipalities of Cauca was also generated. However, due to the lack of nutritional data in several municipalities, the data of the missing municipalities must be imputed in order to calculate the GFSI for all these municipalities. A simple solution to solve this problem of lack of data was by imputing in the empty variables the average value for the 9 municipalities with nutritional data. However, for simplicity and quality of the results, the rest of the article will continue to be explained only for the dataset with 9 municipalities.

## Modeling and Evaluation

In the modeling and evaluation stage, a model to calculate the GFSI was built. Here it is important to bear in mind that the evaluation of the model is not done with the same data in which the model was trained, since the model could be memorizing these data and we expect it to generalize to other datasets.

For the construction of the model, the data available in the GFSI in 2020 by country, published annually by the EIU (3), was used. This data comprises 60 columns or items and 113 rows, each row representing a country. It should be noted that these data were not normalized, therefore for each variable, the

**TABLE 1** | GFSI score by the municipality.

| Model | MAE (normalized) | MAE (not normalized) |
|---|---|---|
| Linear regression | 0.391 | 2.714 |
| Ridge regression | 0.391 | 3.338 |
| Lasso | 0.408 | 3.442 |
| Elastic net | 0.398 | 3.423 |
| Multilayer perceptron | 0.825 | 4.383 |
| Support vector machine | 0.379 | 3.542 |
| Random forest | 3.286 | 3.315 |

*The table shows the mean absolute error of the models tested for the calculation of the GFSI using a test portion of the GFSI dataset by country provided by the EIU, the evaluation was done with normalization and without normalization.*

normalization indicated by the GFSI was carried out to convert the variables into a range from 0 to 100.

With the normalized dataset of GFSI in 2020 by country, a variable to predict was defined, this was the GFSI score for each country, and the other 59 variables will be the data used to train the model to predict the GFSI. The dataset was randomly divided into training and testing using 80% of the data for training and the remaining 20% for testing.

The metric chosen to evaluate the model was the Mean Absolute Error (MAE). Different models were tested on the normalized and non-normalized dataset of GFSI in 2020 by country (see **Table 1**). We obtained better performance with linear regression in both cases.

Once we have the best model, we calculated the food security index in the municipalities of the Department of Cauca for which the dataset was built. To calculate the GFSI, we tried to approximate the required variables with the variables available in the created dataset. In this way, it was possible to reconstruct the GFSI on a departmental or even municipal scale.

For some GFSI items, there were no variables at the municipal level in the dataset created, so variables at the departmental or national level were used. Some of them were taken from the GFSI of Colombia for the year 2014.

# RESULTS

The results described in this article are divided into the same two sections according to the two existing data modalities, satellite images, and metadata for calculating the GFSI.

## Feature Extraction From Satellite Images

Due to the model task, which is a multi-label classification task, the selected metrics were accuracy, precision, recall, and F1-score. With these metrics, it is possible to see the model's performance in each characteristic. Furthermore, we can also see a sample of the model's performance globally using the macro-accuracy, macro-precision, macro-recall, and macro-F1-score, which is an average of the performance of a model for each feature.

Within the tested models, the accuracy, precision, recall, and F1-score results are presented in **Figures 5–8**. The F1-score is selected as the main metric because it is the most balanced one. As a result, the architecture of the Convolutional Neural Network (CNN) ResNet 50 has better performance in most tasks. Therefore, it can be said that the model performs well on most labeled data, even with unknown data. However, these metrics could decrease since the images used for the prediction extracted from Landsat-8 have different resolution and zoom levels than the images in the dataset the model was trained with and evaluated. The reason for selecting images from a different source than the one used in Kaggle, is to ensure that the experiment can be reproduced. Furthermore, the use of Landsat-8 images is open and can be done anywhere in the world.

Finally, the predictions of the characteristics on the 256 x 256 patches for the municipality of Popayán in Cauca can be seen in **Figure 9**. Each image is labeled to obtain the presence of the four characteristics with which the model was trained with.

## Calculation of the GFSI for the Department of Cauca

The results obtained concerning the GFSI for the Department of Cauca, both for the model trained with the normalized and non-normalized datasets of GFSI at a country level, are found in **Table 2**. There is a significant difference between normalized and unnormalized data ($p = 0.15719$). Here we see how the Department of Cauca with an index of 57.444 calculated with normalized data, presents a deficit in food security ($-5.756$ points) concerning the value of the GFSI of Colombia which is 63.2 for the year in which the index was calculated by EIU. The score for the Department of Cauca, which is made up of 42 municipalities, is obtained through the average of the scores of the nine municipalities (Popayán, Balboa, Bolívar, Paez, Patía, San Sebastián, Santander de Quilichao, Timbío, Timbiquí) for which data are available. According to the results, it can be seen that the municipality with a lower food security index is Timbiquí, which has a value of 53.637 with normalized data. When compared with the index calculated for the Department of Cauca which is 57.444 (normalized data), this municipality presents a variation of 3.807 points below the average. In addition, it is observed that this municipality has 9.563 points (normalized) below the national GFSI for 2014.

The municipality of Popayán is the one that obtained the best result in the Department, with scores of 60.569 (normalized). When comparing it with the result of the whole department, a difference of 3.125 points (normalized) is observed above the average. This is an expected value since Popayán is the capital city, while the other municipalities might receive less attention by the national and departmental government. However, although comparing the result with the average value of Colombia, we obtained 2.631 points with normalized data, below the national average.

# DISCUSSION

This article presents the process of building a multi-dimensional dataset of open data and satellite images to characterize food

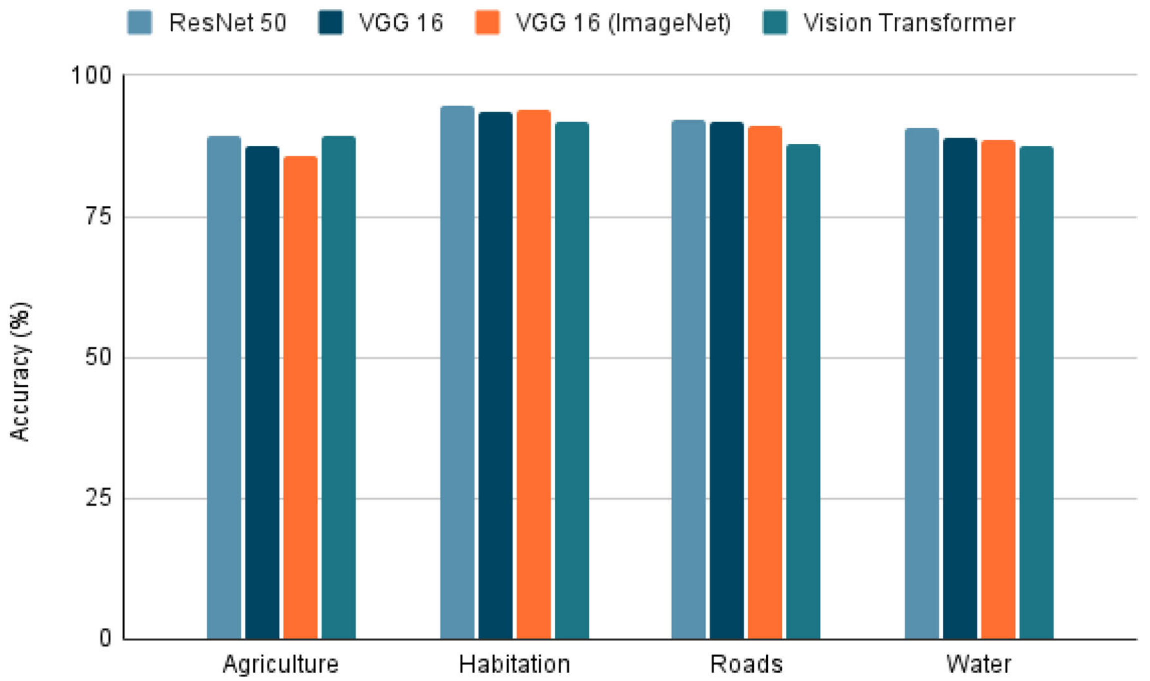


**FIGURE 5 |** Accuracy on tested models. Accuracy of the 4 models tested in the extraction of the 4 characteristics (agriculture, habitation, roads, whater). A highest value indicates a better performance in the classification task.

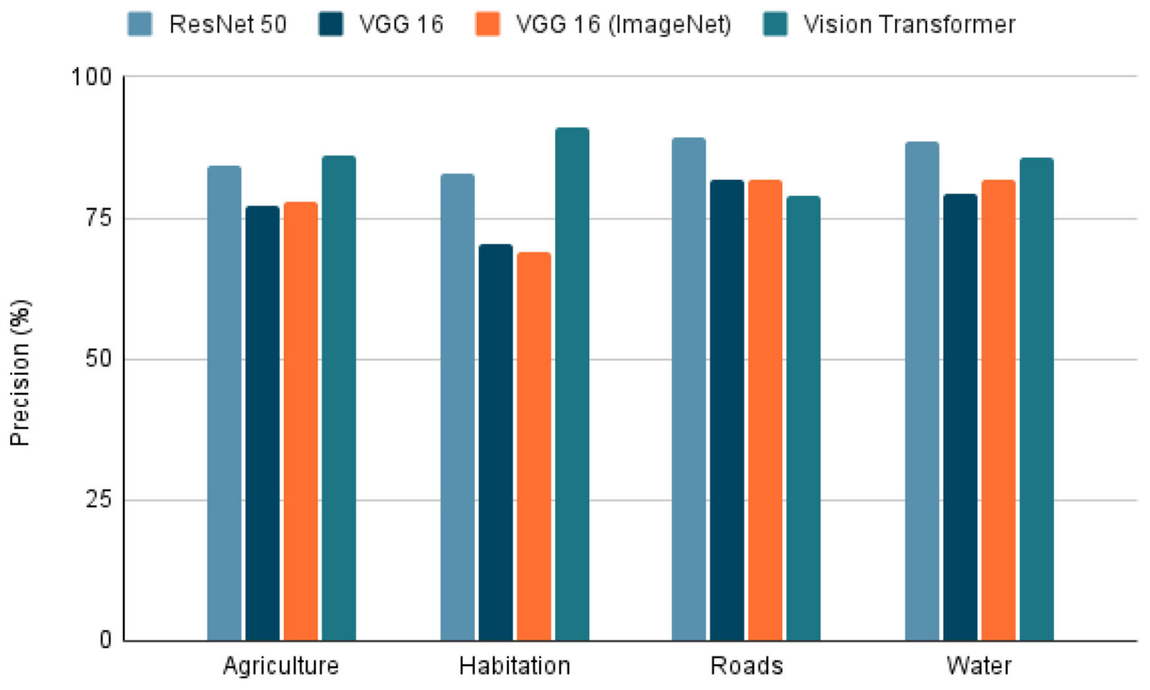


**FIGURE 6 |** Precision on tested models. Precision of the 4 models tested in the extraction of the 4 characteristics (agriculture, habitation, roads, whater). A highest value indicates a better performance in the classification task.

security in the Department of Cauca, Colombia. The dataset is structured according to the four-dimensional model of food security proposed by FAO (2). It allows the calculation of a multidimensional food security index. This research presents the calculation of the index for nine municipalities in the Department of Cauca, using the GFSI model established by the EIU (3).

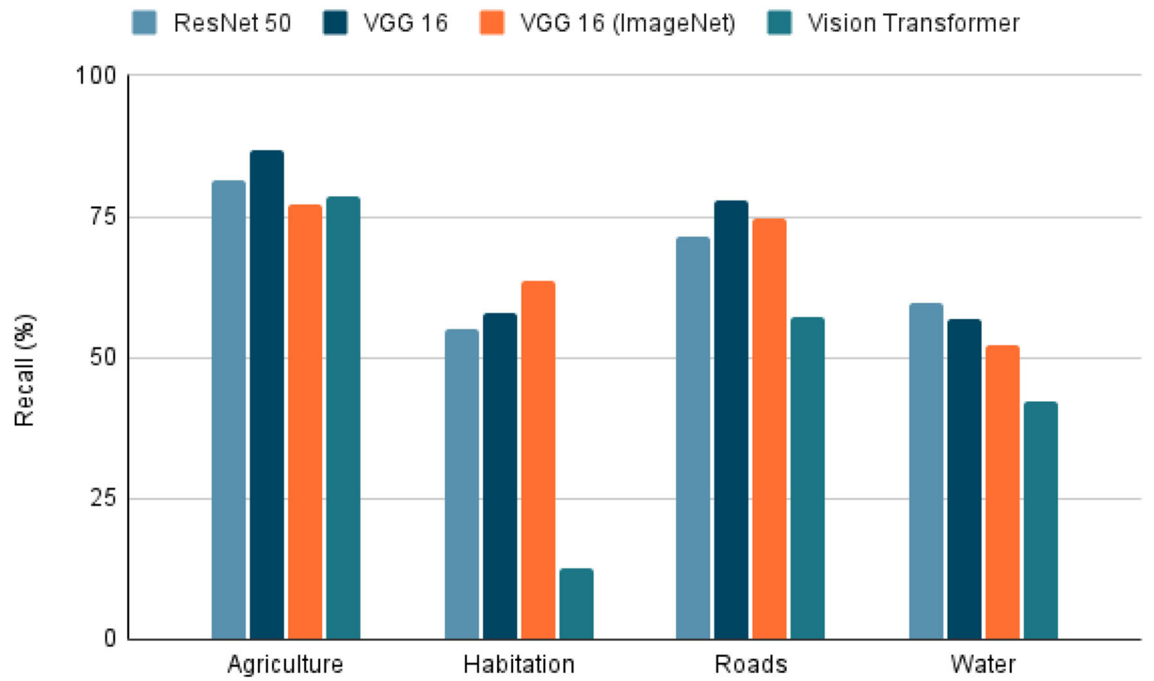


**FIGURE 7 |** Recall on tested models. Recall of the 4 models tested in the extraction of the 4 characteristics (agriculture, habitation, roads, whater). A highest value indicates a better performance in the classification task.

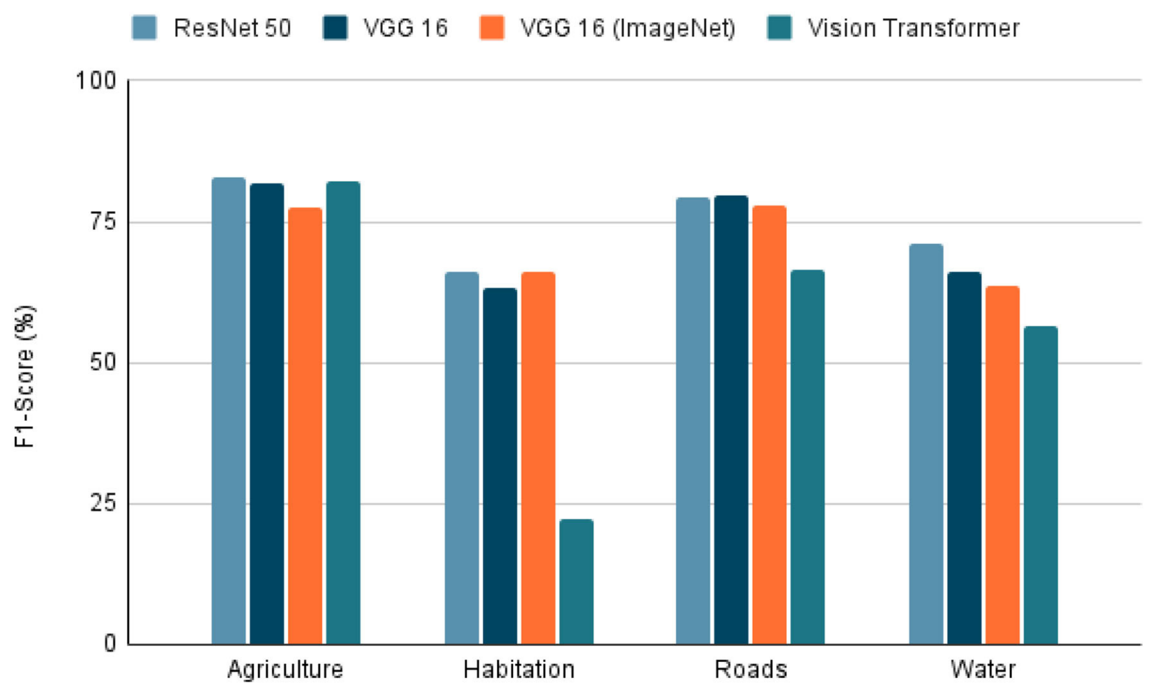


**FIGURE 8 |** F1-score on tested models. F1-score of the 4 models tested in the extraction of the 4 characteristics (agriculture, habitation, roads, whater). A highest value indicates a better performance in the classification task.

