

Finlay, A, Jones, A, Thorp, P, Putra, IGNE, Polden, M, Adams, J, Brealey, J and Robinson, E

Accuracy of menu calorie labelling in the England out-of-home food sector during 2024: Assessment of a national food policy

<https://researchonline.ljmu.ac.uk/id/eprint/27298/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Finlay, A ORCID logoORCID: <https://orcid.org/0000-0002-7509-4772>, Jones, A ORCID logoORCID: <https://orcid.org/0000-0001-5951-889X>, Thorp, P, Putra, IGNE, Polden, M, Adams, J, Brealey, J and Robinson, E Accuracy of menu calorie labelling in the England out-of-home food sector during 2024:

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

Accuracy of menu calorie labelling in the England out-of-home food sector during 2024: Assessment of a national food policy

Amy Finlay^{*a}, Andrew Jones^b, Paula Thorp^a, I Gusti Ngurah Edi Putra^c, Megan Polden^{d,e},
Jean Adams^f, Jane Brealey^a & Eric Robinson^a

^aDepartment of Psychology, University of Liverpool, Liverpool, UK

^bSchool of Psychology, Liverpool John Moores University, Liverpool, UK

^cDepartment of Public Health, Policy and Systems, University of Liverpool, UK

^dDepartment of Primary Care and Mental Health, University of Liverpool, UK

^eNIHR Applied Research Collaboration Northwest Coast, Liverpool, UK

^fMRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK

*Corresponding author: Amy Finlay, Office 2.80 Eleanor Rathbone Building Liverpool L69 7ZA. Email: amy.finlay@liverpool.ac.uk

Short title: *Accuracy of calorie labels out of home*

Key words: calorie labelling; food policy; accuracy; out-of-home

Abbreviations: Out of home (OOH), Local authority (LA), Index of Multiple Deprivation (IMD), Inter-Department Business Register (IDBR).



**The
Nutrition
Society**

This peer-reviewed article has been accepted for publication but not yet copyedited or typeset, and so may be subject to change during the production process. The article is considered published and may be cited using its DOI

10.1017/S0007114525105217

The British Journal of Nutrition is published by Cambridge University Press on behalf of The Nutrition Society

Abstract

Mandatory calorie labelling was introduced in out-of-home (OOH) food sector outlets during 2022 in England. Previous research in North America has found labelled energy content can be underestimated for packaged and quick-serve foods, but no study has evaluated the accuracy of out-of-home food sector menu calorie labelling in response to the mandatory policy introduced in England. N=295 menu items from a range of outlet types (e.g. cafes, pubs, restaurants) and menu categories (e.g., starters and sides, main, dessert) were sampled. Bomb calorimetry was used to quantify energy content and the reported energy content on menus was recorded. Consistency of measured energy was assessed by sampling the same items across outlets of the same business (N=50 menu items). Differences between reported and measured energy content were tested through Wilcoxon Signed Rank tests, and a linear model examined correlates of the difference. Mean measured kilocalories (kcal) were significantly lower than reported kcal (-16.70kcal (± 149.19), $V=16920$, $p<.01$, $r=0.182$). However, both over- (23% of menu items) and under-estimation (11%) by >20% of measured energy content was common and the averaged absolute percentage difference between reported and measured values was 21% ($\pm 29\%$). Discrepancy between measured and reported energy content was more common in some outlet types (pubs) and reported energy content was substantially different (>20%) to measured energy content for 35% of sampled menu items. There were significant inaccuracies in reported energy content of calorie labelled menu items in English food outlets subject to mandatory calorie labelling.

1. Introduction

Consumption of food prepared out-of-home (OOH), for example from cafés, restaurants and fast-food outlets, is associated with increased energy consumption(1) and poorer dietary quality(2). It has been estimated that in the UK, 60% of the population purchase OOH food weekly, with the OOH food sector contributing on average 11% of weekly energy intake(3). In April 2022, calorie labelling on food menus was made mandatory in England for all businesses with over 250 employees. This policy was specific to the OOH food sector, and applied to all non-prepacked foods and drinks made for immediate consumption(4). The aim of the mandatory calorie labelling policy was to promote informed food choices and healthier eating behaviour in the OOH food sector(5). Calorie labelling policies have been implemented outside of the UK, specifically in parts of Australia(6), Canada(7) and in the US nationally(8). Evidence is mixed on the impact that calorie labelling has on consumer behaviour(9, 10), but studies tend to suggest that calorie labelling results in reductions in energy content of menu items(11, 12).

Implementation guidance for the 2022 mandatory calorie labelling policy in England(5) states that businesses can calculate the energy content of menu items by various means. Although laboratory analysis is considered the scientific gold standard, this is expensive and may not be feasible for many businesses. Energy content can be estimated using established and accepted food databases (for example, the McCance and Widdowson's Composition of Foods Dataset), where nutritional composition is based on standard recipes(13). A limitation of this method is that values may not be accurate for recipes that deviate from this standard. Energy values can be averaged based on manufacturer's analysis, or calculated using known or average values of ingredients used and imputing them into nutritional calculators such as MenuCal(5). However, often these approaches will not account for cooking method used, which has a meaningful impact on the final served energy content.

A 20% discretion between the reported and calculated energy content is permitted, and necessary to provide businesses flexibility to account for differences in ingredient composition, portion size, or calculation methods. Guidance acknowledges that accurate testing of energy content may not be viable for local enforcement officers(5). Regarding enforcement, this guidance states that "Local authorities have discretion in how they enforce the Regulations"(5). Specific roles of the enforcement officer include checking that calorie labels are present and that the methods used to estimate energy content are appropriate(5).

However, there is little assistance for enforcement officers in assessing the accuracy of calorie labelling and no explicit instruction or suggestion that accuracy of labelling of menu items should be tested.

In the US, businesses appeared initially to have largely complied with the mandatory provision of calorie labelling, with 79% of businesses fully or partially implementing labels in advance of policy enforcement(14) and similarly in the UK, 80% of businesses provided calorie labelling at any point of choice 6 months post-implementation, compared to only 21% who provided this information prior to implementation(15). The food industry has a history of inadequate compliance and exploitation of loopholes in government policy(16-18). This is likely due to the lack of accountability by the food industry surrounding implemented policy and the lack of resources for compliance monitoring(19).

