

Are food and drink available in online and physical supermarkets the same? A comparison of product availability, price, price promotions and nutritional information

Prachi Bhatnagar^{1,*} , Peter Scarborough¹, Asha Kaur¹, Derya Dikmen², Vyas Adhikari¹ and Richard Harrington

¹Centre on Population Approaches for Non-Communicable Disease Prevention, Nuffield Department of Population Health, University of Oxford, OX3 7LF Oxford, UK: ²Nutrition and Dietetics, Hacettepe University, Ankara, Turkey

Submitted 15 June 2020: Final revision received 17 September 2020: Accepted 21 October 2020: First published online 28 October 2020

Abstract

Objective: Online supermarkets are increasingly used both by consumers and as a source of data on the food environment. We compared product availability, nutritional information, front-of-pack (FOP) labelling, price and price promotions for food and drink products between physical and online supermarkets.

Design: For physical stores, we collected data on price, price promotions, FOP nutrition labels and nutrition information from a random sample of food and drinks from six UK supermarkets. For online stores, we used foodDB, a research-ready dataset of over 14 million observations of food and drink products available in online supermarkets.

Setting: Six large supermarket stores located near Oxford, UK.

Participants: General sample with 295 food and drink products, plus boost samples for both fruit and vegetables, and alcohol.

Results: In the general sample, 85 % (95 % CI 80, 90 %) of products found in physical stores could be matched with an online product. Nutritional information found in the two settings was almost identical, for example, concordance correlation coefficient for energy = 0.995 (95 % CI 0.993, 0.996). The presence of FOP labelling and price promotions differed between the two settings (Cohen's kappa = 0.56(95 % CI 0.45, 0.66) and 0.40 (95 % CI 0.26, 0.55), respectively). Prices were similar between online and physical supermarkets (concordance correlation coefficient > 0.9 for all samples).

Conclusions: Product availability, nutritional information and prices sourced online for these six retailers are good proxies of those found in physical stores. Price promotions and FOP labelling vary between the two settings. Further research should investigate whether this could impact on health inequalities.

Keywords Nutrition Diet Online supermarkets Price Promotion Front-of-pack labelling

Online shopping for groceries was the fastest growing sector of the supermarket industry between 2010 and 2018 in the UK(1) and elsewhere(2); Kantar Worldpanel data indicate that the effects of the Covid-19 pandemic are likely to further support this trend⁽³⁾. In the UK in 2018, 7 % of supermarket shopping was conducted online, representing over £11 billion annually(1). Online grocery sales are particularly strong in the UK, France, the USA, China and South Korea⁽²⁾.

Increasingly, online supermarkets are being used as a source of 'big data' for monitoring the food system and developing novel nutrient composition tables to support public health research. In the UK, data collected from six online supermarkets have been used to monitor the healthiness of ready meals and pizzas⁽⁴⁾ and to evaluate the impact of the Soft Drink Industry Levy on sugar levels in drinks⁽⁵⁾. In Australia, data collected from two online supermarkets have been used to monitor price promotions on sugar-sweetened beverages⁽⁶⁾ and foods⁽⁷⁾. These studies assume that the nutrition information, defined as nutrient declarations and/or supplementary nutrition information⁽⁸⁾, available in online supermarkets is representative of physical (i.e., brick and mortar) supermarkets. Pricing and price promotions were compared between the two

*Corresponding author: Email prachi.bhatnagar@ndph.ox.ac.uk © The Author(s), 2020. Published by Cambridge University Press on behalf of The Nutrition Society





P Bhatnagar et al. 820

settings, but there is limited research in this area. One New Zealand study comparing fruit and vegetable pricing in physical stores with an online supermarket found that many fruits and vegetables were cheaper in the online supermarket⁽⁹⁾; however, it is unclear whether this finding applies outside of New Zealand and whether it applies to other food and drink products found in supermarkets.