Likewise, as an alternative to the lack of some data, the option of using satellite images is presented. A method is proposed to extract characteristics from these and integrate them into a dataset that includes metadata as well. Only four characteristics were extracted (Agriculture, Habitation, Roads, and / or Water). However, this same method can be used to extract other types

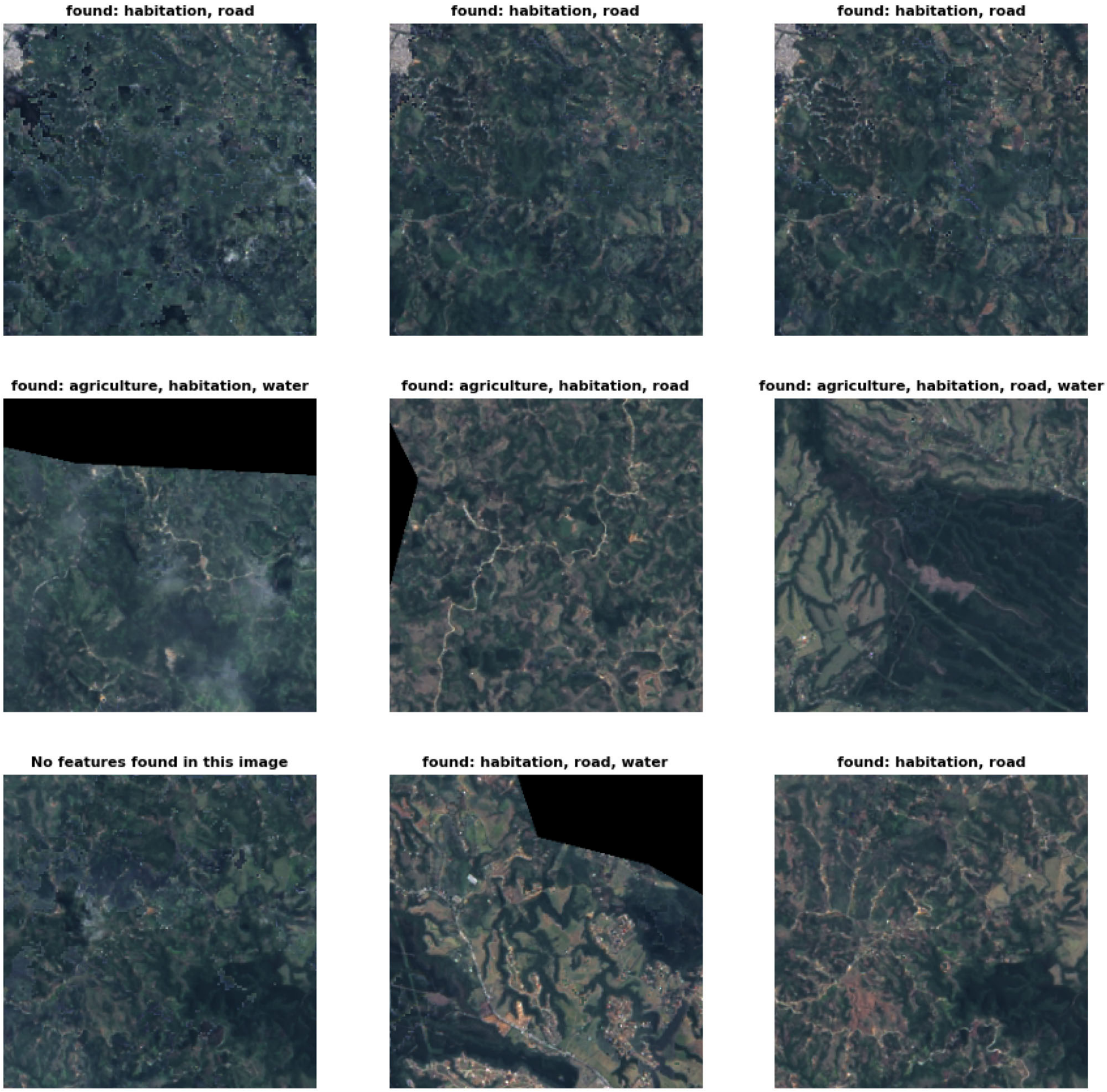


**FIGURE 9 |** Model predictions over Popayán patches. Examples of predictions generated over some patches of Popayán using the ResNet 50 model with highest f1-score in extraction of 4 features from satellite images. Above each patch are the characteristics that were found or if none of the 4 characteristics were found, it is indicated that no characteristics were found in that image.

of characteristics. The characteristics extracted can be used for the dimensions of physical availability (Agriculture, Water) and physical access to food (Roads, Habitation), and sustainability over time, due to the temporal resolution that is better than most surveys and censuses.

The data obtained came from various sources, which differed from the GFSI calculation during their collection. Therefore, an arduous pre-processing stage was required, and data transformations were consistent with those used by the GFSI. This includes data mining, imputation techniques, operations between dataset variables, changes in scales, units, etc. Also, in some cases, the variables for the GFSI calculation were not found within the created dataset, so the models were developed using also data manually imputed in the dataset obtained from reports produced by local, departmental or municipal governments.

The use of the CRISP-DM methodology for the construction of the dataset and calculation of the index gives us a framework for the project phases, from the understanding of the problem

**TABLE 2 |** GFSI score by municipality.

| Municipality | Unnormalized GFSI | Normalized GFSI |
|---|---|---|
| Popayán | 59.500 | 60.569 |
| Balboa | 54.900 | 57.019 |
| Bolívar | 55.583 | 56.787 |
| Paez (Belalcázar) | 56.112 | 57.741 |
| Patía (El Bordo) | 55.052 | 56.722 |
| San Sebastián | 55.857 | 56.821 |
| Santander de Quilichao | 57.812 | 59.856 |
| Timbío | 57.373 | 58.444 |
| Timbiquí | 53.219 | 53.637 |
| Cauca | 56.156 | 57.444 |

*The table shows the value of the GFSI score of 9 municipalities of Cauca using both normalized and unnormalized data. The index was calculated using the models with best performance in both the normalized and unnormalized datasets of the GFSI by country by the EIU used for training.*

to the validation of the results. Furthermore, the CRISP-DM methodology allows us to correct errors promptly and return to early phases of the project. In addition, the stages are appropriate to facing in data mining projects, as it presents a logical order and good practices.

About the model performance, as can be seen in the results regarding the model for calculating the GFSI trained with the dataset provided by the EIU of the GFSI at country scale (**Table 1**), simpler models have better performance than more complex models such as the multilayer perceptron or support vector machines (SVM). Also the Normalization decrease the MAE. These results may occur due to the nature of the data, because the GFSI is usually calculated by a group of experts assigning weights to each of the 59 variables, therefore the linear models such as linear regression should have a good performance. Another reason is the number of columns in relation to the instances (number of countries used in the GFSI). Having so few instances and many columns can cause overfitting in more complex models so they will not generalize well to new data, this problem can be solved by adding the GFSI by country data in other years to the actual dataset used to train the model.

The obtained results reveal a possible food security problem in the Department of Cauca. Historically, this department has been affected by violence, armed conflict, drug trafficking, and corruption, that determine the distribution of equitable resources (37). Therefore, the national government could take this index as a variable to consider in future decision-making in the country, seeking to improve the food security situation of the department and the municipalities that comprise it.

Although the GFSI calculation seeks to identify the food security situation for the Department of Cauca and some of its municipalities, this model can calculate the index in other departments, municipalities, or countries. Therefore, the same process could be followed to build the dataset and ML model with data from the place to be studied.

The other works carried out in the field of food security that makes use of ML focus only on one dimension defined by FAO (9–24), more specifically on the dimension of physical availability of food (primarily focused on predicting crop yield). So it would not give us a total vision of food security problems according to what the FAO proposes. Furthermore, in no case did the studies related to ML in food security calculate an index (9–24). In the studies not using ML techniques to calculate a food security index on a specific regional scale, indices other than GFSI are generally used, not covering the four dimensions FAO proposed (5, 6). This is probably due to the difficulties involved in collecting the data necessary for calculating the GFSI at such a specific level.

Although the methodology has been shown to comply with the work when creating the dataset and calculating the GFSI, some limitations arise in terms of the availability and nature of the data and must be taken into account. Many data are qualitative and must be calculated by experts in the area and other is difficult to find at the municipal and / or departmental level. Given this, the solution to cover these data may be to impute the data used in the GFSI at the country level. Consequently, the index obtained can be biased by external national factors that do not necessarily reflect the food security situation in the department or municipality.

Another limitation of the model is temporality. The data, in some cases, is released with a long delay. This produces results based on outdated data. Also, the frequency with which much of the data is collected is not desired since data is collected every 5 years or more in some cases (i.e., census). This is not enough to support the fourth dimension of FAO, which is stable over time.

In addition, the data are also incomplete on a spatial scale. Some surveys, such as ENSIN, are carried out only for 9 of the 42 municipalities of Cauca. Therefore, a GFSI calculation cannot be obtained with such precision for all the municipalities of Cauca due to the lack of data. This deficiency could be covered by satellite images that have spatial and temporal scale, and are openly accessible to any government, academic institution or civil society. However, the use of these images increases the computational resources required for data processing.

Nevertheless, one limitation for the satellite images used here is that the dataset used to train the model was not created with images from the same satellite, which could add noise to the predictions. This problem could be solved using satellite images from the same satellite and resolution as the Kaggle Planet dataset. Also, using another satellite with higher resolution, such as Sentinel-2, could improve the predictions, although it should be noted that this is only available from 2015. Another alternative could be to use a dataset made with Landsat-8 images to train the model, but it is essential to remember that this will require time and expertise.

## CONCLUSION

This paper proposes the methodology for creating an open dataset combining machine learning techniques to calculate a food security index at a departmental or municipal level, covering the four dimensions described by FAO. Data from different open sources such as censuses, surveys, health records, and satellite images are used to create the dataset, giving an economical alternative and easily accessible to all governments, academic institutions and civil society. It shows how the CRISP-DM methodology can be used to create an open public health data repository. Furthermore, this methodology could be generalized to other types of problems requiring the creation of a dataset. As well as the use of satellite images which presents an alternative for places where data collection is challenging, either for economic or access reasons.

In addition, a dataset with more variables than those used in the GFSI is released, opening research opportunities to areas like nutrition, agriculture, climate change, among others, where problems different to the GFSI and / or the individual dimensions of FAO could be analyzed.

As future work, the investigation of new models and methods for the feature extraction of satellite images is proposed. Also, the quality of the dataset and predictions could be improved by using more bands in satellite images than RGB or with newer sources of satellite images such as sentinel-2 or sentinel-3, as well as novel preprocessing methods due to increased spatial and temporal resolution in the images.

## DATA AVAILABILITY STATEMENT

Datasets has been submitted to Mendeley data. The datasets generated and used for this study can be found with the name "Multidimensional Dataset Of Food Security And Nutrition In Cauca" in Mendeley data at this link: https://data.mendeley.com/datasets/wsss65c885/draft?a=da27eda0-9529-4ad8-9922-6ffe8de79756.

## AUTHOR CONTRIBUTIONS

DL and RV-C contributed to conception and design of the study, as well as in the constant advice during the project. DL contributed in the data identification and acquisition. LP and DR contributed in the data analysis and preprocessing of metadata. DR contributed to the acquisition of satellite images, their preprocessing and the modeling stages using machine learning. JO-V in the review of work and theoretical contributions from the point of view of public health. All authors contributed to manuscript revision, read, and approved the submitted version.

## REFERENCES


1. United Nations. *Transforming Our World: The 2030 Agenda for Sustainable Development*. United Nations (2015). Available online at: https://sdgs.un.org/2030agenda (accessed September 5, 2021).
2. Food and Agriculture Organization. *An Introduction to the Basic Concepts of Food Security*. Food and Agriculture Organization (2008). Available online at: http://www.fao.org/3/al936e/al936e.pdf (accessed August 14, 2021).
3. The Economist Intelligence Unit (EIU). *Global Food Security Index 2020 Addressing Structural Inequalities to Build Strong and Sustainable Food Systems*. London, UK: The Economist Intelligence Unit Limited (2021).
4. Izraelov M, Silber J. An assessment of the global food security index. *Food Sec.* (2019) 11:1135–52. doi: 10.1007/s12571-019-00941-y
5. von Grebmer K, Bernstein J, Nabarro D, Prasai N, Amin S, Yohannes Y, et al. The concept of the global hunger index. In: *2016 Global Hunger Index: Getting to Zero Hunger, Chapter 1*. Bonn Washington, DC and Dublin: Welthungerhilfe, International Food Policy Research Institute, and Concern Worldwide (2016). p. 6–9. doi: 10.2499/9780896292260_01
6. Ballard T, Kepple A, Cafiero C. *The Food Insecurity Experience Scale: Development of a Global Standard for Monitoring Hunger Worldwide*. Rome: FAO. Available online at: http://www.fao.org/fileadmin/templates/ess/voh/FIES_Technical_Paper_v1.1.pdf (accessed August 14, 2021).
7. Instituto Colombiano de Bienestar Familiar. *Presentación, ENSIN: Encuesta Nacional de Situación Nutricional*. Instituto Colombiano de Bienestar Familiar (2017). Available online at: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/PSP/presentacion-lanzamiento-ensin-2015.pdf (accessed October 6, 2021).
8. Yeh C, Perez A, Driscoll A, Azzari G, Tang Z, Lobell D, et al. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nat Commun.* (2020) 11:2583. doi: 10.1038/s41467-020-16185-w
9. Sharifi A. Yield prediction with machine learning algorithms and satellite images. *J Sci Food Agric.* (2021) 101:891–896. doi: 10.1002/jsfa.10696
10. Cao J, Zhang Z, Tao F, Zhang L, Luo Y, Zhang J, et al. Integrating multi-source data for rice yield prediction across china using machine learning and deep learning approaches. *Agric For Meteorol.* (2021) 297:108275. doi: 10.1016/j.agrformet.2020.108275
11. Gómez D, Salvador P, Sanz J, Casanova JL. Modelling wheat yield with antecedent information, satellite and climate data using machine learning methods in Mexico. *Agric For Meteorol.* (2021) 300:108317. doi: 10.1016/j.agrformet.2020.108317
12. Zhang C, Zhang H, Zhang L. Spatial domain bridge transfer: an automated paddy rice mapping method with no training data required and decreased image inputs for the large cloudy area. *Comput Electro Agric.* (2021) 181:105978. doi: 10.1016/j.compag.2020.105978
13. Ramadhani F, Pullanagari R, Kereszturi G, Procter J. Automatic mapping of rice growth stages using the integration of sentinel-2, mod13q1, and sentinel-1. *Remote Sens.* (2020) 12:3613. doi: 10.3390/rs12213613
14. Salvador P, Gómez D, Sanz J, Casanova JL. Estimation of potato yield using satellite data at a municipal level: a machine learning approach. *ISPRS Int J Geo Inform.* (2020) 9:343. doi: 10.3390/ijgi9060343
15. Maimaitijiang M, Sagan V, Sidike P, Hartling S, Esposito F, Fritschi FB. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens Environ.* (2020) 237:111599. doi: 10.1016/j.rse.2019.111599
16. Schwalbert RA, Amado T, Corassa G, Pott LP, Prasad PVV, Ciampitti IA. Satellite-based soybean yield forecast: integrating machine learning and weather data for improving crop yield prediction in southern Brazil. *Agric For Meteorol.* (2020) 284:107886. doi: 10.1016/j.agrformet.2019.107886
17. Wolanin A, Mateo-Garciá G, Camps-Valls G, Gómez-Chova L, Meroni M, Duveiller G, et al. Estimating and understanding crop yields with explainable deep learning in the Indian wheat belt. *Environ Res Lett.* (2020) 15:024019. doi: 10.1088/1748-9326/ab68ac
18. Maimaitijiang M, Sagan V, Sidike P, Daloye AM, Erkbol H, Fritschi FB. Crop monitoring using satellite/UAV data fusion and machine learning. *Remote Sens.* (2020) 12:1357. doi: 10.3390/rs12091357
19. Wang Y, Zhang Z, Feng L, Du Q, Runge T. Combining multi-source data and machine learning approaches to predict winter wheat yield in the conterminous United States. *Remote Sens.* (2020) 12:1232. doi: 10.3390/rs12081232
20. Yang Q, Shi L, Han J, Zha Y, Zhu P. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Res.* (2019) 235:142–53. doi: 10.1016/j.fcr.2019.02.022
21. Ma JW, Nguyen CH, Lee K, Heo J. Regional-scale rice-yield estimation using stacked auto-encoder with climatic and MODIS data: a case study of South Korea. *Int J Remote Sens.* (2019) 40:51–71. doi: 10.1080/01431161.2018.1488291

22. Zhao H, Chen Z, Jiang H, Jing W, Sun L, Feng M. Evaluation of three deep learning models for early crop classification using sentinel-1A imagery time series-a case study in Zhanjiang, China. *Remote Sens.* (2019) 11:2673. doi: 10.3390/rs11222673
23. Van Tricht K, Gobin A, Gilliams S, Piccard I. Synergistic use of radar sentinel-1 and optical sentinel-2 imagery for crop mapping: a case study for Belgium. *Remote Sens.* (2018) 10:1642. doi: 10.3390/rs10101642
24. Wang AX, Tran C, Desai N, Lobell D, Ermon S. Deep transfer learning for crop yield prediction with remote sensing data. In: *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies.* Menlo Park and San Jose, CA (2018). p. 1–5. doi: 10.1145/3209811.3212707
25. Comité Intersectorial De Seguridad Alimentaria y Nutricional Del Cauca. *Plan de Seguridad Alimentaria y Nutricional Para el Departamento de Cauca.* (2009). Available online at: http://anterior.cauca.gov.co/sites/default/files/informacion/cartilla_20cauca_20sin_20hambre_0.pdf (accessed September 5, 2021).
26. Peng W, Dernini S, Berry EM. Coping With Food Insecurity Using the Sociotype Ecological Framework. *Front. Nutr.* (2018) 5:107. doi: 10.3389/fnut.2018.00107
27. Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, et al. *CRISP-DM 1.0: Step-by-step Data Mining Guide*. The CRISP-DM consortium (2000).
28. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens Environ.* (2017) 202:18–27. doi: 10.1016/j.rse.2017.06.031
29. Instituto Colombiano de Bienestar Familiar. *ENSIN: Encuesta Nacional de Situación Nutricional.* (2015). Available online at: https://www.icbf.gov.co/bienestar/nutricion/encuesta-nacional-situacion-nutricional (accessed September 5, 2021).
30. Departamento Administrativo Nacional de Estadística. *COLOMBIA–Tercer Censo Nacional Agropecuario—2014-3er CAN*. Departamento Administrativo Nacional de Estadística (2016). Available online at: http://microdatos.dane.gov.co/index.php/catalog/513/get_microdata (accessed September 5, 2021).
31. Sistema Nacional de Vigilancia en Salud Pública. *Estadísticas de Vigilancia Rutinaria*. Sistema Nacional de Vigilancia en Salud Pública (2015). Available online at: http://portalsivigila.ins.gov.co/Paginas/Vigilancia-Rutinaria.aspx (accessed September 5, 2021).
32. Fick SE, Hijmans RJ. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. *Int J Climatol.* (2017) 37:4302–15. doi: 10.1002/joc.5086
33. Kaggle. *Planet: Understanding the Amazon from Space*. Kaggle (2017). Available online at: https://www.kaggle.com/c/planet-understanding-the-amazon-from-space/data (accessed September 25, 2021).
34. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV (2016). p. 770–8. doi: 10.1109/CVPR.2016.90
35. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: *Proceedings of the 3rd International Conference on Learning Representations*. San Diego, CA: ICLR (2015).
36. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An image is worth 16x16 words: transformers for image recognition at scale. *arXiv*. (2020) arXiv:2010.11929.
37. Ríos-Sierra J, Bula-Galiano P, Morales JA. Departamentos de frontera y violencia periférica en Colombia. *Rev Crim.* (2019) 61:113–32. Available online at: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1794-31082019000200113&lng=en&tlng=es

