

# Adult diet in England: where is more support needed to achieve dietary recommendations?

Dianna Smith (✉ [d.m.smith@soton.ac.uk](mailto:d.m.smith@soton.ac.uk))

University of Southampton <https://orcid.org/0000-0002-0650-6606>

Christina Vogel

MRC Lifecourse Epidemiology Unit

Monique Campbell

University of Southampton

Nisreen Alwan

University of Southampton Faculty of Medicine

Graham Moon

University of Southampton

---

## Research

**Keywords:** Diet, health inequality, small area estimation, England

**Posted Date:** September 3rd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-67773/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

**Background:** Small-area estimation models are regularly commissioned by public health bodies to identify areas of greater inequality and target areas for intervention in a range of behaviours and outcomes. Such local modelling has not been completed for diet consumption in England despite diet being an important predictor of health status. The study sets out whether aspects of adult diet can be modelled from previously collected data to define and evaluate area-level interventions to address obesity and ill-health.

**Methods:** Adults aged 16 years and over living in England. Consumption of fruit, vegetables, and sugar-sweetened beverages (SSB) are modelled using small-area estimation methods in English neighbourhoods (Middle Super Output Areas [MSOA]) to identify areas where reported portions are significantly different from recommended levels of consumption. The selected aspects of diet are modelled from respondents in the National Diet and Nutrition Survey using pooled data from 2008-2016.

**Results:** Estimates indicate that the average prevalence of adults consuming less than one portion of fruit, vegetables or 100% juice each day by MSOA is 6.9% (range of 4.3 to 14.7%, SE 0.06) and the average prevalence of drinking more than 330ml/day of SSB is 11.5% (range of 5.7 to 30.5%, SE 0.03). Credible intervals around the estimates are wider for SSB consumption. The results identify areas including regions in London and urban areas in the North of England which may be prioritised for targeted interventions to support reduced consumption of SSB and/or an increase in portions of fruit and vegetables.

**Conclusion:** These estimates provide valuable information at a finer spatial scale than is presently feasible, allowing for within-country and locality prioritisation of resources to improve diet. Local, targeted interventions to improve fruit and vegetable consumption such as subsidies or voucher schemes should be considered where consumption of these foods is predicted to be low.

## Introduction

Promoting a healthy diet rich in fruit and vegetables is one of the most widespread public health campaigns globally, with five portions a day or at least 400 grams of fruits and vegetables (commonly referred to as 5-A-Day) recommended in the UK and most other high-income countries (1). Although the UK's Eatwell Guide emphasises and strongly recommends adherence to the 5-A-Day target, these guidelines also state that foods high in free sugars (inclusive of sugar sweetened beverages-SSBs) should not exceed five percent of total daily energy intake(2). The latter recommendation is based on the strong association between SSB consumption and weight gain, particularly in adult women (3), and the higher consumption of SSB amongst younger adults compared to older populations (4). A substantial body of literature has identified poor health outcomes, including obesity and type two diabetes, associated with low consumption of fruits and vegetables and high levels of sugar, especially sugar sweetened beverage (SSB) consumption (5-7).

Despite considerable research showing the health benefits of greater fruit and vegetable consumption(8), only 30% of the UK population (9) and 26% of the English population (10) achieve the 5-A-Day target. The UK adult population is also said to consume almost three times the recommended amount of free sugars, with sugar sweetened beverages (SSBs) such as soft/fizzy drinks accounting for a large proportion of the daily sugar intake of both adults and children(11). Among lower income groups the consumption of produce is considerably below recommendations and SSB intake above recommendations which is likely contributor to their poorer health outcomes(12). In addition, recent studies have found obesity and type 2 diabetes to be two of the most prominent risk factors associated with COVID-19. Diet, especially those low in fruits and vegetables and high in saturated fat, refined carbohydrates and sugar (typically referred to as a Western dietary pattern) is identified as one of the main drivers of this risk (13). These findings which when coupled with sub-optimal diets in the UK, help to explain the Government's urgency in tackling the diet and obesity "time bomb" through the recently announced "Better Health" campaign(14).

To address health inequalities, interventions must be considered to support optimal diets for all people. Dietary quality is one area where health policy has the scope to influence substantial changes to overall diet patterns in the population (15). Few studies have explored the potential impacts of higher fruit and vegetable consumption on population health specifically, however one recommended policy intervention for increasing these elements of diet is subsidising fruit and vegetables (15). In June 2020 the House of Lords published recommendations to increase the value of Healthy Start vouchers as part of plans to support a diet high in fruit and vegetables (16). Alongside SSB consumption, it would be beneficial to understand how consumption of low levels of fruit and vegetables vary geographically in the adult population, to consider appropriate interventions to improve dietary behaviours in areas where there is the greatest potential impact.

There is evidence that price increases may reduce consumption of less healthy items, with sugar often a priority area for these interventions. The introduction of a tax on SSB in parts of the US reduced population-level consumption with small levels of taxation,(17) however, a recent systematic review concludes that price increases need to be at least 20% to achieve a reduction in consumption(15). Public health campaigns on reducing sugar consumption have put pressure on the government to enforce reformulation of food products to reduce sugar intake in the UK and led to the creation of a 'sugar tax' implemented in 2018 (18). This tax affects food and drink with high levels of free sugars, with one key category of foods including SSB. If the SSB producers do not reformulate products to reduce the amount of sugar, then a tax is applied to the producers with a higher tax on drinks with 8 grams or more per 100ml and a lower level of tax on drinks with 5 grams or more per 100ml (19). The sugar tax has resulted in some reformulation.

