medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Applying the Nova food classification to food product databases using discriminative ingredients: a methodological proposal

Mariana Fagundes Grilo[1,2*], Beatriz Silva Nunes[1,3*], Ana Clara Duran[1,3,4], Camila Zancheta Ricardo[5], Larissa Galastri Baraldi[1,4], Euridice Martinez Steele[4], Camila Aparecida Borges[1,4].

**Mariana Fagundes Grilo and Beatriz Silva Nunes serve as joint 1st authors.*

1. Center for Food Studies and Research, University of Campinas, Campinas, Brazil.
2. Department of Exercise and Nutrition Sciences, The George Washington University, Washington D.C., US.
3. Graduate Program in Collective Health, Faculty of Medical Sciences, University of Campinas, Campinas, Brazil.
4. Center for Epidemiological Studies in Nutrition and Health, University of São Paulo, São Paulo, Brazil.
5. Doctoral Program in Public Health, School of Public Health, University of Chile, Santiago, Chile.

Abstract

Background: Growing interest in the Nova food classification system surged among various stakeholders, driven primarily by compelling evidence linking the consumption of ultra-processed foods (UPF) to negative health outcomes. However, the conventional classification process often leaves room for uncertainty and operational challenges. **Objective:** This study aimed to develop and test a replicable method to identify UPF based on discriminative food composition ingredients using the 2017 Brazilian Food Labels Database and evaluate the sensitivity of this method in comparison with the identification of UPF by food name and food category. **Methods:** We created six scenarios to identify UPF using food additives and food substances used in the definition of UPF and compare them with the classic methodology of Nova classification based on product name and food categories. We estimated the proportion of foods and beverages identified as UPF according to the different scenarios based on the presence of these discriminative ingredients, total and per food category. Using a diagnostic test and a

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

receiver operating characteristic (ROC) curve, we compared the UPF identified through each of the six scenarios with the ones identified through the classic method. **Results:** We found variations in UPF prevalence from 55% to 72% across scenarios, compared to 70% using the classic method in Brazilian packaged foods. Despite its cautious approach, the scenario using food additives with exclusively cosmetic functions and food substances effectively identified a significant portion of UPF, while maintaining satisfactory sensitivity and specificity, and a better performance on the ROC curve. **Conclusion**: This methodological study emphasizes the importance of detailed criteria to identify UPF, offering researchers alternative and standardized methods for safe decision-making.

Keywords: Ultra-processed foods, methodology, food ingredients, classification, food additives, food substances.

Introduction

The Nova food classification system has, over the past two decades, shifted the single focus of nutrition and food consumption assessment research from nutrient composition to food product level extent and purpose of industrial processing of foods and beverages (1). This system classifies foods into 4 major groups: unprocessed and minimally processed foods, processed culinary ingredients, processed foods, and ultra-processed foods (UPF). UPF are industrial formulations made from substances extracted from food or synthesized from other organic sources, with little or no whole foods, and additives with the purpose of modifying the sensory characteristics of foods. They are usually ready to eat or to heat or reconstituted and require little or no culinary preparation (2).

Since it was developed, the Nova food classification has been proved efficient in

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

predicting nutritional quality of diets and in identifying the increasing consumption of UPF overtime in diverse populations around the world (3-7). The Nova food classification has been applied to observational studies (8, 9), cohort studies (10-12), and randomized trials (13) to assess the link between UPF consumption and health outcomes. Today, there is plenty of evidence on the association of UPF with health outcomes such as weight gain (14), type 2 diabetes (15), cardiometabolic diseases (16), cerebrovascular disease (16), cancer (17), premature deaths (18), all-cause mortality (16), among others (9, 19) including a recent umbrella reviewing showing association with more than 30 health parameters (20).

Many governments are planning policies to discourage consumption of unhealthy food such as UPF and creating policies to regulate their distribution and purchase, such as mandatory front-of-pack nutrition labeling, regulating the advertising of foods high in sugar, sodium and fat, and taxing sugary drinks (21). To implement these regulatory measures, UPF need to be clearly identified. While some strategies (to identify UPF) based on food name, food category or food description (22, 23) have been suggested for food consumption studies when food composition information is lacking, a more standardized approach based on food composition information is needed for regulatory purposes, especially those processes involving food purchases in institutional spaces such as schools, hospitals and other public spaces where it is necessary to specify in detail the food that must be purchased according to current legislation.

