



Article

Deconstructing the Supermarket: Systematic Ingredient Disaggregation and the Association between Ingredient Usage and Product Health Indicators for 24,229 Australian Foods and Beverages

Allison Gaines ^{1,*} , Maria Shahid ², Liping Huang ², Tazman Davies ², Fraser Taylor ², Jason HY Wu ²  and Bruce Neal ^{1,2}

¹ Department of Epidemiology and Biostatistics, Faculty of Medicine, School of Public Health, Imperial College London, London SW7 2AZ, UK; bneal@georgeinstitute.org.au

² The George Institute for Global Health, University of New South Wales, Sydney, NSW 2042, Australia; mshahid@georgeinstitute.org.au (M.S.); hliping@georgeinstitute.org.au (L.H.); tdavies@georgeinstitute.org.au (T.D.); ftaylor@georgeinstitute.org.au (F.T.); jwu1@georgeinstitute.org.au (J.H.W.)

* Correspondence: a.gaines20@imperial.ac.uk



Citation: Gaines, A.; Shahid, M.; Huang, L.; Davies, T.; Taylor, F.; Wu, J.H.; Neal, B. Deconstructing the Supermarket: Systematic Ingredient Disaggregation and the Association between Ingredient Usage and Product Health Indicators for 24,229 Australian Foods and Beverages. *Nutrients* **2021**, *13*, 1882. <https://doi.org/10.3390/nu13061882>

Academic Editor: Daniela Martini

Received: 30 April 2021

Accepted: 27 May 2021

Published: 31 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Unhealthy diets are underpinned by the over-consumption of packaged products. Data describing the ingredient composition of these products is limited. We sought to define the ingredients used in Australian packaged foods and beverages and assess associations between the number of ingredients and existing health indicators. Statements of ingredients were disaggregated, creating separate fields for each ingredient and sub-ingredient. Ingredients were categorised and the average number of ingredients per product was calculated. Associations between number of ingredients and both the nutrient-based Health Star Rating (HSR) and the NOVA level-of-processing classification were assessed. A total of 24,229 products, listing 233,113 ingredients, were included. Products had between 1 and 62 ingredients (median (Interquartile range (IQR)): 8 (3–14)). We identified 915 unique ingredients, which we organised into 17 major and 138 minor categories. ‘Additives’ were contained in the largest proportion of products (64.6%, (15,652/24,229)). The median number of ingredients per product was significantly lower in products with the optimum 5-star HSR (when compared to all other HSR score groups, p -value < 0.001) and significantly higher in products classified as ultra-processed (when compared to all other NOVA classification groups, p -value < 0.001). There is a strong relationship between the number of ingredients in a product and indicators of nutritional quality and level of processing.

Keywords: ingredients; food labelling; health star rating; nutrient profiling; ultra-processing; packaged foods

1. Introduction

Much ill health globally is caused by poor dietary choices [1]. Unhealthy diets are, in turn, underpinned by a food system that tends to promote over-consumption of unhealthy packaged foods and beverages [2,3]. Often these packaged products are high in salt, added sugar and harmful fats which are direct causes of obesity, diabetes and other chronic health conditions [4].

It can be difficult for consumers to pick between more and less healthy packaged food and beverage products. To provide consumers with point-of-sale information that identifies better choices, governments and other organisations are introducing front-of-pack labels. For example, the Health Star Rating (HSR) system [5] developed by the Australian government assigns each food between half a star and five stars using a nutrient profiling

algorithm. The NOVA classification system [6] is another method of distinguishing better-for-you options by defining food and beverage products according to the level of processing involved in manufacturing [7]. A number of studies assessing ultra-processed food intake have found an association between greater consumption of ultra-processed food and increased risks of obesity, cardiovascular disease and mortality [8–10].

Direct assessment of ingredients may be another way to evaluate the healthiness of packaged food products. Ingredient labelling standards and food composition tables in Australia are governed by Food Standards Australia New Zealand (FSANZ). Ingredient labelling is designed to provide consumers with transparency into product composition. FSANZ legislation requires a statement of ingredients to be arranged in descending order by ingoing weight, allowing consumers to discern the primary components of a product's recipe [11]. In addition, explicit percentage labelling is required for 'characterising ingredients' (e.g., strawberry in a strawberry yoghurt) [12]. Food composition tables provide generic reference data describing nutrient content in common foods and can help assess dietary intake [13,14], but do not provide specific information about individual packaged food products.

The numbers and types of ingredients in the packaged food and beverage supply may provide a further means of assessing healthiness [15]. Novel, ingredient-centric work may contribute to improving Australian food composition databases as well as enable studies on ingredient interactions in composite foods [16,17], particularly by better defining ingredient data available on packaged products.

In this study, we sought to define the list of ingredients contained in packaged food and beverage products marketed at major grocery retail outlets in Australia in 2019. We categorised the ingredients identified and explored the associations between the number of ingredients used per product and indicators of product healthiness.

2. Materials and Methods

This was a cross-sectional examination of packaged food and beverage products in the Australian food supply carried out using FoodSwitch, a database developed and operated by the George Institute for Global Health.

2.1. Food and Beverage Product Database

The 2019 FoodSwitch Annual Database, the dataset used in this study, was generated from in-store surveys carried out in Australia at five large supermarkets (IGA, ALDI, Woolworths, Coles and Harris Farm) between August and November 2019 [18]. The dataset represents the majority of supermarket food and beverage products purchased by Australian households [19,20].

The FoodSwitch database holds up to 400 data fields for each included product. Data are obtained direct from packaging (e.g., product name, package size, ingredients, nutritional claims and nutrient content) or derived from data on the packaging (e.g., HSR and NOVA classification). Products held in the FoodSwitch database are classified into a hierarchical categorisation evolved from a system initially developed by the Global Food Monitoring Group [19]. The categorisation system was designed to enable the description and tracking of the nutritional composition of products in groupings meaningful to diverse stakeholders including academic researchers, the food industry, government and consumers [21,22]. There are 18 major food and beverage categories but products are further allocated into hundreds of finer subcategories.

Products were excluded from this analysis if they carried multiple nutrition information panels (e.g., variety packs) or were not in a relevant major food category (i.e., 'Alcohol', 'Vitamins and supplements' or 'Unable to be categorised'). Extensive data cleaning was undertaken but some products were also excluded for having an incomplete or non-standard statement of ingredients.

2.2. Identification of Ingredients

Ingredients were identified from the images of the product label held within the FoodSwitch system and were obtained from either the statement of ingredients or the product name.