Though models have been developed to explore the impact of a sugar tax on population consumption, to date there are no efforts to model local-level consumption and predict which areas may see the biggest reduction in consumption following the 2018 sugar tax. Now that public health is ultimately the responsibility of local authorities underneath the Health and Social Care Act of 2012(20), such knowledge empowers local public health strategy. In addition, consideration could be given to more effective targeting of national voucher schemes such as Healthy Start, which provide funds to purchase fresh or frozen fruit and vegetables to lower income households,(21) or extending local interventions including Rose vouchers which give an additional £3/week/child on top of Healthy Start vouchers(22). Further modelling of diet following modifications to the provision of vouchers, perhaps by increasing the voucher value as

inflation impacts on food prices, could provide an indication of the anticipated impact on population health(23). For such scenario modelling to be feasible, local-level data or reliable estimates must be available.

Evaluations of taxation policy to promote healthier dietary choices recommend a combination of increasing the cost of less healthy options and subsidising healthy items including fruit and vegetables(24). Large-scale models of populations have been used to estimate the cost implications of diet change (25), public health impact of taxation on consumption of fast food in the US (26) and SSB in English Local Authority Districts(27); here we extend the modelling approach to examine smaller areas and identify neighbourhoods where targeted interventions are most needed to reduce inequalities in population-level health outcomes related to diet.

The aim of this paper is to model four variables of diet: i) daily fruit intake, ii) daily vegetable intake, iii) daily fruit & vegetable, and iv) daily SSB intake. Nationally representative data of adult diet ( $\geq 16$  years) from the National Diet and Nutrition Survey were used to model at the neighbourhood level, specifically the Middle Super Output Area (MSOA, n=6791) level which approximates a large neighbourhood with an average population of 7000 and sits within Local and Unitary Authorities. This creates data which may provide estimated baseline consumption to later evaluate implemented interventions where local data are not collected. The estimates will also aid local policy decision-making such as prioritising areas for healthy food subsidies or planning restrictions for unhealthy food outlets.

## Methods

Dietary data were collated from the National Diet and Nutrition Survey (NDNS) datasets from 2008-2016 (28). The NDNS is the foremost diet survey in the UK which captures data on all food and drink consumed over a four-day period, providing a reliable measure of diet quality compared to diet screeners included in less-specific health surveys (29). The pooling of eight years of data provides a larger sample size for the modelled estimates. This annual cross-sectional survey samples individuals aged 1.5 years and over living in private households, with a target of 1000 respondents each year. Portions of SSB were recorded in millilitres (ml) per day. Data on portions of fruit and vegetables consumed included components of prepared dishes as well as pieces of fruit or vegetables whether fresh, frozen or canned. Fruit juice (100%) is only counted as one portion regardless of the total amount consumed each day (28). We modelled the consumption of SSB, portions of fruit, portions of vegetables and the total portions of fruit, vegetables and 100% fruit juice (up to 150 ml) separately. Diet outcomes were modelled for those age 16 years and above to correspond with health outcomes measured in national surveys including the Health Survey for England (10).

The sociodemographic variables most strongly associated with consumption of each of the four dietary variables were identified through chi-square correlation analysis. The sociodemographic variables tested included age, sex, ethnicity, National Statistics Socioeconomic Classification (NSSEC), marital status, housing tenure and educational qualifications. These variables were considered for two reasons. First, they are associated with diet quality in individuals in other datasets and research (11). Second, they are all also collected as part of the decennial Census in the UK, which is required for the method of small-area estimation used here. Variables were considered highly correlated if  $p < 0.001$ . The coding of each variable was specified as follows: SSB were coded at none, up to 330ml (one reflect can of fizzy drink) of SSB and over 330ml; portions of fruit, vegetables and combined fruit, vegetables and 100% fruit juice consumed are coded as 0-0.99, 1-4.99, 5+ portions per day. Each portion is based on an 80g portion size. Binary logistic regression models for four outcomes (< 1 portion fruit, <1 portion vegetables, < 1 portion fruit, vegetable and juice or >330 ml SSB) was the next stage of selecting appropriate predictor variables for these relatively poor diet behaviours, with the final step the creation of multivariate logistic regression models for each of the four diet outcomes to identify the optimal combination of predictor variables in each instance. The results from the multivariate logistic regression model were also used in the creation of credible intervals around the estimated prevalence of diet behaviours, as described later in this section (30). The scale of these estimates, MSOAs, are similar in size to the primary sampling units used to select areas and populations to sample for the national surveys. This MSOA scale is the level used in other small-area estimates created by the Office for National Statistics (31). MSOAs nest within wards, enhancing the utility of the estimates for public health strategists who use ward-level data.

The method of creating the local diet behaviour estimates is one approach to small-area estimation, spatial microsimulation. The spatial microsimulation algorithm used is based on iterative proportional fitting (IPF) with full details of the method provided elsewhere(32). This method has been used successfully in England to estimate common mental disorders and alcohol consumption (33) in England and New Zealand (34). IPF approaches have been applied successfully in England to estimate obesity(35) and smoking (36) and to estimate a range of health outcomes in Scotland (37). The local population data for MSOAs was based on the age-sex mid-year population estimates produced by the Office for National Statistics (ONS) for 2018 (38) and the 2011 Census data (39). Individual socioeconomic classification was identified using the National Statistics Socioeconomic Classification (NSSEC) scheme (40). Data on NSSEC, ethnicity, marital status, housing tenure and educational qualifications were scaled to reflect the 2018 age-sex population estimates. Ethnicity was coded into four categories as White, Asian, Black and Other or Mixed. NSSEC data were collapsed into four categories (High, Moderate, Low, Unclassified [including unemployed or unclassified]) from the original eight, following guidance from ONS (40). Marital status was coded as single, married/civil partner, divorced/widowed/other. Housing tenure is coded as owned, social rented, private rented. Finally, educational qualifications are recoded to degree or above, below degree level/other, no qualifications, in full time education.

The IPF approach calculates the probability of each eligible respondent from NDNS (complete case analysis, those who provided all necessary diet and demographic data) to live in each MSOA across England. The probability is calculated by matching respondents from NDNS to known residents in the Census based on the demographic variables which were shown to best predict the selected dietary behaviours from the correlation analysis. The probabilities for each person to live in each area are then summed to total the local population (based on the ONS 2018 data), and the associated dietary behaviours are attached to these new area populations. The result of the summed probabilities provides an estimated prevalence for each diet behaviour in each MSOA. Figure 1 shows a worked example of the reweighting and subsequent rescaling of the new probabilities to fit ONS population data. For this paper, this process was repeated four times, once for each dietary variable.