Product availability⁽¹⁰⁾ and provision of nutritional information⁽¹¹⁾ are important factors for promoting healthy eating when grocery shopping. There is also some evidence that online grocery shoppers are less price sensitive and have completed higher levels of education⁽¹²⁾, although much more research is needed on this topic, particularly in a UK context. If prices and promotions - two aspects of marketing known to influence purchasing behaviour (13-16) - differ between online and physical stores, then this could influence health inequalities. This is particularly so if categories associated with health outcomes, for example, fruits and vegetables⁽¹⁷⁾ and alcohol⁽¹⁸⁾, differ in how they are marketed online and in physical stores.

We compared a sample of food and drinks (hereafter referred to as 'food' as shorthand for food and drink products) collected in six physical supermarkets with products found in the online offering of those supermarkets to assess coverage, data validity and differentiation of online supermarkets. Specifically, we estimated the proportion of foods available in physical supermarkets that can also be purchased online; we compared nutritional data which should be identical in the two settings and compared provision of front-of-pack (FOP) food labelling, price and price promotions between identical foods available to purchase in online and physical stores.

Methods

Data collection from physical supermarkets

We collected data on price, price promotions, FOP nutrition labels and nutrition information from physical supermarkets by photographing the shelf barkers, and front and back packaging of a random sample of foods from each of Tesco, Sainsbury's, Asda, Waitrose, Morrisons and Iceland. For each supermarket, we selected the nearest superstore to the city of Oxford, UK; we used the supermarkets' own definition of superstore, as listed on their websites. We received permission from store managers to collect data.

Our protocol for random sampling of foods, established before data collection, followed a published method⁽¹⁹⁾. The population that we sampled consisted of all items available for purchase in the supermarket excluding non-food items, items from deli/fresh food counters and items for sale at the checkout. Deli and fresh food counter products were excluded due to difficulty in photographing products without assistance from supermarket staff. Checkout items such as mints or chewing gum were excluded as it was assumed that these items would also be available for purchase in other places in the supermarket, and hence including them in the sampling frame would lead to over-representation. For each supermarket, a researcher visited and identified all of the sections (i.e., aisles and end of aisle displays) where eligible foods were available and assigned each section a number. The number of different foods available in each section was estimated by counting the number of foods in one section; this provided an estimate of the maximum number of products in each section. A random number generator, weighted by the estimated number of foods per section, was used to generate thirty-five randomly selected numbers. Following a standard protocol, these generated numbers were then used to count products in each section; when the randomly generated number was reached, that product was included in the sample. To allow for stratified analyses for two categories strongly associated with health, in each supermarket, an additional seven foods were randomly selected from the fresh fruit and vegetable section, and seven from the alcohol section, making a total of forty-nine foods sampled in each supermarket. We were unable to find any similar previous studies to provide effect sizes for power calculations; therefore, the sample size was based on the number of products needed to estimate Cohen's kappa statistic with 95 % CI of \pm 0.1.

Two researchers used the photographs to build a dataset that consisted of supermarket and product IDs, product name, product size, price observed on the day, whether or not on price promotion (defined as an advertised temporary difference in price), whether or not FOP nutrition labelling of any kind is present and nutritional information per 100 g for the following mandatory⁽²⁰⁾ values: energy, fat, saturated fat, carbohydrates, sugar, fibre, protein and salt. Non-nutritional FOP labels describing dietary types or health claims were not investigated in the current study. Using the FoodEx2 classification system⁽²¹⁾ to categorise the foods sampled, in total fifty-three categories of food were sampled across the six supermarkets (Fig. 1).

Data matching with online supermarkets

We used foodDB to identify equivalent products from online supermarkets. A full description of this dataset is provided elsewhere⁽³⁾. Briefly, foodDB is a research-ready dataset of over 14 million observations of food and drink products available in online supermarkets, collecting weekly 'snapshots' since November 2017. Time-stamped data are collected weekly from seven UK online supermarkets (the six supermarkets included in the current study, plus Ocado which does not have any physical stores) using automated methods which capture, process and store data on over 99% of all food and drink products available to purchase online. Data collected include product price, promotions, product size, ingredients, nutritional information and image. These data are curated in a MySQL database. The presence of FOP labelling was assessed by a researcher manually assessing the images identified in the online database.