There is evidence that nutritional labels can be prone to some degree of inaccuracy for packaged food products. One study found that the energy content of popular US snacks was on average higher than reported on labels(20). However, for most items (96%), reported energy was within 20% of measured energy content. In Canada, over 1000 items from supermarkets, bakeries and restaurants were tested for their nutritional content(21). For the tested foods, sodium and energy were consistently under-reported. For 14% of foods, the measured energy content exceeded label values by more than 20%. A study conducted in the US assessed the accuracy of nutrition labels for low-energy (<500kcal) restaurant and frozen meals(22). For frozen meals, measured energy content was on average 8% higher than reported and for restaurant meals, energy content was on average 18% higher than reported. Several restaurant foods contained up to twice the reported energy content. Similarly, a US study of restaurant menu items in 2011 found that although average measured vs. reported energy content across menu items was similar, 19% of individual menu items had a measured energy content of ≥ 100 kcal per portion more than reported(23).

No research has to date assessed the accuracy of calorie labels in OOH outlets following the implementation of the mandatory calorie labelling policy in England. As customers are encouraged to pay attention to the nutritional labelling of foods(4), it is expected that the provided information will impact behaviour, and inaccurate nutritional labelling may hinder individual efforts to eat healthier. For example, if consumers use calorie labels to factor foods into their daily energy allowance, then underestimation of calories on food menus would lead to consumers unknowingly consuming excess energy. Additionally, if consumers identify

potential inaccuracies in labelling themselves, then they may lose trust and stop using calorie information in the OOH food sector. Furthermore, some research suggests there is little or no enforcement of the policy requirement of providing accurate calorie information(24). Therefore, in the present study we examined the accuracy of menu calorie labelling in OOH sector outlets subject to the 2022 mandatory calorie labelling law in England. As ingredient and cooking method information from businesses is not publicly available, laboratory analysis of menu items' energy content was used to assess accuracy of menu calorie labelling in the present research.

2. Materials and methods

This observational study was pre-registered on the Open Science Framework <https://osf.io/8tfu4/>.

2.1. Outlet selection

Outlets in two local authorities (LAs) in England were sampled to ensure findings were not area specific. Liverpool (North of England) and Milton Keynes (South of England) were selected to ensure mixed geographical coverage and representation of different quintiles of deprivation (Index of Multiple Deprivation (IMD) used at the LA level). LA IMD quintiles (1-5) were used as a local measure of deprivation with IMD1 reflecting the most deprived areas and IMD5 reflecting the least deprived. LAs selected represent quintile 1 (most deprived – Liverpool) and quintile 3 (medium deprivation – Milton Keynes). However, outlets in both LAs were located across quintiles 1-5 as measured at the Lower Super Output Area.

Previous research that evaluated the impact of the calorie labelling legislation on consumer and business behaviour(25) used the Inter-Department Business Register (IDBR) to identify businesses likely to be subject to the mandatory calorie labelling policy within LAs of interest (data sampled June 2021, list produced Autumn 2020). The IDBR is a list of all UK businesses, their core characteristics, number of employees and principal activities defined using the Standard Industrial Classification. Standard Industrial Classification codes likely to include businesses serving food were identified (see the full list of Standard Industrial Classification codes in Supplementary Material 1). Large businesses (defined as 250+ employees) within the relevant classification codes were then located, and their individual outlets were identified using Ordnance Survey Points of Interest data from September 2020. We used the final list of eligible businesses to guide outlet selection in the two LAs.

Through power calculations, we deemed that a minimum of $n=239$ samples (menu items) ($n=240$ rounded to the nearest multiple of 2) would be sufficient to detect small effect sizes through paired samples tests in our primary analysis (see Supplementary Material 2 for further detail). To obtain 240 samples with 4 samples per outlet (not including items deemed potentially inaccurate, which were analysed separately) required 60 outlets in total, or 30 outlets in each LA. For a sub-sample of outlets ($n=5$ in each LA), we planned to match outlets across LAs and conduct paired sample comparisons to assess consistency within chains. Figure 1 depicts the outlet selection process. To obtain a sample with variation in outlet type and item type, we randomly selected (using the RAND function in excel) $N=6$ unique outlets from each of four main outlet categories: café, restaurant, fast food, pubs. These outlet categories were taken from previous research evaluating the national calorie labelling policy (26). If for any outlet category there were not 6 unique outlets (i.e., there were only multiple outlets from the same chain), additional outlets classed as ‘sport & entertainment’, for example cinemas, were sampled. We randomly selected additional outlets from the full database (irrespective of outlet type) until $N=25$ unique outlets were selected for each LA. Duplicates of businesses across the two LAs were permitted.

To assess consistency in calorie labelling between the same businesses across the two local authorities, $n=5$ (20%) selected outlets in each LA were matched to an additional corresponding outlet from the same business in the other LA where the same items would be sampled. This resulted in an additional $n=10$ outlets sampled and a total of $n=50$ matched pairs for our planned secondary analyses. Additional outlets could only be sampled if they were not already selected. For example, if the randomly selected outlet in LA1 was a Starbucks, an additional matched Starbucks would be identified in LA2 for sampling. In total, we sampled $n=300$ menu items from 60 outlets (30 in each LA). For full sample size and power analysis information see Supplementary Material 2.

2.2. Menu item selection

Prior to data collection, online menus for each outlet were used to identify the items on offer in outlets. For each outlet, one menu item was selected for sampling in each of the following categories:

- Starter/side or prepared drink (randomly selected)
- Main meal (randomly selected)
- Dessert (randomly selected)
- Menu item that that was potentially inaccurate (determined by a research staff member and trained nutritionist assessing the full menu)
- Most popular menu item (determined by asking serving staff on the day of data collection)

In instances where menu categories were not distinct (e.g. starter – main – dessert – side dish), two researchers individually considered the categories provided by outlets and re-categorised items into the above groups. Any differences were resolved through discussion. For some outlets, specifically cafés, the categories outlined above did not encompass the majority of items available, and so prepared drinks were sampled in place of starters/sides. This was the case for $n=7$ outlets from $n=3$ unique businesses.

