

Assessing the feasibility of RealityMine to capture digital advertising and elucidating the extent and nature of advertising of unhealthy commodities to children and parents

Simon Russell, Andrew Hamilton, Jessica Packer, Clare Llewellyn, Emma Boyland

Background

The UK Government is committed to restricting online advertising of high fat salt sugar (HFSS) foods to children in the UK. However, there is a limited understanding of the extent to which children are exposed to digital marketing. Digital marketing ecosystems are complex and there is a challenge for researchers, policy makers and advertising standards authorities to know who has been exposed to adverts for HFSS food and drink products and to what extent. Monitoring marketing exposures would inform policymaking and could be helpful in evaluating the impact of policies on reducing children's exposures to digital marketing of HFSS food and drink products. There is a need for an independent monitoring tool that uses a standardised approach to track the implementation and efficacy of potential digital advertising restrictions for HFSS food and drinks to children.

Another key issue for UK policymakers is the advertising of infant formula, which is prohibited in the UK, but many academics, key stakeholders and policy makers suggest the law does not go far enough to curtail coercive marketing by formula companies. Follow-on foods and milks, for example, for older babies and toddlers can be advertised. There is limited monitoring on the extent of digital marketing of breast milk substitutes (BMS) to pregnant and post-natal mothers. There is a need to establish robust and sustainable monitoring mechanisms to implement UK regulations aimed at eliminating inappropriate digital marketing practices surrounding BMS to pregnant and post-natal mothers.

WHO Europe's 'CLICK'¹ framework aims to support monitoring of digital marketing. A pilot study was undertaken using passive metering methods, part of the "investigate exposure" element of this framework, focusing on digital marketing of unhealthy or inappropriate products to children and pregnant/postnatal mothers in the UK. The pilot sought to test the data collection methods, the sample recruitment strategies, the feasibility for a larger study and to obtain preliminary data.

¹ <https://www.unscn.org/uploads/web/file/n.4Joao-CLICK-Monitoring-Framework.pdf>

This report describes exposures of advertising relating to HFSS food and drink products (defined by the Nutrient Profiling Model), breast milk substitute (BMS) products, alcohol products, tobacco products, and gambling products/services.

Aim

The aim of this pilot was to generate a better understanding of the extent and nature of advertising of unhealthy commodities to children, pregnant women, and first-time mothers in the UK, and to assess the feasibility of the RealityMine tool to capture such data.

Methods

The passive metering was provided by RealityMine,² an app that tracks consumers on connected devices, which was licensed to the Office for Health and Improvement Disparities (OHID) by WHO Europe. This UK pilot followed two cohorts, children aged 3-16 years and pregnant women/first-time mothers of children aged ≤ 2 years. Data were collected using Android and Iphone (iOS) mobile devices from the following applications YouTube, Facebook, Instagram, and Twitter. A non-representative sample was taken, which measured exposures to paid-for online advertising.

Ipsos were responsible for gaining ethical approval for the pilot, creating a sampling frame, recruiting participants, coding adverts, and developing an anonymised dataset to share with OHID and the Policy Research Unit for Healthy Weight (HWPRU).

A soft launch was conducted to test the recruitment strategy, technical systems, and coding practices before conducting the main data collection. The soft launch was a non-probability sample of 10 (five children between the ages of 3-16, five pregnant women or first-time mothers with a child ≤ 2 years). Data collection for the soft launch was on Android only and occurred over 14 days for each participant, with rolling start and end dates between 25 November 2021 and 15 December 2021.

The main launch was also a non-probability sample, which aimed to recruit 200 children between the ages of 3-16 years, and 50 pregnant women or first-time mothers with a child ≤ 2 years. Proposed recruitment was not nationally representative, but aimed for balanced representation based on age, sex, ethnicity, social grade, and region; however, data for ethnicity, social grade and region were not provided in the extract. Data collection for the main launch was on Android and iOS and aimed for 80% Android and 20% iOS though device information was not reported. Software was downloaded on participants' primary device, which may have been their own or their parent's device. Android collected primarily social media data (Facebook, Twitter, and Instagram), while iOS collected primarily YouTube data. Android devices had recording issues with YouTube-specific data where adverts were presented but MediaIDs were not detected, therefore there were no usable data. There were other issues with data collection, such as adverts missing media identification numbers, adverts having missing or incomplete meta-data, or adverts not being present in the social media platform library to retrieve relevant data. These issues were present in the soft and main launch.