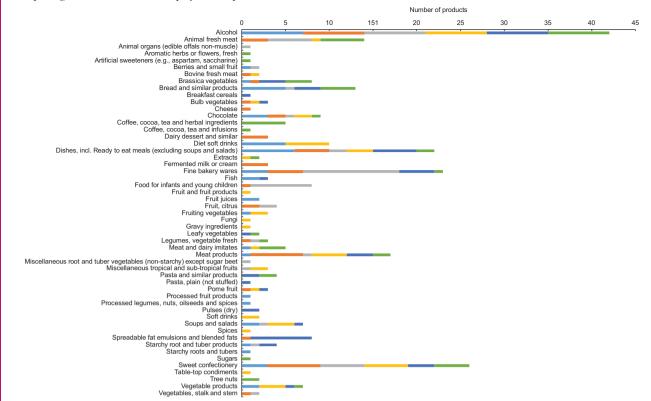


Fig. 1 Food categories sampled in each supermarket. , Asda; , Iceland; , Morrisons; , Sainsburys; , Tesco; , Waitrose

For each of the foods sampled in the physical stores, foodDB snapshots within a maximum of a month of the physical data collection were searched for an equivalent product matched on date of collection, supermarket, product name and product size. Where an equivalent product could not be found, we relaxed the criterion of matching on product size (e.g., if an equivalent food with product size 400 g could not be found in the online supermarket from the same time period, but the same food with product size 200 g was available, we selected the smaller variant). In the cases where product size was not matched, we only included these foods in the comparison of nutritional values. Therefore, this resulted in two datasets. The full match dataset that matched both the name and product size in the physical and foodDB databases was used for comparison of price, price promotions and presence of FOP labels. The product-match dataset was used to compare the nutritional data per 100 g.

Analysis

We assessed the coverage by calculating the proportion of foods collected that were available in online supermarkets for the general, fruit and vegetable, and alcohol samples. Next, we calculated the proportion of price promotions and FOP labelling in online and physical supermarkets. We then assessed the median and interquartile values for price and all nutritional variables using a Wilcoxon

rank-sum test to assess if there were significant differences between the online and physical datasets.

To explore the degree of agreement between the online and physical supermarket samples, we calculated interrater reliability for price, whether or not on price promotion, whether or not contains FOP nutrition labelling and each of the eight nutrients stated on back-of-pack nutrition labelling. For the binary variables, we used Cohen's kappa statistic. For the continuous variables, we used the concordance correlation coefficient and produced Bland-Altman plots to assess evidence of systematic bias. These analyses were conducted on the general sample and also on the two subsamples of fresh fruit and vegetables, and alcoholic drinks separately. For nutritional information, we used data as presented in both physical stores and online supermarkets; the data presented in nutritional tables inconsistently use 0, 1 or 2 decimal places. For purposes of these calculations, we used the value 'x' wherever nutritional information was reported as '<x' (e.g., <0·1 g salt per 100 g).

Results

Matching of products

Table 1 shows the success in matching products identified in physical stores with equivalent foods from online supermarkets found in foodDB. We found that 14.7% (95% CI 10.3%, 20.4%) of the general sample were not found in





P Bhatnagar et al. 822

Table 1 Percentage match between physical store samples and online supermarkets samples

	n	Full match*	95 % CI	Product only match†	95 % CI
General sample	211	85.3 %	79.6, 89.7	90.5%	85.5, 94.0
Fruit and vegetable sample	42	88.1 %	73.6, 95.5	92.9 %	79.4, 98.1
Alcohol sample	42	95.2 %	82.6, 99.2	97.6 %	85.9, 99.9
Total sample	295	257		271	