The latest version of the Nova food classification published in 2019 mentions that a practical way to identify UPF is checking if the list of ingredients contains at least one item characteristic of the UPF group, that include “food substances never or rarely used in kitchens (such as high-fructose corn syrup, hydrogenated or interesterified oils, and hydrolyzed proteins), or classes of additives designed to make the final product palatable or more appealing (such as

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

flavors, flavor enhancers, food colorings, emulsifiers, emulsifying salts, sweeteners, thickeners, and anti-foaming, bulking, carbonating, foaming, gelling and glazing agents)" (2). Previous research has identified UPF in the food supply by considering contributors' elements, such as the presence of cosmetic additives and the Pan American Health Organization nutrient profile model. The results suggest that integrating both elements can help in the identification of UPF, advancing a step-in testing sensitivity to identifying UPF through these markers (24). Therefore, this study aimed to develop and test a replicable method to identify UPF based on discriminative food composition ingredients using the 2017 Brazilian Food Labels Database and test the sensitivity of this method in comparison with the identification of UPF by food name and food category.

Methods

i. Brazilian Food and Beverage Labels Database

In this cross-sectional study, we used data from the 2017 Brazilian Food Labels Database that includes a large sample of foods and beverages sold by the 5 major food retailers in Brazil (25). To identify the 5 major food retailers, commercial data of annual food retail sales was used (26). São Paulo, the largest city in Brazil and located in the southeast region of the country, was chosen as the primary study area. However, because 1 of the top 5 retailers was only present in the northeast of the country, data collection was also held in Salvador, their largest market.

We used companies' websites and/or customer service sites to get information on the locations of all the stores of the selected retailers and then we geo-coded the address and, using a 1-km buffer, Euclidean distance, we determined the neighborhood of each store. We used information on the average income of the heads of household from the 2010 Brazilian Population

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Census (27) in each buffer to determine the neighborhood´s income. Neighborhoods were then classified as low, middle, and high income using terciles of the mean neighborhood incomes. Finally, we selected the largest store in square meters of each of the top 5 retailers in the country in the bottom and top terciles of neighborhood income. One of the top retailers' chains, however, only allowed data collection in its distribution center.

Trained fieldworkers collected data of packaged foods and beverages, except alcoholic beverages, nutrition supplements, and infant formulas or breast milk substitutes, using the methods proposed by Kanter et al (28) between April and July 2017. The fieldworkers photographed all sides of the packages, yielding a total of around 13,000 items.

Then, trained nutritionists entered food composition information available in the nutrition facts panel, package size, list of ingredients, and instructions for reconstitution in "Research Electronic Data Capture" (Redcap) using a previously tested template (29). Items available in more than 1 package size, products without the nutrition facts panel, multipacks with different items, and products with missing values for portion sizes and/or calories were excluded, resulting in a sample of 11,434 foods and beverages.

In this study, because we wanted to classify foods according to Nova based on the list of ingredients, we excluded foods and beverages without list of ingredients on the package (1,574), totaling 9,860 products assessed.

ii. Identification of UPF using the list of ingredients

According to the Nova classification, the UPF are industrial formulations, usually made from components of food, that are modified and recombined, with little or no whole foods, and industrial substances and additives with cosmetic function that modify or enhance the sensory

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

characteristics. They differ from the other Nova food groups not only in terms of ingredient composition, but of the purpose of processing. UPF are developed to be palatable or hyper-palatable and to be convenient by replacing meals cooked from scratch, meaning most are ready to eat or drink. UPF are highly profitable by means of using cheap raw ingredients and being designed to have a strong market appeal. UPF include sodas, artificial juices, packaged cookies, ready-to-eat breakfast cereals, sausages and all ready-to-eat dishes that mimic meals, such as frozen lasagna and nuggets (2).

The Nova classification describes that UPF can be identified by the presence of food substances never or rarely used in kitchens (hereafter referred to as 'food substances'), such as high-fructose corn syrup, hydrogenated or interesterified oils, and hydrolysed proteins, and/or by the presence of food additives with the aim of modifying sensory characteristics of the products (hereafter referred to as 'food additives with cosmetic function'), like "colorants, flavors, artificial sweeteners, emulsifiers, and many other additives" (1).

We created six scenarios to observe to what degree it is possible to identify UPF using the presence/absence of food additives with cosmetic function and food substances, and we also identified UPF through the classic method used when list of ingredients is unavailable (based on product name and food category). To identify and classify food additives we used the references provided by the Codex Alimentarius 2021 and Monteiro, et al., 2019. (2, 30).