Data collection occurred over 14 continuous days for each participant, with rolling start dates (running around 60 devices total at any given time) between August 2022 to November 2022.

² <https://www.realitymine.com/realitymeter/>

The RealityMine app collected row-level data on advert exposure with each row representing an exposure of an advert to a participant. Data collected was for paid-for advertising (investment for reach through search/display advertising networks) on specific social media platforms (Facebook, Instagram, Twitter) and YouTube. Owned (where media is owned by the brand) or earned advertising (typically influencers promoting brands and products within their online content) was not collected, nor was paid-for advertising on other social media platforms (e.g. TikTok), general websites, direct email advertising, and other mediums not previously listed. It is unclear if RealityMine identified adverts that were time-limited, such as adverts between Instagram stories which disappear after viewing.

For each advertising exposure, where available, the following data were extracted: the advertising company, the promotion type (paid partnerships or sponsored), media duration, title, description, whether static (without media/moving elements) or video, whether product (a tangible item) or brand (overall image of company or service), and whether the advert was for an included commodity. For each included commodity, additional data were extracted, where available: for HFSS foods, the category of food/drink (e.g., biscuits); for alcohol, the type of drink (e.g., beer), the strength by volume; for tobacco/nicotine products, the product type; for gambling products/services, the category (e.g., online bingo), the presence of odds or VIP schemes, the presence of health information or a warning; for alcohol and gambling products, the nature of the business (e.g., alcohol producer); and for BMS, the type of formula, the format, the type of meal, the type of drink, the type of snack, the type of equipment, and the age range of weaning product. Many of these data fields were empty but captured and extracted data were cleaned and coded.

RealityMine pulled primary fields from the raw metadata from adverts into columns, including the media ID, the media duration and a description. Coding from metadata that was not automatically pulled through was first done by Ipsos and subsequently by analysts within OHID/Global Obesity Evidence and Delivery team, Department of Health and Social Care (DHSC), who also undertook data cleaning. An investigator (SR) on the HWPRU undertook further cleaning, categorisation (including retail/sector and food group categorisations) and labelling of the data, before running descriptive statistics in StataMPv17.

Results

Description of the sample

Tables 1 and 2 show the child and parent samples with age and sex. It is not clear why the numbers in the child and parent samples do not match the intended sampling allocations (200 and 50). Of the female parents, 17 (32.7%) were pregnant at the time of participation.

Table 1. Child sample with age and sex

Sex	N	%	Mean age (SD)	Age range
Female	110	55.8	9.8 (4.1)	3-16
Male	87	44.2	8.9 (4.0)	
Total	197	100	9.4 (4.1)	

Table 2. Parent sample with age and sex

Sex	N	%	Mean age (SD)	Age range
Female	43	81.1	31.2 (3.8)	25-39
Male	10	18.9	31.9 (6.3)	20-40
Total	53	100	31.3 (4.3)	20-40

Advertising exposures

There were 106,443 total advertising exposures (2,896 and 103,547 from the soft and main launch respectively). Of all exposures, 5,905 were for food/drink products/brands, 1,409 were for alcohol products/brands, 20 were for nicotine products/brands, 1,654 were for gambling products/brands, and 299 were for breast milk substitute products/brands (BMS).

However, 186 exposures were coded as being for both food and alcohol products, and 27 exposures as both alcohol and gambling products; reviewing other variables that provided further detail indicated these exposures could be exclusively categorised as alcohol and gambling respectively. Table 3 shows the exclusive categorisations, which have been used in subsequent descriptive analysis.