Items categorised as starter/side or prepared drink, main meal and dessert were numbered and a random number generator was used to select items for sampling prior to data collection. To determine items sampled under the category of potentially inaccurate, two researchers considered the available description of menu items and whether the reported energy content looked inaccurate using menu information alone (either too high or too low). We adopted this approach to reflect how enforcers of the policy may be expected to assess the accuracy of calorie labels on food menus (i.e., relatively subjectively). If there were no items that appeared to be inaccurate at face value (e.g., very high or low energy content based on type of menu item), researchers considered instances where potential variability in serving size (e.g., inconsistency in the serving size of a dish of pasta) could result in differences in energy content from reported. Once each researcher had determined a potentially inaccurate item for each outlet, a final item was decided upon through discussion. Due to the absence of sales data to assess popularity, in each outlet, the researcher asked the serving staff what the most popular item on the menu was, and this item was sampled. If the most popular menu item or potentially inaccurate item from an outlet was a drink, this was eligible for sampling.

For instances where menu items required customisation (e.g. choosing a side to go with a dish), researchers randomly selected from all available customisation options.

2.3. Energy content of sampled items and procedure

The researcher responsible for data collection recorded the energy content reported on menus for each sampled menu item on the day of sampling. Samples were collected (Monday-Thursday) between 11am-7pm during April-May 2024. For each outlet, the researcher responsible for data collection ordered food to dine in. All sampled items ($n=300$ from $n=60$ outlets) were individually weighed and packaged in the restaurant. Calibration weights were used to ensure scales were accurate. All items were sent via a courier to an external laboratory accredited by the UK Accreditation Service (SGS Cambridge) for nutritional analysis through bomb calorimetry and energy (kcal) data were analysed.

2.4. Analysis

2.4.1. Primary analyses:

For all analyses, the potentially inaccurate items were considered separately from all other items tested as we anticipated their inclusion may overestimate average difference between measured and reported energy content. For the main analyses, one item was missing due to lab error (spoiled food due to storage error), and three items were excluded from analysis due to implausibility of energy values. This resulted in $n=236$ samples. Similarly, for potentially inaccurate items, one menu item was missing due to lab error, resulting in $n=59$ samples.

Throughout the results, relative differences are calculated as *measured kcal – reported kcal* for both mean and percentage difference. Absolute differences are calculated as the mean or percentage difference from measured kcal regardless of the direction of the difference, calculated using the ‘abs()’ function in R Studio.

Absolute percentage differences were used to explore whether differences between measured and reported energy content differed from 0% (perfect accuracy) and 20% (permitted inaccuracy). As data were not normally distributed, one-sample Wilcoxon signed rank tests were conducted against 0 and 20% to examine if the observed absolute difference values were significantly different from 0% and 20%.

We also conducted Wilcoxon signed rank tests to assess whether there was a significant difference between reported energy and lab-measured energy content (relative difference).

2.4.2. Secondary analyses:

For the subsample of restaurants where a corresponding outlet from each LA was tested, resulting in $n=50$ matched pairs, we examined whether the measured energy content differed between the two locations using a paired Wilcoxon signed rank test. We created a new variable [measured kcal – reported kcal] and fitted a linear regression model to examine potential predictors (e.g., outlet characteristics) of the relative difference between measured and reported energy. We explored outlet type, menu category, item energy content, IMD score for outlet and LA. If any predictor variables were significant, we conducted subgroup analyses to explore differences. We also used this method to explore predictors of absolute mean percentage difference.

Results for primary analyses were considered significant at $p < .05$. To account for multiple comparisons, results for secondary and any exploratory analyses were considered significant at $p < .01$. For explanation of deviation to the pre-registered analysis plan, see Supplementary Material 3.

3. Results

A total of $n=295$ menu items were sampled from $N=60$ outlets across two LAs in England. Figure 2 displays the mean reported and measured energy content for each category of item. Menu items were categorised by outlet type (pubs, cafes, fast food, restaurants and entertainment) and menu category (starter/side, main, dessert, popular item, potentially inaccurate item and drinks). Mean energy and percentage differences between reported and measured energy for each category are reported in Table 1. For the most popular menu items and items that were potentially inaccurate, approximately 2/3 of sampled items were main meal items. Overall, the mean relative difference between reported and measured energy (measured-reported) was -16.70 kcal (± 149.19) and -8.75% ($\pm 35.27\%$) expressed as a relative percentage of the menu item energy content. This was 98.76 kcal ± 112.88 and 21.30% ($\pm 29.45\%$) when the absolute energy and percentage difference were examined (i.e. mean difference regardless of direction of difference).

In the main analyses (all menu items except for those that were potentially inaccurate), 56% of items had a lower measured energy content than reported on menus. For over one third (35%) of items, the energy content reported on menus was outside of the 20% leeway permitted; for 66% of these (23% of total) this was due to reported energy being substantially higher than measured and for 31% of these (11% of total) this was due to reported energy being substantially lower than measured. Findings were similar in the potentially inaccurate items (See supplementary Material 4).

Two Wilcoxon signed rank tests found that the absolute percentage difference between reported and measured energy ($n=236$) was significantly greater than 0% ($V=27261$, $p<.001$) but not greater than 20% ($V=0$, $p>0.99$). However, reported and measured energy content were significantly different from each other, whereby on average energy content reported on menus was significantly greater than measured energy content ($V=16920$, $p<.01$, $r = 0.182$). A boxplot of mean differences is shown in Figure 3. Reported and measured energy were not significantly different for potentially inaccurate items.

To explore identified significant predictors, paired samples tests (Wilcoxon signed ranks) were conducted which examined differences between measured vs reported energy for cafes, restaurants and pubs, and the starter and popular item categories. Pubs were the only outlet type where reported and measured energy were significantly different from each other ($V=943$, $p<.01$). For pubs, mean reported kcal were 725.00 ± 326.89 and mean measured kcal were 674.10 ± 338.36 . There was no significant difference between measured and reported energy for the starter or popular menu categories. Mean differences, percentage differences and the proportion of items outside of the 20% leeway for each outlet type and item category are shown in Table 1.

4. Discussion

This study of large business owned OOH food outlets subject to the mandatory calorie labelling law in England during 2024 found that averaged across all sampled menu items, the energy content (kcal) reported on OOH outlet menus was significantly greater than measured energy. However, it was common for reported energy content of menu items to be greater and less than measured energy content. For 35% of menu items, reported energy was over 20% greater or less than measured energy. Pubs had a particularly pronounced difference between reported and measured energy content compared to other outlet types. The measured energy content of the same menu items sampled from different outlets of the same chain were largely similar. Collectively, these findings suggest that reported energy content for significant numbers of menu items in the OOH sector in England differ substantially to measured energy content.