Because food additives can have more than one function, we identified and distinguished food additives with an exclusive cosmetic function from food additives with a potential cosmetic function and tested both scenarios to help in the identification of UPF. In addition to their role in preservation, vitamins and minerals can also play a role in fortification and finally, as food additives with cosmetic function, for changing or enhancing the color of foods. In the latter

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

instance, they may be indicative of ultra-processing (31). Due to their multifaceted functions, we have developed scenarios that either incorporate or exclude vitamins and minerals for their potential cosmetic benefits, to examine whether this type of additive could influence the identification of UPF.

In agreement with the Nova classification, we considered as food additives with cosmetic function: flavors, flavor enhancers, food colorings, emulsifiers, emulsifying salts, artificial sweeteners, thickeners, and anti-foaming, bulking, carbonating, foaming, gelling and glazing agents (2). Besides searching in the list of ingredients for the food additives with cosmetic function provided by the Codex Alimentarius, we searched for non-technical terms corresponding to flavorings because in the Brazilian legislation (RDC nº 259) the use of technical terms for the flavoring additives is not mandatory in the list of ingredients (32). For food substances, we included varieties of derivatives of added sugar, derivatives of carbohydrates, modified oils, derivatives of protein and fiber sources based on Zancheta et al., 2023 and described in Supplementary 1 (2). Below is a description of the six scenarios as well as the method based on food names and categories (or classic method):

Scenario 1: Presence of at least one food additive with an exclusive cosmetic function (not including any vitamins and minerals) (Supplementary 1).

Scenario 2: Presence of at least one food additive with an exclusive cosmetic function (including vitamins and minerals with a potential cosmetic function) (Supplementary 1 and 2).

Scenario 3: Presence of at least one food additive with an exclusive cosmetic function (not including any vitamins and minerals), *i.e. scenario 1* OR at least one food substance used in the definition of UPF (Supplementary 1 and 3).

Scenario 4: Presence of at least one food additive with an exclusive cosmetic function

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

(including vitamins and minerals with a potential cosmetic function), *i.e. scenario 2* OR at least one food substance used in the definition of UPF (Supplementary 1, 2 and 3).

Scenario 5: Presence of at least one food additive with a potential cosmetic function (including vitamins and minerals with a potential cosmetic function) (Supplementary 2 and 4).

Scenario 6: Presence of at least one food additive with a potential cosmetic function (including vitamins and minerals with potential cosmetic function), *i.e., scenario 5* OR at least one food substance used in the definition of UPF (Supplementary 2, 3 and 4).

Classic method: Considering the product name and food categories.

iii. Food categories

In order to facilitate the analytical process, the foods and beverages available in the Brazilian Food and Beverage Labels Database were categorized into the 25 following food categories commonly used in other studies on UPF and NOVA food classification (25, 33): Breakfast cereals and granola bars; Bakery products; Convenience foods; Unsweetened dairy products; Sweetened dairy products; Salty snacks; Cookies; Canned vegetables; Oils and fats; Sauces and dressings; Coffee and tea; Candies and desserts; Cereals, beans, other grain products; Packaged fruits and vegetables; Meat, poultry, seafood, and egg; Sugar and other noncaloric sweeteners; Processed meats; Juices; Nectars; Fruit-flavored drinks; Sodas; Other beverages; Nuts and seeds; Cheeses; Fruit preserves.

iv. Statistical Analysis

We described the prevalence of UPF using the six scenarios and the classic method, overall and by food category (31). Subsequently, we conducted diagnostic tests to assess the sensitivity, specificity, positive predictive value, and negative predictive value for each scenario in comparison with the classic method. In addition, we developed the receiver operating

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

characteristic (ROC) curve to evaluate the performance of the scenarios. The ROC curve is a graphical representation of a binary classifier's performance, plotted by sensitivity (true positive rate) against 1 – specificity (false positive rate). Its effectiveness is primarily gauged by the Area Under the Curve (AUC), with values closer to 1 indicating better performance. A superior model's ROC curve approaches the top-left corner, reflecting a high sensitivity without a significant increase in false positives. The curve's initial steepness and its concave shape towards the top-left are also signs of a robust classifier, indicating an effective balance between sensitivity and specificity. In essence, the closer and more bowed towards the top-left the curve is, the better the model is at distinguishing between the two classes.

Analyses were performed with Stata/MP 16.1, College Station, TX: StataCorp LLC.

Results

The prevalence of UPF across the six scenarios varied between 55.3% in scenario 1 to 72.5% in scenario 6, compared to 70.5% of UPF using the classic method (Table 1). The greater use of discriminatory ingredients, such as the combination of food additives and food substances, reinforces how the layers of details added to the process can enhance the identification of UPF. In Table 1,it is noted that the majority of UPF was captured with just one criterion – presence of at least one food additive with an exclusive cosmetic function in scenario 1 - and adding other criteria individually, such as the presence of food substances managed to identify more UPF. Scenarios 5 and 6 identified a percentage UPF that was the closest to the percentage obtained through the classic method.