^{*}Both a product and size match between the physical and online supermarket

Table 2 Description and comparison of price in £ per product and nutritional data in g/100 g in physical and online supermarket total samples

	Physical supermarkets		Online supermarkets		Difference between physical and online supermarkets		P			
	Median	First quartile	Third quartile	Median	First quartile	Third quartile	Median	First quartile	Third quartile	Wilcoxon rank sum
Price	£1.87	£1.16	£3.00	£1.85	£1.19	£3.00	£0.02	-£0·03	£0.00	0.8928
Energy (kcal)*	218.0	78.5	384.5	206.5	70.75	386.25	11.50	7.75	−1 ·75	0.5751
Protein	5⋅1	1.8	9.2	4.8	1.5	9.70	0.30	0.30	-0.50	0.8001
Carbohydrates	16.0	53.0	2.0	13.6	2.85	52.12	2.40	50.15	-50.12	0.8933
Sugar	2.6	0.6	16.0	2.55	0.6	17.90	0.05	0.00	−1 ·90	0.724
Fat	6.7	1.0	17.9	6.4	0.5	16.50	0.26	0.48	1.43	0.5404
Saturated fat	1.8	0.2	6.0	1.6	0.175	6.03	0.20	0.03	-0.03	0.5804
Fibre	1.6	0.5	2.9	1.8	0.6	3.00	-0.20	-0.10	-0.15	0.5826
Salt	0.4	0.1	0.9	0.3	0.03	0.79	0.07	0.06	0.12	0.2728

^{*}To convert kcal to kJ, multiply by 4.184.

online supermarkets, but this reduced to 9.5% (6.0%, 14.5%) when including product matches (where the product name matched, but the size of the product differed between physical and online supermarkets). Confidence intervals for both full matches and product only matches overlapped for the general sample and the fruit and vegetables, and alcohol samples (Table 1). Full and product only matches by supermarket are available in the online Supplementary Tables. Prices in physical and online supermarkets had similar medians (£1.87 v. £1.85); nutritional value medians and interquartile ranges were also similar between the two settings (Table 2).

Comparing front-of-pack labelling, price and promotions

The price displayed for products was very similar between the physical and online stores, with little indication of difference for the general sample, alcohol or fruit and vegetables samples (concordance correlation coefficient > 0.9 for all samples). Bland-Altman plots showed no evidence of bias (Fig. 2). The presence of FOP labelling was only assessed in the general sample as alcoholic drinks are only allowed to display an energy label on a voluntary basis (19), and fruit and vegetable products are often not packaged. There was moderate agreement between physical stores and online supermarkets (Cohen's kappa = 0.56 (95 % CI 0.45, 0.66)), with 74% of FOP labels present on products from physical stores, compared with 42% on products in online stores (P < 0.001). There was fair agreement in the general sample for the presence of price promotions (Cohen's kappa = 0.40 (95 % CI 0.26, 0.55)) with more price promotions present in physical supermarkets (32% of products) as compared with online supermarkets (24% of products); however, CI were wide for the alcohol and fruit and vegetable samples (Table 3).

Nutritional information validation

In the general sample, the nutrient composition information displayed on online supermarkets was almost identical to the nutrient composition information displayed on the packaging found in physical supermarkets (Table 4). Bland-Altman plots for all nutrients showed no evidence of bias (Fig. 2).

Discussion

In a random sample of foods selected from six UK supermarkets, we found that there was a good correlation in the availability of products between online and physical supermarkets. The nutritional information displayed on packaging is almost identical to the nutritional information displayed on online supermarket websites. The presence of FOP labelling was somewhat inconsistent between the two settings. Prices were very similar between online



[†]The same product was found in both the online and physical supermarkets, but the size of the product differed.