Expressed as a relative calorie value and a percentage of each item's measured energy content, average reported energy of items was 17kcal and 9% greater than measured energy content. However, expressed as an absolute value (size of deviation from measured energy content irrespective of direction), the mean difference between measured and reported energy was 99 kcal and 21%, due to both over- and under-estimation of measured energy content being common (although over-estimation of energy content was observed more frequently). This suggests that on average, calories reported for menu items in the English OOH food sector may differ substantially to their measured energy content. The overall pattern of results differs somewhat to research conducted in the US (22) which identified that measured energy content tended to be greater than reported on labelling and on restaurant menus. This difference in our findings may be a result of differences in types of outlets or menu items examined. Previous research has largely sampled packaged foods which may be less prone to human variability during the preparation and cooking process compared to items made in the OOH food sector. There may also be differences in the sampling methods used that could inadvertently impact the measured energy content of items (e.g. sampling and packaging hot food vs. sampling cold, packaged foods). One US study considered OOH foods, but limited investigation to menu items under 500kcal(22), which may help to explain differences as our findings would suggest there is less scope for inaccuracy at lower levels of energy content. However, our findings are comparable to those of a study conducted in Canada sampling a range of outlet types (21), whereby items had an overall greater mean labelled energy content

than mean measured energy content, with evidence of both over and under-estimation of reported energy content.

In the previous North American studies discussed, reported energy content was largely within the 20% leeway of measured content (70%(20), 86%(21), 59%(22)) and this is comparable to the present study (65%). However, across all of these and the present study, a proportion of menu item labels appear to be significantly different from measured values (>20%). There are multiple possible explanations for the observed error in reported energy content in the present and prior research, for example errors in ingredient information or human error in calculation. It would now be informative to study what the causes of inaccuracy are. In England, the calorie labelling legislation outlines several accepted methods to calculate energy content, and for this research we chose to use the gold standard of measurement (i.e. bomb calorimetry)(27). It is plausible that the observed error comes from businesses utilising alternative methods (e.g., calculation using standardised ingredient information) to estimate energy content(5).

While five different outlet types were explored, the largest difference between mean reported and measured energy content was observed for pubs. This outlet type had 46% of items outside of the 20% legislation leeway, while all other outlet types had between 30-36% of items outside of the leeway. Observed differences between measured and reported energy may have been greater in pubs due to differences in the use of ingredients which are more prone to fluctuations in water and energy content through cooking. This, along with potentially greater autonomy for chefs in how foods are prepped and cooked (e.g. compared to fast food outlets or the café chains sampled, where many foods are pre-packaged, and serving sizes more standardised) may explain why discrepancies were greater in pubs. Alternatively, differences between outlet types, and individual outlets may exist if there are differences in how the energy content of menu items is calculated.

When items deemed potentially inaccurate were examined, results were similar to the main analysis. Items in this category were selected based on two researcher's assessments of all items on a menu with consideration of the description and composition of the dish and the potential for inconsistencies in portion sizes made by servers in outlets. As these findings were not largely different from the items in the main analyses, this suggests that reliably identifying likely inaccurately labelled menu items based on menu information alone may be difficult.

Our findings highlight the difficulty that enforcers of the policy may have if attempting to assess accuracy of calorie labelling in OOH outlets without expending substantial resources on laboratory measurement. At present there appears to be minimal enforcement training supplied, although how enforcement is monitored is at the discretion of the LA(5). To aid improved accuracy of a menu's calorie information and enforcement of the policy, laboratory analysis of menu items would preferably determine reported energy content, however this would result in significant financial cost to businesses. Despite this, a one-off or annual cost may be minimal for larger food chains such as those explored in this study. Greater support for businesses in calculating energy content, as well as greater training and assistance for the individuals responsible for assessing compliance would likely help to reduce the discrepancy between measured and reported energy content. Additionally, clearer pathways for compliance monitoring and sanctions for observed inaccuracy may be necessary to improve accuracy of labelling by businesses. Our results showed that the same menu items from different outlet locations of the same business tended to have similar measured calorie content, so analysis at the chain level is likely representative of all outlets in the chain. However, this should be explored in greater depth, and particularly within pubs, where the greatest inconsistencies in reported vs. measured energy content were observed.

4.1. Strengths and Limitations

There are a number of limitations of this work that should be considered alongside findings. This study explored a large sample of menu items in the OOH food sector; however, this sample is not representative of all menu items in outlets in England and instead provides a snapshot of the accuracy of calorie labels in the OOH food sector among large businesses. The analyses relating to matched samples and exploration of potentially inaccurate items can be considered as exploratory only, as the study was not powered for these smaller samples. We used a gold standard measurement of energy content, and this is a strength of the present work. However, we collected one sample per menu item from outlets for our main analyses and future research would benefit from taking multiple samples of the same item. This would improve measurement accuracy for estimated energy content and also allow for examination of consistency of menu item energy content within the same outlet.

5. Conclusion

There were significant inaccuracies in reported energy content of calorie labelled menu items in English food outlets subject to mandatory calorie labelling and this appears to be caused by both over- and under-estimation of reported energy content.

Financial support

This piece of work was primarily funded by the University of Liverpool Policy Support Fund. AF and ER's salary is supported by an ESRC grant (ES/W007932/1). ER is funded by the National Institute for Health and Care Research (NIHR) Oxford Health Biomedical Research Centre (BRC). JA is supported by the Medical Research Council [grant number MC_UU_00006/7]. MP receives support from the NIHR Applied Research Collaboration ARC NWC and Alzheimer's Society and is funded through a Post-Doctoral Fellowship. IGNEP receives support from the NIHR Development and Skill Enhancement Award (DSE) [grant number NIHR305076].

Declaration of Interest

Declaration of interest: ER has previously received research funding from Unilever and the American Beverage Association for unrelated research projects. Other authors have no competing interests.