Table 3. Proportion of exposures for included commodities

Relevant advert category	N	% relevant adverts	% all adverts
Food/drink	5719	63.0	5.4
Gambling	1654	18.2	1.6
Alcohol	1382	15.2	1.3
BMS	299	3.3	0.3
Nicotine	20	0.2	0.0
Total included commodities	9074	100	8.5
Total advertising exposures	106443	-	100

Table 4 shows the proportion of relevant adverts by category for children and parents, by age and sex. Females aged 11-16 and young adult males saw the greatest proportion of food/drink

product adverts. Older adult males saw the greatest proportion of alcohol adverts, males aged 11-16 saw the greatest proportion of nicotine product adverts. Interpretation of these findings is limited as platform use could not be compared by demographic group (iOS only recorded exposures on YouTube, while Android recorded exposures for social media platforms). There were also no data on social media use/duration. Children aged 3-10 were exposed to a high proportion of gambling and BMS adverts but we should be careful in how we interpret the data around children's exposures. We don't know whether the child or the parent was using the device, or whether the targeted advert was the result of the child or parents' media use. There were also very small numbers for some groups, particularly adult males. All adults were either pregnant mothers or parents of small children aged 0-2; comparisons are based on small numbers (pregnant women n=17, parents of small children n=36), but adult females were exposed to BMS adverts to a greater extent than males in terms of proportions (Table 4) and exposures per person (Table 5).

Table 4. Proportion of exposures for included commodities by age and sex groups

	Age	Sex	N	Advert category				
				Food %	Alcohol %	Nicotine %	Gambling %	BMS %
Child	3 - 10	Male	53	55.5	14.7	0.1	24.1	5.7
		Female	57	61.3	20.3	0.0	16.8	1.5
	11 - 16	Male	34	66.9	12.5	1.3	18.5	0.8
		Female	53	76.3	8.9	0.1	13.8	1.0
Parent	20 - 29	Male	3	76.5	19.6	0.0	3.9	0.0
		Female	15	69.2	10.8	1.1	13.7	5.1
	30 - 40	Male	7	68.9	23.9	0.0	6.8	0.4
		Female	28	65.7	17.9	0.0	10.1	6.3

Table 5 shows the mean number of exposures per person for relevant groups and total exposures. Overall, males aged 3-10 years were exposed to the most adverts.

Table 5. Mean exposures for included commodities and all exposures per person by age and sex groups

	Age	Sex	N	Advert category (N)					Included commodities (N)	All exposures (N)
				Food	Alcohol	Nicotine	Gambling	BMS		
Child	3 - 10	Male	53	34.2	9.0	0.0	14.9	3.5	61.7	566.9
		Female	57	25.0	8.3	0.0	6.8	0.6	40.8	558.5
	11 - 16	Male	34	14.7	2.7	0.3	4.1	0.2	21.9	255.8
		Female	53	20.7	2.4	0.0	3.8	0.3	27.2	400.8
Parent	20 - 29	Male	3	13.0	3.3	0.0	0.7	0.0	17.0	300.3
		Female	15	25.2	3.9	0.4	5.0	1.9	36.4	406.1
	30 - 40	Male	7	24.7	8.6	0.0	2.4	0.1	35.9	432.6
		Female	28	10.5	2.9	0.0	1.6	1.0	15.9	164.3

Where recorded (5161/9074 - missing data were across all included commodities), BMS adverts were found to have the highest proportion of product adverts, while gambling adverts had the highest proportion of brand adverts, although proportions were based on low numbers,

particularly for BMS (Table 6). Overall, 38% of adverts were for brands, which is broadly consistent with evidence suggesting 40% of food advertising is for brands.³

Table 6. Proportion of exposures that were for product vs brand for included commodities

	Product		Brand	
	N	%	N	%
Food/drink	1923	64.4	1064	35.6
Gambling	514	48.6	544	51.4
Alcohol	686	67.1	337	32.9
BMS	70	75.3	23	24.7
Nicotine	-	-	-	-
Total	3193	61.9	1968	38.1

Where recorded (2752/9074 – there was no obvious pattern to the missing data and occurred across all platforms), 57.9% of relevant adverts were in video rather than static format; food/drink products had the largest proportion of adverts that were in video format, when compared to other relevant advert types (Table 7).