Table 1. Prevalence of ultra-processed food (UPF) by scenarios of UPF identification, overall and by food category. Brazilian Food Labels Database, 2017.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Food category	Total	Scenario 1		Scenario 2		Scenario 3		Scenario 4		Scenario 5		Scenario 6		Classic	
		Ultra-processed foods													
		n	%	n	%	n	%	n	%	n	%	n	%	n	%
Breakfast cereals and granola bars	308	231	75.0	237	76.9	270	87.7	272	88.3	288	93.5	294	95.5	308	100.0
Bakery products	594	305	51.3	306	51.5	487	82.0	487	82.0	451	75.9	509	85.7	594	100.0
Convenience foods	795	453	57.0	460	57.9	553	69.6	557	70.1	520	65.4	591	74.3	795	100.0
Unsweetened dairy products	174	0	0.0	85	48.9	6	3.4	91	52.3	114	65.5	115	66.1	2	1.1
Sweetened dairy products	483	432	89.4	435	90.1	476	98.6	476	98.6	466	96.5	477	98.8	483	100.0
Salty snacks	356	239	67.1	239	67.1	257	72.2	257	72.2	264	74.2	272	76.4	346	97.2
Cookies	747	590	79.0	600	80.3	661	88.5	661	88.5	668	89.4	683	91.4	747	100.0
Canned vegetables	345	27	7.8	27	7.8	31	9.0	31	9.0	46	13.3	50	14.5	0	0.0
Oils and fats	294	68	23.1	70	23.8	73	24.8	75	25.5	100	34.0	103	35.0	46	15.6
Sauces and dressings	791	427	54.0	429	54.2	489	61.8	489	61.8	501	63.3	542	68.5	751	94.9
Coffee and tea	68	30	44.1	30	44.1	30	44.1	30	44.1	30	44.1	30	44.1	5	7.4
Candies and desserts	1,218	999	82.0	1,001	82.2	1,144	93.9	1,144	93.9	1,113	91.4	1,167	95.8	1,210	99.3
Cereals, Beans, other grain products	463	73	15.8	76	16.4	91	19.7	94	20.3	108	23.3	119	25.7	62	13.4
Packaged fruits and vegetables	299	4	1.3	4	1.3	4	1.3	4	1.3	4	1.3	4	1.3	0	0.0
Meat, Poultry, Seafood, and egg	49	0	0.0	2	4.1	0	0.0	2	4.1	3	6.1	3	6.1	0	0.0
Sugar and other noncaloric sweeteners	66	46	69.7	46	69.7	47	71.2	47	71.2	47	71.2	47	71.2	44	66.7
Processed meats	810	532	65.7	567	70.0	574	70.9	590	72.8	581	71.7	595	73.5	617	76.2
Juices	144	34	23.6	34	23.6	55	38.2	55	38.2	36	25.0	55	38.2	46	31.9

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Nectars	160	133	83.1	133	83.1	139	86.9	139	86.9	143	89.4	145	90.6	160	100.0
Fruit-flavored drinks	220	214	97.3	214	97.3	216	98.2	216	98.2	217	98.6	218	99.1	220	100.0
Sodas	106	104	98.1	104	98.1	104	98.1	104	98.1	105	99.1	105	99.1	106	100.0
Other beverages	286	221	77.3	221	77.3	239	83.6	239	83.6	252	88.1	259	90.6	269	94.1
Nuts and seeds	72	8	11.1	8	11.1	12	16.7	12	16.7	10	13.9	12	16.7	8	11.1
Cheeses	607	174	28.7	252	41.5	241	39.7	291	47.9	499	82.2	508	83.7	126	20.8
Fruit preserve	405	104	25.7	105	25.9	231	57.0	232	57.3	217	53.6	248	61.2	1	0.2
Total	9,860	5,448	55.3	5,685	57.7	6,430	65.2	6,595	66.9	6,783	68.8	7,151	72.5	6,946	70.4

*Scenario 1: Presence of at least one food additive with an exclusive cosmetic function (not including vitamins and minerals); Scenario 2: Presence of at least one food additive with an exclusive have cosmetic function (including vitamins and minerals with cosmetic function); Scenario 3: Presence of at least one food additive with an exclusive cosmetic function (not including vitamins and minerals) OR food substance; Scenario 4: Presence of at least one food additive with an exclusive cosmetic function (including vitamins and minerals with cosmetic function) OR food substances; Scenario 5: Presence of at least food additive with potential cosmetic function (including vitamins and minerals with cosmetic function); Scenario 6: Presence of at least one food additive with potential cosmetic function (including vitamins and minerals with cosmetic function) OR food substance.