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Nutrition

BRIEF RESEARCH REPORT
published: 03 February 2022
doi: 10.3389/fnut.2022.794468

# Verifying the Use of Food Labeling Data for Compiling Branded Food Databases: A Case Study of Sugars in Beverages

*Edvina Hafner[1,2], Živa Lavriša[1], Maša Hribar[1], Sanja Krušič[1], Anita Kušar[1], Katja Žmitek[1,3], Mihaela Skrt[2], Nataša Poklar Ulrih[2] and Igor Pravst[1,2,3]**

*[1] Nutrition and Public Health Research Group, Nutrition Institute, Ljubljana, Slovenia, [2] Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia, [3] VIST–Faculty of Applied Sciences, Ljubljana, Slovenia*

OPEN ACCESS

**Edited by:**
*Alessandra Durazzo, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Tony Kuo, UCLA Fielding School of Public Health, United States*
*Avi Magid, Peres Academic Center, Israel*

***Correspondence:**
*Igor Pravst*
*igor.pravst@nutris.org*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** *13 October 2021*
**Accepted:** *10 January 2022*
**Published:** *03 February 2022*

**Citation:**
*Hafner E, Lavriša Ž, Hribar M, Krušič S, Kušar A, Žmitek K, Skrt M, Poklar Ulrih N and Pravst I (2022) Verifying the Use of Food Labeling Data for Compiling Branded Food Databases: A Case Study of Sugars in Beverages. Front. Nutr. 9:794468. doi: 10.3389/fnut.2022.794468*

Branded food composition databases are an important tool for research, education, healthcare, and policy making, amongst others. Such databases are typically compiled using food labeling data without chemical analyses of specific products. This study aimed to verify whether the labeled sugar content in sugar-sweetened beverages (SSBs) corresponds to the actual sugar content in these products, thus enabling food monitoring studies to be conducted. A secondary objective was to determine the specific types of sugars in these SSBs. A case study was conducted using market share-driven sampling of these beverages from the Slovenian food supply. On the basis of nationwide yearly sales data, 51 best-selling products were sampled in 2020 and analyzed using high-performance liquid chromatography. This sales-driven approach to sampling has been shown to be very useful for conducting food monitoring studies. With the careful selection of a small proportion of available products, we finished with a manageable sample size, reflecting the composition of a majority (69%) of the national market share volume. The analyzed total sugar content was compared with labeled data, within the context of the European Union's regulatory labeling tolerances. In all samples, the sugar content was within the tolerance levels. The most common ($N = 41$) deviation was within ±10% of the labeled sugar content. In the subcategories, the differences between the analyzed and labeled median sugar contents were not statistically significant. Sucrose was most commonly ($N = 36$; 71%) used for sweetening, suggesting that the proportion of fructose in most SSBs was around 50%. A higher fructose content was only observed in beverages with fructose–glucose syrup or a higher content of fruit juice. The study results show that the labeled sugar content information in SSBs is reliable and can be used to compile branded food databases and monitor the nutritional quality of foods in the food supply.

**Keywords: beverages, reformulation, added sugar, food composition, HPLC, database**

## INTRODUCTION

Food composition databases are an important tool for research, education, healthcare, and policy making, amongst others (1). Epidemiological nutritional studies mostly convert food consumption data into nutrient intake data based on food composition databases (2, 3). While databases for generic foods are typically built based on complex laboratory analyses, such datasets lack details of the nutritional composition of specific processed (branded) foods (4). Because branded foods are gaining importance in people's diets, branded food composition datasets have become a common data source used in nutrition research (5, 6). The standard methodology for compiling branded food databases is cross-sectional data collection in food stores, where information is extracted from food labels (7–9). Trustworthy information on branded foods can improve the accuracy of nutrient intake data, support the monitoring of food reformulation progress, and empower consumers to make informed food choices.

Data in branded food databases, however, are often limited in terms of availability and accuracy due to the information provided by manufacturers on the food labels. In the European Union (EU), food labeling is regulated by Regulation (EU) 1169/2011, i.e., certain food information must be provided to the consumer (10). This mandatory nutritional declaration typically consists of energy value and contents of fats, saturated fats, carbohydrates, sugar, protein, and salt; the provision of other information such as fiber, vitamins, and minerals is optional. According to the regulation, food manufacturers are responsible for aligning the information provided in the nutritional declaration with the actual composition of the food. The regulation also allows manufacturers to either provide the results of laboratory analyses or to estimate the nutrient contents based on the known composition of the ingredients used in the production process (11). While laboratory analyses using valid procedures are the optimal choice for this purpose, these often involve significant costs and are therefore not commonly performed. Because the nutritional composition of foods is affected by various factors, including production and storage, often products do not contain the exact nutrient values stated on their labels. For this reason, the European Commission (EC) issued guidance in order to establish tolerances for specific nutrients in the nutritional declaration so that labeled and actual contents do not substantially differ (12).

In EU countries, the regulatory control of food labeling depends on local laws and each country's law enforcement bodies; i.e., each country implements their own enforcement rules and fines as they relate to Regulation (EU) 1169/2011 (10). To support harmonized food control, the EC issued guidance on how the responsible authorities should address compliance with the EU legislation (12). Food control plans of these authorities include the verification of nutrient values declared on a label. While all countries are encouraged to share these verification reports with the EC, to the best of our knowledge, such reports are not readily available to the public. The authority responsible for this food control function in Slovenia is the Administration for Food Safety, Veterinary Sector and Plant Protection and Health Inspectorate of the Republic of Slovenia. Penalties for violating the Regulation (EU) No. 1169/2011 vary from country to country and depend on the type of violation. For example, in Slovenia, the penalty for providing inadequate nutritional value information ranges from 500 € to 15,000 € (e.g., for foods not labeled with nutritional information or that provide inaccurate health claims) (13), while in Italy, penalties for violating Regulation (EU) No. 1169/2011 are in a range from 500 € to 40,000 € (14).

Although food control in Europe is carefully regulated, the process is mostly focused on areas with pre-identified food safety risks, making the information provided in the nutritional declaration a low priority for the authority responsible (15). As food manufacturers are responsible for the accuracy of the information on food labels, questions frequently arise as to whether this information is sufficiently accurate and suitable for compiling branded food databases and for use in nutrition research. Very few studies have compared the actual (analyzed) nutritional composition of foods with the nutrient information provided on food labels, and the two sets of information are sometimes contradictory. For example, in 2011, a study from the United States (US) showed notable differences between labeled and analyzed sugar contents in soft drinks (16). Similarly, in Malaysia, they found that only 66% of the products complied with their legislative limits as required for the nutritional declaration (17). In Europe, studies have shown better compliance with EU tolerance ranges, as stated in Regulation (EU) 1169/2011 (10), but the results still varied. In Portugal, the highest compliance was found for fats (88%), while it was lower for salt (74%) and saturated fats (73%) (18). A study in Ireland found that labeled sugar contents in yogurts was among the least reliable information on the label (19). Much better compliance was demonstrated in a recent Spanish study (20), where 98.4% of the analyzed processed (branded) foods met the EU tolerance range for sugars. A key challenge in all of these studies is sample selection, since there are thousands of foods on the market and sampling/analysis capacities are typically much more limited. While a common randomized sample selection is a common approach, sampling driven by market share is a more practical and relevant approach, i.e., it better reflects the overall food supply.

Sugar-sweetened beverages (SSBs) are one of the key food categories addressed in public health policies (21). SSBs have been shown to affect oral health, weight gain, and increase the risk of chronic conditions such as obesity, diabetes, fatty liver disease, and cardiovascular diseases (21–25). The main health risk of SSBs is that they contain high amounts of free sugar and are consumed in large quantities by vulnerable population groups. Although the World Health Organization (WHO) recommends that dietary intake of free sugars should be $<$10% of a person's daily energy intake (26, 27), a large proportion of Europeans easily exceed this intake limit. For example, SSBs have been identified as a major contributor to free sugar intake among several population groups, particularly children/adolescents (2, 28, 29). The results of several epidemiological studies suggest that sugars in beverages can lead to a greater risk of developing metabolic syndrome than sugar in other foods. A plausible reason for this observation is that fructose and its unique associated metabolism have specific negative health effects (30, 31). SSBs often have a higher fructose content, which is better absorbed by the body and results in higher concentrations stored in the liver (32). Higher amounts of fructose are often found in SSBs because manufacturers increasingly use concentrated fruit juice and fructose–glucose syrup (FGS). The use of FGS is growing in popularity (33) due to its cheap production, long shelf life, and sweeter taste (34). Since SSBs are one of the primary contributors to increased fructose intake (35), monitoring the amount of this sugar in SSB products is crucial for investigating dietary fructose intake and its related health outcomes.

Given the challenges described above, our study sought to verify whether the labeled sugar content in SSBs corresponds to the actual sugar content contained in these products, and if so, how the information can be utilized to enable the reliable compilation of branded food databases for use in national food monitoring studies. A secondary objective of our study was to determine the specific types of sugars found in SSBs in the Slovenian food supply, particularly the proportion of fructose.

# MATERIALS AND METHODS

## Sample Selection

Our sample included the most consumed SSBs in Slovenia, which were selected using a market share approach. The selection of these beverages was thus based on the yearly nationwide (Slovenia) sales data provided by the NielsenIQ agency. The sales data were provided in Microsoft™ Excel worksheets in the universal form, which included barcode number, product name, and quantity of products sold from the year 2019. Information was available for the following selected NielsenIQ food categories: energy drinks, fruit juices, iced tea, mineral water, syrup, and soft drinks. In the next step, products contributing to 95% of the nationwide volume sales (N = 380) were re-categorized to tease out the different types of SSBs, i.e., sugar-sweetened colas, iced-tea drinks, sugar-sweetened energy drinks, flavored waters, and others, such as fruit and other carbonated drinks. In each of these subcategories, we summed the quantities of the same type of products sold, differing only in the package quantity/form. For example, sugar-sweetened Coca-Cola, which was available in plastic bottles (0.5; 1; 1.5; 1.75, 2 L), cans (0.25; 0.33 L), and glass bottles (0.25 L), was assigned to the same SSB type. From each of the subcategories, we then selected the top six SSBs sold. This sequential approach helped assure that the most representative products were sampled from each subcategory. Additional samples were also selected in descending order regardless of subcategory, to reach the total laboratory analysis capacity; this was capped at 51 analyses. The final sample consisted of 7 sugar-sweetened colas, 7 iced-tea drinks, 7 sugar-sweetened energy drinks, 10 flavored waters, and 20 other SSBs. Altogether the sample represented 69% of the national market share for the selected categories. The chosen SSBs were purchased from different retailers located in Ljubljana, Slovenia, in 2020 for laboratory analyses.

## Labeled Composition and Sugar Content

To provide insights into the feasibility of the market share-driven sampling approach, we collected data on the labeled sugar content for the SSBs in the original NielsenIQ dataset. Using the 2019 edition of the Slovenian branded foods database (4), which was compiled using labeled nutrition declarations, we were able to ascertain the labeled sugar contents for 309 SSBs linked to available yearly sales data.

For the selected products, the data were extracted directly from the labels of the purchased beverages. The product labels were photographed and used to extract the barcode numbers, product names, and nutritional declarations, including total sugar, the ingredients list, package quantity, expiration date, and manufacturer. Sources of sugars were identified using the ingredient lists.

## Laboratory Determination of Sugars

Laboratory analyses were performed in the Biotechnical faculty (University of Ljubljana, Slovenia). We analyzed the presence of free fructose, free glucose, and sucrose, using high-performance liquid chromatography (HPLC). Analyses of each sample were replicated, and average values were used for further calculations. Glucose (anhydrous for biochemistry), fructose (for biochemistry, purity HPLC ≥99.0 %), and sulphuric acid were obtained from Merck, Darmstadt, Germany; analytical grade sucrose was obtained from Kemika, Zagreb, Croatia. A mixed standard solution of glucose, fructose, and sucrose was prepared with double-deionised water (Milli-Q, Millipore Corp., Milford, MA, USA) of 18.2 MΩ/cm resistivity in the range from 1 to 20 g/l. Peak identification was based on HPLC retention times as compared with the standards. Peak integration was performed with ChemStation software (revision B.04.03-SP2). Quantitation was based on the external standard method using seven-point calibration curves fitted by linear regression analysis with the Data Analysis Tools in Excel. Samples were centrifuged at 3,000 × g for 10 min, and the supernatant was filtered through a 0.45 μm Chromafil® RC membrane (Macherey-Nagel, Düren, Germany). The filtrate was appropriately diluted before direct injection into the HPLC.

The HPLC system used in this study (Agilent 1260 Infinity, Agilent Technologies, Germany) was equipped with a G1322A degasser, a G1312B binary pump, a G1367E Hip ALS autosampler with G1330B FC/ALS autosampler thermostat, a G1316A thermostated column compartment, and a G1362A refraction index detector (RID). A total of 20 μl of the sample was injected on a column Aminex HPX-87H (BioRad, Richmond, CA) with a length of 300 mm x 7.8 mm i.d. and a particle size of 9 μm. The analysis was performed at 35 °C with a flow rate of 0.6 ml/min using isocratic elution with 5 mM $H_2SO_4$ as a mobile phase.

## Data Processing and Analysis

Data were processed and analyzed using Microsoft Excel 2019 (Microsoft, Redmond, WA, USA) and R 2020 (R Core Team, Vienna, Austria). An assessment of the market share sampling approach was completed using the Wilcoxon rank-sum test. The average labeled total sugar content in our sample of beverages with the highest sales ($N = 51$) was compared with all beverages with available sales and composition data ($N = 309$). The level of significance was set at $p < 0.05$.

To assess the labeled vs. the analyzed total sugars, we first determined legislative boundaries for each product based on the

EC regulatory guidance on tolerances (12). For drinks with a sugar content below 10 g/100 mL, the acceptable deviation was ±2 g/100 ml, and for drinks with a sugar content above 10 g/100 mL, the acceptable deviation was ±20 %. We calculated the discrepancy percentage between the labeled and the analyzed total sugar contents and compared it with the acceptable deviation. Descriptive statistics were used to describe medians, 25th/75th percentiles of labeled/analyzed total sugar content, and the difference between them. To verify whether labeled sugar content data can be used to research and monitor the food supply, we compared the labeled and analyzed sugar contents as two independent samples. As the data were not normally distributed, differences were investigated with Mood's median test with the significance level set at $p < 0.05$.

We further applied sales-weighing to examine the differences between our sample and all beverages and between labeled and analyzed total sugar contents. Sales-weighted average total sugar content was calculated based on the quantity of products sold per year (L) and their total sugar content.

To investigate the different types of sugars in SSBs, we estimated the percentage of fructose in the products. This was calculated based on the products' analyzed free fructose and fructose from sucrose (sucrose content divided by two).

## RESULTS

Although our study only sampled a small proportion of available SSBs, we were able to compare labeled total sugar content in the original sample ($N = 309$) with that of the laboratory analysis sample ($N = 51$), employing a sampling strategy by market share. In doing so, we found no statistically significant differences ($p < 0.05$) in the average labeled sugar content between the two samples— both for SSBs in general (8.2 vs. 7.9 g/100 ml, respectively) and across selected beverage categories (**Supplementary Table 1**)—suggesting that the composition of the study sample was consistent with a majority of the national market share volume. Additionally, the sales-weighted average labeled sugar contents were the same for both samples (8.8 g/100 ml; **Supplementary Table 1**). We observed that, in general, the sales-weighted sugar content was higher than the non-sales-weighted average.

### Compliance of Labeled Sugar Content With Laboratory Results

Compliance of the labeled sugar content with the analytically determined sugar content (as required by the EU regulatory tolerance range) for specific samples is shown in **Figure 1**. All 51 SSB samples were within the regulatory tolerance range of the labeled sugar content. The differences between the labeled and analytically determined sugar contents deviated almost equally each way, i.e., positively and negatively, with 24 drinks containing less sugar than labeled and 26 drinks containing more sugar than labeled. One sample contained the same amount of sugar as labeled. The observed difference range was from −2.0 g to +1.9 g per 100 mL (from −18% to 35%). Most samples ($N = 41$, 80%) had a deviation of ±10%. Only two sugar-sweetened colas and one energy drink had a deviation close to the regulatory limits. Interestingly, flavored waters always contained more sugar than labeled and also had higher deviations (2–35%). As a result of the lower sugar content in this subcategory, the regulatory relative tolerance ranges were also wider than for other soft drinks. A deviation in a positive direction was also observed in iced-tea drinks, where six out of seven beverages had higher analytically determined sugar than the labeled values. Moreover, sugar-sweetened colas (6 out of 7) and energy drinks (5 out of 7) contained less sugar than labeled.

**Table 1** shows medians with 25th−75th percentile values for the analytically determined and labeled total sugar contents from the whole sample and the subcategories. The whole sample median for labeled and analyzed sugar contents was very similar (8.9 vs. 8.5 g/100 mL); the difference (−4%) was not statistically significant. Additionally, none of the five subcategories showed significant differences between labeled and analyzed total sugar contents. In absolute terms (g/100 mL), the largest differences in medians were observed in sugar-sweetened colas and energy drinks (−0.7 g). This coincided with the regulatory tolerance ranges for these SSBs. Drinks with a higher sugar content (>10 g/100 mL; sugar-sweetened colas and energy drinks) have a tolerance range defined by percentage (±20%), which allows for greater deviations in g/100 mL, while drinks with less sugar (<10 g/100 mL; flavored water) have a tolerance range defined by the amount (±2 g/100 mL), which allows for less deviations in g/100 mL.