To provide validation of model estimates, previous researchers have aggregated estimates to a higher-level geography where the prevalence of an outcome or behaviours is known (33) or compared estimates to related outcomes, such as the prevalence of diabetic amputations to validate diabetes prevalence estimates (41). Earlier models of SSB consumption in England ran a series of sensitivity analyses to provide more confidence in estimates (27). Another consideration is the internal goodness-of-fit in the modelled populations, with the preference being standardised Total Absolute Error (TAE) between the observed and estimated demographic variables used to predict the outcomes (here, age, sex, and so on)(42). Observed values are gained from the Census and ONS and the mean TAE for each constraint variable is calculated across the 6791 MSOAs in England. As there are no readily available comparable datasets of SSB or fruit and vegetable consumption in England at the MSOA level (29), the focus of model fit is on the internal comparisons.

Credible intervals for the estimated prevalence of diet behaviours were calculated using the outputs from multiple regression models for each of the diet outcomes and using a multilevel modelling framework. Details of this process are explained fully in Whitworth et al 2017 (30). Briefly, the data from NDNS has already been analysed in single-level multiple regression models for each diet outcome. The associated sampling data from NDNS for each respondent (sample cluster, region) is used to build a multilevel model for each outcome where individuals are nested within the sample clusters, nested within regions. The outputs from these multilevel models indicate the variance at level two (sample clusters). This captures the uncertainty present at the cluster level, units that are of a similar size to an MSOA. The variance is replicated stochastically to represent a range of 1000 possible values for each MSOA and the upper and lower credible intervals are taken from the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles of this distribution.

All statistical analysis was completed using SPSS v 25 except for the credible interval multilevel models which were calibrated using MLwiN (3.01). The spatial microsimulation model is written in R code and the maps were created using ArcGIS 10.6.

## Results

The estimated patterns of SSB, fruit, vegetables and the portions of fruit and vegetables, including 100% juice, consumed each day from the NDNS datasets are summarised in Table 1 along with respondent demographics. Here, the predictor variables that were most significant following the binary and multivariate logistic regression analysis are retained (sex, age, marital status, housing tenure and educational qualifications) while those which were less significant were dropped from the analysis (NSSEC, ethnicity). Tenure and educational qualifications were highly correlated so were not used in the same models.

Table 1: Summary of respondents and crosstabulations with outcome variables. P values for Chi-square tests shown

	SSB consumption		n=5160			Total portions (F&V)		n=5150			Fruit portions		n=5157	
	no SSB	1-329ml	330ml +	p value	< 1 portion	1-4.99	five +	p value	< 1 portion	1-4.99	five +	p value	< 1 portion	
<b>Sex</b>				<0.001				0.353				<0.001		
Male	51.1	32.2	16.7		6.4	69.9	23.7		59.9	38.1	2		18.1	
Female	55	35.6	9.4		6.5	68.1	25.4		53	45.4	1.6		18.1	
<b>Age</b>				<0.001				<0.001				<0.001		
16-24	26.2	41.9	31.9		10.8	78.5	10.7		76.3	23.1	0.6		30.8	
25-34	43.5	38.5	18		6	70.1	23.9		60.8	38	1.1		18.1	
35-44	52.2	37.3	10.5		5.4	70.4	24.2		57.8	41.1	1.1		15.8	
45-54	61.3	32.3	6.5		5.6	66.9	27.5		51.3	46.2	2.5		14.1	
55-64	68	28	4		4.1	61.2	34.7		42.4	54.9	2.7		12.9	
65+	70.6	27.1	2.3		5.7	64.2	30.1		44.4	53.1	2.5		15.9	
<b>Education</b>				<0.001				<0.001				<0.001		
Degree or higher	57.7	34.1	8.2		1.8	57.8	40.4		38.2	59.4	2.3		8.1	
Below degree level	52.8	34.3	12.9		5.5	71.1	23.4		55.7	42.6	1.7		17.1	
No qualifications	67.3	27.2	5.5		11.2	70.7	18.2		64.1	34	1.8		24.6	
Full time student	27.4	44.5	28.1		10.4	76.4	13.2		73	26.3	0.7		30.6	
<b>Housing tenure</b>				<0.001				<0.001				<0.001		
Owned	54.7	35.2	10.1		4.4	66.7	28.8		49.1	48.8	2.1		14.1	
Social Rent	52.5	30.3	17.2		13.1	75	11.9		73.8	25.1	1.1		31	
Private rent	48.5	34.4	17.1		7.5	70.8	21.6		64.6	34.5	0.9		21.1	
<b>Marital status</b>				<0.001				<0.001				<0.001		
Single	38.6	38	23.3		9.3	72.6	18.1		67.9	30.9	1.2		25.8	
Married	60.7	33.5	5.7		3.4	65.3	31.3		45.8	52.1	2.2		11.2	
Separated/divorced/widowed	65.5	28.6	6		7.2	68.9	23.9		53.9	44.2	1.9		19.1	

The results of the chi-square test for significance show the potential predictor variables (age, sex, marital status, tenure and educational qualifications) were all strongly correlated with diet outcomes with the exception of sex and fruit and vegetable portions and vegetable portions alone. All significant associations were confirmed with the logistic regression models. On the basis of these observed relationships, the SAE models for total fruit and vegetable portions and the model for vegetable portions only were created using age, tenure and marital status as the predictors. The model estimating the portions of fruit used age, tenure, marital status, and sex to predict this behaviour. The final model, estimating SSB consumption, used age, sex, marital status, and educational qualification.