Within food categories (Table 1), we observe that sweetened dairy products, candies and desserts, packaged fruits and vegetables, sugar, and other non-caloric sweeteners, as well as processed meats, juices, fruit-flavored drinks and sodas, present more similar prevalence of UPF across the different scenarios. However, for some food categories such as convenience foods, unsweetened dairy products, salty snacks, sauces and dressings, and coffee and tea the prevalence of UPF varied more between scenarios (Table 1).

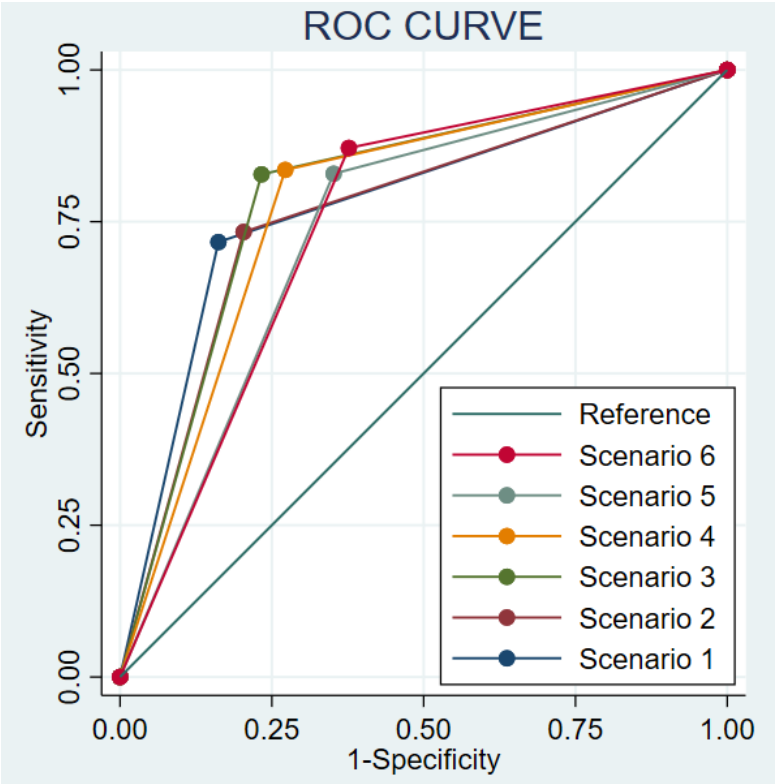
Table 2. Diagnostic tests comparing the ultra-processed foods identified using each of the 6 scenarios with the ones identified through the classic method. Brazilian Food Labels Database, 2017.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Statistical metrics	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Sensitivity	71.6%	73.3%	82.8%	83.5%	82.9%	87.1%
Specificity	83.8%	79.6%	76.7%	72.8%	64.8%	62.3%
Positive predictive value	91.3%	89.6%	89.4%	88.0%	84.9%	84.6%
Negative predictive value	55.3%	55.6%	65.2%	65.0%	61.4%	67.0%

Using a diagnostic test to compare the six scenarios of UPF identification with the classic method, we found an increase in the sensitivity from scenario 1 (71.6%) to scenario 6 (87.1%), and a decrease in specificity (from 83.8% in scenario 1 to 62.3% in scenario 6) (Table 2).

Figure 2. ROC curve for 6 scenarios using the classical method as a classifier.



*Scenario 1: Presence of at least one food additive with an exclusive cosmetic function (not including vitamins and minerals); Scenario 2: Presence of at least one food additive with an exclusive have cosmetic function (including vitamins and minerals with cosmetic function); Scenario 3: Presence of at least one food additive with an exclusive cosmetic function (not including vitamins and minerals) OR food substance; Scenario 4: Presence of at least one food additive with an exclusive cosmetic function (including vitamins and minerals with cosmetic

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

function) OR food substances; Scenario 5: Presence of at least food additive with potential cosmetic function (including vitamins and minerals with cosmetic function); Scenario 6: Presence of at least one food additive with potential cosmetic function (including vitamins and minerals with cosmetic function) OR food substance.; Reference: classic method.

The Figure 2 describes the ROC curve comparing the 6 scenarios. We observe that scenario 3 demonstrates a larger area under the curve (AUC), indicating better performance compared to the classic method that was used as the classifier. It is possible to observe in this figure that considering the presence of vitamins with potential cosmetic function (scenarios 2 and 4) does not significantly alter the AUC as much as considering the presence of food substances or additives with potential cosmetic function.