Table 7. Proportion of exposures that were for static vs video for included commodities

	Static		Video	
	N	%	N	%
Food/drink	580	32.4	1212	67.6
Gambling	331	60.1	220	39.9
Alcohol	206	59.9	138	40.1
BMS	43	66.2	22	33.8
Nicotine	-	-	-	-
Total	1160	42.2	1592	57.8

Where recorded (1718/9074), alcohol products had the longest advert duration (20 seconds) compared to other relevant adverts (Table 8).

Table 8. Duration of exposures for included commodities

	N	Duration (secs)	
		Mean	SD
Food/drink	1395	15.8	59.9
Gambling	193	16.3	9.5
Alcohol	107	20.1	28.7
BMS	12	13.5	3.5
Nicotine	11	12.4	7.3
Total	1718	16.1	54.6

³ <https://www.statista.com/statistics/1415176/food-ad-spend-product-type-uk/>

Of the most frequently recorded companies or brands (>20 exposures), supermarkets and retailers accounted for the most frequent exposures (Table 9); by category (Figure 1) and for supermarkets (Figure 2).

Table 9. Companies/Brands that accounted for the most frequent exposures (>20)

Brands/companies	Exposures N
Amazon (including Fresh, Prime)	175
Ocado	160
Marks and Spencer	154
Asda	150
Morrisons	141
Tesco	118
Sainsbury's	108
McDonald's	79
Waitrose & Partners	52
Cadbury UK	49
Coca-Cola	48
Las Iguanas	46
Pret A Manger	45
TCS London Marathon	42
KFC	41
Uber Eats	41
CLIF Bar	39
HelloFresh	38
Beefeater	32
Costa Coffee	32
Gopuff - Grocery Delivery	30
Just Eat	27
Lidl	24
Influencer	24
Starbucks	24
Follow	23
Order YOYO	21
Ella's Kitchen	20