### Types of Sugars in SSBs

Nutrition declarations on food labels only contain information concerning the total sugar content and not information about specific types of sugars. In **Table 2**, we present the results of the laboratory analysis of the contents of different types of sugars. In general, sampled beverages contained a similar amount of sucrose (median 22.2 g/L), free glucose (27.6 g/L), and free fructose (26 g/L), with notable differences between different subcategories. Sugar-sweetened colas contained the highest content of free glucose (42.1 g/L) and fructose (46.9 g/L), and lower amounts of sucrose (2.9 g/L). In contrast, flavored waters contained more sucrose (22.7 g/L), and less free glucose (9 g/L) and fructose (9.5 g/L). The overall median for the proportion of fructose (note: monosaccharides and fructose are present in sucrose) was 50% (range: 40–66%). An above-average proportion of fructose was found in sugar-sweetened colas (51%; range: 50–59%). Interestingly, sugar-sweetened energy drinks, which had the highest total sugar content, had the lowest proportion of fructose (42%; range: 40–50%).

To generate further context and insights, these results were assessed by considering the sources of sugar, as provided in the ingredient lists (labeling information). Sucrose was found to be a key ingredient in the sweetening of selected SSBs ($N = 36$; 71%). We found four drinks with labeled fructose-glucose syrup (FGS) and seven with glucose-fructose syrup (GFS). FGS appeared in one sugar-sweetened cola and in other (carbonated) soft drinks ($N = 3$). Among the products with a high fructose content/proportion (defined by a minimum of 3 g total fructose per 100 mL and a proportion of fructose above 52%), there

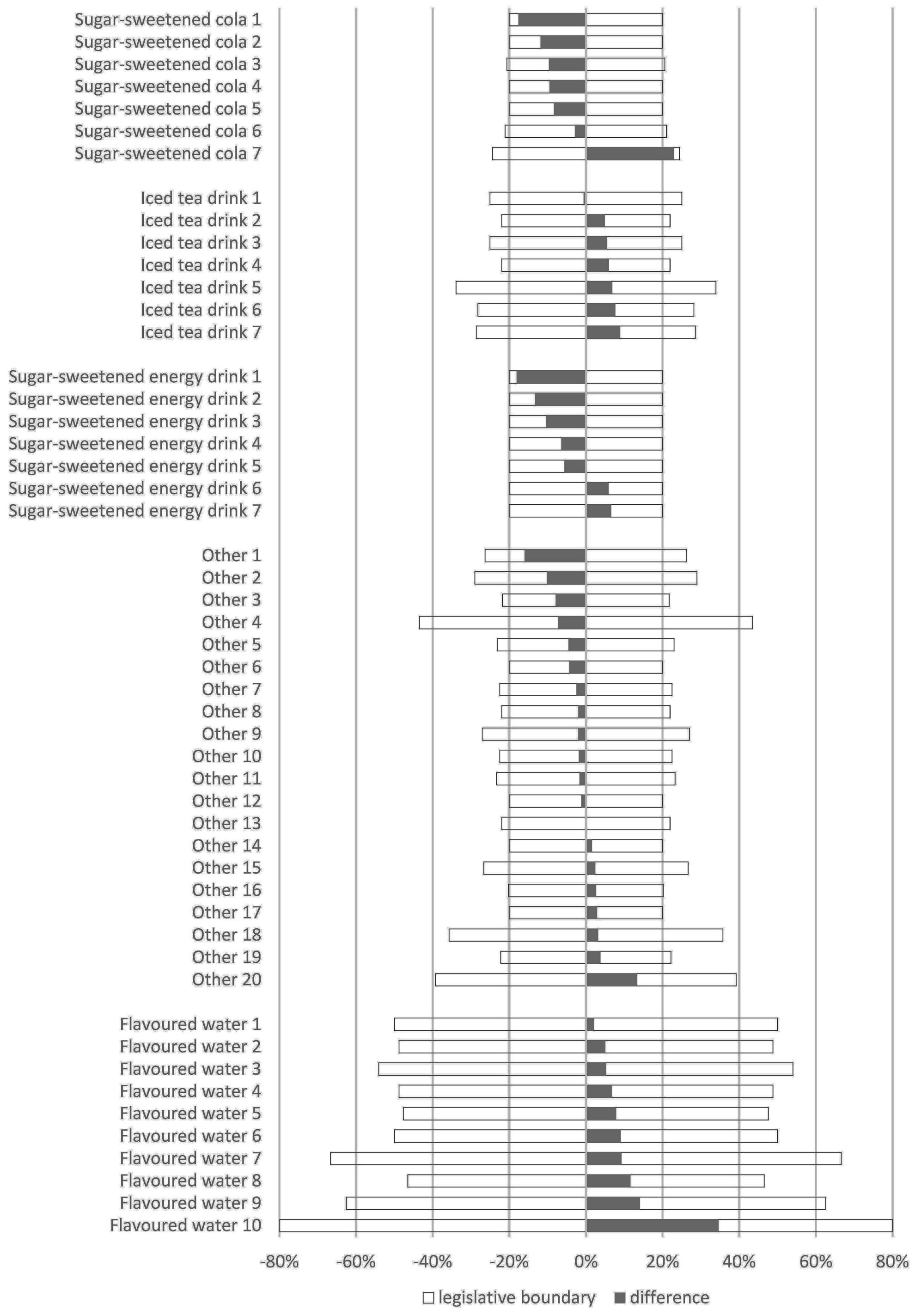


**FIGURE 1 |** Difference (%) between labeled and analyzed total sugar content in sampled sugar-sweetened beverages and calculated EU tolerance ranges (%).

were two products with fruit juice and four products with FGS. Altogether, the proportion of fructose in FGS-containing beverages was between 59 and 61%. GFS was found in iced-tea drinks ($N = 2$) and other beverages ($N = 5$) containing fruit juice. We observed that the use of FGS or GFS is commonly associated with a specific manufacturer. When FGS or GFS appeared in one product, it was commonly used in other products from the same manufacturer. Interestingly, only one sugar-sweetened cola labeled the use of FGS, while others only listed the use of sucrose. Nevertheless, laboratory analysis showed high levels of free glucose and fructose in the cola subcategory. The use of FGS was not labeled in flavored waters and sugar-sweetened energy

**TABLE 1 |** Labeled and analyzed values of total sugar content in different (sub)categories of sugar-sweetened beverages.

| | *N* | Labeled total sugar (g/100 mL) | Analyzed total sugar (g/100 mL) | Difference | | |
|---|---|---|---|---|---|---|
| | | Median (P25–P75) | Median (P25–P75) | g/100 mL | % | *p* |
| Total | 51 | 8.9 (5.6–10.1) | 8.5 (5.8–9.9) | −0.4 | −4% | ns |
| Flavored waters | 10 | 4.0 (3.2–4.1) | 4.2 (3.6–4.4) | 0.2 | 5% | ns |
| Sugar-sweetened cola | 7 | 10.3 (9.5–11.0) | 9.6 (8.8–10.1) | −0.7 | −7% | ns |
| Iced tea drinks | 7 | 8.0 (7.0–9.1) | 8.0 (7.6–9.6) | 0 | 0% | ns |
| Sugar-sweetened energy drinks | 7 | 11.0 (10.3–11.0) | 10.3 (9.5–10.6) | −0.7 | −6% | ns |
| Other sugar-sweetened beverages | 20 | 8.9 (7.4–9.7) | 8.6 (6.6–9.7) | −0.3 | −3% | ns |

*N, number of products; P25, 25th percentile; P75, 75th percentile; ns, not significant.*

**TABLE 2 |** Analyzed values of sucrose, glucose, and fructose content in different (sub)categories of sugar-sweetened beverages.

| | Sucrose (g/L) Median (P25–P75) | Glucose (g/L)[a] Median (P25–P75) | Fructose (g/L)[a] Median (P25–P75) | % Fructose[b] (min–max) |
|---|---|---|---|---|
| Total | 22.2 (10.2–38.3) | 27.6 (14.1–37) | 26 (14.5–36.5) | 50 (40–66) |
| Other | 16.9 (10.1–33.0) | 28.2 (23.0–35.8) | 33.1 (23.1–36.8) | 50 (49–66) |
| Flavored waters | 22.7 (17.9–29.4) | 9.0 (6.6–10) | 9.5 (7.1–10.1) | 50 (49–61) |
| Sugar-sweetened cola | 2.9 (0–3.8) | 42.1 (41.3–49.7) | 46.9 (42.6–51.1) | 51 (50–59) |
| Iced tea drinks | 40.8 (26.3–51.7) | 27.6 (13.1–28.4) | 24.9 (11.5–27) | 49 (48–50) |
| Sugar-sweetened energy drinks | 43.4 (23.3–52.6) | 37.7 (26.4–43.4) | 24.9 (20.2–28.4) | 42 (40–50) |

*P25, 25th percentile; P75, 75th percentile.*
[a] *Measured in monosaccharide form.*
[b] *Percentage (%) of fructose was calculated with consideration of free fructose and fructose present in sucrose.*

drinks. In addition to sucrose, most energy drinks (5 out of 7) also had glucose, or glucose syrup (GS) added, which tilted the fructose-to-glucose ratio in favor of glucose.

## DISCUSSION

Studies comparing the labeled quantities of specific nutrients in food products with results from laboratory analyses are scarce. To the best of our knowledge, this is the first study to examine the sugar contents of beverages in the Slovenian food supply in this context. Additionally, the accuracy of nutrition labels is rarely the subject of concern for the authorities tasked with regulating food safety and quality. Food inspections, for example, are typically focused on food safety issues, microbiological and chemical safety, the presence of additives and undeclared allergens, etc. (15), rather than on the nutritional quality of food. As such, the responsibility for ensuring that the nutritional information on food labels is accurate lies with the manufacturers, who must follow regulatory tolerance ranges (12). Our study showed that the analyzed SSBs were generally within the EU tolerance ranges for sugars. This finding is similar to the results from a recent Spanish study (20) in which only 5% of SSBs exceeded the regulatory tolerance range, i.e., in all cases, the analytically determined sugar was below the labeled content. The Spanish study also reported similar results for other food categories, documenting an overall compliance rate with EU regulation of 98.4% (20). In Ireland, the results differed. O'Mahony et al. (19) reported that in yogurts, the sugar content was more likely to be non-compliant with the EC guidance, differing from the labeled value. Out of 200 sampled yogurts, 19% did not meet EU tolerance ranges, and significant differences were seen in all types of yogurts (natural, flavored, and luxury). Much better compliance, however, was reported for other nutrients, particularly for fats and saturated fats (3 and 5%, respectively). Albuquerque et al. investigated the compliance of mandatorily labeled nutrients in Portugal (18), but they focused primarily on fats, salt, and saturated fats; the observed compliance for these three nutrients was 88, 74, and 73%, respectively. Their study highlighted notable differences between different food categories, e.g., nutrient contents in fast food and potato products were typically overestimated, while they were generally underestimated in sauces. The reasons for the observed deviations from labeled values can be attributed to a variety of issues in the production process. For example, manufacturers commonly use calculations based on the food ingredients to estimate the nutrient content (11). These calculations frequently contain errors due to limited data on the nutritional composition of the ingredients; these values are then further confounded by the production process, batch-to-batch variability, and issues related to processing and product stability (18). Errors in calculations can also occur when a manufacturer employs laboratory analyses, e.g., problems related to inappropriate sampling and sample preparation (including homogenisation) or an inappropriate analytical method for the selected food matrix are potentially common occurrences (17).

Deviations, even within regulatory tolerance ranges, can also limit the reliability of the data in branded food databases. However, the results of our study did not show any significant differences between the labeled and analyzed sugar contents, either in the whole sample or among the specific subcategories, affirming previous findings concerning the sugar contents of SSBs and other food categories (20). While our results suggest that differences may occur when market share differences are considered, our sales-weighted figures should be taken with some caution since the beverage with the largest market share in Slovenia (cola-type drink) contains added FGS and can contain up to 5% of maltose and other sugars (36), and maltose and other sugars were not quantified in our laboratory analyses.

Our study results mostly indicate that the labeled sugar content information in SSBs is reliable for compiling branded food databases and for use in nutrition research. Branded food databases contain a large amount of data based on food labeling information, which in turn (if accurate) can help the responsible authorities and researchers compile a more comprehensive picture of food intake in the population and monitor food reformulation progress. For example, in Slovenia, such data are collected as part of the Composition and Labeling Information System (CLAS) (4). SSBs are a particularly important category because they are subject to reformulation activities. Recent studies have shown that reformulation changes can happen quickly in this category (37–40). Sugar reduction and the use of non-caloric sweeteners represent another cause of significant differences in the compositions of non-alcoholic beverages, even within the same subcategories (41). As a result, in nutrition research, it has become increasingly important for researchers to accurately assess sugar intake without under- or overestimating the sugar content of food products. Finally, branded food databases can play a vital role in policy development. When such databases are compiled and connected in a cross-sectional manner and across different time points, they can be very useful for generating insights on the effectiveness of food reformulation initiatives (6). Similarly, when these databases are combined with sales and/or consumer habits data, they can be used to inform emerging national and local food policies (42, 43).

Historically, data on the composition of SSBs, as they relate to the content of specific types of sugars, are very limited in the scientific literature. The results of our study showed that, in most beverages, the content of total fructose is about 50 %, which is consistent with the labeled ingredients for these products (most beverages were mainly sweetened with disaccharide sucrose, which is composed of fructose and glucose). This is a different situation than that observed in the U.S., where FGS (also commonly known as high fructose corn syrup) almost completely replaced sugar in SSBs (44). Beverages in the U.S. contain higher amounts of fructose, around 55%, reflecting the use of a standard version of FGS containing 55% of fructose (36). In our study sample, products sweetened with FGS included a few carbonated drinks ($N = 4$) with fructose proportions between 59 and 61%. The use of FGS is less common in beverages in Europe than in the US; in Europe, GFS with 42% fructose is sold as a standard (45). In our study sample, GFS was present in seven beverages, mainly in fruit drinks, in which the GFS offset the higher fructose content from fruit juice so that the final proportion of fructose was again around 50%, which is comparable with SSBs sweetened with sucrose only. Higher fructose levels were only seen in beverages with either FGS or higher fruit juice contents. Beverages with a higher percentage of fruit and 100% fruit juice are often perceived as a healthier choice. However, as a result of their high levels of naturally occurring sugar (especially fructose), studies suggest that their metabolic effects are very similar to beverages with added sugars (46). Our study showed that sugar-sweetened colas mainly consist of free glucose and fructose, regardless of whether sucrose or FGS is used for sweetening. Similar findings were reported in other studies, in which researchers hypothesized the potential usage of unlabelled FGS (16, 47). In aqueous solutions, sucrose can be subject to natural hydrolysis, which occurs in acidic conditions over time (45). Birkhed reported that in SSBs, the majority of sucrose can be hydrolysed after 5 months of storage at room temperature (48). A recent case study reported that this process is much faster in cola drinks than in other carbonated fruit drinks (49). For this reason, the content of monosaccharides in any final products cannot be used as a reliable indicator of specific sweetening ingredient use in SSBs. Since cola drinks with sucrose labeled in our study sample contained around 50 % fructose, and inverted sugar is rarely used in soft drinks (45), we believe that sucrose hydrolysis most likely took place in these samples. Meanwhile, energy drinks had the lowest proportion of fructose (42%) due to the use of glucose or GS. Glucose is a common ingredient in sport and energy drinks since it provides more energy per unit of sweetness and allows faster use of input energy (45). Nevertheless, even with a lower proportion of fructose, energy drinks still contained the highest amount of total sugar from the SSB drinks analyzed in the study, and, consequently, the amount of fructose (g/100 ml) in these drinks was similar to, and sometimes even higher than, other SSBs (50).

SSBs contribute significantly to fructose intake, generally due to the high sugar content and FGS use. In the US, the use of FGS is increasing (33). In the EU, the use of both FGS and GFS in SSBs was restricted by the Common Agricultural Policy (CAP), which contained a quota that strictly limited the quantities in SSB manufacture (45). In 2017, this quota was removed, and manufacturers were allowed to add caloric sweeteners in SSBs during the following years. The EC, however, does not expect these policy changes to cause an increase in the use of FGS and/or GFS (51). Further monitoring of these policy changes and their effects is essential, as fructose intake has already increased greatly in recent decades (52). Various studies have found associations between high fructose intake and increased risks for major non-communicable chronic diseases, such as metabolic syndrome, heart disease, obesity, diabetes, and dementia (31, 53–55). Various studies also suggest that high fructose consumption can negatively impact less physically active people (56). However, information on the negative health effects of the current levels of fructose consumption in Europe remains limited (57). The EC encourages the monitoring of fructose and its intake for the aforementioned reasons and requires that manufacturers inform consumers about the type(s) of sugar (sucrose, FGS, GFS, etc.) that is/are contained in food products (58).

### Strengths and Limitations

A major strength of our study is that we were able to analyse the most relevant SSBs in the food supply using the market share sampling approach. Although we only analyzed 51 beverages, the sampled products represented 69% of the national market share volume. However, this sampling approach has an important limitation: it does not provide insights into niche products with a low market share. Another limitation of the approach is that the samples were purchased in 2020, based on 2019 market share information, i.e., yearly market share data were only available for the previous year. While food labeling data are commonly used in nutrition research, very few studies have investigated the concordance of food labeling data to actual product composition. In this regard, our study represents a major contribution to this field of research, with results that can be used to inform the future efforts of food control authorities and policymakers in Slovenia and across the EU. The tolerances used in our study corresponded to the official EC guidance on regulatory tolerances, which are required for the nutritional declaration on food labels. In certain cases (i.e., for nutrition and health claims), this guidance provides stricter criteria, but they were not applied, as the study sampling approach was not designed in such a way as to capture a sufficient number of products labeled with such claims. Finally, we should note that, in the laboratory analyses, we only investigated the content of sucrose, glucose, and fructose, not other sugars. Other sugars may have been present in smaller quantities, and they could have affected the overall sugar content, sales-weighting, and proportion of fructose, especially in products with FGS, where maltose and other saccharides can represent up to 5% of the sugar content (36).