Discussion

In this methodological study, we found that the choice of food classification method can significantly influence the proportion of UPF identified in a Food Label Database, varying as much as 20%, ranging from 55% in the first scenario to 72% in the sixth scenario. This is in comparison to a 70% prevalence rate identified using the traditional method. Only the least conservative scenario among the six proposed exceeded the UPF prevalence found using the traditional method. Therefore, this study presents viable and still cautious scenarios like the classic method that could assist in future studies using the Nova classification and public policies that need to distinguish UPF from other foods and beverages. Thus, despite its more cautious approach, scenario 3 (that considered the inclusion of food additives with an exclusive cosmetic function, excluding vitamins and minerals with potential cosmetic function, along with food substances), was still effective in capturing a high proportion of UPF while maintaining a satisfactory level of specificity in its classification and presented the best result in the ROC curve when using the classic method as the classifier.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

In view of the urgency of having standard methods to identify UPF, other studies have used the list of ingredients with this purpose, especially to assess UPF consumption (31), and more recently to identify the presence of nutrients of public health concern in the food supply (24). In our case, the reliance on discriminatory ingredients, such as a combination of food additives and substances, introduces a level of detail and complexity in the classification process that can help research for regulatory process. The fact that scenario 3 aligned more closely with the classic method according to the ROC curve, suggests that this scenario may provide more accurate UPF estimates. Scenario 6 also showed alignment, in which it was included food additives with a potential cosmetic function in combination with food substances for UPF identification. Despite industries being able to claim on the packaging that an additive was used for a specific purpose, it may be serving more than one function in the product.

Although the list of ingredients does not provide information about the food matrix, using discriminatory ingredients that correspond to parts of foods and additives with the function of altering the physical characteristics of food allows a systematic identification of UPF (2). It can counteract inconsistencies or difficulties in reaching consensus between evaluators (34), allowing a more robust and functional classification.

The largely preserved UPF prevalence across scenarios by specific food categories proves the consistency of the scenario's identification methods across different types of products. The findings show that certain categories, like sweetened dairy products and candies, exhibit more stable prevalence rates across scenarios, suggesting that these frequently contain at least one food additive of cosmetic function and/or food substance that characterize them as UPF. In fact, another study showed that 100% of sweet cookies, savory biscuits, margarine, cakes, sweet pies, chocolate, dairy beverages and ice cream classified as UPF in the same the Brazilian

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

database of foods and beverages contained at least one additive with a cosmetic function or an excess of certain nutrients of interest to public health, such as sugars, salt, oils and fats (24).

On the contrary, categories like convenience foods and unsweetened dairy products showed higher variability in UPF prevalence between the proposed scenarios in our study, reinforcing the existence of higher variation in the ingredient composition. In the case of coffee and tea, the prevalence of UPF was the same across all 6 scenarios (44.1%), but differed from the classic method (7.4%), suggesting that the classic method may underestimate the prevalence of UPF in this beverage category. For instance, they may be considered fresh or minimally processed by name and food group, but they may contain markers of ultra-processing in the list of ingredients, such as food additives with cosmetic function.

Regarding the trade-off between sensitivity and specificity in the diagnostic test, while an increase in sensitivity from scenario 1 to scenario 6 is promising in accurately identifying true positive cases (decreasing the false negative UPF), there is an expected decrease in specificity, due to increasing false positive UPF. We highlight that even with a more cautious approach for identifying UPF, scenario 3 still demonstrated a sensitivity of 82.8% and a specificity of 76.7%, identifying a significant portion of UPF while maintaining an acceptable level of specificity in its categorization.

The present study has some limitations. Even though there is the function of some additives in the list of ingredients, they may still have non-listed functions in certain foods and beverages (30). To address this, we created different types of scenarios, including scenarios that identify additives that only have a cosmetic function and additives which may also have a non-cosmetic function and also, to enhance reproducibility, we relied on references from international guidelines such as the Codex Alimentarius for food additives (35) and studies conducted in other

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

countries (31). In addition, when applying this methodology to studies in other countries, local legislation regarding the listing of ingredients should be considered when choosing additives to identify UPF, such as flavorings that are not listed in Codex. For example, in Brazil, it is not mandatory to include the composition of compound ingredients if it is less than 25% on the list of ingredients, which may lead to an underestimation of the presence of certain food substance (36). Despite this, it is widely recognized that UPF have a similar ingredient composition worldwide (24).