For products with a statement of ingredients, data were transcribed into the database as delimited text strings reflecting the form of information provided on pack. The text strings were disaggregated using an automated process creating separate fields for each ingredient and extracting compound ingredient information when present (Table A1 in Appendix A).

Compound ingredients are ingredients made up of two or more sub-ingredients [11]. Sub-ingredients require declaration if they constitute more than 5% of the final food or beverage product or contain a known allergen [23]. There were three forms of compound ingredient presentation: (1) those for which the listed sub-ingredients together made up the entire compound ingredient (complete sub-ingredients), for example, where fortified wheat flour sub-ingredients were listed as wheat flour, thiamine and folic acid; (2) those for which the listed sub-ingredients comprised only the additives or fortificants (incomplete sub-ingredients), for example, where fortified wheat flour sub-ingredients were listed as just thiamine and folic acid; and (3) those for which sub-ingredients were not listed.

Extensive cleaning of the disaggregated ingredient data was done using a semi-automated process (Figure A1). Key elements of the process were the removal of symbols (e.g., reduced-fat milk became reduced fat milk), the correction of spelling errors (e.g., watermelon became watermelon) and the alignment of plurality (e.g., raspberries became raspberry). For ingredients with various common names, a manual process to align terminology was undertaken (e.g., garbanzo bean became chickpea). Descriptive terms were removed from ingredient names (e.g., natural freshly squeezed orange juice became orange juice). When a term relating to an agricultural technique was identified (e.g., organic or free-range) it was removed and stored separately. Approximately 20% of ingredient extractions were manually validated.

For products with ingredients in the product name, a manual ingredient identification and extraction process was performed. The FSANZ Ingredient Labelling of Food standard precludes products from requiring a statement of ingredients in three cases: (1) where the product is labelled with the name of the food and the name of the food includes all ingredients, (2) where the product is water in a packaged form and (3) where the product is contained in a small package (surface area less than 100 cm²) [11]. For products without a statement of ingredients but whose ingredient information could potentially be determined from the product name, a text-matching algorithm was used to extract ingredient information. For example, the ingredient 'blueberry' was extracted from the product name 'Fresh Blueberries'. In some cases, multiple ingredients were extracted from the product name, for example, the ingredients 'pork mince' and 'veal mince' from the product 'Pork & Veal Mince'.

2.3. Categorisation of Ingredients

Ingredients were categorised into a single, hierarchical structure with each ingredient assigned a major and minor category. Category naming conventions and the category structure were determined by discussion amongst the authors and guided by the terminology used by two specialty ingredient companies: Cargill and ADM [24,25]. Minor categories for 'Sweeteners' were verified against prior publications on sugar types contained in packaged food products [26]. A 'Compound ingredients' category was created to capture any main ingredients which, by name, were clearly composed of multiple sub-ingredients (e.g., chocolate).

Ingredients were flagged depending upon whether they constituted a substantive part of any product (substantive ingredients) or were always a small component of the products (non-substantive ingredients). Non-substantive ingredients were defined as ingredients that are usually less than 2% of a product formula, such as ingredients with

maximum legal usage rates such as certain preservatives, vitamins and minerals; or ingredients that are always used at low levels to provide functional benefits such as gums and other stabilisers [27]. Colours, flavours and most additives were considered non-substantive ingredients.

E numbers were cleaned and listed under the ingredient's common designation, such as the chemical name or additive type. For example, E100 through E199 were listed under the ingredient name colour. Similarly, rare substantive ingredients were grouped under a more common ingredient name where possible (e.g., za'atar was listed as dried herb).

2.4. Analysis

The total number of ingredients was calculated by tallying the number of ingredients at the most granular disaggregation level after the removal of duplicates (e.g., 'canola oil, butter (cream, water, salt), eggs, white vinegar, salt' contains six ingredients since salt was listed twice). Ingredients that were different prior to cleaning but the same after cleaning (e.g., E101 and E102 both classified as colour) were treated as distinct ingredients. Compound ingredients that contained an incomplete or missing sub-ingredient list were considered singular ingredients.

Descriptive statistics were presented for the average number of ingredients per product overall and for main food categories. We assessed the association between the number of ingredients per product and HSR, as well as NOVA classification. Differences in the number of ingredients across HSR scores and NOVA classifications were evaluated using a Kruskal–Wallis H test with post-hoc Dunn pairwise analyses. A subsidiary analysis was done using ingredient numbers based on main ingredients alone (i.e., excluding sub-ingredients). All data preparation and statistical analyses were conducted using Stata IC 15 (StataCorp) and Excel. A two-sided p -value < 0.05 was considered statistically significant. Ethical approval was not required for this study.

3. Results

A total of 27,366 food products were included in the Australian 2019 FoodSwitch Annual Database. Of these, 655 were in excluded product categories or were multi-pack products with more than one nutrition information panel, and 441 had no available means of identifying an ingredient list. A further 2041 products had a disorganised statement of ingredients format so disaggregation of the ingredients was not possible (Figure A2). This left 24,229 products for analysis. Of these, 22,493 (92.8%) displayed the statement of ingredients on pack and 1736 (7.2%) required extraction of ingredients from the product name. The majority of products requiring ingredient extraction from the product name had one ingredient ($n = 1732$) and the remaining four products had two ingredients.

3.1. Identified Ingredients

The 24,229 products listed a total of 233,113 ingredients, producing a list of 31,934 unique ingredient names. Semi-automated cleaning reduced this number to 915 distinct ingredients (Figure A3). The 915 distinct ingredients were categorised into 17 major categories (Table 1) and 138 minor categories (Tables S1 and S2 in the Supplementary Materials). Of the ingredients, 835 were identified as nutritionally substantive and 80 were identified as nutritionally non-substantive. The majority of unique ingredients belonged to the major category 'Fruit and vegetables' ($n = 229$), followed by 'Compound ingredients' ($n = 128$), 'Additives' ($n = 86$) and 'Liquids' ($n = 72$).

A total of 9283 (38.3%) products contained at least one compound ingredient with sub-ingredients and 4631 (19.1%) products contained at least one compound ingredient without sub-ingredients. In total, 21,420 compound ingredients were listed. Across all compound ingredient types, 10,650 (49.7%) contained a complete sub-ingredient list, 6139 (28.7%) contained an incomplete sub-ingredient list and 4631 (21.6%) did not contain sub-ingredients which could be disaggregated.

Table 1. Ingredient frequencies by major ingredient category.