## CONCLUSIONS

Our study showed that the labeled sugar content in SSBs in Slovenia corresponded to the actual sugar content found in the food products, suggesting the usefulness and reliability of nutrition label information for compiling branded food databases and monitoring the food supply. Our case study also provided an example of how market share-driven sampling can be used for these types of verification studies. The results from the laboratory analyses and the assessment of the ingredients in SSBs showed that sucrose was the main sweetening component and that the proportion of fructose was typically around 50% across the sampled products. Given the possible changes in the use of caloric sweeteners, further monitoring of this area for food policy and nutrition research purposes is recommended, with a particular focus on FGS usage.

## DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because restrictions apply. The raw data supporting the conclusions of this article will be made available by the authors without undue reservation (without disclosure of specific brands). Sales data were obtained for internal use only and can be ordered directly from NielsenIQ agency. Requests to access the datasets should be directed to Igor Pravst (igor.pravst@nutris.org).

## AUTHOR CONTRIBUTIONS

IP: conceptualization. ŽL, MH, and SK: sample collection. MS: laboratory analyses. EH: data analyses, formal analysis, and writing—original draft preparation. IP and EH: methodology. IP, AK, KŽ, MS, and NP: manuscript writing—review and editing. All authors have read and agreed to the published version of the manuscript.

## FUNDING

Data collection for this study was supported by the national research programmes Nutrition and Public Health (P3-0395, funded by the Slovenian Research Agency) and Biochemical and biophysical characterization of natural compounds (P4-0121, funded by the Slovenian Research Agency), Infrastructure programme for monitoring of the composition and labeling of foods (IO-0054, funded by the Slovenian Research Agency), and research project L3-9290, funded by the Ministry of Health of Republic of Slovenia and Slovenian Research Agency. The study was conducted within the Food Nutrition Security Cloud project (FNS-Cloud), which received funding from the European Union's Horizon 2020 Research and Innovation Programme (H2020-EU.3.2.2.3.—A sustainable and competitive agri-food industry) under grant agreement no. 863059. Neither European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use of the information contained herein. The funders had no role in the design of the study, in the collection, analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2022.794468/full#supplementary-material

## REFERENCES

1. Kapsokefalou M, Roe M, Turrini A, Costa HS, Martinez-Victoria E, Marletta L, et al. Food composition at present: new challenges. *Nutrients*. (2019) 11:1714. doi: 10.3390/nu11081714
2. Zupanič N, Hristov H, Gregorič M, Blaznik U, Delfar N, Koroušić Seljak B, et al. Total and free sugars consumption in a Slovenian population representative sample. *Nutrients*. (2020) 12:1729. doi: 10.3390/nu12061729
3. Seljak BK, Valenčič E, Hristov H, Hribar M, Lavriša Ž, Kušar A, et al. Inadequate intake of dietary fibre in adolescents, adults, and elderlies:

results of slovenian representative SI. Menu study. *Nutrients*. (2021) 13:3826. doi: 10.3390/nu13113826
4. Pravst I, Hribar M, Žmitek K, BlaŽica B, Korousic Seljak B, Kušar A. Branded foods databases as a tool to support nutrition research and monitoring of the food supply: insights from the Slovenian Composition and Labelling Information System (CLAS). *Front Nutr*. (2021) 8:798576. doi: 10.3389/fnut.2021.798576
5. Harrington RA, Adhikari V, Rayner M, Scarborough P. Nutrient composition databases in the age of big data: foodDB, a comprehensive, real-time database infrastructure. *BMJ Open*. (2019) 9:e026652. doi: 10.1136/bmjopen-2018-026652
6. Black R. Begin to imagine: thoughts and considerations following the 39th NNDC. *J Food Compos Analy*. (2017) 64:143–4. doi: 10.1016/j.jfca.2017.07.021
7. Neal B, Sacks G, Swinburn B, Vandevijvere S, Dunford E, Snowdon W, et al. Monitoring the levels of important nutrients in the food supply. *Obes Rev*. (2013) 14:49–58. doi: 10.1111/obr.12075
8. Rayner M, Vandevijvere S. *INFORMAS Protocol: Food Labelling Module,* Auckland: The University of Auckland (2017).
9. Food Monitoring Group. Progress with a global branded food composition database. *Food Chem*. (2013) 140:451–7. doi: 10.1016/j.foodchem.2012.10.065
10. *European Commission. Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the Provision of Food Information to Consumers.* (2011). Available online at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32011R1169 (accessed September 13, 2021).
11. Machackova M, Giertlova A, Porubska J, Roe M, Ramos C, Finglas P. EuroFIR Guideline on calculation of nutrient content of foods for food business operators. *Food Chem*. (2018) 238:35–41. doi: 10.1016/j.foodchem.2017.03.103
12. European Commission. *Guidance Document for Competent Authorities for the Control of compliance With EU Legislation on: Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the Provision of Food Information to Consumers* (2012). Available online at: https://ec.europa.eu/food/system/files/2016-10/labelling_nutrition-vitamins_minerals-guidance_tolerances_1212_en.pdf (accessed September 13, 2021).
13. The Government of the Republic of Slovenia. *Decree Implementing the Regulation (EU) on the Provision of Food Information to Consumers*. The Official Gazette of the Republic of Slovenia (2014). Available online at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32011R1169 (accessed September 13, 2021).
14. Klaus B, Busani C. Italy adopts legislation establishing sanctions for the violation of the provisions of regulation (EU) No. 1169/2011 on food information to consumers. *Eur Food Feed Law Rev*. (2018) 13:148–50. https://www.jstor.org/stable/90021317
15. European Commission. *The Rapid Alert System for Food and Feed - Annual Report 2020*. Luxembourg (2021).
16. Ventura EE, Davis JN, Goran MI. Sugar content of popular sweetened beverages based on objective laboratory analysis: focus on fructose content. *Obesity*. (2011) 19:868–74. doi: 10.1038/oby.2010.255
17. Kok SC, Radzi CwJBwM. Accuracy of nutrition labels of pre-packaged foods in Malaysia. *Br Food J*. (2017) 119:230–41. doi: 10.1108/BFJ-07-2016-0306
18. Albuquerque TG, Nunes MA, Oliveira MBPP, Costa HS. Compliance of declared vs. analysed values with EU tolerance limits for mandatory nutrients in prepacked foods. *Food Chem*. (2020) 302:125330. doi: 10.1016/j.foodchem.2019.125330
19. O'Mahony S, O'Donovan CB, McDonald K, Lyons OC, Quinn SF, Creane RJ, et al. The accuracy of nutrition labels as a means of monitoring the reformulation of food in Ireland - the case of yoghurts. *Proc Nutr Soc*. (2020) 79:E713. doi: 10.1017/S002966512000662X
20. Yusta-Boyo MJ, Bermejo LM, García-Solano M, López-Sobaler AM, Ortega RM, García-Pérez M, et al. Sugar content in processed foods in Spain and a comparison of mandatory nutrition labelling and laboratory values. *Nutrients*. (2020) 12:1078. doi: 10.3390/nu12041078
21. Zupanič N, Mis NF, Pravst I, Galanakis CM. *Chapter 11 - Soft Drinks: Public Health Perspective. Trends in Non-alcoholic Beverages*. London: Academic Press (2020). p. 325–69. doi: 10.1016/B978-0-12-816938-4.00011-2
22. Malik VS, Pan A, Willett WC, Hu FB. Sugar-sweetened beverages and weight gain in children and adults: a systematic review and meta-analysis. *Am J Clin Nutr*. (2013) 98:1084–102. doi: 10.3945/ajcn.113.058362
23. Malik VS, Hu FB. Sugar-sweetened beverages and cardiometabolic health: an update of the evidence. *Nutrients*. (2019) 11:1840. doi: 10.3390/nu11081840
24. Yoshida Y, Simoes EJ. Sugar-sweetened beverage, obesity, and type 2 diabetes in children and adolescents: policies, taxation, and programs. *Curr Diab Rep*. (2018) 18:31. doi: 10.1007/s11892-018-1004-6
25. Asgari-Taee F, Zerafati-Shoae N, Dehghani M, Sadeghi M, Baradaran HR, Jazayeri S. Association of sugar sweetened beverages consumption with non-alcoholic fatty liver disease: a systematic review and meta-analysis. *Eur J Nutr*. (2019) 58:1759–69. doi: 10.1007/s00394-018-1711-4
26. World Health Organization. Diet, Nutrition and the Prevention of Chronic Diseases—Report of the Joint WHO/FAO Expert Consultation. Geneva, Switzerland: World Health Organization; 2003.
27. World Health Organization. *Guideline: Sugars Intake for Adults and Children*. Geneva: World Health Organization (2015).
28. Azaïs-Braesco V, Sluik D, Maillot M, Kok F, Moreno LA. A review of total & added sugar intakes and dietary sources in Europe. *Nutr J*. (2017) 16:6. doi: 10.1186/s12937-016-0225-2
29. Graffe MIM, Pala V, De Henauw S, Eiben G, Hadjigeorgiou C, Iacoviello L, et al. Dietary sources of free sugars in the diet of European children: the IDEFICS Study. *Eur J Nutr*. (2020) 59:979–89. doi: 10.1007/s00394-019-01957-y
30. Hannou SA, Haslam DE, McKeown NM, Herman MA. Fructose metabolism and metabolic disease. *J Clin Invest*. (2018) 128:545–55. doi: 10.1172/JCI96702
31. Taskinen M-R, Packard CJ, Borén J. Dietary fructose and the metabolic syndrome. *Nutrients*. (2019) 11:1987. doi: 10.3390/nu11091987
32. Sundborn G, Thornley S, Merriman TR, Lang B, King C, Lanaspa MA, et al. Are liquid sugars different from solid sugar in their ability to cause metabolic syndrome? *Obesity*. (2019) 27:879–87. doi: 10.1002/oby.22472
33. Marriott BP, Cole N, Lee E. National estimates of dietary fructose intake increased from 1977 to 2004 in the United States. *J Nutr*. (2009) 139:1228S–35. doi: 10.3945/jn.108.098277
34. Mayes PA. Intermediary metabolism of fructose. *Am J Clin Nutr*. (1993) 58(5 Suppl.):754s−65. doi: 10.1093/ajcn/58.5.754S
35. Sluik D, Engelen AI, Feskens EJ. Fructose consumption in the Netherlands: the Dutch national food consumption survey 2007–2010. *Eur J Clin Nutr*. (2015) 69:475–81. doi: 10.1038/ejcn.2014.267
36. White JS, Hobbs LJ, Fernandez S. Fructose content and composition of commercial HFCS-sweetened carbonated beverages. *Int J Obes*. (2015) 39:176–82. doi: 10.1038/ijo.2014.73
37. Silva PD, Cruz R, Casal S. Sugars and artificial sweeteners in soft drinks: a decade of evolution in Portugal. *Food Control*. (2021) 120:107481. doi: 10.1016/j.foodcont.2020.107481
38. Goiana-da-Silva F, Morais Nunes A, Miraldo M, Bento A, Breda J, Ferreira Araújo F. Using pricing policies to promote public health: the sugar sweetened beverages taxation experience in Portugal. *Acta Med Portuguesa*. (2018) 31:191–5. doi: 10.20344/amp.10222
39. Hafner E, Hribar M, Hristov H, Kušar A, Žmitek K, Roe M, et al. Trends in the use of low and no-calorie sweeteners in non-alcoholic beverages in Slovenia. *Foods*. (2021) 10:387. doi: 10.3390/foods10020387
40. Hafner E, Pravst I. The sharp rise in the use of low- and no-calorie sweeteners in non-alcoholic beverages in slovenia: an update based on 2020 data. *Front Nutr*. (2021) 8:778178. doi: 10.3389/fnut.2021.778178
41. Vin K, Beziat J, Seper K, Wolf A, Sidor A, Chereches R, et al. Nutritional composition of the food supply: a comparison of soft drinks and breakfast cereals between three European countries based on labels. *Eur J Clin Nutr*. (2020) 74:17–27. doi: 10.1038/s41430-019-0442-9
42. Zupanic N, Hribar M, Fidler Mis N, Pravst I. Free sugar content in pre-packaged products: does voluntary product reformulation work in practice? *Nutrients*. (2019) 11:2577. doi: 10.3390/nu11112577
43. Pravst I, Lavriša Ž, Kušar A, Miklavec K, Žmitek K. Changes in average sodium content of prepacked foods in Slovenia during 2011–2015. *Nutrients*. (2017) 9:952. doi: 10.3390/nu9090952
44. Ng SW, Slining MM, Popkin BM. Use of caloric and noncaloric sweeteners in US consumer packaged foods, 2005–2009. *J Acad Nutr Diet*. (2012) 112:1828–34.e1–6. doi: 10.1016/j.jand.2012.07.009

45. Ashurst PR, Hargitt R, Palmer F. Chapter 3 - ingredients in soft drinks. In: Ashurst PR, Hargitt R, Palmer F, editors. *Soft Drink and Fruit Juice Problems Solved*. 2nd ed. Sawston: Woodhead Publishing (2017). p. 29–66. doi: 10.1016/B978-0-08-100918-5.00003-5
46. Choi HK, Curhan G. Soft drinks, fructose consumption, and the risk of gout in men: prospective cohort study. *BMJ*. (2008) 336:309–12. doi: 10.1136/bmj.39449.819271.BE
47. Walker RW, Dumke KA, Goran MI. Fructose content in popular beverages made with and without high-fructose corn syrup. *Nutrition*. (2014) 30:928–35. doi: 10.1016/j.nut.2014.04.003
48. Birkhed D. Sugar content, acidity and effect on plaque pH of fruit juices, fruit drinks, carbonated beverages and sport drinks. *Caries Res*. (1984) 18:120–7. doi: 10.1159/000260759
49. Pero M, Ershadi A, Gholami R, Atrchi P, Jafari Asl M. Physicochemical analysis of carbonated fruit drink and cola beverage prepared from fructose syrup 55 and sucrose. *Alborz Univ Med J*. (2021) 10:49–54. http://aums.abzums.ac.ir/article-1-1268-en.html
50. Idris AM, Vani NV, Almutari DA, Jafar MA, Boreak N. Analysis of sugars and pH in commercially available soft drinks in Saudi Arabia with a brief review on their dental implications. *J Int Soc Prev Community Dent*. (2016) 6(Suppl. 3):S192–6. doi: 10.4103/2231-0762.197190
51. European Commission. *The End of the Sugar Production Quotas in the EU 2017*. Available online at: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_17_3488 (accessed September 21, 2021).
52. Lowette K, Roosen L, Tack J, Vanden Berghe P. Effects of high-fructose diets on central appetite signaling and cognitive function. *Front Nutr*. (2015) 2:5. doi: 10.3389/fnut.2015.00005
53. Malik VS, Hu FB. Fructose and cardiometabolic health. *J Am Coll Cardiol*. (2015) 66:1615–24. doi: 10.1016/j.jacc.2015.08.025
54. Siervo M, Wells JCK, Brayne C, Stephan BCM. Reemphasizing the role of fructose intake as a risk factor for dementia. *J Gerontol Ser A*. (2011) 66A:534–6. doi: 10.1093/gerona/glq222
55. Johnson RJ, Segal MS, Sautin Y, Nakagawa T, Feig DI, Kang D-H, et al. Potential role of sugar (fructose) in the epidemic of hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease. *Am J Clin Nutr*. (2007) 86:899–906. doi: 10.1093/ajcn/86.4.899
56. Tappy L, Rosset R. Health outcomes of a high fructose intake: the importance of physical activity. *J Physiol*. (2019) 597:3561–71. doi: 10.1113/JP278246
57. Zafar MI, Frese M, Mills KE. Chronic fructose substitution for glucose or sucrose in food or beverages and metabolic outcomes: an updated systematic review and meta-analysis. *Front Nutr*. (2021) 8:647600. doi: 10.3389/fnut.2021.647600
58. European Commission. *Consumption and Impact of High Fructose Syrups*. Luxembourg (2018).


**Author Disclaimer:** Information and views in this report do not necessarily reflect the official opinion or position of the European Union.

**Conflict of Interest:** IP has led and participated in various other research projects in the area of nutrition, public health, and food technology, which were (co)funded by the Slovenian Research Agency, Ministry of Health of the Republic of Slovenia, the Ministry of Agriculture, Forestry, and Food of the Republic of Slovenia, and, in the case of specific applied research projects, also by food businesses.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Nutrition

ORIGINAL RESEARCH
published: 23 March 2022
doi: 10.3389/fnut.2021.795802

# Responsible Governance for a Food and Nutrition E-Infrastructure: Case Study of the Determinants and Intake Data Platform

*Lada Timotijevic[1]*, Indira Carr[2], Javier De La Cueva[3], Tome Eftimov[4], Charo E. Hodgkins[1], Barbara Koroušić Seljak[4], Bent E. Mikkelsen[5], Trond Selnes[6], Pieter Van't Veer[6] and Karin Zimmermann[6]*

[1] School of Psychology, University of Surrey, Guildford, United Kingdom, [2] School of Law, University of Surrey, Guildford, United Kingdom, [3] Javier de la Cueva and Asociados, Madrid, Spain, [4] Jožef Stefan Institute, Ljubljana, Slovenia, [5] Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark, [6] Wageningen Economic Research, Wageningen University and Research Centre, Wageningen, Netherlands

OPEN ACCESS

**Edited by:**
*Massimo Lucarini, Council for Agricultural Research and Economics, Italy*

**Reviewed by:**
*Bo Zhao, Tilburg University, Netherlands*
*Darija Vranesic Bender, University Hospital Centre Zagreb, Croatia*

***Correspondence:**
*Lada Timotijevic*
*l.timotijevic@surrey.ac.uk*

**Specialty section:**
*This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition*

**Received:** *15 October 2021*
**Accepted:** *23 December 2021*
**Published:** *23 March 2022*

**Citation:**
*Timotijevic L, Carr I, De La Cueva J, Eftimov T, Hodgkins CE, Koroušić Seljak B, Mikkelsen BE, Selnes T, Van't Veer P and Zimmermann K (2022) Responsible Governance for a Food and Nutrition E-Infrastructure: Case Study of the Determinants and Intake Data Platform. Front. Nutr. 8:795802. doi: 10.3389/fnut.2021.795802*

The focus of the current paper is on a design of responsible governance of food consumer science e-infrastructure using the case study Determinants and Intake Data Platform (DI Data Platform). One of the key challenges for implementation of the DI Data Platform is how to develop responsible governance that observes the ethical and legal frameworks of big data research and innovation, whilst simultaneously capitalizing on huge opportunities offered by open science and the use of big data in food consumer science research. We address this challenge with a specific focus on four key governance considerations: data type and technology; data ownership and intellectual property; data privacy and security; and institutional arrangements for ethical governance. The paper concludes with a set of responsible research governance principles that can inform the implementation of DI Data Platform, and in particular: consider both individual and group privacy; monitor the power and control (e.g., between the scientist and the research participant) in the process of research; question the veracity of new knowledge based on big data analytics; understand the diverse interpretations of scientists' responsibility across different jurisdictions.