These findings have implications for policymakers, public health professionals, and researchers. Understanding the impact of different identification scenarios on UPF prevalence and diagnostic test performance allows for more informed decision-making when establishing criteria for UPF classification. It also emphasizes the importance of considering the diverse nature of food categories, as certain products may require tailored criteria for accurate classification according to Nova food system.

Finally, this step-by-step approach can help a greater recognition of the various compounds and types of ingredients used in food manufacturing in different countries around the world. It enables the monitoring of reformulation processes and modifications in the nutritional composition of products, especially after the implementation of labeling policies such as front-of-pack labeling.

Conclusion

In conclusion, this study not only contributes to the understanding of UPF identification but also underscores how the layers of details added to the process can enhance the identification of UPF. In addition, it presents researchers and experts with alternative methods to the classical ones, allowing safe decision-making, especially when it involves regulatory processes and the

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

construction of laws that must clearly define healthy and unhealthy foods based on the Nova classification.

Competing Interest

The authors have no conflict of interest to disclose.

Funding Statement

M.F.G., A.C.D., and C.B. received funding from Bloomberg Philanthropies. B.S.N. was granted a master scholarship from the Brazilian Coordination of Superior Level Staff Improvement. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors Contribution

M.F.G and B.S.N. were responsible for conceptualization, data curation, formal analysis, investigation, methodology and writing – original draft preparation. A.C.D. contributed with conceptualization, methodology, project administration, resources, supervision, and writing – review & editing., C.Z. contributed with conceptualization, data curation, methodology, and writing – review & editing. E.S. and L.G.B. contributed with formal analysis, methodology, and writing – review & editing. C.B. contributed with conceptualization, methodology, supervision, data curation, and writing – review & editing.

References

1. Scrinis G, Monteiro C. From ultra-processed foods to ultra-processed dietary patterns. Nature Food. 2022;3(9):671-3.
2. 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(5):936-41.
3. Baker P, Machado P, Santos T, Sievert K, Backholer K, Hadjikakou M, et al. Ultra-processed foods and the nutrition transition: Global, regional and national trends, food systems transformations and political economy drivers. Obesity Reviews. 2020;21(12):e13126.
4. Juul F, Parekh N, Martinez-Steele E, Monteiro CA, Chang VW. Ultra-processed food consumption among US adults from 2001 to 2018. The American Journal of Clinical Nutrition. 2022;115(1):211-21.
5. Juul F, Hemmingsson E. Trends in consumption of ultra-processed foods and obesity in Sweden between 1960 and 2010. Public Health Nutr. 2015;18(17):3096-107.
6. Louzada MLdC, Ricardo CZ, Steele EM, Levy RB, Cannon G, Monteiro CA. The share of ultra-processed foods determines the overall nutritional quality of diets in Brazil. Public Health Nutrition. 2018;21(1):94-102.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