Major Ingredient Category	No. of Unique Ingredients	Proportion of Unique Ingredients	No. of Non-Substantive Ingredients ³	Frequency across All Products ¹	Proportion across All Products ²
Additives	86	9.4%	67	15,652	64.6%
Herbs and spices	35	3.8%		13,766	56.8%
Sweeteners	30	3.3%	8	12,437	51.3%
Liquids	72	7.9%		11,767	48.6%
Fruits and vegetables	229	25.0%		10,772	44.5%
Oils and fats	33	3.6%		9914	40.9%
Flours and starches	17	1.9%		7245	29.9%
Dairy	56	6.1%	1	7203	29.7%
Compound ingredients	128	14.0%	1	4980	20.6%
Baking ingredients	10	1.1%	1	4977	20.5%
Grains	48	5.2%		4059	16.8%
Nuts and seeds	31	3.4%		2809	11.6%
Meat	60	6.6%		2652	10.9%
Legumes	27	3.0%		2579	10.6%
Eggs	8	0.9%		1699	7.0%
Seafood	41	4.5%		1052	4.3%
Dietary fibres	4	0.4%	2	914	3.8%
	915	100%	80		

¹ Frequency and ² proportion across all ingredient lists describes the number/proportion of the 24,229 foods for which one or more of the ingredients in each major ingredient category was included in the ingredient list on pack. ³ Non-substantive ingredients were defined as ingredients usually less than 2% of a product formula, such as preservatives, vitamins and minerals, gums and other stabilisers.

3.2. Most Frequently Used Ingredients

As seen in Table 1, at the major ingredient category level, ‘Additives’ were contained in the largest proportion of products (64.6%, (15,652/24,229)), followed by ‘Herbs and spices’ (56.8% (13,766/24,229)) and ‘Sweetener’ (51.3% (12,437/24,229)). A total of 44.5% (10,720/24,229) of products had at least one ingredient categorised under ‘Fruits and vegetables’ and only 10.6% (2579/24,229) had at least one ingredient under ‘Legumes’.

Across all included products, salt (11,182/24,229) and sugar (9872/24,229) were the most frequently used ingredients. ‘Additives’ (i.e., flavours, colours, emulsifiers, acidity regulators, preservatives and thickeners) represented 6 of the 10 most frequently used ingredients (Table 2).

Table 2. The top 15 ingredients in the 2019 Australian packaged food and beverage supply.

Ingredient Rank ¹	Ingredient	Ingredient Major Category	Substantive Ingredient	Frequency ²	Proportion ³
1	Salt	Herbs and spices	N	11,182	46.15%
2	Sugar	Sweeteners	N	9872	40.74%
3	Water	Liquids	N	7971	32.90%
4	Flavour	Additives	Y	7533	31.09%
5	Colour	Additives	Y	4392	18.13%
6	Wheat flour	Flours and starches	N	3780	15.60%
7	Emulsifier	Additives	Y	3771	15.56%
8	Acidity regulator	Additives	Y	3124	12.89%
9	Vegetable oil	Oils and fats	N	3102	12.80%
10	Thickener	Additives	Y	3015	12.44%
11	Preservative	Additives	Y	3013	12.44%
12	Yeast	Baking	N	2763	11.40%
13	Milk powder	Dairy	N	2676	11.04%
14	Dried herbs	Herbs and spices	N	2470	10.19%
15	Vitamins	Additives	Y	1369	5.65%

¹ Rank = ingredients are ranked according to how many of the ingredient lists they are included on out of the 24,229 included packaged food and beverage products. ² The frequency is the number of products for which the ingredient is listed. ³ The proportion is the percent of the 24,229 total products evaluated.

3.3. Number of Ingredients Per Product across Product Categories

Products had between 1 and 62 ingredients, though about 90% of products had fewer than 20 ingredients and single-ingredient products made up 18.1% (4381/24,229) of the total (Figure 1). The median (interquartile range (IQR)) number of ingredients across all products was 8 (3–14) (Table 3). ‘Convenience foods’ had the greatest median number of ingredients per product (20 (14–27)), followed by ‘Foods for specific dietary use’ (12 (7–22)) and ‘Bread and bakery products’ (13 (9–19)) (Table 3).

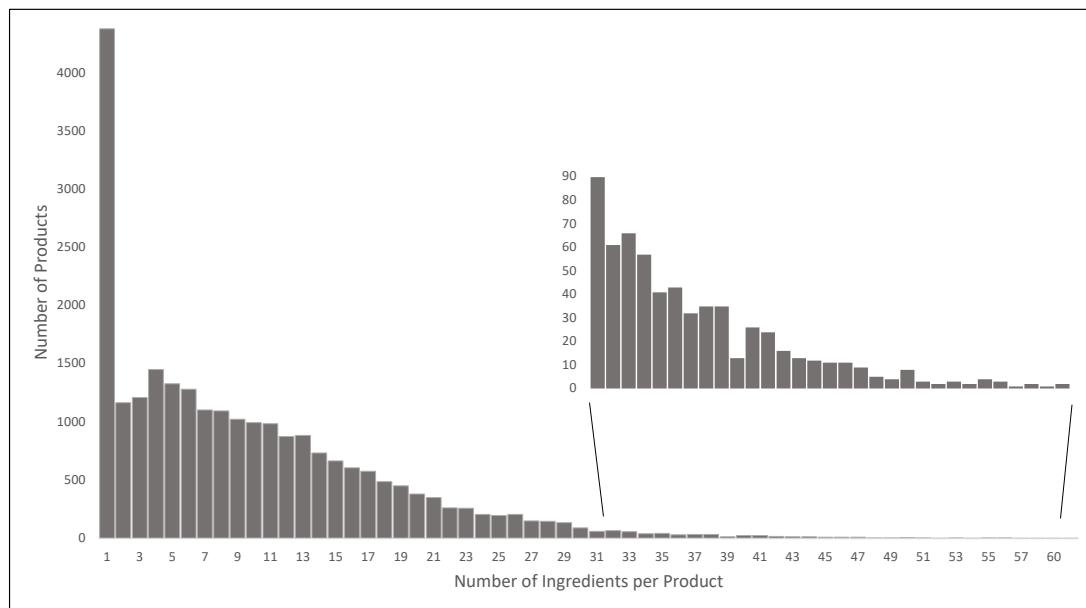


Figure 1. Distribution of the number of ingredients per product. The bar graph displays how many products contain the specified number of disaggregated ingredients in the statement of ingredients. The inset magnified graph displays the number of ingredients per product where product numbers are < 100.

Table 3. The mean number and ranges of ingredients, overall and in major food categories.