**Keywords: food consumer behavior, food consumer choice, data quality, interoperability, standardization, big data, ethical, machine learning**

## INTRODUCTION

Big data provides immense opportunities to radically alter the way in which science is done, fostering cross-fertilization between disciplines, and providing connectivity between disparate data-sets. There has been a huge surge in recent years of initiatives to develop structures and networks that foster data gathering, connectivity and large data analytics to ensure advancements in core areas of science. Distributed computing infrastructures—commonly known as e-infrastructures—have been created that provide researchers shared access to large data collections enabled through advanced Information Communications, large-scale computing resources, and high-performance visualization (1). Research e-infrastructures have received considerable EU funding, with over 20 billion Euros currently being invested for their development (2). However, the scientific domain of food consumption and its relation to health and sustainability, is not yet facilitated by a well-supported research infrastructure (RI). In the area of food, there are well-established research infrastructures in the omics and nutritional systems biology domain [e.g., ELIXIR[1]], medical sciences domain [e.g., clinical research: ECRIN ERIC[2]; translational medicine: EATRIS ERIC[3], biobanking: BBMRI, ERIC[4]], and health research is relatively well funded by the public and private sector (3). However, it is increasingly recognized that the focus on diseases brings a curative bias and, as the COVID pandemic has amply demonstrated, prevention is a crucial aspect of well-functioning social and health care (4). Similarly, agriculture has traditionally been a domain for public investment in research, with the emergence of research infrastructures in agri-ecosystems such as Analysis and Experimentation on Eco-systems[5]. Indeed, in the food domain, environmental issues are paramount. But there is increasing recognition that the current global challenges associated with food and environmental sustainability cannot only be solved within a productivist paradigm and the associated technical solutions, instead, that we need to explore the consumption process and its interaction with the production (5–7).

The case for a dedicated food nutrition e-infrastructure has been made in the past several years with the emergence of a network of scientists set to develop Food Nutrition Health Research Infrastructure[6]. A fundamental part of this international initiative is the development of an e-infrastructure that enables innovative science in the domain of dietary determinants and intake. Using ICT and new technology such as smart phones, APPs, sensors, internet of things and big data offers new ways of exploring food consumption in the context of the food chain. To harness these new technologies and address this gap, the Determinants and Intake (DI) Data Platform has been forged from the work of two European projects – EuroDISH (2012–2014) and Richfields (2015–2018). The DI Data Platform emerged in recognition that the food chain and consumer food choices are of direct relevance to public health, prevention, health promotion, environmental sustainability and socioeconomic impacts of the food system, but the domain of food consumption has been omitted from research funding (3, 8). Intensifying the research in this domain would therefore be necessary (9).

The focus of the current paper is to examine how to develop responsible governance of a food nutrition e-infrastructure using the case study Determinants and Intake (DI) Data Platform. The proposed DI Data Platform that is being developed will ostensibly utilize big data in promoting research and innovation in this domain. Hereby rests the challenge of how to develop responsible governance that observes the ethical and legal frameworks of big data research and innovation, whilst simultaneously capitalizing on huge opportunities offered by big data in food nutrition and health research. To reach this ambitious goal and to counteract some of the reservations that actors traditionally have about big data, agreements and consensus among a broad range of stakeholders is needed and as a result a fair and accepted governance structure is necessary.

This paper first describes the unique features of big data and the broad challenges for ethical governance it poses when harnessed within a research e-infrastructure. Following the description of the vision and mission of DI Data Platform, we address the core challenges for the development of responsible governance with a specific focus on four key governance considerations: data type and technology; data privacy and security; data ownership; and institutional arrangements for ethical governance. The paper concludes with a synthesis of ethical challenges linking these to the responsible research governance principles that may help the development of the research e-infrastructures, and in particular, highlight the relevance of Responsible Research and Innovation (RRI) as a form of meta-responsibility that brings together ethical and legal aspects of governance under a single framework. RRI aims to "shape, maintain, develop, coordinate and align existing and novel research and innovation-related processes, actors and responsibilities with a view to ensuring desirable and acceptable research outcomes" (10).

## ETHICAL CHALLENGES OF BIG DATA IN SCIENCE: DATA, DATA PROCESSING AND DATA MANAGEMENT

The meaning of big data has been widely discussed (11–20) as efforts have been made to delineate the concept. Fothergill et al. (11) summarize the literature that has attempted to define and explain big data. Big data is often described as "large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of the information" (p. 11) (21). It's value to research is derived from its specific properties: it is indexical in nature, relational, flexible, scalable, re-purposable, continuously updatable and easily removable from the context of data collection (3).

The sheer scale of big data poses ethical, legal and societal challenges to e-infrastructures. There is a growing literature concerned with addressing these as part of the projects for research e-infrastructure building workable data governance frameworks [e.g., (2, 11)]. Common to most of this literature is the realization that responsible research e-infrastructure governance is a matter of contextualized and deeply embedded decision-making that departs to a large extent from the principles of the conventional research ethics and governance enshrined in professional practice and law. Governance specifies how decision-making within an organization should be structured

[1] https://elixir-europe.org/about-us (accessed October 13, 2021).
[2] https://ecrin.org/ (accessed October 13, 2021).
[3] https://eatris.eu/ (accessed October 13, 2021).
[4] https://bbmri.eu (accessed October 13, 2021).
[5] https://www.anaee.eu/about/missions (accessed October 13, 2021).
[6] https://fnhri.eu/ (accessed October 13, 2021).

and implies allocation of responsibility (22), in terms of who is responsible for what and under what conditions. Responsibility toward an individual participant has been the main focus of legal frameworks developed to regulate traditional scientific processes. For instance, in the EU, the General Data Protection Regulation (23) is developed with an explicit remit to protect individual[7] privacy rights, whilst at the same time to remove the obstacles to flows of personal data within the Union through harmonization of the law across the member states. Outside of Europe, the regulation is seemingly much more fragmented—for instance the US has only recently introduced Information Transparency and Personal Data Control Act[8], which protects personal information and institutes the Federal Trade Commission responsible for the development and oversight of the requirementsts for collecting, processing, storing and sharing sensitive personal information. The bill however does not include the right of an individual to access, correct or delete the data stored about them, which is included within the EU GDPR.

Big data has revealed new challenges for our conceptualization of researchers' and scientists' responsibilities when utilizing big data. Researching with big data makes it incumbent upon researchers to re-think their own responsibilities vis-à-vis both the participants but also society at large. This has resulted in the explicit need to broaden out the scope of responsible research governance to include considerations of diverse data types (not just human and animal data), analytical processes, group protections and long term implications of research (3). It has provided new challenges to considerations of the rights of the research participants and the rights and responsibilities of researchers.

This paper will highlight these challenges through the case of developing DI Data Platform. In the sections that follow we present the core considerations that have informed the development of the responsible governance approach for this e-infrastructure, and which are analyzed in light of the principles of RRI.

## THE CASE OF DETERMINANTS AND INTAKE DATA PLATFORM

A conceptual design of DI Data Platform was built on the vision of a Food, Nutrition and Health Research Infrastructure (hereafter: FNH-RI), which would connect the data and science in the domains that link food consumption, sustainability and health. It was first proposed by the EU project EuroDISH [www.eurodish.eu; (9)], which defined the main pillars of the proposed FNH-RI: **D**eterminants of food choice (why do we buy and eat what we eat?), **I**ntake of food (what do we actually eat, when, where and how?), **S**tatus of the body (how overweight, obese are we?) and **H**ealth. One finding of the EuroDISH project was the need for better data in the pillars of Determinants and Intake where there is notable absence of standardized concepts, methods and tools for data collection, and no international data depositaries. This led to an EU-funded project RICHFIELDS (acronym for Research Infrastructure on Consumer Health and Food Intake for E-science with Linked Data Sharing) with an explicit remit to prepare a design for a Determinants Intake (DI) Data Platform. The idea was to implement this platform into a larger FNH RI, with a unique infrastructure providing data services ("DATA"), facilities and tools ("FACT"), and training, education and dissemination ("TED") to scientists based on standardized collection and integration of data by a DI "Richfields" Consumer APP drawn from consumer-donated data; as well as research and business data repositories, drawn from public research institutions, business and non-governmental organizations. The provision of these scientific services will ultimately enable policymakers, NGOs, food industries, SMEs, farmers and consumers to make more responsible decisions.

The DI Data Platform combines different types of food consumer science data: consumer-generated data, mostly real-time and *in situ* (e.g., food consumption data generated via APPs); business-generated data (e.g., sales data, food composition data); and research-generated data from research laboratories, experimental facilities and from existing and developing RIs (24). The DI Data Platform is summarized in **Figure 1**.

Ultimately, the DI Data Platform will enable research with real impact on our current food system, addressing, through the use of innovative technologies and big data, the current challenges of consumption and production within the food system. **Table 1** summarizes the areas of research and the associated food system domains the DI Data Platform will address.

## CASE STUDY RESULTS: DEVELOPING RESPONSIBLE GOVERNANCE FOR DI DATA PLATFORM

Responsible research governance of the DI Data Platform has been a fundamental part of the data platform design. For information about our approach to the platform development, please refer to the published articles (25–28). The sections below address responsible research governance in the context of data type and technology; data ownership; privacy and security; and institutional arrangements.

### Data Type and Technology Relevant to DI Data Platform

Development of responsible governance requires clear understanding of the food consumer science data being used, its purpose and value, technologies for data harvesting, and the social and legal implications of their use.

[7]The EU is the process supplementing this regulation with a Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act) {SEC(2020) 405 final} - {SWD(2020) 295 final} - {SWD(2020) 296 final}. On 30 November 2021 the negotiators from the Council and the European Parliament reached provisional agreement on the Data Governance Act. This Act will enable reuse of certain categories of public-sector data that are subject to the rights of others, such as personal data and data protected by intellectual property rights. However public sector bodies that allow safe reuse of data need to ensure that privacy and confidentiality are fully preserved. See section *Privacy concerns and security* below.

[8]https://www.congress.gov/bill/116th-congress/house-bill/2013 (accessed October 13, 2021).

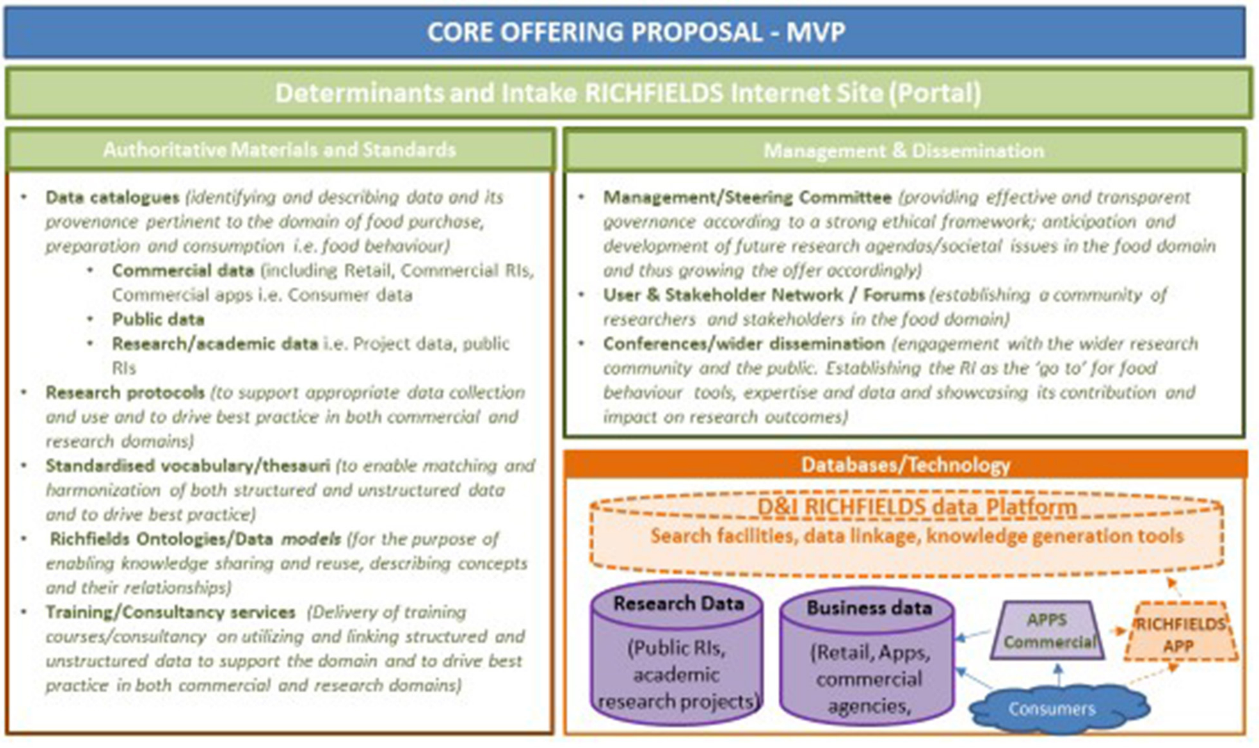


**FIGURE 1 |** Core offering of DI "Richfields" Data Platform - a minimum viable product.

The scientific relevance of the proposed data platform is dependent on the diversity of data available to it and these include:

- **Research data** from other research infrastructures, laboratories and experimental facilities. This is mostly structured data.[9]
- **Business data** (e.g., data from retailers, public procurement companies, statistical institutions and market organizations). This can be structured, semi-structured[10] and unstructured data.[11]
- **Consumer-generated data** from APPS (smartphone and tablet applications) and sensors. This is mostly unstructured data.

The data types relevant to the DI Data Platform can be referred to as reference data (e.g., food composition data), observational data (e.g., food intake data, physical activity), or data that is transformed into output data (e.g., nutrient intake, dietary patterns). Each of these types can relate to three key domains of human behavior that are relevant to food intake: purchase, preparation and consumption, each presenting particular issues and considerations for the DI Data Platform governance.

## Research Data

Linking between research data held within the **existing RIs (e.g., ELIXIR,**[12] **ECRIN ERIC)** in the food and health domain and the DI Data Platform is possibly the most accessible form of research data. However, the development of a DI Data Platform ontology and the harmonization of entities, food classification and description systems is fundamental to facilitate future data access/exchange between existing and new RIs. The development of authoritative materials and standards is also fundamental in order to establish best practice and to help shape the research community moving forwards thus making future data sharing activities easier. This will to a large extent depend on developing a community of experts in the field engaged to provide best practice guidelines and standards for data ontologies and structured data capture. **Laboratories/Data facilities (public and private)** also represent an opportunity to harvest data for research—of the 39 labs and facilities in Europe involved in consumer research in the food and health domain, many collect the data that is proprietary and typically not formatted, standardized or stored in a manner conducive to sharing outside the original purposes of the research study undertaken. In addition, the diversity of data-generating

[9]Structured or relational data concerns all data, which can be stored in a relational database.

[10]Semi-structured data is a form of structured data that does not conform with the formal structure of data models associated with relational databases; however, it may have information associated with it, such as metadata tagging, that allows elements contained to be addressed (e.g., XML or JSON data).

[11]Unstructured data, such as text, PDF documents, media posts, photos, audio files etc., does not have a pre-defined data model, thus it is not a good fit for a relational database.