Figure 1. Most frequently recorded exposures by category/sector

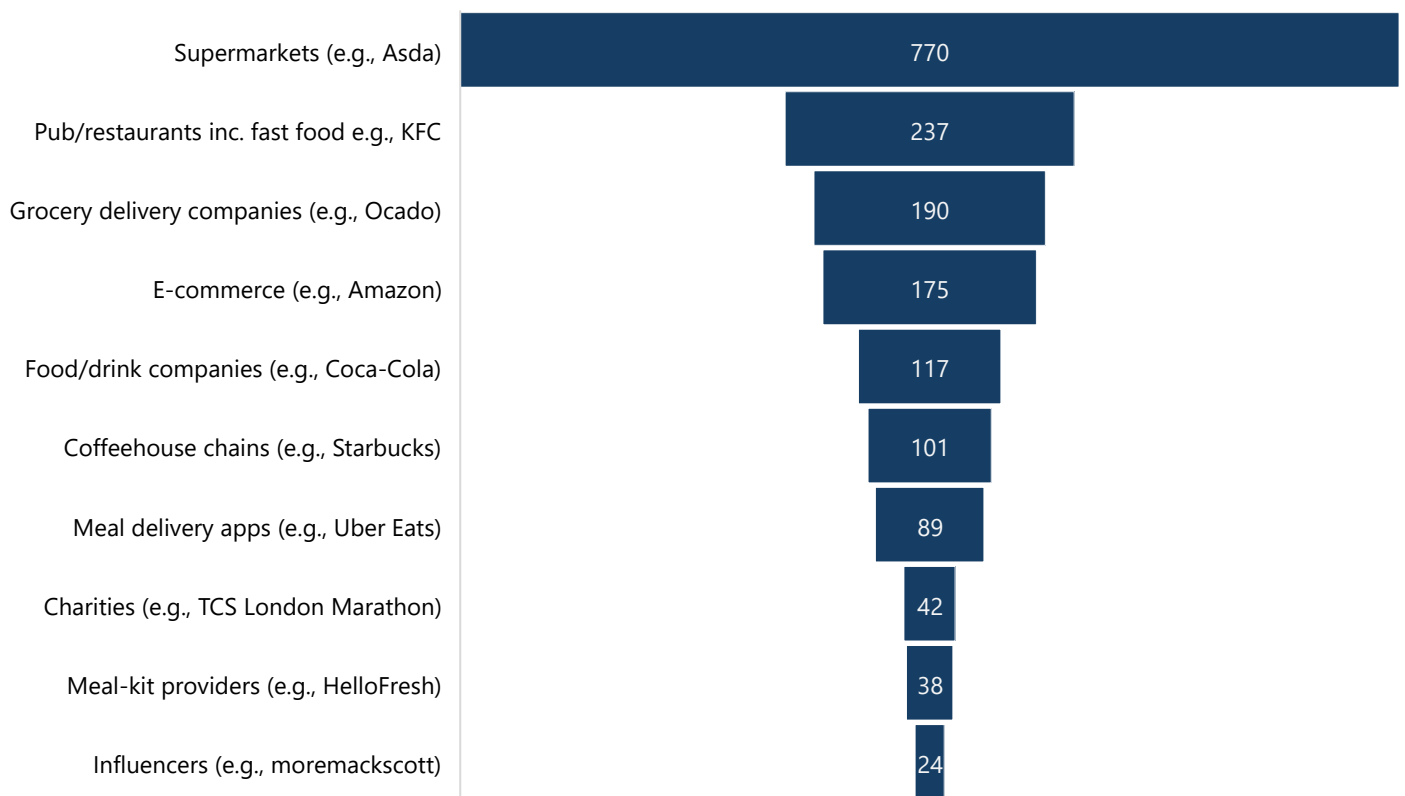
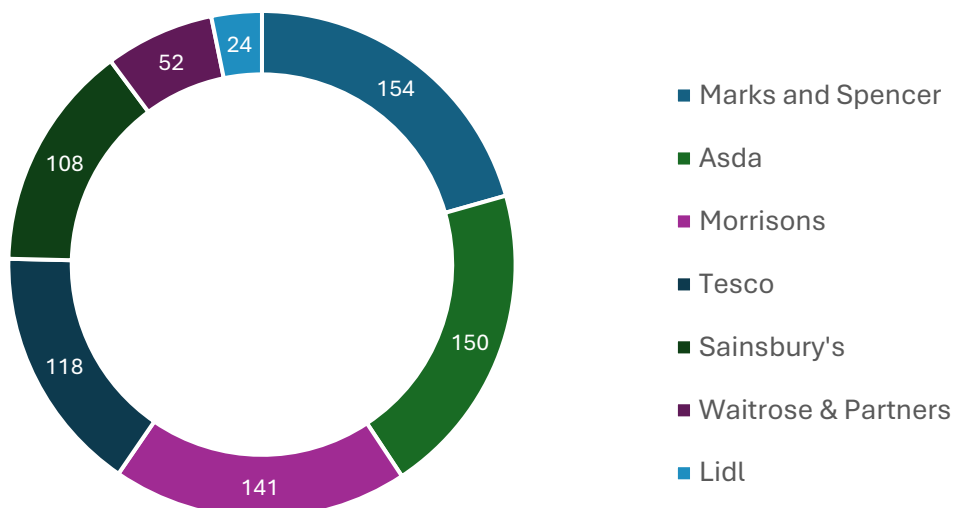


Figure 2. Proportion of exposures by supermarket chain



Of all exposures, Facebook comprised the biggest proportion (42.4%, n=43,955), followed by YouTube (33.9%, n=35,149), Instagram (16.8%, n=17,462) and Twitter (7.0, n=7,214); however, not all platforms were captured by both Android and iOS. Assuming YouTube was exclusively captured by iOS and Facebook, Instagram and Twitter by Android; 33.9% of all exposures were captured by iOS and 66.1% by Android. Of included commodities, YouTube had the biggest proportion of food/drink product adverts, Instagram had the biggest proportion of alcohol product adverts, Twitter had the biggest proportion of gambling adverts; frequencies of nicotine and BMS adverts by platform were small (Table 10).