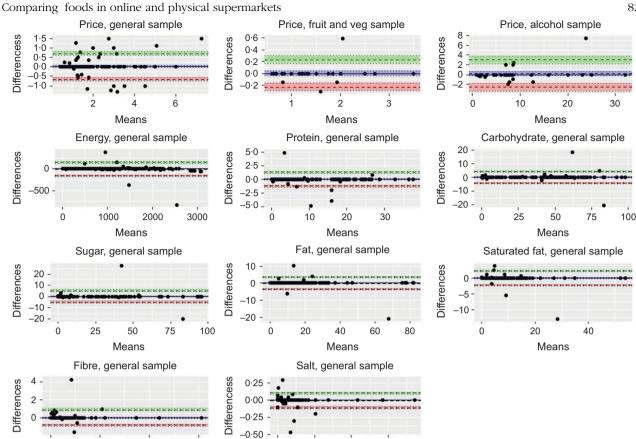


Fig. 2 (colour online) Bland-Altman plots

Means

20

30

Table 3 Comparison of price, presence of front-of-pack labelling and price promotions between online supermarkets and physical supermarkets, using full match dataset

6

Means

9

0

<i>n</i> in analytic sample	Estimate	Lower 95 % CI	Upper 95 % CI	
	Cohen's kappa statistic			
180	0.56	0.45	0.66	
179	0.40	0.26	0.55	
36	0.67	0.40	0.93	
40	0.28	-0.02	0.57	
	C	oncordance correlation co	efficient	
180	0.95	0.94	0.97	
37	0.98	0.97	0.99	
40	0.98	0.97	0.99	
	179 36 40 180 37	179 0.40 36 0.67 40 0.28 C 180 0.95 37 0.98	179 0.40 0.26 36 0.67 0.40 40 0.28 -0.02 Concordance correlation co 180 0.95 0.94 37 0.98 0.97	

and physical supermarkets, but there is evidence of marked differences in price promotions between the two settings.

Nutrient composition tables on packaging are required by law to be accurate⁽²²⁾ and previous research has suggested that nutritional labels are a reasonably accurate representation of the food contained in the packaging⁽²³⁾. Our study shows that the nutritional information available online and in physical stores in the UK correlates very

highly; therefore, it is viable to use data from online supermarket databases such as foodDB⁽⁴⁾ for nutrition studies and interventions.

The difference in price promotions between online and physical stores is interesting and particularly relevant in light of the UK government's 2018 childhood obesity plan⁽²⁴⁾ which outlines a policy to ban price promotions on unhealthy foods. As this policy may also be applied





P Bhatnagar *et al.*

Table 4 Comparison of nutritional information between online supermarkets and physical supermarkets, using the product-match dataset and general sample

Nutrition information		C	Concordance correlation coefficient			
	n in analytic sample	Estimate	Lower 95 % CI	Upper 95 % CI		
Energy	191	0.995	0.993	0.996		
Protein	191	0.997	0.996	0.998		
Carbohydrates	191	0.997	0.996	0.998		
Sugar	191	0.993	0.991	0.995		
Total Fat	191	0.993	0.991	0.995		
Saturated fat	191	0.991	0.988	0.993		
Fibre	191	0.994	0.991	0.996		
Salt	191	0.999	0.999	0.999		

online, it is important to identify differences between physical and online stores to determine how effective the regulations may be and to prevent exacerbation of inequalities. The report also indicates a plan to ban promotions of unhealthy items by location in physical retail stores; although it states an aim to extend policy to online shopping, the precise details of this are unclear.

As use of online supermarkets increases, understanding how differences in price promotions impact dietary inequalities is becoming increasingly important. The higher number of price promotions in physical stores means that it may be easier to get better value for money when grocery shopping in physical stores. Although our survey was underpowered to assess whether price promotions varied according to the healthiness of products, there is some evidence that higher sugar foods are more likely to be price promoted⁽²⁵⁾ and shoppers whose baskets contain high proportion of foods on price promotion buy higher numbers of less healthy foods⁽²⁰⁾. Price promotions are made use of similarly by people with differing incomes and family structures in physical retail stores in the UK⁽²⁶⁾; however, it is unknown if this translates to online supermarket shopping. Further research into promotions online is needed to better understand patterns in promotions, the impact on purchasing behaviour and the implications for policy.