[12]https://elixir-europe.org/news/elixir-launches-new-food-and-nutrition-community

**TABLE 1 |** Examples of research and innovation domains of food nutrition and health research that can be supported by determinants and intake data platform.

| | | Scientific research domains (9) | | | | |
|---|---|---|---|---|---|---|
| **Food system** | Building blocks of the food system | Food chain: food security, safety, quality, environmental sustainability | Consumer behavior: Determinants and intake of foods and nutrients | Consumer health: Status and function of the body, up to risk for health and disease | Development into coherent research domain | |
| | Consumers, foods and diets | Food reformulation toward energy-poor and nutrient- rich food supply | Innovative assessments by apps, sensors, wearables; ambulatory monitoring Communication of environmental sustainability of food supply to consumers and stakeholders. | Sensors and wearables for e.g., heart rate monitoring, blood glucose, lipids, etc. | Food Consumer science | Food systems science |
| | | Sustainable products and replacers of animal proteins design according to consumer needs | Food choices, preferences, hunger and satiety, behavior. | Biomolecular, (bio) chemical mechanisms and (patho-) physiological disease pathways for major chronic diseases and nutritional deficiencies. | | |
| | | High (nutritional) quality foods for acceptable prices | Interoperable EU-nutrition surveillance system, incl physical activity and psycho-social determinants. | Personalized nutrition for clinical settings and high risk groups; standard for dieticians, available to citizens. | | |
| | | Data from social surveys, nutritional epidemiology and community interventions are interoperable and link to pan-EU multicentre studies; they include health and safety issues and can line up with other food systems outcomes (social, environmental, economic). | | | | |
| | Consumers and the food environment | Portion sizes and labeling | Communication of health and nutritional quality of food products in food environments to consumers | Using big data to link consumer behavior and health risks. | | |
| | | Access and affordability to foods for all socio-economic groups | Communication of environmental sustainability of food products and food chain to consumers and stakeholders in food environment | Precision nutrition – linking genetics, food environment and behavior. | | |
| | Consumers and food supply chain | Standardized and valid LCAs on GHGe, LU and FFU. From farmgate, regional distributed center, consumer, waste. | Recipes/food composition enriches FCDBs | Food producers, retailers, restaurants and catering can evaluate the health and sustainability of their products, recipes and menus through transparent and standardized procedures, benchmarking their corporate responsibility | Agri-food science | |
| | | Closed nutrient cycles, e.g., for carbon and nitrogen, eutrophication and acidification minimized. | | | | |
| | | Sourcing of commodities respects social justice, equity, animal welfare and biodiversity | Alternatives for animal protein | Production quantities, nutrients and food processing meet health requirements. | | |

devices including video and audio results in a wide variety of data types increasing the difficulty of *post-hoc* data integration. The task of transforming data into information that can be integrated can be solved using various technologies, depending on the type of data. For instance, information extraction based on Natural Language Processing (NLP) can be used for dealing with textual data that is one the major sources of data (e.g., from literature, media and social media). Another modern technology is Deep Learning (DL) for automatically recognizing entities (information) from images. Transforming information into knowledge requires Machine Learning/Artificial Intelligence (ML/AI) techniques, such as Named Entity Recognition (NER) and Representative Learning for describing information in a form of knowledge (e.g., as a knowledge graph). Each of these technologies, however, carry a potential risk of introducing errors as a result of improper digitization or storage of information, and the errors introduced due to the geographic and cultural differences in the meaning of the data and its interpretation (29). Furthermore, by its very nature, ML can perpetuate the existing biases in the data, mis-representing aspects of reality.

Similarly with structured public research data, it would be paramount to harmonize Standard Operating Procedures (SOPs), data management protocols, including calibration/standardization protocols and improved approaches to obtaining ethical consent at the outset of the studies for future sharing with the wider research community. Processes and procedures need to be in place that would provide a full account of the nature of the ML/AI and AI used in the case of DI Data Platform data linking and extrapolation, reflecting on the consequences of the actions enacted by algorithms, on a case-by-case basis. This will be an explicit focus of the DI Data Platform organizational and data governance.

### Business Data

Here, the ICT landscape is fast-paced, driven by an increasing number and connectivity of mobile devices used by consumers, and cheaper and better sensors. **Table 2** gives an overview of modern IC technology either being used or with the potential for future data collection. It is clear that the DI Data Platform will be flexible enough to be able to respond to this dynamic ICT environment, however, careful consideration is needed on a case by case basis about the extent to which the data captured is reflective of the proposed research concepts and the assumptions underpinning the ML algorithms utilized in business and the DI Data Platform, to ensure that the data captured and linked is of sufficient quality to be treated as a useful variable for the DI Data Platform. Data collection may be significantly impacted by business purpose (e.g., policies to control suppliers or for organic procurement) which may limit the potential usefulness of the data for scientific purposes within the proposed DI Data Platform. Re-purposing of data needs to be carefully scrutinized and controlled such that ethical compliance with the original participants' consent is always maintained. Appropriate meta-data must be assigned to data such that the possibility of non-compliant sharing from either a legal/ethical or data owner requirement is eliminated.

### Consumer Generated Data

Whilst we typically talk about data collected via APPS and sensors (e.g., Fitbit) as being consumer-generated, in reality, unless the data is being shared directly from the consumer to the DI Data Platform, this type of data must also be considered business data. There are three domains of behavior that could be relevant for the platform: purchase, preparation and consumption. There are a number of limitations associated with the APPS collecting this data from the consumer, as summarized in **Table 3**.

From a scientific perspective, the unknown quality and validity of the food composition databases used to underpin these APPS and the non-standardized procedures for portion size estimation means that conclusions with respect to the relationship between food consumption and nutrition-related diseases may be limited. Detailed research on the associations between specific nutrients and health outcomes may also be limited since majority of APPS in this domain focus only on energy and macronutrients.

Consumer-generated food purchase, preparation and consumption data are not typically collected in isolation from other potentially relevant data. A vital source for better understanding the possible drivers and barriers for people's food purchase, preparation and consumption behavior is likely to come from associations between these data and other relevant social, health and lifestyle data. This undoubtedly has the potential to give a more valid picture whereby different data sets corroborate each other to create a fuller, more accurate picture overall and the interconnectedness of APPS/tools now presents new opportunities to further enrich the food-related data from external sources. For example, it may be useful to gain domestic food purchase, preparation and consumption data from dedicated APPS and link this with health and lifestyle APPS for an individual. This combined data could be further enriched with demographic, situational and social context data collected through APPS such as Facebook, Twitter, TikTok, Instagram. However, the issue of ML-induced errors, biases and breeches of rights—e g., to privacy or human autonomy—are at the forefront of this endeavor. The extent to which users would find this interlinkage acceptable and be willing to share this type of extensive data with the proposed data platform will need to be carefully considered and governed, as will be discussed below. Due to the lack of available legal documents related to the terms and conditions and privacy statements linked to various APPS, there is often insufficient information available to assess the terms users must accept in order to use a service and the ways in which each APP gathers, uses, discloses, and manages their users' data. Hence, the legal limitations, organizational restrictions, confidentiality and privacy concerns related to collection, integration and dissemination of this consumer-generated data remain difficult to navigate other than on a specific case by case basis/detailed exploration with each individual APP of interest.

In short, the variety of data sources potentially involved and the varying levels of consent they carry with them present significant challenges to the open access vision of the DI Data Platform. This is especially important in a public-private business model scenario when there are often differing drivers and a

**TABLE 2 |** Review of ICT used by retail and market research organizations.

| Sector | Type of technology | Data capturing technology | Devices facilitating data capture | Type of data collected | Case studies |
|---|---|---|---|---|---|
| Retail | Consumer location sensing technologies | Geo-fencing | Smartphones, GPS-devices | Location data involving a location-sensitive device (eg. smartphones with GPS) | RetailNext (Aurora, Mobile Engage), Euclid (Traffic, Insight), Shopkick (shopBeacon), Brickstream (Brickstream 3D+), Axper (3D vision, Sentinel), PathTracker |
| | | Wi-Fi | Smartphones, tablets | Location data of smartphones connected to Wi-Fi | |
| | | Bluetooth low energy (BLE) | iBeacon-compatible transmitters, smartphones | Proximity data to Bluetooth beacons of enabled smartphones | |
| | | Visual systems | Analog or IP cameras, infrared cameras | Visual tracking data | |
| | | RFID technology | Smartphone RFID reader, RFID sensors | Consumer real-time product choice and purchasing data. Aggregated shopper tracking data to determine shopping speed, purchasing speed, and geography of trips. | |
| | | Combination of technologies mentioned above | Several sensors available that combines different data capturing technologies. E.g., Aurora from Retailnext combines video technology with BLE and WIFI. | | |
| | E-commerce and m-Commerce | Online analytic tools for personal computers | Smartphone, personal computer, tablet | Web browsing patterns and online shopping patterns (Cookie data), online purchasing data | Adobe marketing cloud (Adobe), Virtual stores (Walmart) |
| | | Online analytic tools for mobile devices | Smartphone, personal computer, tablet | Mobile phone data | |
| | Social media | | | Social media sentiment analysis data | Kellogg's tweet shop |
| | Point of sale technologies | Barcode technology | Digital barcode scanner, Smartphone barcode app (mobile point of sale), self-service checkouts, tablets, NFC tags | Consumer grocery shopping data | GfK ConsumerScan "Mini-Danmark, Mobile Point-of-Sale (SCANDIT), NFC tags in Casino supermarkets (France) |
| | | Other point of sale hardware | Payment terminals, weighing sensors, cash registers | Amount owed, weight, money transactions | |
| | | Cloud based Point-of-sale software | Uses data from devices mentioned in barcode technology and other point of sale hardware | | Epos Now, Lightspeed Retail, Revel Systems, Lavu iPad POS |

*(Continued)*

**TABLE 2 |** Continued

| Sector | Type of technology | Data capturing technology | Devices facilitating data capture | Type of data collected | Case studies |
|---|---|---|---|---|---|
| | | Traditional point of sale software | Uses data from devices mentioned in barcode technology and other point of sale hardware (except smartphone barcode scanners) | | AIMsi, AmberPOS, RetailSTAR |
| Market research organization | Automated voice response and voice recognition | Interactive voice response survey | Touchscreen, freephone, post-call transfer to survey line, computer aided telephone interviews, web, email and SMS | Consumer feedback on product purchased and used | Vision OneTotalRecall |
| | Digital observation and video | Digital diary and video recording | Webcam, smartphone, tablets, video camera, or some other type of digital audio/video recording device. | Consumer can either speak into the camera to describe a situation or feeling, or can take us on a tour, so to speak. | Olinger digital video diary |
| | Geo-location | GPS technology | Smart phone using apps with image, video capturing and survey questionnaire and integrated location | Photograph and record in-the-moment data in a specific location. | SSI's mobile QuickThoughts® 2.0 app. Geo-Intercepts app with features such as: GeoValidation, GeoIntensity and GeoNotification®. |
| | Neuromarketing research | Neuromarketing techniques | Smart phone, tablet and laptops using facial recognition and other neuro analytics software | Captures the expressions and emotions people exhibited toward using a product | Face Reader- Noldus IREACT and eye tracking- one vision |

**TABLE 3 |** Potential opportunities and associated limitations for the scientific use of purchase, preparation, and consumption consumer-generated data.

| Domain | Potential opportunities | Limitations |
|---|---|---|
| Purchase | • Inferences about the trends at the population level linked to purchase intention/food spend etc<br>• Trends linked to C2B interactions (which retailers/restaurants/outlets are most visited)<br>• Trends in how preferences in different food groups /products are shifting i.e., atttitudinal changes re purchase intention | • Cannot directly link to an individual purchase<br>• Cannot directly link to the individual's consumption<br>• Cannot identify the unit of analysis (i.e. does the data refer to the individual or household?) |
| Preparation | • People's search behavior online<br>• Trends in recipe generation<br>• Trends in social networking facilitated by food preparation knowledge/recipe sharing etc. | • Links to individual preparation behavior<br>• Cannot directly link to purchase or consumption at an individual level |
| Consumption | • People's individual food intake profiles<br>• Understanding of habitual food consumption behaviors across groups of interest | • Quality/completeness of the underlying food composition databases questionable<br>• Quality and completeness of the self-reports through diet intake/physical activity APPS<br>• Level of detail of the estimated food composition values is low, with APPS typically focusing on energy and macronutrients.<br>• Lack of information regarding the procedures for estimating portion sizes<br>• High prevalence of behavioral change objective which might pose a barrier toward a better understanding of the real determinants of food consumption behaviors as well as the ability to provide an unbiased insight in peoples' habitual food consumption behaviors |

different set of guiding principles in terms of ethics. This would require a fully transparent governance structure where the roles and responsibilities within it are well-defined, and which allows for an on-going, cross-sectoral and cross-disciplinary reflexivity on the role of ML/AI, the nature of the research questions and their potential impact (positive and negative) on the challenges and limits to the accuracy and validity of scientific insights.

## Data Ownership and Intellectual Property

The DI Data Platform operates in a relational context and participates in a *data cycle* composed of three scenarios where data related activities are held: the first scenario implies collecting data, that may come from a third party or be built by the activities held by the organization (surveys, sensors, personal interviews, etc.); a second scenario exists within the DI Data Platform, composed of activities such as transforming, ordering, cataloging, analyzing or deleting data; and a third scenario, where public data dissemination or its private delivery takes place.

The traditional question of the ownership of these data is losing intensity due to the different nature of the digital domain. Hess and Ostrom (30), studied the different needs of scholars when using information and concluded that property, as understood in the material world, has a different significance when applied to the digital world, and concluded that the most relevant activities are in building on a previous knowledge of where to access, extract, manage, exclude and alienate information. There is no need to "own" it in a traditional sense but rather to be able to exercise certain activities on it. Data may be owned by anyone as long as these activities are specifically allowed.

This also seems to be the view of the European Open Science Cloud (EOSC) (31), which focuses on FAIR data, where none of its 15 guiding principles includes the "ownership" of the data only on its reuse. Nevertheless, as pointed out by Labastida and Margoni (32), "Data can be covered by different layers of copyright protection making the relationship between data and copyright particularly complex", a peculiarity that forces the need of all research organizations to avoid risks related to copyright infringement through a proactive attitude toward legal interoperability (33). In addition, the new paradigm of Open Science must be taken into account. The UNESCO's Draft Recommendation on Open Science[13] proposed to its November 2021 General Assembly establishes as one of its key objectives adherence to Open Science thus "maximizing access to scientific knowledge and the reuse and combination of data and software, including source code, and thereby maximizing the common good achieved through public investment in scientific resources and infrastructures".

As reviewed in Section Data type and Technology Relevant to DI Data Platform the data types relevant to the DI Data Platform come from different sources: research data, business data and consumer generated data. Practice shows that each source is associated with different terms and conditions, so it will be necessary to analyse the legal conditions that will be applicable to each dataset, hopefully in a license and not in an agreement. The tendency in ICT is to automatise to the maximum this analysis, designing tools to check interoperability (34) or standardization (35) of the licenses but in the current stage a manual check is needed.

The DI Data Platform will be able to use third parties' intellectual property works and create derivative works over

[13]https://unesdoc.unesco.org/ark:/48223/pf0000374837 (accessed October 13, 2021).

them, if their license allows it, or to use immutable works due to a non-permissive license. At the same time, it will be able to create intellectual property works so, as an author, will be able to decide how to license it if the funding agent does not force specific terms and conditions.

## Privacy Concerns and Security

Adhering to the principles of Open Science poses an important question for researchers using big data of how they can harness the richness of the data whilst meeting the legal (e.g., EU GDPR) standards. The recent draft proposal for Data Governance Act 2020[14] wishes to address the potential contradiction emerging from the clear need for data sharing for public benefit such as scientific advancement (made particularly relevant in the context of COVID-19), and the need to protect the interests of data-subjects (Article 16&17). The proposal aims to facilitate data sharing of digital data across the EU member states through the creation of new infrastructures for data sharing. It explicitly recognizes the value of data sharing for public benefit and anticipates governance measures to make it easier to re-use sensitive public sector data, including clearly specified role for data intermediaries, the role of the European Data Innovation Board with the focus on "altruistic" use of data (Article 27&28). However, whilst the need to enable easier data-sharing for scientific research in public interest is important, in practice, it is not clear what will be the process of transparently deciding on what sort of data analysis can be classified as "for public benefit", or how to ensure appropriate oversight of data provenance. There are many semi-public organizations, private-public partnerships and affiliations, and the division between the two is not always clear in practice. It follows that without significant scrutiny on a case-by-case basis of each of the existing publicly held datasets, the data they hold is not readily useable by researchers.

The requirements imposed by the GDPR may be seen as (legally) onerous by researchers who wish to use business data and/or transfer data for research purposes. This may ultimately limit the value of any commercial data that the DI Data Platform incorporates into the proposed data platform for scientific purposes unless the issues associated with consent are fully addreessed. The proposed solution to achieve the required level of ethical consent for the re-use of consumers' data across all their APPS, is for the DI Data Platform to develop a proprietary APP that could not only act as an aggregator to link with other APPS used by an individual, but also as a means of collecting additional standardized data from a cohort of individuals that are of interest for research purposes. In this way direct consent will be obtained from the consumer for the use of their data either for general research or even for specific purposes and that consent held as meta-data within the DI Data Platform which from a governance perspective is the most desirable scenario. Providing different levels of consent options to research participants would allow them to specify exactly those stakeholder categories they are willing for their data to be shared with, in a time-limited manner. In this way direct consent could be obtained from the consumer for the use of their data either for research, policy development or commercial activities and that consent held as meta-data within the data platform, which, from a governance perspective, is the most desirable scenario.

This will be facilitated and ensured through the transparent processes of organizational governance that will ultimately serve to ensure trust of the consumers whose data is fundamental to the scientific work enabled by the DI Data Platform.

## Organizational and Institutional Governance

Organization and institutional governance of the DI Data Platform has been designed to address the aforementioned challenges with a primary aim to balance the interests of using open data vs. protecting sensitive data. The DI Data Platform is designed as a distributed research infrastructure with an independent legal status as a Foundation under the Dutch law. It will be, in fact, an intermediary for data sharing. It is managed by the Board made up of public and private stakeholder who have responsibility for ethics policy, and to safeguard compliance of all relevant laws and regulations when handling, storing, or processing personally identifiable data resulting from research and from APPS. This is built on the notion that being ethical means recognizing that privacy is "contextual" and "situational", and that decision-making should be on a case by case basis, not reducible to a simple public/private distinction. Three advisory boards are important to the design of an ethical responsible governance: (1) the Scientific Advisory Committee (SAC); (2) Ethical Legal and Societal Issues Board ELSI); and (3) the Industry Forum (IF).

*The SAC* consists of the scientists appointed in their own right, and the Ethical Legal and Societal Issues Board (ELSI) also includes the scientists and representatives of society. They do not represent their own organization or country, and they represent a variety of disciplines relevant to the study of food consumption and ethics within the broader agenda of food systems research. Both SAC and ELSI can fill in their own vacancies and decide about the rules of engagement. An important ethical arrangement is that the DI Data Platform will be periodically evaluated by an independent visiting commission. All scientific research, especially in relation to persons and to health, has to be cleared by SAC and ELSI on the basis of a research plan before it can start. SAC and ELSI report to the highest level in the governance and have the duty to publish decisions it takes. ELSI will also advice on the protocols relating to data security, transfer of data to third countries, assessing the genuineness of a request by data users and the rules of operation in the event of requests that may be ethically dubious or questionable, data subjects' requests, and complaints procedures, and periodically reviewing privacy measures. In addition, ELSI will be making judgments whether any data sharing is permissible based on the assessments of the proposed research as being for public benefit. Crucially, the SAC and ELSI will ensure that a data protection officer is employed on a permanent basis, and, in line with the GDPR and has an oversight of all the necessary procedures, acts as a first point of

[14]https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0767&from=EN (accessed October 13, 2021).

call for all the interested parties, especially the data-subjects with an interest to exercise their data rights.