The results in Table 2 summarise the least healthy dietary behaviour categories including high SSB consumption (more than one 330ml portion/day), lower fruit and vegetable consumption (less than one portion of fruit or one portion of vegetables, or less than one combined fruit, vegetables and 100% fruit juice portion/day).

Table 2: Mean prevalence of consumption for each diet variable by MSOA

	Mean (%)	Minimum (%)	Maximum (%)	Std. Error	Std. Dev
>1 portion of fruit	54.25	44.15	76.10	0.06	5.24
< 1 portion of veg	17.32	12.78	30.74	0.04	2.93
< 1 portion of F&V	6.85	4.34	14.67	0.02	1.61
330 + ml SSB	11.47	5.69	30.51	0.03	2.63

The mean prevalence of adults consuming less than one portion of fruit, vegetables or 100% fruit juice is low overall, at 6.9% (range 4.3-14.7) of the population age 16+ years living in an MSOA. Comparing low levels of consumption of fruit or vegetables, estimates based on the NDNS dataset predict that fewer adults eat, on average, less than one portion of vegetables (17.3%, range 12.8-30.7) compared to fruit (54.3% [range 44.1-76.1]) per day. On average,

11.5% of the adult population in an MSOA are high SSB consumers (at least 330ml/day [range 5.7-30.5]). There is greater variation in the three single component variables (fruit, veg or SSB only) compared to the combined fruit and vegetable consumption (Table 2).

The results for each of these estimated dietary variables were mapped to MOSAs in quintiles with the results showing a distinct geography of higher SSB consumption, with greater prevalence (13.4-30.5% of the adult population drinking more than 330m/day) in urban areas (Figure 2). Several rural areas along the south coast, Midlands and North also have relatively high prevalence of higher SSB consumption.

The estimated prevalence of adults eating less than one portion of fruit, vegetables, or 100% fruit juice each day varies compared to the pattern of higher SSB consumption (Figure 3). Notably, northern regions and the eastern boroughs of London have higher rates of the poorest level of fruit and vegetable consumption. This is defined as MSOAs where 8.2-14.7% of the adult population consume less than a portion of fruit or veg/day, which is top quintile of MSOAs for this measure of diet quality.

A subset of areas showed low fruit and vegetable consumption predicted alongside higher prevalence of greater SSB consumption. There are 806 MSOAs (11.9% of England) where more than 8.4% of the population are estimated to consume less than a portion of fruit, vegetables or juice a day and in each of these MSOAs more than 13.4% of the population are estimated to drink >300ml of SSB a day. These MSOAs include areas from Birmingham, Leeds, Liverpool and Manchester in the north and Southampton, Portsmouth and Brighton in the south. The areas included in this combined higher risk group are predominantly urban, as clearly illustrated in the London inset map.

The variables that were used to estimate the dietary behaviours such as age and sex are also estimated in the model to provide an indication of internal model goodness-of-fit by comparing to the known population distribution from the ONS mid-year estimates. The mean TAE (mTAE) for each variable was low. For SSB the mTAE were 2.1 E-15, 0.00067, 0.009 and 0.014 with age being the variable with the highest average percent error. In the total portion variables, the mTAE values were 2.01 E-15, 0.0001 and 0.002 with age showing the most error. The values were similarly very low for portions of fruit and portions of vegetables (data not shown). This internal error is well within the suggested limits for internal goodness-of-fit, where error is less than 20% in at least 80% of the study areas(43).

Data on adult fruit and vegetable consumption at the 2011 local authority district level (n=326, with an average of 21 MSOAs in each) was available from the Active Lives Survey 2016/17 (44), however, analysis of this survey data showed that respondents were far more likely to have five portions of fruit and vegetables a day (58.1% of those age 16+ years) than respondents from the NDNS (24.7%). This disparity in descriptive statistics between the NDNS used to model the outcomes and data collected in the Active Lives survey would incorrectly indicate that there was a high proportion of error in the estimated data based on NDNS respondents. The Active Lives Survey asks only the screener question about the portions of fruit and vegetables which is prone to error when a more in-depth dietary assessment method is used such as the four-day food diary in NDSNS. This may explain the disparity between the reported prevalence of people achieving five a day between the two surveys.

The credible intervals reflect the uncertainty in the estimated values. They are wider for the outcomes where the original regression model results showed a weaker relationship between the predictor variables and the diet outcome. An example of this is the relatively tight credible intervals for low levels of fruit and vegetable consumption (less than one portion/day). On average, the estimated prevalence of adults age 16+ who did not have at least one portion a day was 6.9% with an average credible interval of 3.1% - 14.6%. In contrast, the average value for drinking at least 330ml of SSB/day was 11.5% with a wider credible interval of 5.4-22.8% (average values). The credible intervals are not symmetrical around the point estimates as they have a lower limit of zero, so the upper limit will be greater from the point estimate to reflect this value constraint (Figure 5).

## Discussion

The estimated diet behaviours in the adult population show that in a small number of MSOAs, two poor diet behaviours (more than 330ml/SSB/day or less than portion of fruit, vegetable or 100% fruit juice/day) are more prevalent. Relatively high SSB consumption (at least 330 ml/day) was more common than estimated prevalence of low daily fruit and vegetable intake (< 1 portion of fruit, vegetables or 100% juice/day). These areas could benefit from closer consideration by local public health teams for area-based interventions to promote healthier diet behaviours.

Results of the SSB model highlight inner London and Northern areas where there is a higher proportion of more socially and materially deprived populations. Here, more of the adult population are likely to experience a greater impact of the sugar tax on the cost of their beverage consumption. The data from this study show high (>330 ml/day) SSB consumption is more common in areas with greater numbers of younger adults as seen in earlier research (4) and higher proportions of full-time students.