Authorship

AF, ER, MP, IGNEP, AJ and JA conceptualised the study. AF and PT were responsible for data curation. IGNEP and AF conducted formal analyses. ER and AF acquired funding. AF, ER, MP and PT were responsible for investigation and methodology. AF was responsible for project admin. JB conducted data validation. AF and ER wrote the original draft of the manuscript, and all authors reviewed and edited the manuscript.

References

1. Goffe L, Rushton S, White M, Adamson A, Adams J. Relationship between mean daily energy intake and frequency of consumption of out-of-home meals in the UK National Diet and Nutrition Survey. *International Journal of Behavioral Nutrition and Physical Activity*. 2017;14(1):131.
2. Taher AK, Evans N, Evans CEL. The cross-sectional relationships between consumption of takeaway food, eating meals outside the home and diet quality in British adolescents. *Public Health Nutrition*. 2019;22(1):63-73.
3. Garbutt J, Townsend N, Johnson L, Jones A, O'Flaherty M, Colombet Z, et al. The contribution of the out-of-home food (OOHF) sector to the national diet: a cross-sectional survey with repeated 24-hour recalls of adults in England (2023-2024). *medRxiv*. 2025:2025.06.30.25330369.
4. Department of Health and Social Care. New calorie labelling rules come into force to improve nation's health. UK: Gov.uk; 2022.
5. Department of Health and Social Care. Calorie labelling in the out of home sector: implementation guidance. UK: GOV.UK; 2021.
6. Wellard-Cole L, Chapman K, Dunford EK, Goldsbury D, Havill M, Hughes C, et al. Monitoring the changes to the nutrient composition of fast foods following the introduction of menu labelling in New South Wales, Australia: an observational study. *Public Health Nutrition*. 2018;21(6):1194-9.
7. Goodman S, Vanderlee L, White CM, Hammond D. A quasi-experimental study of a mandatory calorie-labelling policy in restaurants: Impact on use of nutrition information among youth and young adults in Canada. *Preventive Medicine*. 2018;116:166-72.
8. Greenthal E, Sorscher S, Pomeranz JL, Cash SB. Availability of calorie information on online menus from chain restaurants in the USA: current prevalence and legal landscape. *Public Health Nutrition*. 2023;26(12):3239-46.
9. Bleich SN, Economos CD, Spiker ML, Vercammen KA, VanEpps EM, Block JP, et al. A systematic review of calorie labeling and modified calorie labeling interventions: impact on consumer and restaurant behavior. *Obesity*. 2017;25(12):2018-44.
10. Robinson E, Marty L, Jones A, White M, Smith R, Adams J. Will calorie labels for food and drink served outside the home improve public health? *BMJ*. 2021;372.
11. Zlatevska N, Neumann N, Dubelaar C. Mandatory Calorie Disclosure: A Comprehensive Analysis of Its Effect on Consumers and Retailers. *Journal of Retailing*. 2018;94(1):89-101.
12. Essman M, Burgoine T, Huang Y, Jones A, Polden M, Robinson E, et al. Changes in energy content of menu items at out-of-home food outlets in England after calorie labelling policy implementation: a pre-post analysis (2021-2022). *medRxiv*. 2024:2024.08.09.24311741.
13. Food Standards Agency. Guidance on voluntary energy labelling for out of home businesses in Northern Ireland. UK; 2021.
14. Cleveland LP, Simon D, Block JP. Compliance in 2017 with Federal Calorie Labeling in 90 chain restaurants and 10 retail food outlets prior to required implementation. *American Journal of Public Health*. 2018;108(8):1099-102.

15. Polden M, Jones A, Essman M, Adams J, Bishop T, Burgoine T, et al. Point-of-choice kilocalorie labelling practices in large, out-of-home food businesses: a preobservational versus post observational study of labelling practices following implementation of The Calorie Labelling (Out of Home Sector) (England) Regulations 2021. *BMJ Open*. 2024;14(4):e080405.
16. Boyland EJ, Harris JL. Regulation of food marketing to children: are statutory or industry self-governed systems effective? *Public Health Nutrition*. 2017;20(5):761-4.
17. Roberts M, Pettigrew S, Chapman K, Miller C, Quester P. Compliance with children's television food advertising regulations in Australia. *BMC Public Health*. 2012;12(1):846.
18. van der Westhuizen B, Frank T, Abdool Karim S, Swart REC. Determining food industry compliance to mandatory sodium limits: successes and challenges from the South African experience. *Public Health Nutrition*. 2023;26(11):2551-8.
19. Swinburn B, Kraak V, Rutter H, Vandevijvere S, Lobstein T, Sacks G, et al. Strengthening of accountability systems to create healthy food environments and reduce global obesity. *The Lancet*. 2015;385(9986):2534-45.
20. Jumpertz R, Venti CA, Le DS, Michaels J, Parrington S, Krakoff J, et al. Food label accuracy of common snack foods. *Obesity*. 2013;21(1):164-9.
21. Fitzpatrick L, Arcand J, Abbe M, Deng M, Duhaney T, Campbell N. Accuracy of Canadian Food Labels for Sodium Content of Food. *Nutrients*. 2014;6(8):3326-35.
22. Urban LE, Dallal GE, Robinson LM, Ausman LM, Saltzman E, Roberts SB. The Accuracy of Stated Energy Contents of Reduced-Energy, Commercially Prepared Foods. *Journal of the American Dietetic Association*. 2010;110(1):116-23.
23. Urban LE, McCrory MA, Dallal GE, Das SK, Saltzman E, Weber JL, et al. Accuracy of Stated Energy Contents of Restaurant Foods. *JAMA*. 2011;306(3):287-93.
24. Essman M, Bishop T, Burgoine T, Jones A, Polden M, Robinson E, et al. Implementation and enforcement of mandatory calorie labelling regulations for the out-of-home sector in England: qualitative study of the experiences of business implementers and regulatory enforcers. *medRxiv*. 2024:2024.02.18.24302990.
25. Polden M, Jones A, Adams J, Bishop T, Burgoine T, Essman M, et al. Kilocalorie labelling in the out-of-home sector: an observational study of business practices and consumer behaviour prior to implementation of the mandatory calorie labelling policy in England, 2022. *BMC Public Health*. 2023;23(1):1088.
26. Polden M, Jones A, Essman M, Adams J, Bishop TRP, Burgoine T, et al. Evaluating the association between the introduction of mandatory calorie labelling and energy consumed using observational data from the out-of-home food sector in England. *Nature Human Behaviour*. 2025;9(2):277-86.
27. Basolo A, Parrington S, Ando T, Hollstein T, Piaggi P, Krakoff J. Procedures for Measuring Excreted and Ingested Calories to Assess Nutrient Absorption Using Bomb Calorimetry. *Obesity*. 2020;28(12):2315-22.