	Total Products	Number of Ingredients						
		Mean ¹	SD ²	Min	25%	50%	75%	Max
All	24,229	9.6	8.3	1	3	8	14	62
By food category:								
Convenience foods	1425	21.0	10.1	1	14	20	27	61
Foods for specific dietary use	2476	15.0	10.6	1	7	12	22	50
Bread and bakery products	2393	14.6	7.4	1	9	13	19	49
Meat and meat alternatives	726	11.9	9.6	1	1	11	18	53
Snack foods	1930	11.8	6.5	2	6	12	17	38
Sauces, dressings, spreads and dips	685	11.7	6.5	1	7	12	16	43
Confectionery	1351	10.8	6.0	1	7	9	13	62
Cereal and grain products	2556	9.2	8.7	1	1	6	15	58
Dairy	2990	8.3	6.7	1	4	6	11	56
Seafood and seafood products	1978	7.2	6.1	1	3	5	11	41
Non-alcoholic beverages	768	5.5	4.6	1	1	5	8	37
Fruit, vegetables, nuts and legumes	3923	4.2	4.4	1	1	3	6	50
Sugar, honey and related products	441	4.0	4.1	1	1	2	6	26
Edible oils and oil emulsions	498	3.4	3.9	1	1	1	3	16
Egg and egg products	89	1	0	1	1	1	1	1

¹ Mean = the average number of ingredients in products based on the on-pack statement of ingredients. All main and first-level sub-ingredients are included in the total count of ingredients. ² SD = standard deviation.

Excluding sub-ingredients reduced the total count of ingredients across all products to 165,011. The range in number of ingredients was decreased to 1 to 42, and the median number of main ingredients across all products was decreased to 6 (2–10). The median number of ingredients was reduced across all major food categories and markedly for ‘Convenience foods’ (from 20 to 10).

3.4. Associations of Number of Ingredients with Healthiness and Level of Processing

The number of ingredients per product was lower in healthier products with higher HSR. Products with 4-star, 4.5-star or 5-star ratings had a significantly lower median ingredient count than all other HSR values (all p -values < 0.001). The highest HSR, a 5-star rating, had the lowest median number of ingredients (1 (1–4)). The association was, however, not linear, with numbers of ingredients across HSR values from 0.5 to 3.5 being approximately similar. A progressive decline in numbers of ingredients used per product was observed across HSR values from 3.5 to 5.0 (Figure 2A).

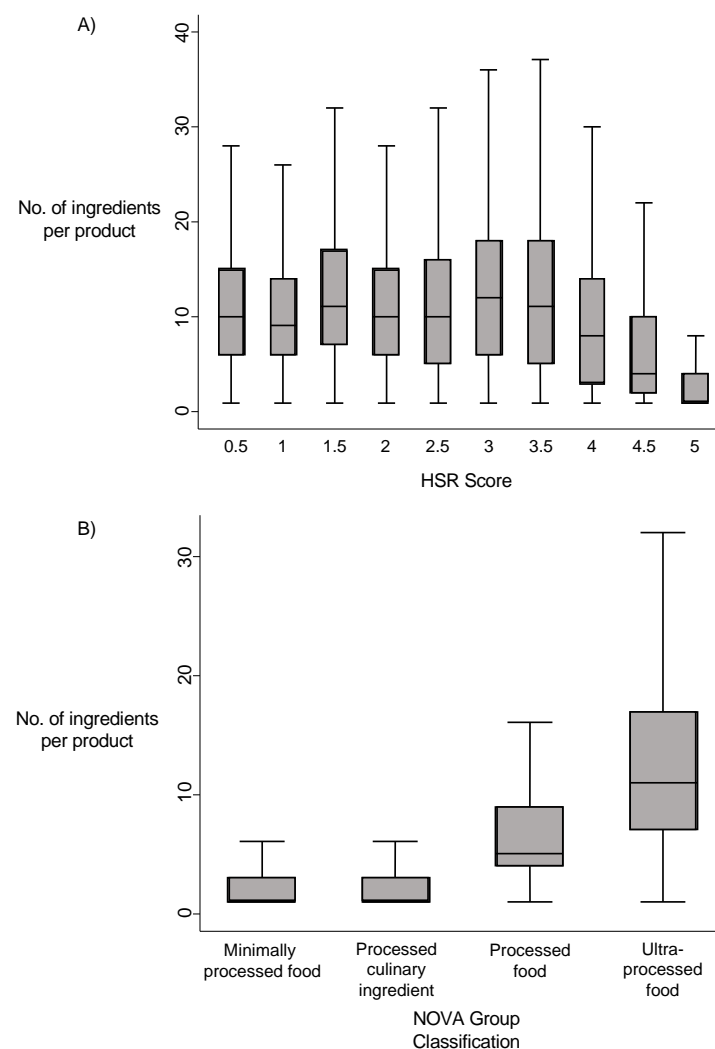


Figure 2. The association of the number of ingredients with (A) Health Star Rating and (B) NOVA classification. Box plot (A) shows the median and interquartile range (IQR) number of ingredients in products assigned different Health Star Rating (HSR) scores. The whiskers are defined as the lower quartile – 1.5 * IQR and the upper quartile + 1.5 * IQR. Outliers are excluded from the graph. Box plot (B) shows the median and interquartile range (IQR) number of ingredients in products assigned different NOVA classification groups. The whiskers are defined as the lower quartile – 1.5 * IQR and the upper quartile + 1.5 * IQR. Outliers are excluded from the graph.

A similar pattern was seen across the four NOVA classification groups. Ultra-processed foods (NOVA group 4) used the highest number of ingredients per product (8 (3–14)) while minimally processed foods (NOVA group 1) used the lowest number of ingredients (1 (1–3)). The median number of ingredients per product was significantly higher in ultra-processed products (NOVA group 4) compared to all other NOVA groups (all p -values < 0.001); similarly, the median number of ingredients in processed products (NOVA group 3) was significantly higher compared to NOVA groups 1 and 2 (all p -values < 0.001). The association between number of ingredients per product and NOVA group was non-linear, with similarly low numbers of ingredients in both minimally processed foods (NOVA group 1) and processed culinary ingredients (NOVA group 2) (Figure 2B).

The difference in the number of ingredients used across HSR and NOVA groups was attenuated but the same pattern remained in analyses based on main ingredients alone.

4. Discussion

There are many tens of thousands of packaged food and beverage products available in Australia, but these products are comprised of a much smaller number of ingredients. Of these ingredients, the most commonly used are salt, sugar, flavours, colours and other additives which are widely considered to be indicative of unhealthy foods [28,29]. At the same time, a higher number of ingredients in a product is strongly associated both with lower (i.e., healthier) HSR values and NOVA classifications indicating greater degrees of food processing.