The impact of the sugar tax, a regressive approach to taxation would usually be argued to affect the lowest income groups most. Indeed, the data from the NDNS data confirms higher prevalence of >330ml consumption in the lowest social group as classified by NSSEC (data not shown). Although the sugar tax has promoted some reformulation of products to reduce sugar, in some situations the higher costs for producers is passed on to consumers. The results of this study suggest that areas of inner London, Leeds, Birmingham and Liverpool would experience the greatest impact. These results agree with the geographic pattern identified by Collins et al (27) when reductions of kcal/day following a 20% duty on SSB were estimated at the LAD level. One further option to reduce adult SSB consumption is to target retailers in the areas where high consumption is estimated to be greatest to promote non-SSB alternatives.

Greater prevalence of low consumption of fruit and vegetables show a similar geography to high SSB with overlapping areas in London, however, there are parts of Liverpool, Manchester and Sheffield where a higher proportion of the adult population are estimated to consume <1 portion of fruit, vegetables or fruit

juice each day. There are 806 MSOAs identified where the model estimates predict both relatively high SSB consumption and low fruit and vegetable consumption. Most of these areas are located in the northern part of the country and inner London, which is supported by previous research showing greater inequalities in health between the North and South of England (45, 46). Stronger local interventions to actively support more fruit and vegetable consumption, such as local voucher schemes or targeted redistribution of fresh produce to these neighbourhoods through existing organisation (FareShare) may have the greatest benefit for addressing diet-related health inequalities in these populations.

These data will be particularly helpful for Local Authorities who hold responsibility for localised public health campaigns (20), monitoring trends over time or proposing interventions. In combination with enhanced provision of fruit and vegetable vouchers or redistribution in areas identified by the modelling, local interventions may be prioritised to localities where the impact can be greatest. This illustrates one strength of small-area estimation in the absence of comprehensive local surveys, as they allow public health teams in local authorities to identify priorities for local population health inequalities. For this reason, Public Health England regularly commissions small-area estimation models of behaviours such as smoking and outcomes like obesity or depression; this model is an extension to include additional behaviours which would benefit from focused interventions. The results of this model will provide health strategists with data to support local intervention development or help evaluate the potential effectiveness or best location for interventions. Previous scenario modelling using spatial microsimulation provided insight for the optimal locations of smoking cessation services (36) and a similar approach may be applied for diet interventions as described above.

The estimates for diet are built using the demographic profiles of respondents to the NDNS survey and attributing the observed diet behaviours to local area populations based on the known population characteristics from the Census and ONS updates. One limitation to this method is the local population data is not as current as we would prefer. In addition, this model assumes that all people replicate the behaviours of respondents to the NDNS. As the NDNS data are used for annual updates of population-level diet behaviours, this is a consistent assumption and the same SAE method has been successful in modelling other health outcomes and behaviours such as obesity (35) common mental disorders(33) and smoking (34).

The width of credible intervals reflect the relatively small sample size of respondents from the NDNS, which was the most appropriate dataset to capture a range of diet outcomes; the Health Survey for England (average annual sample of approximately 10,000) stopped consistently collecting data on dietary behaviours beyond a screener to collect consumption of fruit and vegetables in 2012 (29). In addition, NDNS collects data using a four-day food diary which is more reliable than a screener based on the last 24 hours of consumption as seen in Health Survey for England and the Active Live Survey (47).

The results of this study provide an indication of geographic areas where diet-related policy interventions such as voucher schemes or redistribution services may have the greatest impact; urban areas in the north of England and some southern coastal cities should be prioritised for dietary improvements. Population-based policy interventions such as taxes to reduce SSB consumption can be evaluated using these local estimates as a baseline. There is also capacity to use the results for optimally locating targeted interventions such as increased Healthy Start or similar voucher values in selected areas. This modelling is a valuable tool which may be easily adopted by local authorities as they maintain responsibility for prioritising health campaigns to reduce health inequalities in their populations, and further validated as more data becomes available. The methods used in this study can be replicated in any country with national survey data on dietary variables.

## Declarations

**Conflict of Interest:** The authors declare no conflict of interest. CV has a non-financial research relationship with a food retail company and maintains independence in all evaluation activities. This article, however, is not related to this relationship.

**Funding:** MC is funded by the ESRC South Coast Doctoral Training Programme (ES/P000673/1).

**Ethics:** Ethical approval for this project was granted by the University of XXXXXX ethics committee (blinded for peer review).

**Consent for Publication:** Not applicable

**Availability of data and materials:** The datasets generated and/or analysed during the current study are available in the UKDS repository (<https://ukdataservice.ac.uk/>) and on our project website at xxxxxxxx (blinded for peer review).

**Acknowledgements:** Not applicable

**Author Contribution:** DS conceived the study, completed the analysis, drafted the manuscript. CV and NA helped to refine the research question and study scope. MC assisted with the analysis and prepared figures. GM contributed to the analysis. All authors edited subsequent versions of the manuscript and read and approved the final version.