The presence of FOP labelling differed somewhat between online and physical supermarkets. As the aim of FOP nutritional labelling is to convey the nutritional content of a food in a simple manner, this may mean that those doing their grocery shopping online lose out on this information. However, many online supermarkets provide traffic light labelling on the individual product pages which may overcome this issue somewhat.

We made an assumption that products of the same type with different sizes have the same ingredients and nutrient composition; this was the basis for the 'product only' match group. If the nutritional composition per 100 g was to change with the size of the product, then this would have an impact on our findings for the 'product only' matches.

Our study was underpowered to detect differences between the sub-categories of alcohol and fruit and vegetables and to produce supermarket-level results. Future research should use these data for sample size calculations, to enable studies to be adequately powered to detect differences in price promotions between online and physical stores, and whether these differences are associated with the healthiness of a product.

Our data collection was limited to Oxford and its surrounding areas; therefore, it is possible that regional variation in price and price promotions may be present which could not be investigated in our study. Supermarkets are known to practice price flexing or varying process by geographic location⁽²⁷⁾. Given that there are variations in price by geographic location, it is possible that there may also be variation in the offering of online supermarkets according to the location of the shopper; however, to date this is unknown.

Overall, we found there is good correlation between online supermarkets and physical stores. The current study provides some validation for using online and big data methods to monitor food systems, as well as using online supermarket data for public health interventions.

Acknowledgements

Acknowledgements: This paper was supported by the NIHR Biomedical Research Centre at Oxford. Financial support: This work was supported by the Medical Research Council (MR/R015465/1), British Heart Foundation (FS/15/34/ 31656), University of Oxford Nuffield Department of Population Health Early Career Fellowship (Award date Jan 2019), The Scientific and Technical Research Council of Turkey (TUBITAK 2219), NIHR Project Grant (16/130/01), NIHR Biomedical Research Centre at Oxford (IS-BRC-1215-20008). These funders had no role in the design, analysis or writing of this article. Conflicts of interest: There are no conflicts of interest. Authorship: P.S., A.K., R.H. and P.B. were involved in the design of the study. D.D. and P.B. conducted the data collection. V.A., P.S. and P.B. analysed the data. All authors contributed to writing the manuscript. Ethics of human subject participation: The current study did not involve human participants.

Supplementary material

For supplementary material accompanying this paper visit https://doi.org/10.1017/S1368980020004346