The association between the median number of ingredients and nutritional quality was clear, though only apparent for products at the healthier end of the HSR spectrum. At HSR below 3.5 stars, the relationship between the number of ingredients and healthiness plateaued. The reason for this is not clear. A more detailed investigation of the healthiness of individual ingredients in a product with respect to the product HSR was beyond the scope of this initial report but might provide further insights in the future.

Using this large and representative dataset, we ascertained that the median number of ingredients used in products was positively correlated with higher NOVA classifications for levels of processing. More extensive processing done to enhance palatability, convenience and large-scale manufacturing often results in end products comprised of many components that have little in common with the base ingredients. Complex food chemistry may be involved, requiring multiple different additives, albeit usually in small quantities. Extensive processing is also associated with products being less healthy [10] and this correlates well with the observed associations between HSR and number of ingredients using this comprehensive dataset.

The NOVA classification system defines ultra-processed foods as formulations made mostly or entirely from substances derived from foods and additives ‘typically with five or more and usually many ingredients’ [30]. Thus, as previously suggested, products classified as ultra-processed were expected to contain a larger number of ingredients [6,31]. However, this is not always the case. For example, rice cakes contain only one to two ingredients (rice and, in some cases, salt), yet undergo numerous heat treatments and product modifications to cook, puff and compress the ingredients into the final product. This raises the question of which factors, including chemical, biological and cultural, should be considered when defining foods as ultra-processed. In fact, terminology around food processing and usage in policy making is of growing concern [32,33]. Currently, there is no single, accepted standard for classifying a food product’s level of processing [34]. Our results support the need for a more comprehensive classification system as well as further work to understand how processing impacts diet-related research and policy.

The research also identified several inconsistencies in ingredient labelling that may mitigate the goal of providing consumers with transparency into product composition, a pattern which has also been recognised in an analysis of United States product ingredient labelling [35]. The vast range of terms used in ingredient lists as compared to the actual number of different ingredients almost certainly introduces errors in consumer interpre-

tation and understanding. Likewise, the use of compound ingredients, with inconsistent requirements for reporting of sub-ingredients, is another point at which misinterpretation of product ingredients is likely. While the ‘Compound ingredients’ category had a large number of unique ingredients, the usage of each was small, and it would be relatively little work for the affected companies to disclose those sub-ingredients.

There is a clear opportunity for better standardisation of ingredient reporting to enhance consumer understanding, as well as nutrition research. For example, presently, ingredient naming can be done using common names or generic names—if a product includes pear, apple and peach, a valid statement of ingredients can use either the discrete list of all three ingredients or the generic ingredient fruit [11]. It is also currently permissible for an ingredient to be omitted if it makes up a small proportion of the final product and is not a common allergen, or is a substance used as a processing aid or product flavouring [11]. This may not be a substantial issue in many cases, but in some circumstances it may be important. For example, vitamin and mineral fortification of packaged products, as well as the use of additives, is of increasing interest to consumers and governments around the world [36], and only with full reporting of all ingredients will such data be generally available. There is a strengthening case for both better specification of the regulations governing ingredient reporting as well as their more rigorous enforcement.

The novel ingredient dataset developed for this study enables future research in food composition analysis. Understanding the interactions between food ingredients, particularly when ingredients are processed or cooked in different ways, is of increasing interest in nutrition and food chemistry research [37]. In addition, our enhanced ingredient-level data defining the packaged food and beverage supply can contribute to the new nutrition science concept, which combines biological, environmental and social sciences [38]. For example, our methods can be applied to incorporate ingredient list data into widely used databases to better analyse the packaged food and beverage supply (e.g., FoodEx and LanguaL). There is also potential to address sustainable diets using ingredient breakdowns [39].

A key strength of our study is the size of the dataset on which we have based our analyses. The FoodSwitch platform systematically collects product information from all the main retailers, and these products represent the majority of all packaged food and beverages sold in Australia each year. The extensively cleaned and concise final list of categorised ingredients provides for a unique new description of the main constituents of the Australian packaged food and beverage supply. Future work using ingredient descriptors such as ‘organic’ and ‘free-range’ will provide additional opportunities for research and greater insight into the health and environmental characteristics of Australian packaged foods. The large number of carefully defined data points also enabled us to reliably describe the numbers of ingredients in different food categories and how they relate to widely used indicators of healthiness and food processing.

There are also some limitations. For about one-in-ten products, it was not possible to obtain an analysable statement of ingredients, and this was likely a more pronounced issue for products with larger numbers of ingredients. While extensive work was done to standardise ingredient terminology, this work was mostly done by one author and there may be some misclassifications. Workload limitations did not allow us to fully categorise different E numbers, which is a growing area of interest in nutrition research and will be a future task for us [40].

5. Conclusions

The assessment of constituent ingredients provides a novel means of evaluating packaged food and beverage products. Our initial findings related to HSR and NOVA classifications suggest that further exploration of food products based upon their ingredients may provide new insights into their effects on both human health and planetary health.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/nu13061882/s1>, Table S1: Major and minor ingredient categories and their definitions, Table S2: Final list of 915 ingredients in the 2019 Australian packaged food and beverage supply.

Author Contributions: A.G. and B.N. designed the research; M.S., L.H. and T.D. assisted in developing the data disaggregation and cleaning code and validating the categorisation and results; F.T. and J.H.W. provided essential data and materials; A.G. conducted the research and performed statistical analyses; A.G. and B.N. were responsible for final content. All authors have read and agreed to the published version of the manuscript.

Funding: A.G. was supported by the President’s PhD Scholarship provided by Imperial College London. B.N. is supported by a Principal Research Fellowship provided by the National Health and Medical Research Council of Australia (APP1106947). J.H.W. is supported by a Scientia Fellowship provided by the University of New South Wales.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy policies.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Formatting requirements for successful disaggregation of a statement of ingredients.