## Abbreviations

GIS: Geographic Information Systems

IPF: Iterative Proportional Fitting

MSOA: Middle Super Output Area

NDNS: National Diet and Nutrition Survey

NSSEC: National Statistics Socioeconomic Classification

ONS: Office for National Statistics

SAE: Small-Area Estimation

SSB: Sugar-Sweetened Beverages

TAE: Total Absolute Error

## References

1. World Health Organisation. Healthy Diet 2020 [Available from: <https://www.who.int/news-room/fact-sheets/detail/healthy-diet>].
2. Public Health England. The Eatwell Guide. 2018.
3. Lavin R, Timpson H. Exploring the Acceptability of a Tax on Sugar-Sweetened Beverages: Brief Evidence Review. Centre for Public Health: Liverpool John Moores University; 2013 April 2013.
4. Gibson S, Shirreffs SM. Beverage consumption habits "24/7" among British adults: association with total water intake and energy intake. Nutr J. 2013;12(9):1–9.
5. Malik VS, Popkin B, Bray G, Despres J, Hu F. Sugar-sweetened beverages and the risk of metabolic syndrome and type 2 diabetes: a meta-analysis. Diabetes Care. 2010;33:2477–83.
6. Popkin B. Sugary beverages represent a threat to global health. Trends in Endocrinology Metabolism. 2012;23(12):591–3.
7. British Dietetic Association. Sugar:Food Fact Sheet. 2017.
8. Department for Environment Food and Rural Affairs. Family Food 2015. National Statistics; 2017.
9. Public Health England. New National Diet and Nutrition Survey shows UK is eating too much sugar, saturated fat and salt. 2014.
10. NHS Digital. Health Survey for England 2015 Trend Tables. 2016.
11. Public Health England  
Public Health England. Agency FS. National Diet and Nutrition Survey. Results from Years 7–8 (combined) of the Rolling Programme (2014/15 to 2015/16). 2018.
12. Pechey R, Jebb SA, Kelly MP, Almiron-Roig E, Conde S, Nakamura R, et al. Socioeconomic differences in purchases of more vs. less healthy foods and beverages: Analysis of over 25,000 British households in 2010. Soc Sci Med. 2013;92(0):22–6.
13. Butler MJ, Barrientos RM. The impact of nutrition on COVID-19 susceptibility and long-term consequences. Brain Behav Immun. 2020;87:53–4.
14. Department for Health and Social Care. New obesity strategy unveiled as country urged to lose weight to beat coronavirus (COVID-19) and protect the NHS. 2020.
15. Wright A, Smith KE, Hellowell M. Policy lessons from health taxes: a systematic review of empirical studies. BMC Public Health. 2017;17(1):583.
16. Select Committee on Food Poverty Health and the Environment. Hungry for change: fixing the failures in food. House of Lords; 2020.
17. Colantuoni F, Rojas C. The impact of soda sales taxes on consumption: Evidence from scanner data. Contemp Econ Policy. 2015;33:714–34.
18. Scarborough P, Adhikari V, Harrington RA, Elhussein A, Briggs A, Rayner M, et al. Impact of the announcement and implementation of the UK Soft Drinks Industry Levy on sugar content, price, product size and number of available soft drinks in the UK, 2015–19: A controlled interrupted time series analysis. PLOS Medicine. 2020;17(2):e1003025.
19. HM Treasury. Soft Drinks Industry Levy: 12 things you should know 2016 [Available from: <https://www.gov.uk/government/news/soft-drinks-industry-levy-12-things-you-should-know>].
20. Heath S. Local authorities' public health responsibilities (England). House of Commons Library; 2014.
21. NHS England. Healthy Start Vouchers: What to buy 2016 [Available from: <https://www.healthystart.nhs.uk/healthy-start-vouchers/what-to-buy-with-the-vouchers/>].
22. Alexandra Rose Charity. Rose Vouchers 2017 [Available from: <http://www.alexandrarose.org.uk/rose-vouchers>].
23. Lee Y, Mozaffarian D, Sy S, Huang Y, Liu J, Wilde PE, et al. Cost-effectiveness of financial incentives for improving diet and health through Medicare and Medicaid: A microsimulation study. PLOS Medicine. 2019;16(3):e1002761.
24. Mytton OT, Clarke D, Rayner M. Taxing unhealthy food and drinks to improve health. BMJ: British Medical Journal. 2012;344.
25. Scarborough P, Kaur A, Cobiac L, Owens P, Parlesak A, Sweeney K, et al. Eatwell Guide: modelling the dietary and cost implications of incorporating new sugar and fibre guidelines. BMJ Open. 2016;6(12):e013182.
26. Zhang D, Giabbani PJ, Arah OA, Zimmerman FJ. Impact of different policies on unhealthy dietary behaviors in an urban adult population: an agent-based simulation model. Am J Public Health. 2014;104(7):1217–22.
27. Collins B, Capewell S, O'Flaherty M, Timpson H, Razzaq A, Cheater S, et al. Modelling the Health Impact of an English Sugary Drinks Duty at National and Local Levels. PLOS ONE. 2015;10(6):e0130770.
28. NatCen Social Research UCLU. National Diet and Nutrition Survey Years 1–6, 2008/09–2013/14. [data collection]. 8th Edition. UK Data Service. SN: 6533. MRC Elsie Widdowson Laboratory, University College London Medical School, 2017.
29. Campbell M, Smith D, Baird J, Vogel C, Moon EG. A critical review of diet-related surveys in England, 1970–2018. Archives of Public Health. 2020;78(1):66.

30. Whitworth A, Carter E, Ballas D, Moon G. Estimating uncertainty in spatial microsimulation approaches to small area estimation: A new approach to solving an old problem. *Comput Environ Urban Syst.* 2017;63:50–7.
31. Office for National Statistics. Model-Based Estimates of households in poverty for Middle Layer Super Output Areas, 2013/14, Technical Report. 2017.
32. Smith DM, Clarke GP, Harland K. Improving the synthetic data generation process in spatial microsimulation models. *Environment Planning A.* 2009;41(5):1251–68.
33. Riva M, Smith D. Generating small-area prevalence of psychological distress and alcohol consumption: validation of a spatial microsimulation method. *Soc Psychiatry Psychiatr Epidemiol.* 2012;47(5):745–55.
34. Smith DM, Pearce JR, Harland K. Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? An example of smoking prevalence in New Zealand. *Health Place.* 2011;17(2):618–24.
35. Edwards KL, Clarke GP. The design and validation of a spatial microsimulation model of obesogenic environments for children in Leeds, UK: SimObesity. *Soc Sci Med.* 2009;69(7):1127–34.
36. Tomintz MN, Clarke GP, Rigby JE. The geography of smoking in Leeds: estimating individual smoking rates and the implications for the location of stop smoking services. *Area.* 2008;40(3):341–53.
37. Campbell M, Ballas D. SimAlba: A Spatial Microsimulation Approach to the Analysis of Health Inequalities. *Frontiers in Public Health.* 2016;4(230).
38. Office for National Statistics. Middle Super Output Area Population Estimates. In: Office for National Statistics, editor. 2018.
39. Office for National Statistics. 2011 Census Data 2011 [Available from: <https://www.ons.gov.uk/census/2011census/2011censusdata>.
40. National Statistics. National Statistics Socioeconomic Classification User Manual. Palgrave macmillan; 2015.
41. Congdon P. Estimating diabetes prevalence by small area in England. *J Public Health.* 2006;29(1):71–81.
42. Voas D, Williamson P. Evaluating Goodness-of-Fit Measures for Synthetic Microdata. *Geographical Environmental Modelling.* 2001;5(2):177–200.
43. Birkin M, Clarke GP, Clarke M. Retail geography and intelligent network planning. Chichster: Wiley; 2002.
44. Sport England. Active Lives Survey, 2016–2017. UK Data Service.
45. Bambra C. Health Divides: Where You Live Can Kill You. Bristol: Policy Press; 2016.
46. Buchan IE, Kontopantelis E, Sperrin M, Chandola T, Doran T. North-South disparities in English mortality 1965–2015: longitudinal population study. *Journal of Epidemiology and Community Health.* 2017.
47. Cade JE, Warthon-Medina M, Albar S, Alwan NA, Ness A, Roe M, et al. DIET@NET: Best Practice Guidelines for dietary assessment in health research. *BMC Med.* 2017;15(1):202.