7. Cediel G, Reyes M, Corvalán C, Levy RB, Uauy R, Monteiro CA. Ultra-processed foods drive to unhealthy diets: evidence from Chile. Public Health Nutrition. 2021;24(7):1698-707.
8. Lane MM, Davis JA, Beattie S, Gómez-Donoso C, Loughman A, O'Neil A, et al. Ultraprocessed food and chronic noncommunicable diseases: A systematic review and meta-analysis of 43 observational studies. Obesity Reviews. 2021;22(3):e13146.
9. Srour B, Kordahi MC, Bonazzi E, Deschasaux-Tanguy M, Touvier M, Chassaing B. Ultra-processed foods and human health: from epidemiological evidence to mechanistic insights. The Lancet Gastroenterology & Hepatology. 2022;7(12):1128-40.
10. Gómez-Donoso C, Sánchez-Villegas A, Martínez-González MA, Gea A, Mendonça RdD, Lahortiga-Ramos F, et al. Ultra-processed food consumption and the incidence of depression in a Mediterranean cohort: the SUN Project. European Journal of Nutrition. 2020;59(3):1093-103.
11. Blanco-Rojo R, Sandoval-Insausti H, López-Garcia E, Graciani A, Ordovás JM, Banegas JR, et al. Consumption of Ultra-Processed Foods and Mortality: A National Prospective Cohort in Spain. Mayo Clinic Proceedings. 2019;94(11):2178-88.
12. Levy RB, Rauber F, Chang K, Louzada MLdC, Monteiro CA, Millett C, et al. Ultra-processed food consumption and type 2 diabetes incidence: A prospective cohort study. Clinical Nutrition. 2021;40(5):3608-14.
13. 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(1):67-77.e3.
14. Askari M, Heshmati J, Shahinfar H, Tripathi N, Daneshzad E. Ultra-processed food and the risk of overweight and obesity: a systematic review and meta-analysis of observational studies. International Journal of Obesity. 2020.
15. Costa de Miranda R, Rauber F, Levy RB. Impact of ultra-processed food consumption on metabolic health. Current Opinion in Lipidology. 2021;32(1).
16. Pagliai G, Dinu M, Madarena MP, Bonaccio M, Iacoviello L, Sofi F. Consumption of ultra-processed foods and health status: a systematic review and meta-analysis. Br J Nutr. 2020:1-11.
17. 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:k322.
18. Nilson EAF, Ferrari G, Louzada MLC, Levy RB, Monteiro CA, Rezende LFM. Premature Deaths Attributable to the Consumption of Ultraprocessed Foods in Brazil. American Journal of Preventive Medicine. 2023;64(1):129-36.
19. Hajmir MM, Shiraseb F, Hosseininasab D, Aali Y, Hosseini S, Mirzaei K. The mediatory role of inflammatory markers on the relationship between the NOVA classification system and obesity phenotypes among obese and overweight adult women: a cross-sectional study. Frontiers in Nutrition. 2023;10.
20. Melissa ML, Elizabeth G, Shutong D, Deborah NA, Amelia JM, Sarah G, et al. Ultra-processed food exposure and adverse health outcomes: umbrella review of epidemiological meta-analyses. BMJ. 2024;384:e077310.
21. Popkin BM, Barquera S, Corvalan C, Hofman KJ, Monteiro C, Ng SW, et al. Towards unified and impactful policies to reduce ultra-processed food consumption and promote healthier eating. Lancet Diabetes Endocrinol. 2021;9(7):462-70.
22. Steele EM, O'Connor LE, Juul F, Khandpur N, Galastri Baraldi L, Monteiro CA, et al. Identifying and Estimating Ultraprocessed Food Intake in the US NHANES According to the NOVA Classification System of Food Processing. The Journal of Nutrition. 2022.
23. Costa CDS, Faria FR, Gabe KT, Sattamini IF, Khandpur N, Leite FHM, et al. Nova score for the consumption of ultra-processed foods: description and performance evaluation in Brazil. Rev Saude Publica. 2021;55:13.
24. Canella DS, Pereira Montera VdS, Oliveira N, Mais LA, Andrade GC, Martins APB. Food additives and PAHO's nutrient profile model as contributors' elements to the identification of ultra-processed food products. Scientific Reports. 2023;13(1):13698.
25. Duran AC, Ricardo CZ, Mais LA, Bortoletto Martins AP. Role of different nutrient profiling models in identifying targeted foods for front-of-package food labelling in Brazil. Public Health Nutr. 2020:1-12.
26. Retail P. *Brazil: Retail Sales 2013–2022* ; Planet Retail: Frankfurt, Germany. 2020.
27. IBGE. Instituto Brasileiro de Geografia e Estatística. Operação censitária. Censo 2010. 2020.
28. Kanter R, Reyes M, Corvalán C. Photographic Methods for Measuring Packaged Food and Beverage Products in Supermarkets. Curr Dev Nutr. 2017;1(10):e001016.
29. REDCap. REDCap Brasil. 2021.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.12.24305721; this version posted April 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

30. Commission JFWCA. Codex General Standard for Food Additives (GSFA) Online Database. 2017.
31. Zancheta Ricardo C, Duran AC, Grilo MF, Rebolledo N, Díaz-Torrente X, Reyes M, et al. Impact of the use of food ingredients and additives on the estimation of ultra-processed foods and beverages. Frontiers in Nutrition. 2023;9.
32. ANVISA. Agência Nacional de Vigilância Sanitária. RDC Nº 259, DE 20 DE SETEMBRO DE 2002. 2002.
33. Fagundes Grilo M, Taillie LS, Zancheta Ricardo C, Amaral Mais L, Bortoletto Martins AP, Duran AC. Prevalence of low-calorie sweeteners and related front-of-package claims in the Brazilian packaged food supply. J Acad Nutr Diet. 2021.
34. Braesco V, Souchon I, Sauvant P, Haurogné T, Maillot M, Féart C, et al. Ultra-processed foods: how functional is the NOVA system? European Journal of Clinical Nutrition. 2022;76(9):1245-53.
35. Codex Alimentarius. Guidelines for the Simple Evaluation of Dietary Exposure to Food Additives. CAC/GL-1989. Adopted 1989. Revision 2014. 2014.
36. ANVISA. RESOLUÇÃO DA DIRETORIA COLEGIADA - RDC Nº 727, DE 1° DE JULHO DE 2022 (Publicada no DOU nº 126, de 6 de julho de 2022). 2022.