Required Format for Statements of Ingredients	Example Product Name	Example Ingredient List
<i>General</i>		
Statements of ingredients must contain a maximum of one colon (:), as two or more colons indicate a multi-pack product	Arnott’s Sao Biscuits	Biscuits: wheat flour vegetable oil salt yeast malt extract (from barley) sugar baking powder emulsifier (soy lecithin) antioxidant (e307b from soy)
Statements of ingredients must contain an equal number of opening bracket ([) and/or parentheses (()) as closing brackets (]) and/or parentheses (')	Mountain Bread Corn Wraps	Pork water cure [salt sugar mineral salts (451 450) antioxidant (316) preservative (250)]
<i>Main ingredients</i> Ingredients are separated by a comma (,) or a pipe delimiter ()	McKenzie’s Italian Style Soup Mix	Peas beans lentils
Ingredient weight percentages are included directly after the ingredient in parentheses (())	Val Verde Traditional Pasta Sauce	Tomatoes (96.6%) onion salt onion powder acidity regulator (citric acid (330))
<i>Sub-ingredients</i> Sub-ingredients are indicated using brackets ([]) or parentheses (()) directly after the compound ingredient Each sub-ingredient is separated by a comma (,) or a pipe delimiter ()	Mountain Bread Corn Wraps	Corn flour (70%) wheat flour (flour folate thiamin) filtered water iodised salt

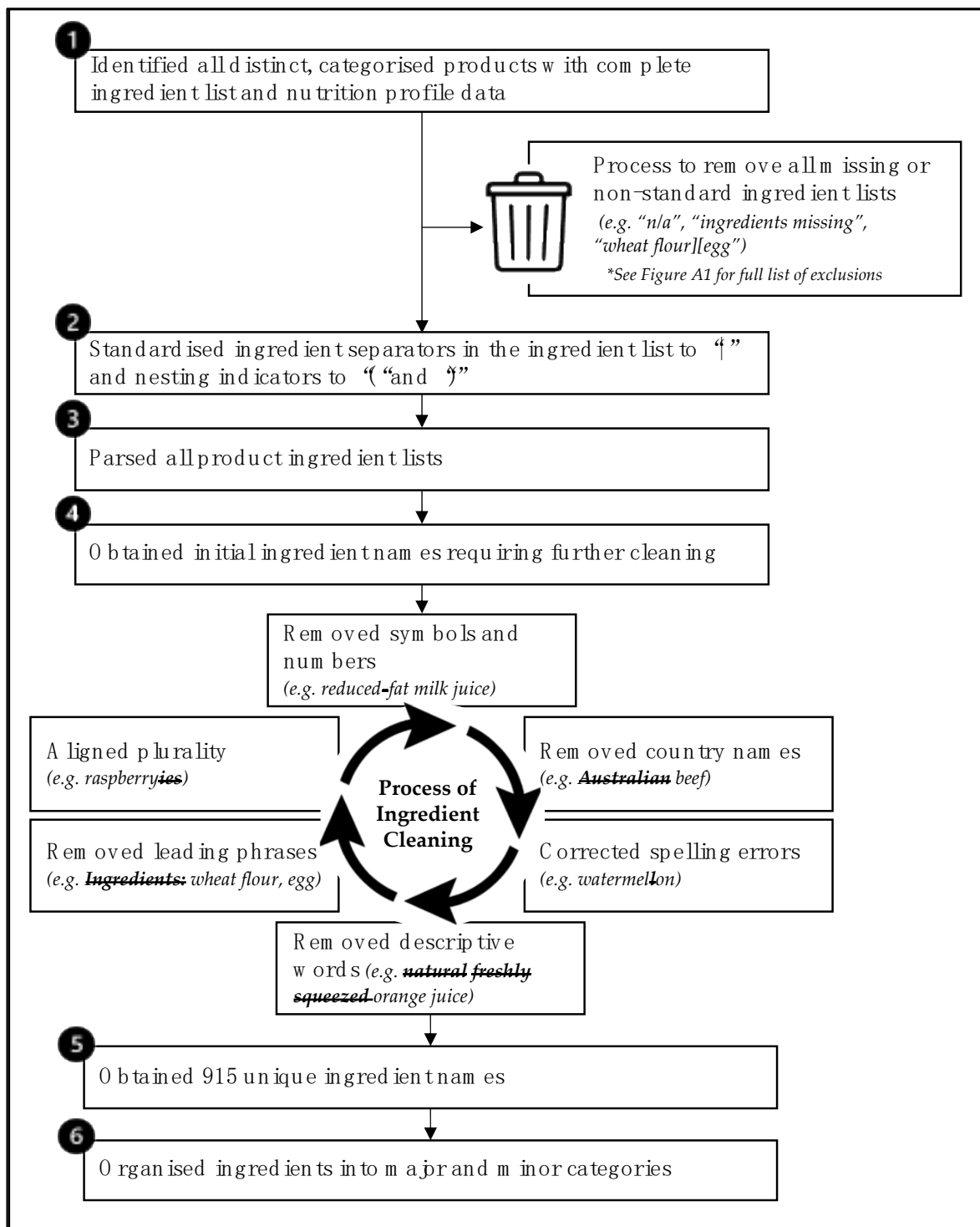


Figure A1. Workflow for developing the list of unique ingredients in Australian food and beverage products. Struck-through words and descriptors that were removed during the process of ingredient cleaning.

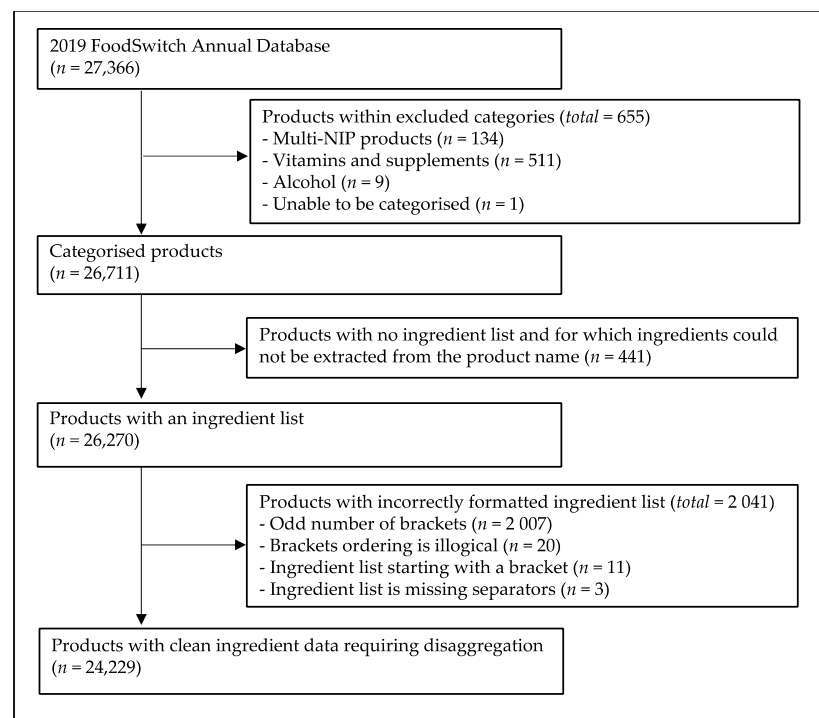


Figure A2. Flowchart describing product inclusion. The 2019 FoodSwitch Annual Database extract is comprised of products scanned during the in-store surveys done between August and November 2019 in Sydney locations of the following large supermarkets: IGA, ALDI, Woolworths, Coles and Harris Farm.