## Figures

$$n_i = w_i s_{ij} / m_{ij}$$

where

$n_i$  = the new weight of individual i

$w_i$  = is the old weight of individual i

$s_{ij}$  = is the element of the MSOA ONS table for individual i and variable j

$m_{ij}$  = is the element of the NDNS data table for individual i and variable j

NDNS respondents		
ID	sex	Weight ( $w_i$ )
001	Male	0.5
002	Male	0.2
003	Female	0.6

MSOA 001 ( $s_{ij}$ values)	
35 males	
40 females	

NDNS totals ( $m_{ij}$ values)	
1995 males	
2220 females	

Survey respondents: new weight calculation			
ID	sex	weight	$w_i s_{ij} / m_{ij} = n_i$
001	Male	1	$0.5 \times 35 / 1995 = 0.0088$
002	Male	1	$0.2 \times 35 / 1995 = 0.0035$
003	Female	1	$0.6 \times 40 / 2220 = 0.0108$

Then the new weights are scaled to fit known ONS populations:

$$n_i^c = w_i^c \times \left( t_{i,c}^c / t_{i,c} \right)$$

$$N_{i,c}^c = \sum n_i^c$$

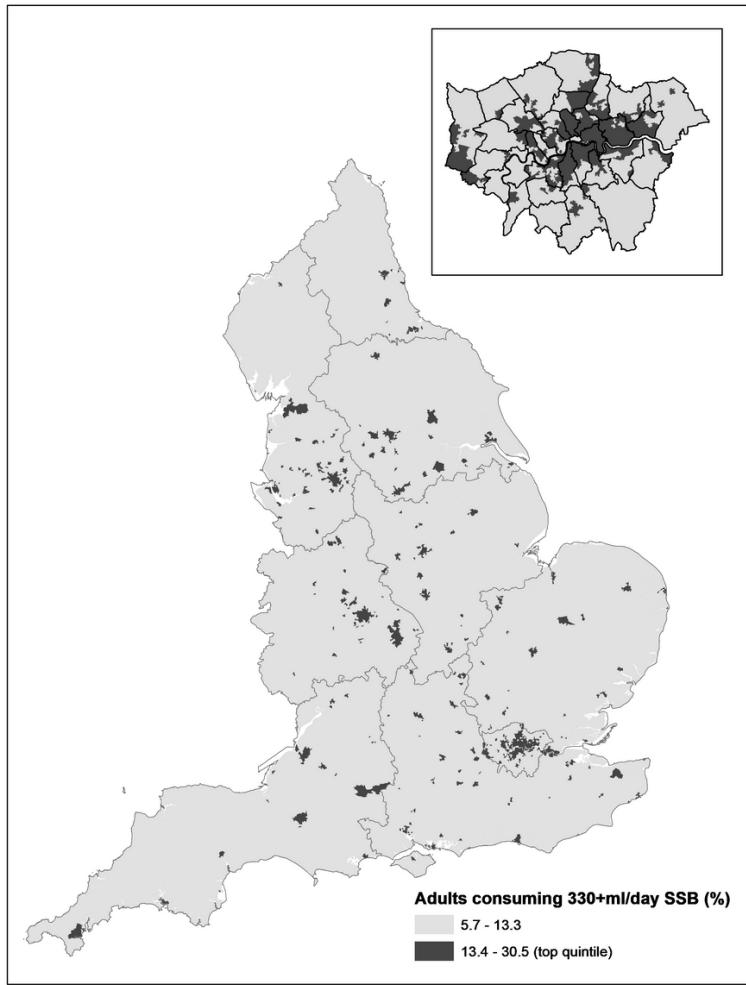
with scaling factor

$$t_{i,c}^c / N_{i,c}^c$$

In this example, where  $n_i$  is new weight for individual in MSOA<sub>001</sub>,  $N_o$  the total weight for all individuals in MSOA<sub>001</sub>,  $c$  the variable subcategory (e.g. male or female for the sex category),  $t_{i,c}^c$  the total population in variable c in MSOA o (ONS totals used here),  $t_{i,c}$  the total population in variable c in survey data (NDNS used here) and  $w_i$  is new weight for an individual. The sum of all the new weights after the sex variable is reweighted is calculated for MSOA<sub>001</sub> ( $N_o$ ), and will sum to the total population of MSOA<sub>001</sub>. Then, the sum of the new weights for all males is calculated ( $N_{i,c}^c$ ). The ratio of the number of males reweighted by the model ( $t_{i,c}^c$ ) to the reweighted population for that area ( $N_{i,c}^c$ ) is used as a scaling factor on all of the new weights generated by this variable reweighting process.