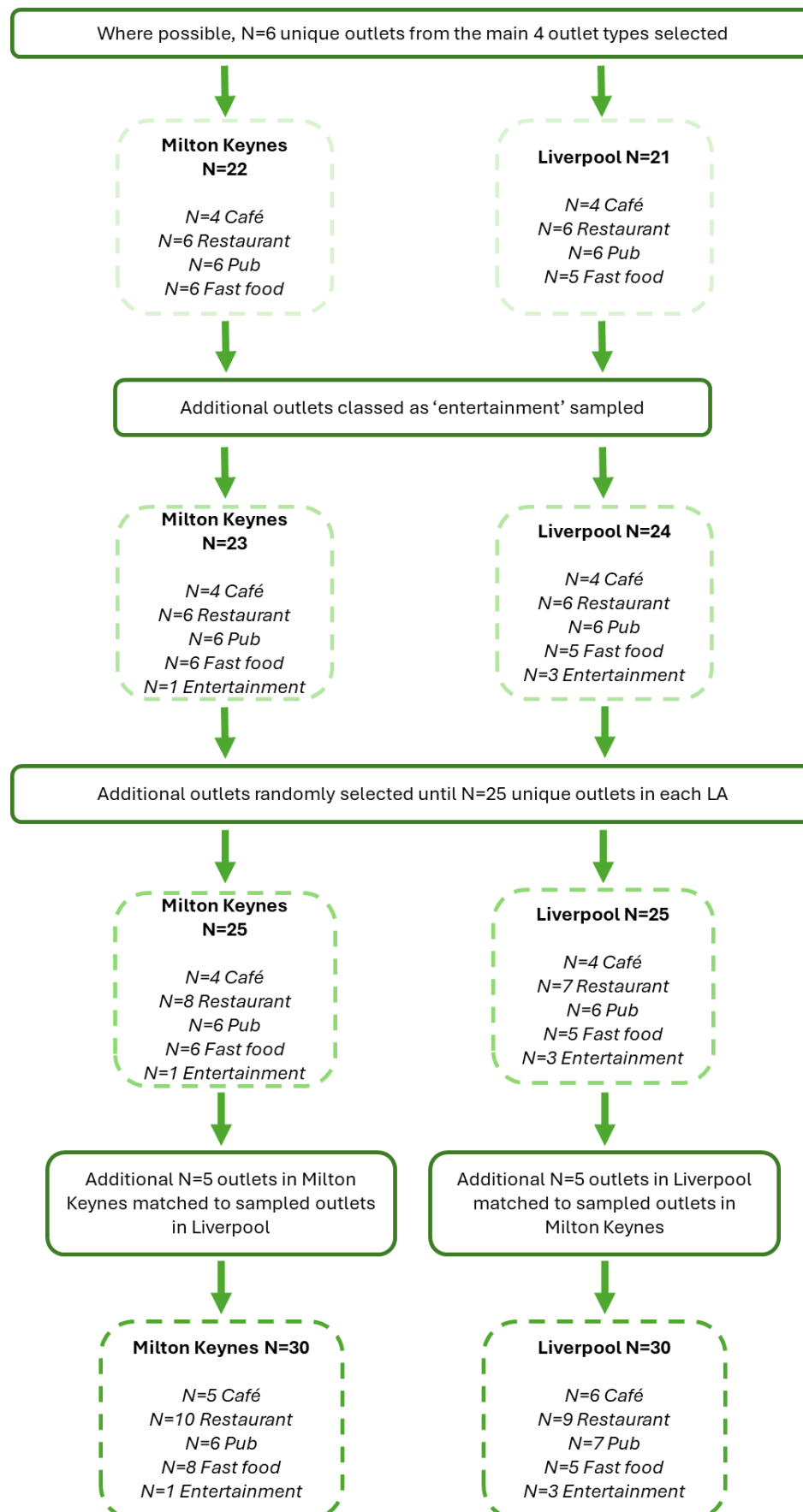
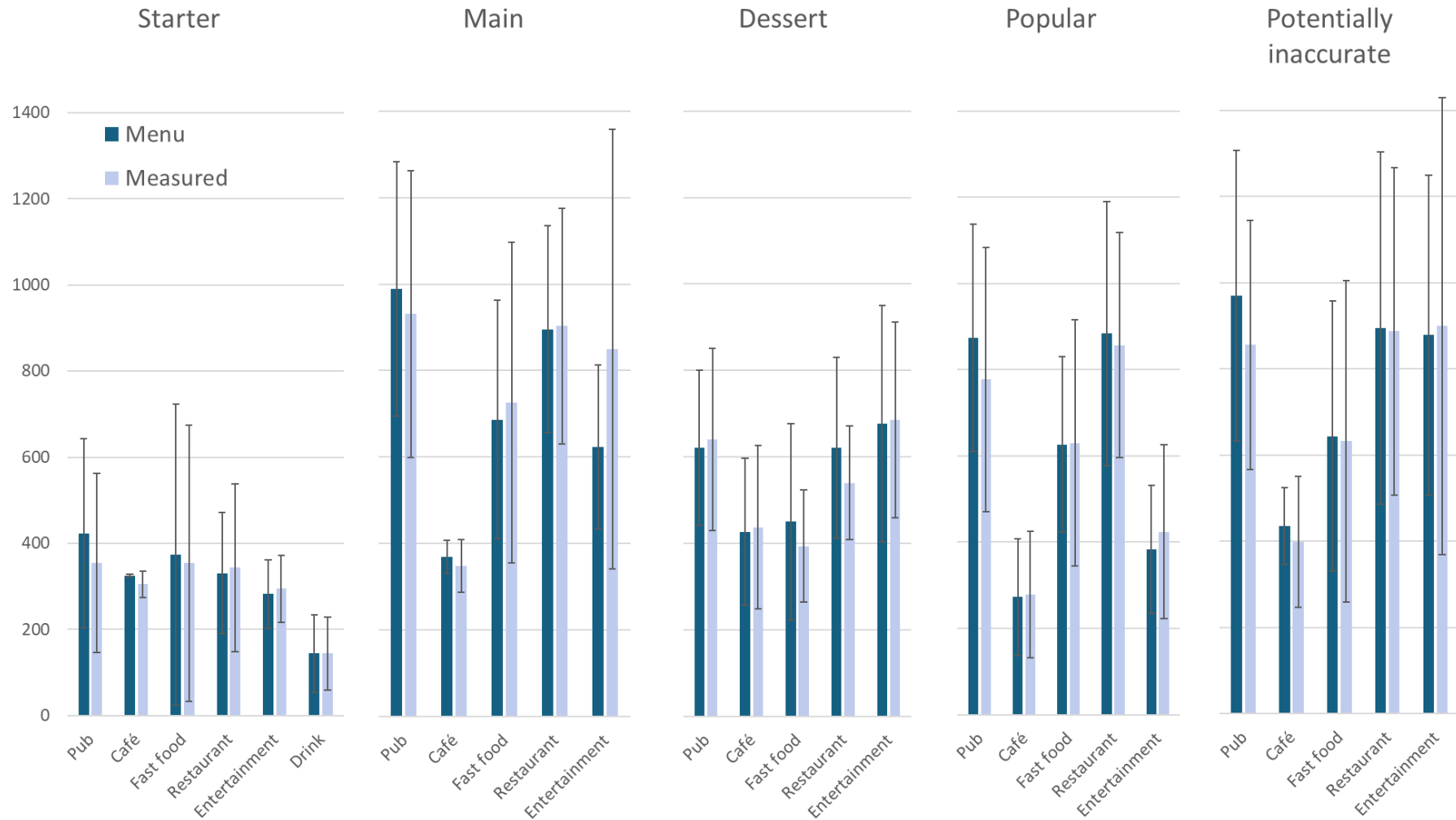
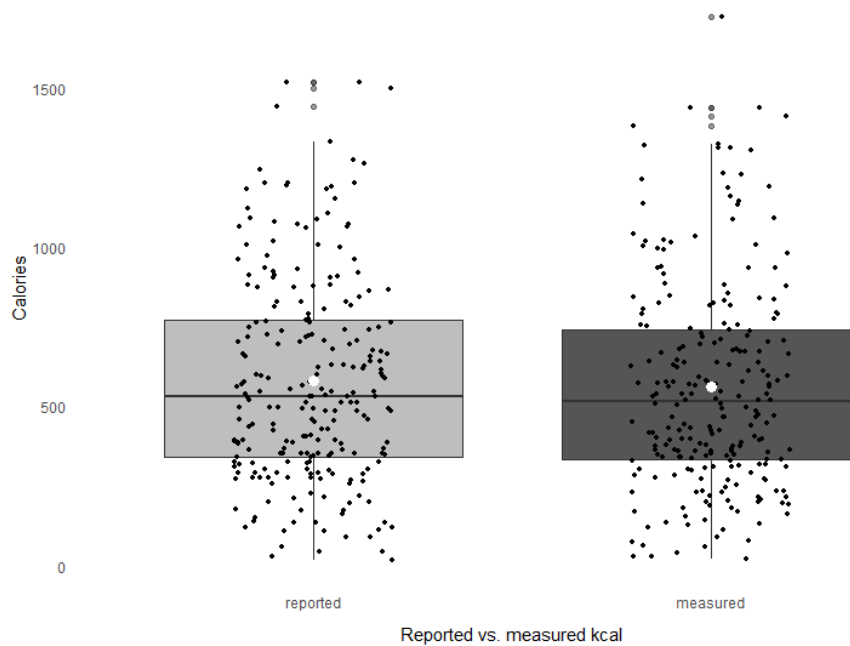
Figure 1: Outlet selection process