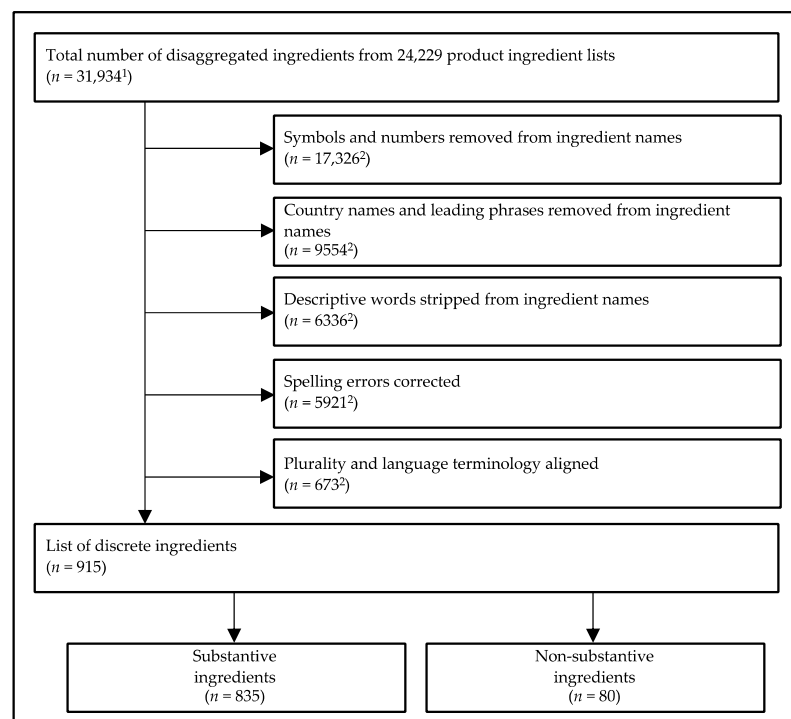


Figure A3. Flowchart describing ingredient cleaning and classifications. ¹ The original 31,934 ingredients were a result of the original ingredient disaggregation and consolidation into unique ingredient names. ² A number of ingredients required multiple cleaning steps, so the number of ingredients cleaned in each step sums to more than the total number of disaggregated ingredients removed.

References

1. Burden of Disease. Australian Institute of Health and Welfare (AIHW). Available online: <https://www.aihw.gov.au/reports/australias-health/burden-of-disease> (accessed on 2 February 2021).
2. Hawkes, C.; Jewell, J.; Allen, K. A food policy package for healthy diets and the prevention of obesity and diet-related non-communicable diseases: The NOURISHING framework. *Obes. Rev.* **2013**, *14*, 159–168. [CrossRef] [PubMed]
3. Popkin, B.M. Contemporary nutritional transition: Determinants of diet and its impact on body composition. *Proc. Nutr. Soc.* **2011**, *70*, 82–91. [CrossRef]
4. Swinburn, B.A.; Sacks, G.; Hall, K.D.; McPherson, K.; Finegood, D.T.; Moodie, M.L.; Gortmaker, S.L. The global obesity pandemic: Shaped by global drivers and local environments. *Lancet* **2011**, *378*, 804–814. [CrossRef]
5. Health Star Rating System. Australian Department of Health. Available online: <http://www.healthstarrating.gov.au/internet/healthstarrating/publishing.nsf/content/home> (accessed on 28 January 2020).
6. Monteiro, C.A.; Cannon, G.; Moubarac, J.-C.; Levy, R.B.; Louzada, M.L.C.; Jaime, P.C. The UN Decade of Nutrition, the NOVA food classification and the trouble with ultra-processing. *Public Health Nutr.* **2018**, *21*, 5–17. [CrossRef]
7. Lusk, J.L.; Briggeman, B.C. Food Values. *Am. J. Agric. Econ.* **2009**, *91*, 184–196. [CrossRef]
8. Monteiro, C.A.; Moubarac, J.-C.; Levy, R.B.; Canella, D.S.; Louzada, M.L.D.C.; Cannon, G. Household availability of ultra-processed foods and obesity in nineteen European countries. *Public Health Nutr.* **2018**, *21*, 18–26. [CrossRef] [PubMed]
9. Kim, H.; Hu, E.A.; Rebholz, C.M. Ultra-processed food intake and mortality in the USA: Results from the Third National Health and Nutrition Examination Survey (NHANES III, 1988–1994). *Public Health Nutr.* **2019**, *22*, 1777–1785. [CrossRef]
10. Elizabeth, L.; Machado, P.; Zinöcker, M.; Baker, P.; Lawrence, M. Ultra-Processed Foods and Health Outcomes: A Narrative Review. *Nutrients* **2020**, *12*, 1955. [CrossRef]
11. Food Standards in Australia and New Zealand (FSANZ). *Australia New Zealand Food Standards—Code-Standard 1.2.4—Ingredient Labelling of Food*; Australian Government Federal Register of Legislation: Canberra, Australia, 2015.
12. Food Standards in Australia and New Zealand (FSANZ). *Australia New Zealand Food Standards—Code-Standard 1.2.10—Characterising Ingredients and Components of Food*; Australian Government Federal Register of Legislation: Canberra, Australia, 2018.
13. Elmadfa, I.; Meyer, A.L. Importance of food composition data to nutrition and public health. *Eur. J. Clin. Nutr.* **2010**, *64*, S4–S7. [CrossRef]
14. Australian Food Composition Database. Food Standards in Australia and New Zealand. Available online: <https://www.foodstandards.gov.au/science/monitoringnutrients/afcd/Pages/default.aspx> (accessed on 10 August 2020).
15. Heller, M.C.; Keoleian, G.A.; Willett, W.C. Toward a Life Cycle-Based, Diet-level Framework for Food Environmental Impact and Nutritional Quality Assessment: A Critical Review. *Environ. Sci. Technol.* **2013**, *47*, 12632–12647. [CrossRef]
16. Jacobs, D.R.; Tapsell, L.C.; Temple, N.J. Food Synergy: The Key to Balancing the Nutrition Research Effort. *Public Health Rev.* **2011**, *33*, 507–529. [CrossRef]
17. Tapsell, L.C.; Neale, E.P.; Satija, A.; Hu, F.B. Foods, Nutrients, and Dietary Patterns: Interconnections and Implications for Dietary Guidelines. *Adv. Nutr.* **2016**, *7*, 445–454. [CrossRef]
18. Dunford, E.; Trevena, H.; Goodsell, C.; Ng, K.H.; Webster, J.; Millis, A.; Goldstein, S.; Hugueniot, O.; Neal, B. 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. [CrossRef] [PubMed]
19. Jones, A.; Rådholm, K.; Neal, B. Defining ‘Unhealthy’: A Systematic Analysis of Alignment between the Australian Dietary Guidelines and the Health Star Rating System. *Nutrients* **2018**, *10*, 501. [CrossRef]
20. Coyle, D.H.; Shahid, M.; Dunford, E.K.; Mhurchu, C.N.; Mckee, S.; Santos, M.; Popkin, B.M.; Trieu, K.; Marklund, M.; Taylor, F.; et al. Contribution of major food companies and their products to household dietary sodium purchases in Australia. *Int. J. Behav. Nutr. Phys. Act.* **2020**, *17*. [CrossRef]
21. Dunford, E.; Webster, J.; Metzler, A.B.; Czernichow, S.; Mhurchu, C.N.; Wolmarans, P.; Snowdon, W.; L’Abbe, M.; Li, N.; Maulik, P.K.; et al. International collaborative project to compare and monitor the nutritional composition of processed foods. *Eur. J. Prev. Cardiol.* **2012**, *19*, 1326–1332. [CrossRef]
22. Neal, B.; Sacks, G.; Shahid, M.; Taylor, F.; Huffman, M. FoodSwitch: State of the Food Supply. The George Institute for Global Health. 2019. Available online: https://www.georgeinstitute.org/sites/default/files/food_supply_report.pdf (accessed on 13 August 2020).
23. Food Standards in Australia and New Zealand (FSANZ). *Australia New Zealand Food Standards—Code-Standard 1.2.1—Requirements to Have Labels or Otherwise Provide Information*; Australian Government Federal Register of Legislation: Canberra, Australia, 2016.
24. Ingredients, Resources, and Expertise for Creating Successful Food and Beverage Products. Available online: <https://www.cargill.com/food-beverage> (accessed on 27 August 2020).
25. Food and Beverage Solutions. Available online: <https://www.adm.com/products-services/food> (accessed on 27 August 2020).
26. Bernstein, J.; Schermel, A.; Mills, C.; L’Abbé, M. Total and Free Sugar Content of Canadian Prepackaged Foods and Beverages. *Nutrients* **2016**, *8*, 582. [CrossRef]
27. Food Standards in Australia and New Zealand (FSANZ). *Australia New Zealand Food Standards Code—Standard 1.3.1—Food Additives*; Australian Government Federal Register of Legislation: Canberra, Australia, 2019.