*Industry Forum* (IF) comprises of the industry, technology developers and other relevant business organizations. The IF will contribute to the decision-making process together with the SAC that will also include algorithmic auditing within the software-development as a means of minimizing biases and errors intrinsic to ML/AI-enabled big data research. The decisions of this forum will be subject to open and transparent rules of decision-making. The Industry Forum will need to strike a balance between many potentially competing and contradictory values such as achieving transparency, whilst protecting privacy or ensuring openness whilst respecting proprietary rights.

The most important aim of the organizational governance is to ensure a fair, transparent and inclusive data governance: the type of data and their provenance; the nature, use and limitations of ML that underpins the data brokerage within the DI Data Platform; the way in which parameters are identified and treated, and the way in which inference about the data will be performed.

## DISCUSSION

In this paper we presented a case study exploring the decision-making in the development of responsible governance of an e-infrastructure—DI Data Platform. As illustrated above, big data is fundamentally altering the nature of research and changing the nature of the relationship between the research participant and the researcher (36). In the traditional research governance process, there is a direct relationship between the two. Research with traditional data sets is characterized by a great deal of certainty of research governance, managed through formal ethical reviews carried by university research ethics committees, on a specific type of data that is deemed to be "personal", and which focuses on the procedures to guarantee protections of human subjects participating in research through tools such informed consent, offer of access and withdrawal of data (unless the right is waved by the data subject) and specified analytical purpose. Under the current legal frameworks [e.g., the EU GDPR (23)] it is necessary to contend with the original purpose and context of data collection and the individual implications of any alteration and re-purposing of the data.

However, as demonstrated in this case study, big data is often not collected for research purposes—for instance, retail data or APPS-generated data (which offer possibilities of new insights in food consumption research) exist to support key business operations. Furthermore, the data thus generated is not stored in discrete locations and does not allow for easy retrieval and removal of the data upon participants' request. Finally, efforts to link and integrate the data will require that scientists contend with the difficult issue of the nature and role of algorithms that enable linkages and inter-operability between diverse data sets. This poses new challenges of the meaning of participants' informed consent for data protection and data sharing under this new regime, the nature of inferences from the new approaches to data analysis and how to assess the veracity of the findings, how best to protect data-subjects' rights (e.g., rights to remove their data), and how to monitor the re-purposing and re-use of the data.

For instance, the narrow focus of the GDPR upon the risks to individual data privacy through its emphasis upon the protection of personal and special category of data and the need for de-identification of an individual does not recognize that big data research affects not only rights of individuals, but also those of groups. Identifiability is not a binary issue and the disclosure risks increase with number of data points for an individual case. A person may be anonymous yet identifiable in terms of the kind of person that he or she represents based on their behaviors, inferred attitudes, and purported identities. Identifying a person as representative of a group that can be characterized, described and located, can also result in harm – both at the personal level and at the societal level. Thus, privacy concerns are not only linked to personal data. "**Group privacy**" (18) has emerged as an important consideration specific to big data. Information about an individual, even if stripped of person-identifying information, which nevertheless links them with a group, can induce harm and pose complex ethical challenges. As highlighted in the current case study, responsible governance of big data broadens its enquiry to go beyond the narrow consideration of privacy, consent and anonymity. Instead, the meaning of data and the nature of harm it can pose must also be inferred from the way in which data is collected, the way in which it is connected to other data sets, processed, managed and controlled.

The nature of big data collection and processing fundamentally departs from the traditional scientific research. Traditional research methods in science are built on hypotheses-testing or observations based on *apriori* set of assumptions and reasonably articulated set of goals. Its primary aim is uncovering causal relationships. Big data analytics, on the other hand, uncovers patterns and correlations observed across large swathes of data and often requires that the data originally collected for a specific purpose be re-purposed and linked to other data sets that may alter its meaning encapsulated by the original aims of data collection (18). For instance, the data on people's consumption practices may not only serve the purpose of assessing consumption trends but be linked to the person's health data to provide estimates of cancer risks. This has implications for the way in which informed consent is sought, but ultimately, it removes the **power and control** over the uses of data from those who have donated it.

The processing that occurs through big data analysis is different from that of the traditional research. It is based on identifying patterns of large-scale data sets and through numerous data points it can identify patterns in the data that are not part of the initial research and scientific enquiry (18). Big data enables finding random commonalities based on incidental co-occurrence, which raises the question of **veracity of findings** because the larger the data set, the more connections can be identified through random process (18). Indeed, big data science may inadvertently create a version of the world that has little bearing on reality and provide interpretations and shape policies in a way that creates or perpetuates biases and injustices.

There is however another important ethical issue associated with the randomness of big data analytics, and it pertains to the consequences of the knowledge thus generated. One of the core issues, which will be of relevance to bio-medical and public health scientific domains, is the consequence of incidental findings, the duty of care and safeguarding. For instance, big data science can inadvertently identify certain groups which may be at more immediate risk of developing a disease (for instance, at risk to develop cancer). Whilst the ethical role of the scientist in traditional research has been defined in terms of the specific aims of the research study that made it possible to anticipate and mitigate the risks, this role is hugely complicated in the context of big data science. The possibility of **incidental findings and randomly generated knowledge** can have dual consequence: either removing scientists from any responsibility toward their research participants or broadening out the frame within which the responsibility for research is interpreted and allocated beyond the traditional scope of scientific work. The issue of what constitutes **responsible big data analytics** is a moot point, and one that requires anticipation, reflexivity and contextual decision-making, as well as greater engagement with legal frameworks other than those typically guiding the scientific process.

This brings us to the final point about the requirements for data management that must take into consideration **diverse jurisdictions**, national and regional level policies and economic realities that can be relied upon to enable smooth research governance. Big data and Open Science (37) presuppose free and interoperable/standardized flow of data across different data platforms, which may be located within diverse jurisdictions. The legal constraints of data transfers and processing in these different contexts are widely acknowledged and to an extent dealt with by the existing legal frameworks [e.g., GDPR, Chapter 5 (23)]. However, what is less often discussed is how this translates into the actual research practice in which diverse cultural norms and moral codes are in place. Whilst the Western cultural and philosophical tradition is based on the concept of autonomy, agency, self-determination and individualism, and encapsulated in the concept of human rights and respect for privacy (38), this is not necessarily the case across other cultures and philosophical traditions [e.g., Eastern European (39)]. The consequence of this may be that scientists' and data subjects' interpretation of consent and harm may vary across different cultures and therefore, that implementation of privacy laws, even if ostensibly harmonized across jurisdictions, may be variable. This poses a fundamental question of how these divergencies should be addressed within not only legal, but also ethical codes of research. Is it enough for a scientist in question to accept different readings of consent and diverse implications of autonomy, thus effectively hiding behind the (often inadequate) legal provisions in different jurisdictions? Or does a scientist have a **moral responsibility to reflect on the conditions of use of the data collected in diverse jurisdictions**, and impose own ethical standards of data use? This question has been addressed by Metcalf and Crawford (24) in their analysis of the US treatment of Big Data appearing in public domain with a problematic premise that publicly available data such as Twitter or Facebook feeds poses minimal risk to human subjects. The authors demonstrate the flaw in this position adopted by the regulator and enacted by the scientists in the US. In line with the principles of RRI, a more nuanced deliberation about the way in which the data in public domain is re-purposed for science is crucial in order to achieve responsible governance of e-infrastructures.

## CONCLUSIONS

Numerous ethical challenges are posed to scientists in the context of research e-infrastructures that connect big data across diverse regulatory, societal and epistemic regimes. The issues of privacy – both individual and group, the relationship between the researcher and the data subject, power and control in the process of research, veracity of findings and the changing nature of scientist responsibility for the generation of knowledge, represent key ethical issues that must be reflected in the context of responsible e-infrastructure governance. Addressing these requires a flexible, adaptable, responsive and transparent organizational and data governance process, open to scrutiny and in step with the technological, business and socio-political changes. RRI will form a basis of the governance of this e-infrastructure as a process guided by the principles of anticipation, reflexivity, engagement and responsiveness. Such an approach will ensure an optimal framework for responsible governance of the complex, uncertain and impactful technologies with far-reaching consequences that are likely to revolutionize the domain of food, nutrition and health (10).

## DATA AVAILABILITY STATEMENT

The authors of this article conducted this research whilst working on the Richfields Project, which was used as the case study described herein. The data supporting this study should be requested from the RICHFIELDS consortium (https://fnhri.eu/contact/) and will be granted upon reasonable request.

## AUTHOR CONTRIBUTIONS

LT contributed to the ethical design of DI data platform, synthesized the results in the article, and wrote the article. IC contributed to the data privacy and security section and reviewed the article. JD contributed to the data ownership and intellectual property section and reviewed the article. TE contributed to the data type section. CH, BK, and BM contributed to the data type section and reviewed the article. TS contributed to the organizational and institutional governance section and reviewed the article. PV was the scientific lead of the DI data platform design process and reviewed the article. KZ was the coordinator of the overall DI data platform design process. All authors contributed to the article and approved the submitted version.

## FUNDING

The RICHFIELDS project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 654280 (www.richfields.eu).

## REFERENCES

1. European Strategy Forum on Research Infrastructures ESFRI: Strategy Report on Research Infrastructures Roadmap 2010. (2011). Available online at: http://ec.europa.eu/research/infrastructures/pdf/esfri-strategy_report_and_roadmap.pdf (accessed February 18, 2022).
2. Available online at: https://sciencebusiness.net/news/call-additional-eu4b-investment-european-research-infrastructures#:~:text=The%20budget%20for%20research%20infrastructures,spent%20in%202021%20and%202022 (accessed January 18, 2022).
3. Brown KA, Timotijević L, Geurts M, Arentoft JL, Dhonukshe-Rutten RA, Fezeu L, et al. Concepts and procedures for mapping food and health research infrastructure: new insights from the EuroDISH project. *Trends Food Sci Technol.* (2017) 63:113–31. doi: 10.1016/j.tifs.2017.03.006
4. European Commission (2018). Recipe for change: An agenda for a climate-smart and sustainable food system for a healthy Europe. EC Food 2030 Expert Group. Available online at: https://ec.europa.eu/research/bioeconomy/pdf/publications/ES_recipe_for_change.pdf
5. European Commission (2011). Sustainable food consumption and production in a resource-constrained world - The 3rd SCAR Foresight Exercise Luxembourg: Publications Office of the European Union.
6. De Schutter O. The political economy of food systems reform. *Eur Rev Agric Econ.* (2017) 44:705–31. doi: 10.1093/erae/jbx009
7. Fresco LO, Poppe KJ. *Towards a Common Agricultural and Food Policy*. Wageningen University & Research. (2016). doi: 10.18174/390280
8. Medical Research Council. *Review of Nutrition and Human Health Research*. July 580 (2017).
9. Snoek HM, Eijssen LM, Geurts M, Vors C, Brown KA, Bogaardt MJ, et al. Advancing food, nutrition, and health research in Europe by connecting and building research infrastructures in a DISH-RI: Results of the EuroDISH project. *Trends Food Sci Technol*. (2018) 73:58–66. doi: 10.1016/j.tifs.2017.12.015
10. Stahl BC. Responsible research and innovation: The role of privacy in an emerging framework. *Sci Public Policy.* (2013) 40:708–16. doi: 10.1093/scipol/sct067
11. Fothergill B, Knight W, Stahl BC, Ulnicane I. Responsible data governance of neuroscience big data. *Front Neuroinform.* (2019) 13:28. doi: 10.3389/fninf.2019.00028
12. Crawford K, Gray ML, Miltner K. Big Data| critiquing Big Data: politics, ethics, epistemology| special section introduction. *Int J Commun.* (2014) 8:10.
13. Ekbia H, Mattioli M, Kouper I, Arave G, Ghazinejad A, Bowman T, et al. Big data, bigger dilemmas: A critical review. *J Assoc Inf Sci Technol.* (2015) 66:1523–45. doi: 10.1002/asi.23294
14. Fox S, Do T. Getting real about big data: applying critical realism to analyse Big Data hype. *Int J Manag Project Business*. (2013) 6:739–60. doi: 10.1108/IJMPB-08-2012-0049
15. Kitchin R. Big data and human geography: opportunities, challenges and risks. *Dial Human Geogr.* (2013) 3:262–7. doi: 10.1177/2043820613513388
16. Wang J, Crawl D, Purawat S, Nguyen M, Altintas I. Big data provenance: challenges, state of the art and opportunities. In: *2015 IEEE international conference on big data (Big Data).* (2015). p. 2509–16. IEEE. doi: 10.1109/BigData.2015.7364047
17. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. In: *The ethics of biomedical big data.* (2016). p. 445–80. doi: 10.1007/978-3-319-33525-4_19
18. Zwitter A. Big data ethics. *Big Data Soc.* (2014) 1:2053951714559253. doi: 10.1177/2053951714559253
19. Salerno J, Knoppers BM, Lee LM, Hlaing WM, Goodman KW. Ethics, big data and computing in epidemiology and public health. *Ann Epidemiol.* (2017) 27:297–301. doi: 10.1016/j.annepidem.2017.05.002
20. Lepri B, Staiano J, Sangokoya D, Letouzé E, Oliver N. The tyranny of data? the bright and dark sides of data-driven decision-making for social good. In: *Transparent data mining for big and small data*. Springer, Cham. (2017). p. 3–24. doi: 10.1007/978-3-319-54024-5_1
21. Mills S, Lucas S, Irakliotis L, Rappa M, Carlson T, Perlowitz B. *Demystifying big data: a practical guide to transforming the business of government.* Tech America Foundation, Washington. (2012).
22. Löfmarck E, Uggla Y, Lidskog R. Freedom with what? Interpretations of "responsibility" in Swedish forestry practice. *Forest Policy Econ*. (2017) 75:34–40. doi: 10.1016/j.forpol.2016.12.004
23. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). Available online at: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng (accessed October 13, 2021).
24. Bogaardt MJ, Copani G, de la Cueva J, Finglas P, Geelen MM, Hodgkins C, et al. Position and final paper of RICHFIELDS: deliverable D1. 2. In: *EU.* (2018).
25. Bogaardt MJ, Geelen A, Zimmermann K, Finglas P, Raats MM, Mikkelsen BE, et al. Designing a research infrastructure on dietary intake and its determinants. *Nutr Bul.* (2018) 43:301–9. doi: 10.1111/nbu.12342
26. Maringer M, van't Veer P, Klepacz N, Verain MC, Normann A, Ekman S, et al. User-documented food consumption data from publicly available apps: an analysis of opportunities and challenges for nutrition research. *Nutr J.* (2018) 17:1–3. doi: 10.1186/s12937-018-0366-6
27. Seljak BK, Popp K, Finglas P, Timotijević L, van't Vee P, Zimmerman K. Infrastructure for innovative research on healthy food choice, preparation and consumption: a position paper on the RICHFIELDS project. In: *2019 IEEE International Conference on Big Data (Big Data).* (2019). p. 5183–5. IEEE. doi: 10.1109/BigData47090.2019.9006393
28. Timotijevic L, Astley S, Bogaardt MJ, Bucher T, Carr I, Copani G, et al. Designing a research infrastructure (RI) on food behaviour and health: Balancing user needs, business model, governance mechanisms and technology. *Trends Food Sci Technol*. (2021) 116:405–14. doi: 10.1016/j.tifs.2021.07.022
29. Van Erp M, Reynolds C, Maynard D, Starke A, Martín RI, Andres F, et al. Using natural language processing and artificial intelligence to explore the nutrition and sustainability of recipes and food. *Front Artif Intel*. (2020) 3. doi: 10.3389/frai.2020.621577
30. Hess C, Ostrom E. Artifacts, facilities, and content: information as a common-pool resource. In: *Conference on the Public Domain, Duke University Law School, Durham*, *NC*. (2001). p. 44–79.
31. Hong NC, Cozzino S, Genova F, Hoffman-Sommer M, Hooft R, Lembinen L, et al. Six recommendations for implementation of FAIR practice. (2020).
32. Labastida I, Margoni T. Licensing FAIR data for reuse. *Data Intel*. (2020) 2:199–207. doi: 10.1162/dint_a_00042
33. Graber-Soudry O, Minssen T, Nilsson D, Corrales M, Wested J, Illien B. *Legal Interoperability and the Fair Data Principles*. (2020). Available online at: https://zenodo.org/record/4471311 (accessed February 18, 2022).
34. Joinup Licence Assistant - Find and compare software licenses. Available online at: https://joinup.ec.europa.eu/collection/eupl/solution/joinup-licensing-assistant/jla-find-and-compare-software-licenses (accessed October 13, 2021).
35. See the technology used by the Linux trunk to standardize the references of the different licences that are used in the development of the Operative System: "The Software Package Data Exchange". Available online at: https://spdx.dev/ (accessed October 13, 2021).
36. Metcalf J, Crawford K. Where are human subjects in big data research? The emerging ethics divide. *Big Data Soc.* (2016) 3:2053951716650211. doi: 10.1177/2053951716650211
37. Burgelman JC, Pascu C, Szkuta K, Von Schomberg R, Karalopoulos A, Repanas K, et al. Open science, open data, and open scholarship: European policies to make science fit for the twenty-first century. *Front Big Data.* (2019) 2:43. doi: 10.3389/fdata.2019.00043
38. Bennett CJ, Raab CD. *The Governance of Privacy: Policy Instruments in Global Perspective.* Routledge. (2017). doi: 10.4324/9781315199269

39. Pew Research Center, Oct. 29, 2018. *Eastern and western europeans differ on importance of religion, views of minorities, and key social issues*. (accessed October 13, 2021).

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.