Figure 2: Reported vs. Measured energy (kcal) split by outlet type and food category

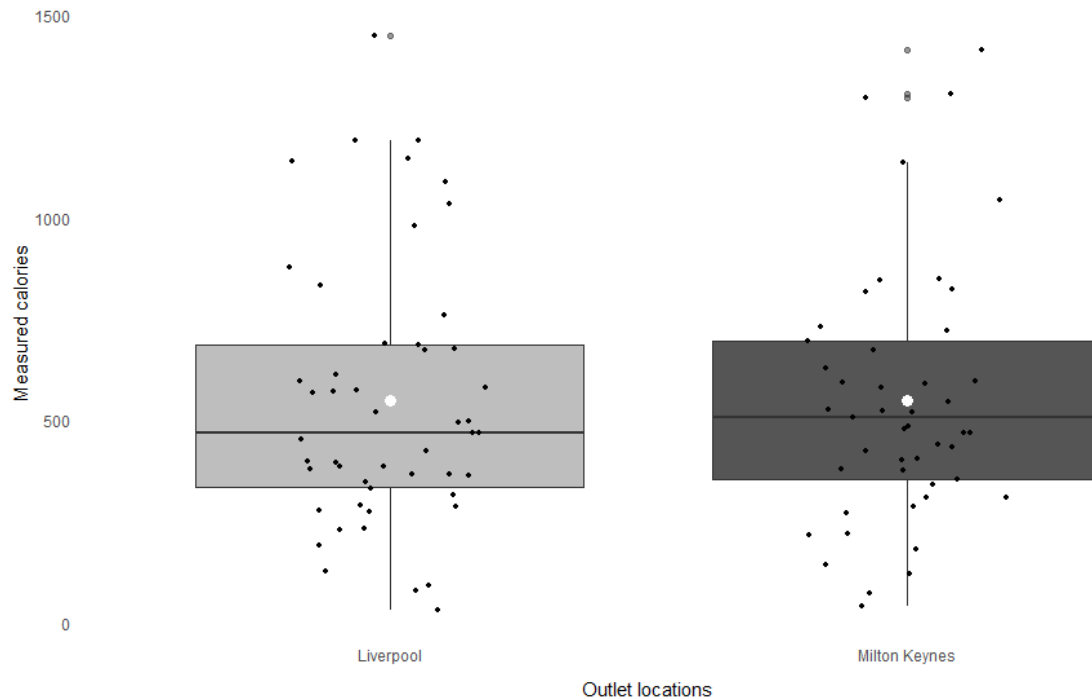
*Error bars represent standard deviation

Figure 3: Reported and measured energy content of n=236 items.