28. Elliott, P.; Stamler, J.; Nichols, R.; Dyer, A.R.; Stamler, R.; Kesteloot, H.; Marmot, M. Intersalt revisited: Further analyses of 24 hour sodium excretion and blood pressure within and across populations. *BMJ* **1996**, *312*, 1249–1253. [[CrossRef](#)]
29. Yang, Q.; Zhang, Z.; Gregg, E.W.; Flanders, W.D.; Merritt, R.; Hu, F.B. Added Sugar Intake and Cardiovascular Diseases Mortality Among US Adults. *JAMA Intern. Med.* **2014**, *174*, 516. [[CrossRef](#)]
30. Monteiro, C.; Cannon, G.; Levy, R.; Moubarac, J.-C.; Jaime, P. NOVA. The star shines bright. Position paper 2. *World Nutr.* **2016**, *7*, 28–38.
31. Monteiro, C.A.; Cannon, G.; Levy, R.B.; Moubarac, J.C.; Louzada, M.L.; Rauber, F.; Khandpur, N.; Cediel, G.; Neri, D.; Martinez-Steele, E.; et al. Ultra-processed foods: What they are and how to identify them. *Public Health Nutr.* **2019**, *22*, 936–941. [[CrossRef](#)] [[PubMed](#)]
32. Gibney, M.J. Ultra-Processed Foods: Definitions and Policy Issues. *Curr. Dev. Nutr.* **2019**, *3*. [[CrossRef](#)] [[PubMed](#)]
33. Knorr, D.; Watzke, H. Food Processing at a Crossroad. *Front. Nutr.* **2019**, *6*. [[CrossRef](#)]
34. Sadler, C.R.; Grassby, T.; Hart, K.; Raats, M.; Sokolović, M.; Timotijevic, L. Processed food classification: Conceptualisation and challenges. *Trends Food Sci. Technol.* **2021**, *112*, 149–162. [[CrossRef](#)]
35. Ahuja, J.K.C.; Li, Y.; Bahadur, R.; Nguyen, Q.; Haile, E.; Pehrsson, P.R. IngID: A framework for parsing and systematic reporting of ingredients used in commercially packaged foods. *J. Food Compos. Anal.* **2021**, *100*, 103920. [[CrossRef](#)]
36. Chadare, F.J.; Idohou, R.; Nago, E.; Affonfere, M.; Agossadou, J.; Fassinou, T.K.; Kénou, C.; Honfo, S.; Azokpota, P.; Linne-mann, A.R.; et al. Conventional and food-to-food fortification: An appraisal of past practices and lessons learned. *Food Sci. Nutr.* **2019**, *7*, 2781–2795. [[CrossRef](#)]
37. Durazzo, A.; Lisciani, S.; Camilli, E.; Gabrielli, P.; Marconi, S.; Gambelli, L.; Aguzzi, A.; Lucarini, M.; Maiani, G.; Casale, G.; et al. Nutritional composition and antioxidant properties of traditional Italian dishes. *Food Chem.* **2017**, *218*, 70–77. [[CrossRef](#)]
38. Cannon, G.; Leitzmann, C. The new nutrition science project. *Public Health Nutr.* **2005**, *8*, 673–694. [[CrossRef](#)]
39. Traka, M.H.; Plumb, J.; Berry, R.; Pinchen, H.; Finglas, P.M. Maintaining and updating food composition datasets for multiple users and novel technologies: Current challenges from a UK perspective. *Nutr. Bull.* **2020**, *45*, 230–240. [[CrossRef](#)]
40. Partridge, D.; Lloyd, K.A.; Rhodes, J.M.; Walker, A.W.; Johnstone, A.M.; Campbell, B.J. Food additives: Assessing the impact of exposure to permitted emulsifiers on bowel and metabolic health—Introducing the FADiets study. *Nutr. Bull.* **2019**, *44*, 329–349. [[CrossRef](#)]