References

- Statista (2019) UK Consumers: Online Grocery Shopping. London: Statista.
- 2. Nielsen (2018) Connected Commerce. New York: Nielsen.
- Kantar (2020) UK grocery growth accelerates as retailers and shoppers look to next stage of lockdown. https://www. kantar.com/inspiration/fmcg/uk-grocery-growth-acceleratesas-retailers-and-shoppers-look-to-next-stage-of-lockdown (accessed May 2020).
- Harrington RA, Adhikari V, Rayner M et al. (2019) Nutrient composition databases in the age of big data: foodDB, a comprehensive, real-time database infrastructure. BMJ Open 9, e026652.
- Scarborough P, Adhikari V, Harrington RA et al. (2020) The impact of the announcement and implementation of the UK Soft Drinks Industry Levy on sugar content of soft drinks in the UK, 2015–19: controlled interrupted time series analysis. PLoS Med 17, e1003025.
- Zorbas C, Gilham B, Boelsen-Robinson T et al. (2019) The frequency and magnitude of price-promoted beverages available for sale in Australian supermarkets. Aust N Z J Public Health 43, 346–351.
- Riesenberg D, Backholer K, Zorbas C et al. (2019) Price promotions by food category and product healthiness in an Australian supermarket chain, 2017–18. Am J Public Health 109, 1434–1439.
- 8. Rayner M, Wood A, Lawrence M *et al.* (2013) Monitoring health-related food labelling. *Obes Rev* **14**, 70–81.
- Pearson AL, Winter PR, McBreen B et al. (2014) Obtaining fruit and vegetables for the lowest prices: pricing survey of different outlets and geographical analysis of competition effects. PLoS One 9, e89775.
- Hollands GJ, Carter P, Anwer S et al. (2019) Altering the availability or proximity of food, alcohol, and tobacco products to change their selection and consumption. Cochrane Database Syst Rev 2019, CD012573.
- Crockett RA, King SE, Marteau TM et al. (2018) Nutritional labelling for healthier food or non-alcoholic drink purchasing and consumption. Cochrane Database Syst Rev 2, CD009315.
- Jilcott Pitts SB, Ng SW, Blitstein JL et al. (2018) Online grocery shopping: promises and pitfalls for healthier food and beverage purchases. Public Health Nutr 21, 3360–3376.
- NIebylski ML, Redburn KA, Duhaney T et al. (2015) Healthy food subsidies and unhealthy food taxation: a systematic review of the evidence. Nutrition 31, 787–795.
- Afshin A, Peñalvo JL, Del Gobbo L et al. (2017) The prospective impact of food pricing on improving dietary consumption: a systematic review and meta-analysis. PLoS One 12, e0172277.

- Gittelsohn J, Trude ACB & Kim H (2017) Pricing strategies to encourage availability, purchase, and consumption of healthy foods and beverages: a systematic review. *Prev Chronic Dis* 14, E107.
- 16. Martin L, Bauld L & Angus K (2017) Rapid Evidence Review: The Impact of Promotions on High Fat, Sugar and Salt (HFSS) Food and Drink on Consumer Purchasing and Consumption Behaviour and the Effectiveness of Retail Environment Interventions. Edinburgh: NHS Health Scotland.
- Aune D, Giovannucci E, Boffetta P et al. (2017) Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality—a systematic review and dose-response meta-analysis of prospective studies. *Int J Epidemiol* 46, 1029–1056.
- Rehm J, Baliunas D, Borges GL et al. (2010) The relation between different dimensions of alcohol consumption and burden of disease: an overview. Addiction 105, 817–843
- Hieke S, Kuljanic N, Pravst I et al. (2016) Prevalence of nutrition and health-related claims on pre-packaged foods: a five-country study in Europe. Nutrients 8, 137.
- Population Health Division/ 10800 (2016) Technical Guidance on Nutrition Labelling. London: Department of Health.
- European Food Safety Authority (2019) Data Standardisation. https://www.efsa.europa.eu/en/data/data-standardisation (accessed December 2019).
- REGULATION (EU) No 1169/2011. https://eur-lex.europa. eu/legal-content/EN/TXT/?uri=CELEX%3A02011R1169-20180101 (accessed November 2020).
- Albuquerque TG, Nunes MA, Oliveira MBPP et al. (2020) Compliance of declared v. analysed values with EU tolerance limits for mandatory nutrients in prepacked foods. Food Chem 302, 125330.
- Department of Health and Social Care (2018) Chapter 2, Global public health directorate: obesity, food and nutrition / 10800. Childhood Obesity: A Plan for Action. London: HM Government.
- 25. Smithson M, Kirk J & Capelin C (2015) An Analysis of the Role of Price Promotions on the Household Purchases of Food and Drinks High in Sugar. London: Public Health England.
- Coker T, Rumgay H, Whiteside E et al. (2019) Paying the Price: New Evidence on the Link Between Price Promotions, Purchasing of Less Healthy Food and Drink, and Overweight and Obesity in Great Britain. London: Cancer Research UK.
- Competition Commission (2008) The supply of groceries in the UK market investigation. https://www.gov.uk/cmacases/groceries-market-investigation-cc (accessed November 2020).