*white circle indicates the mean, ** outliers are shown twice (see grey points)

For the subsample of menu items from outlets matched across the two LAs, a Wilcoxon signed rank test found no significant differences in measured energy between matched items ($V=583$, $p=.775$). See Figure 4. The mean energy content of these items in Liverpool was $547.39 \text{ kcal} \pm 327.96$ and for Milton Keynes this was $549.45 \text{ kcal} \pm 312.02$. The mean difference (Liverpool – Milton Keynes) was -14.15 kcal , and the relative percentage difference was -7.99% . 73% of items sampled in Milton Keynes were within 20% of the value in Liverpool.

Figure 4: Matched pairs of menu items across the two local authorities (n=50 matched pairs)



*white circle indicates the mean

A linear model explored predictors of the mean difference between reported and measured energy content (Table 2). Variance Inflation Factor (VIF) for all predictors were below 2, indicating minimal multicollinearity. Items with a higher measured energy content were more likely to have a measured energy higher than reported energy ($B = 0.22$, 99% CI 0.12 to 0.32, $p < .001$) whereby every measured 1kcal increase, was associated with an increased difference of 0.22kcal. Main effects were also observed for outlet type and item category, whereby pubs and restaurants (vs. cafes), and popular dishes (vs. starters/sides) had a greater difference between reported and measured energy. There were no significant predictors of difference between measured and reported energy for potentially inaccurate items (full results in Supplementary Material 4). We also explored predictors of absolute mean percentage difference between reported and measured energy content, and results were similar to main analyses (Supplementary Material 5).

Table 1: Difference between reported and measured energy (kcal) for the different categories

	Mean kcal content	Relative mean difference (measured-reported) (kcal)*	Relative percentage difference**	Absolute mean difference (kcal)***	Absolute percentage difference***	Percent of items outside 20% leeway
Overall (N=236)	560.40 (329.11)	-16.70 (149.19)	-8.75% (35.27%)	98.76 (112.88)	21.30% (29.45%)	34.75%
Outlet category						
Cafes (N=44)	313.03 (158.74)	-2.60 (48.96)	-3.44% (20.52%)	36.29 (32.51)	15.02% (14.30%)	29.55%
Pubs (N=51)	674.00 (338.37)	-51.00 (154.94)	-20.08% (50.54%)	126.26 (101.97)	29.78% (45.46%)	45.10%
Fast food (N=51)	523.75 (323.79)	-8.62 (134.77)	-4.97% (27.13%)	82.89 (105.98)	17.73% (21.01%)	29.41%
Restaurants (N=76)	660.67 (318.30)	-21.97 (175.63)	-8.71% (31.73%)	123.30 (126.22)	21.38% (24.96%)	35.53%
Entertainment (N=14)	552.79 (330.88)	63.14 (204.14)	1.91% (43.10%)	119.50 (174.89)	22.71% (35.88%)	28.57%
Item category						
Starter (N=52)	342.85 (220.37)	-17.61 (127.99)	-19.42 (55.76)	84.88 (96.71)	33.44% (48.60%)	48.08%
Main (N=59)	763.98 (360.26)	6.10 (184.10)	-3.77 (22.36)	122.21 (136.88)	16.15% (15.84%)	28.81%
Dessert (N=60)	520.73 (190.61)	-31.23 (132.03)	-6.66 (28.66)	85.35 (104.95)	17.77% (23.34%)	23.33%
Popular (N=57)	658.03 (330.78)	-26.45 (154.66)	-7.06 (27.33)	110.64 (110.33)	18.75% (20.95%)	35.09%
Drink (N=8)	144.18 (84.81)	-0.57 (38.06)	-3.74 (30.19)	31.86 (16.99)	25.00% (15.00%)	75.00%
Potentially inaccurate (N=59)****	738.16 (378.56)	-35.45 (169.36)	-9.93% (31.56%)	121.96 (121.78)	20.88% (25.64%)	35.59%
Local authority						
Liverpool (N=119)	538.15 (298.88)	-27.03 (130.10)	-8.93% (29.63%)	87.51 (99.70)	19.62% (23.91%)	34.45%
Milton Keynes (N=117)	587.77 (354.39)	-6.19 (166.28)	-8.56% (40.33%)	110.20 (124.26)	23.01% (34.19%)	33.33%

*Negative values indicate that reported kcal were greater than measured,

**Percentage difference for reported – measured kcal,

***Mean difference between reported and measured kcal regardless of direction (- or +) of difference

****Numbers in brackets are standard deviations

*****Potentially inaccurate items not included in mean and standard deviation values for ‘overall’, ‘item category’ and ‘outlet category’

Table 2: Linear model exploring predictors of mean relative difference between reported and measured energy content (calculated as measured – reported)

<i>Predictors</i>	Mean kcal difference		
	<i>Estimates</i>	<i>99% CI</i>	<i>p</i>
(Intercept)	-16.87	-111.00 – 77.27	0.642
Pub*	-134.08	-222.36 – -45.80	<0.001
Fast Food*	-61.09	-144.82 – 22.64	0.059
Restaurant*	-99.88	-185.36 – -14.39	0.003
Entertainment*	8.66	-111.15 – 128.47	0.851
Main**	-79.16	-161.72 – 3.39	0.013
Dessert**	-64.15	-136.38 – 8.08	0.022
Popular**	-88.89	-166.98 – -10.79	0.003
Drink**	-22.60	-174.43 – 129.23	0.699
IMD quintile [2]***	-5.59	-84.61 – 73.43	0.854
IMD quintile [3] ***	3.08	-55.14 – 61.30	0.891
IMD quintile [4] ***	3.75	-84.58 – 92.08	0.912
IMD quintile [5] ***	-54.97	-246.44 – 136.50	0.456
Outlet location [MK]****	15.66	-34.90 – 66.21	0.422
Total energy kcal	0.22	0.12 – 0.32	<0.001
Observations	236		
R ² / R ² adjusted	0.181 / 0.129		

Reference categories are: *Cafes, ** Starters/sides, ***IMD quintile 1, ****Liverpool

Negative values indicate a decreased relative difference