Table 10. Proportions of included commodities by platform

	Food/drink		Alcohol		Nicotine		Gambling		BMS		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Facebook	2383	58.9	803	19.8	9	0.2	731	18.1	122	3.0	4048	45.2
YouTube	1841	81.2	116	5.1	11	0.5	287	12.7	12	0.5	2267	25.3
Instagram	888	68.2	277	21.3	0	0.0	114	8.8	24	1.8	1303	14.5
Twitter	535	39.8	168	12.5	0	0.0	509	37.9	132	9.8	1344	15.0
Total	5647	63.0	1364	15.2	20	0.2	1641	18.3	290	3.2	8962	100

Of the 3,109 food advert exposures that were categorised, 216 (6.9%) were coded as 'other/not in scope of HFSS advertising restrictions'. Of the remaining 2,893, the proportion of food adverts for product categories are shown in Table 11 and Figure 3. These categories included brand-categorised adverts, where a brand sold one food/drink group exclusively (e.g., Coca-Cola).

Table 11. Frequency and proportion of exposures for food and drink categories

Food and Drink Groups	N	%
Ready meals, pizza	653	22.6
Out of home meals	604	20.9
Chocolate and confectionery	480	16.6
Sugar sweetened beverages	457	15.8
Cakes, biscuits, ice cream, desserts	286	9.9
Potato snacks	229	7.9
Cereal and morning goods	116	4.0
Breaded/battered products	18	0.6
Yoghurts	50	1.7
Total	2893	100

Figure 3. Proportion of exposures for food and drink categories

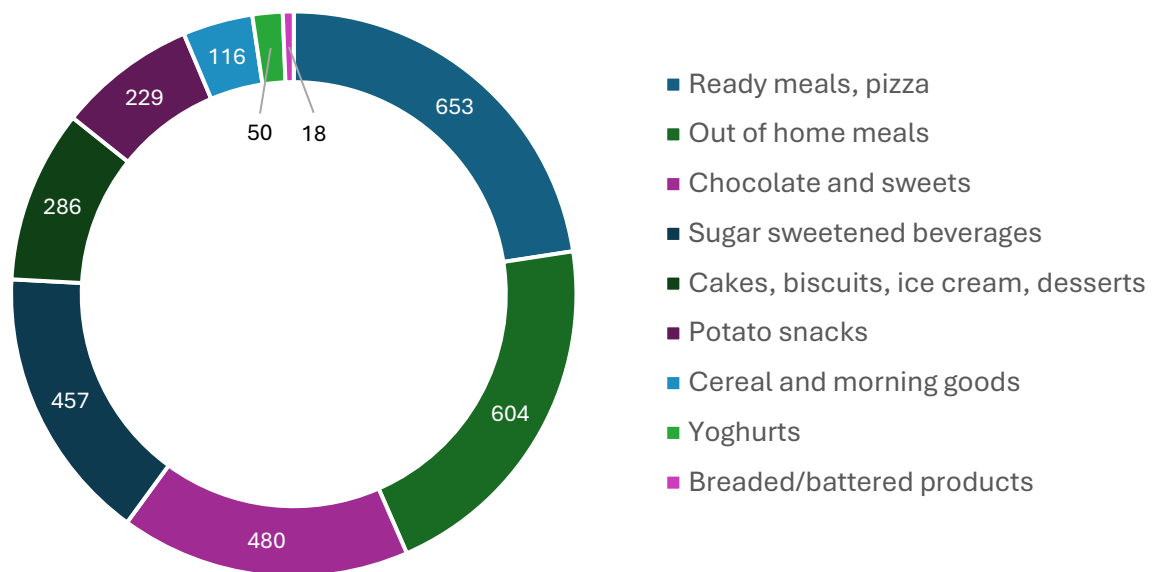


Table 12. Frequency and proportion of exposures for alcohol categories

Alcohol Groups	N	%
Spirits	487	47.6
Beer	252	24.6
Wine	113	11.0
Other	97	9.5
Alcopops	52	5.1
Cider	23	2.3
Total	1024	100