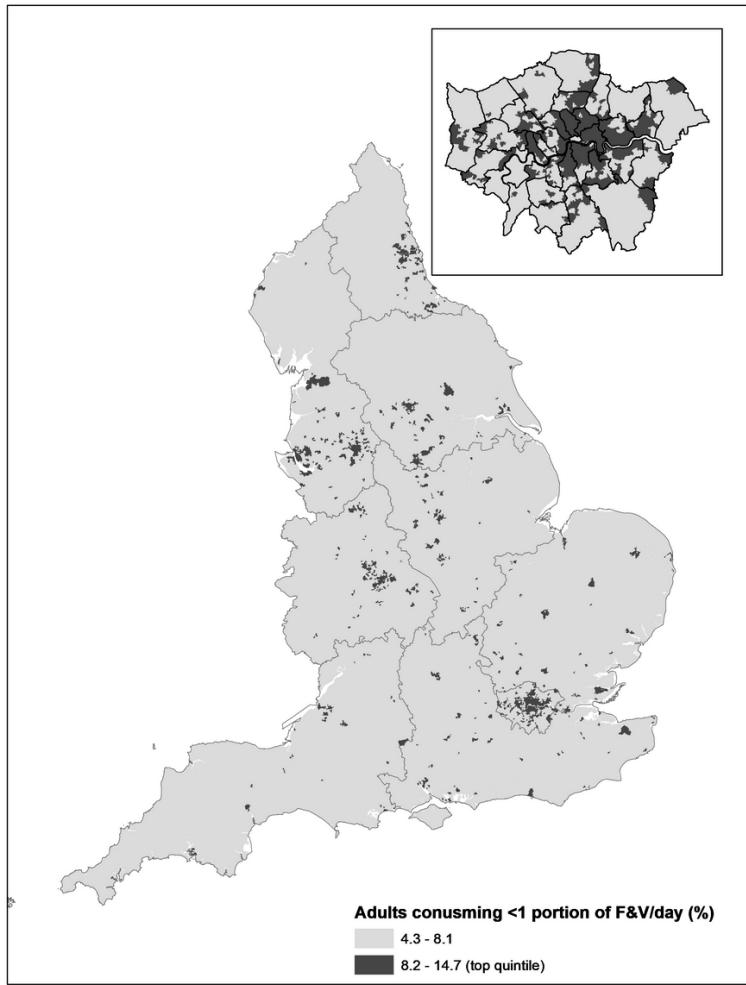
**Figure 1**

Worked example of IPF routine for SAE



**Figure 2**

Estimated prevalence consuming more than 330ml SSB per day for adults by MSOA (2018 population) with inset of London



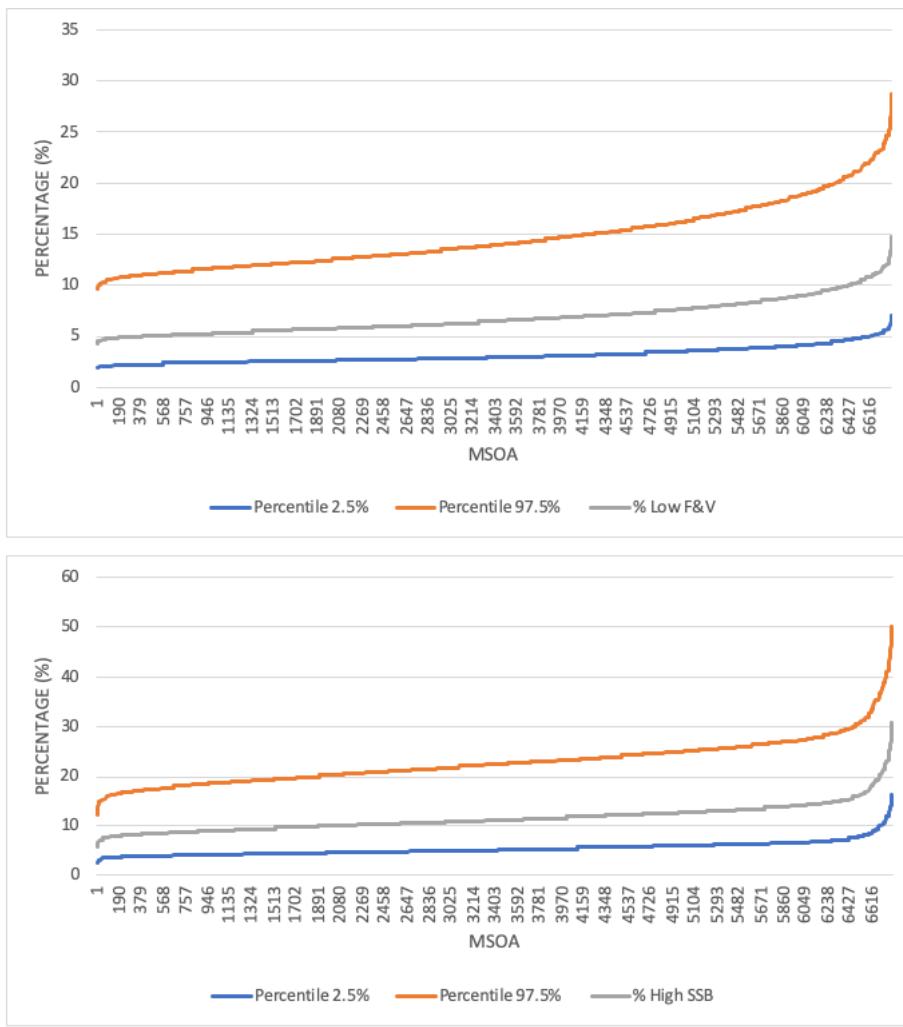
**Figure 3**

Estimated prevalence of consuming less than 1 portion of 'five a day' including 100% juice for adults by MSOA (2018 population) with inset of London



**Figure 4**

Combined poor diet prevalence (quintiles of highest SSB consumption and highest percentage of the population eating less than one portion of fruit and vegetable a day) by MSOA (2018 population) with inset of London



**Figure 5**

Credible intervals around a sample of the estimates of SSB and overall fruit and vegetable consumption