Table 13. Frequency and proportion of exposures for gambling products and services

Gambling Advert Groups	N	%
Casino	478	43.2
Bingo	312	28.2
Sports	189	17.1
Other	114	10.3
Videogame/loot	13	1.2
Total	1106	100

Table 13. Frequency and proportion of exposures for BMS products

BMS Advert Groups	N	%
Equipment	79	66.4
Baby meal	34	28.6
Formula milk	5	4.2
Baby finger food	1	0.8
Total	119	100

There were a number of further variables relating to BMS that had high proportions of missing data, coding or had ambiguous categories. These included formula type (five exposures categorised), formula format (ambiguous categories i.e., yes/no), meal type (mixed fruit/veg n=27, savoury meal n=4), baby drink type (none categorised), finger-food type (fruit bars n=8, rice cakes n=3), equipment type (bottles n=53, preparation machine n=9), and the age range for the weaning product (1 year n=1, 18 years n=10).

Limitations and Data Issues

Sampling and data collection

- Sampling was non-probability, which is convenient but the sample was not representative of target groups.
- Owing to sampling methods, more than one child from the same family may have participated in the study.
- There is a major limitation that neither Android or iOS recorded data comprehensively from all platforms (iOS recorded primarily YouTube and Android primarily social media platform). The type of device used by participants was also not indicated meaning the resulting data is potentially incomplete. Comparison between participants and platforms could not be drawn.
- During the soft launch, Android devices had YouTube-specific data recording issues where YouTube adverts were presented but MediaIDs were not detected. Most advert exposures on YouTube during the soft launch did not contain descriptive information; therefore, entries in the dataset were mostly blank. These issues resulted in changes to the main data collection, where both iOS and Android devices were used to capture data from different platforms. Android collected primarily but not exclusively social media data (Facebook, Twitter, and Instagram), while iOS collected primarily but not exclusively YouTube data.
- There were no data on social media use or duration meaning advertising exposure per time period of use could not be determined.
- Care should be taken in interpreting the high exposure of adverts to children, especially young children given that mobile devices they have access to may also be used by parents, carers and/or siblings. There were no data to indicate who was using the device. The high rate of exposure of alcohol and BMS adverts to young children suggest either children were using their parents' devices (and received adverts targeted to their parents) or the advertising was contextual and targeting a mixed audience (e.g., if a child was watching sports and exposed to gambling adverts).
- There were other issues with data collection, such as adverts missing media identification numbers, adverts having missing or incomplete meta-data, or adverts not being present in the social media platform library to retrieve relevant data. These issues were present in the soft and main launch.

Coding and analysis

- Coding for relevant adverts was undertaken using free text fields and searches of relevant terms. This process was not comprehensive so we cannot be confident that all relevant adverts were captured. Food advertising may be underrepresented in these data.
- There was inconsistency and a lack of information around methods of coding. Coding was undertaken by Ipsos and OHID analysts with a mixture of automated and manual coding. The Ipsos approach and methods are unclear.
- HWPRU researchers were unsure how brands were defined and coded.
- It is challenging to know how to interpret the data on young children as it is highly likely that devices were also being used by parents.

- Analyses for BMS was particularly limited owing to low numbers; only 87 were categorised as milk, meal or equipment (of those, 79 were for equipment).
- Descriptive analyses have limited utility given low numbers in some categories.
- Companies and/or brands were only counted where their name was present/recorded explicitly, meaning they were not systematically or comprehensively counted. For some brands, products were listed separately; for example, Diet coke and Buxton were listed separately to Coca-Cola.
- There were adverts that were of interest but did not meet the inclusion criteria for a relevant advert. For example: If an advert was non-relevant but mentioned relevant products (e.g., Shell petrol advertising Waitrose wine); or if an advert used a relevant food product as part of a recipe (as detailed in the metadata).

This study was funded by the NIHR Policy Research Programme (PRP-PRU-02-15 Healthy Weight). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.