Cochrane Library

**Cochrane** Database of Systematic Reviews

# Calorie (energy) labelling for changing selection and consumption of food or alcohol (Review)

Clarke N, Pechey E, Shemilt I, Pilling M, Roberts NW, Marteau TM, Jebb SA, Hollands GJ

Clarke N, Pechey E, Shemilt I, Pilling M, Roberts NW, Marteau TM, Jebb SA, Hollands GJ.
Calorie (energy) labelling for changing selection and consumption of food or alcohol.
*Cochrane Database of Systematic Reviews* 2025, Issue 1. Art. No.: CD014845.
DOI: 10.1002/14651858.CD014845.pub2.

**www.cochranelibrary.com**



WILEY

## TABLE OF CONTENTS

[Intervention Review]

# Calorie (energy) labelling for changing selection and consumption of food or alcohol

Natasha Clarke[1,2], Emily Pechey[1], Ian Shemilt[3], Mark Pilling[1], Nia W Roberts[4], Theresa M Marteau[1], Susan A Jebb[5], Gareth J Hollands[1,3]

[1]Behaviour and Health Research Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [2]School of Sciences, Bath Spa University, Bath, UK. [3]EPPI Centre, UCL Social Research Institute, University College London, London, UK. [4]Bodleian Health Care Libraries, University of Oxford, Oxford, UK. [5]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

**Contact:** Gareth J Hollands, gareth.hollands@ucl.ac.uk.

**Editorial group:** Cochrane Central Editorial Service.
**Publication status and date:** New, published in Issue 1, 2025.

**Citation:** Clarke N, Pechey E, Shemilt I, Pilling M, Roberts NW, Marteau TM, Jebb SA, Hollands GJ. Calorie (energy) labelling for changing selection and consumption of food or alcohol. *Cochrane Database of Systematic Reviews* 2025, Issue 1. Art. No.: CD014845. DOI: 10.1002/14651858.CD014845.pub2.



## ABSTRACT

### Background

Overconsumption of food and consumption of any amount of alcohol increases the risk of non-communicable diseases. Calorie (energy) labelling is advocated as a means to reduce energy intake from food and alcoholic drinks. However, there is continued uncertainty about these potential impacts, with a 2018 Cochrane review identifying only a small body of low-certainty evidence. This review updates and extends the 2018 Cochrane review to provide a timely reassessment of evidence for the effects of calorie labelling on people's selection and consumption of food or alcoholic drinks.

### Objectives

– To estimate the effect of calorie labelling for food (including non-alcoholic drinks) and alcoholic drinks on selection (with or without purchasing) and consumption.

– To assess possible modifiers – label type, setting, and socioeconomic status – of the effect of calorie labelling on selection (with or without purchasing) and consumption of food and alcohol.

### Search methods

We searched CENTRAL, MEDLINE, Embase, PsycINFO, five other published or grey literature databases, trial registries, and key websites, followed by backwards and forwards citation searches. Using a semi-automated workflow, we searched for and selected records and corresponding reports of eligible studies, with these searches current to 2 August 2021. Updated searches were conducted in September 2023 but their results are not fully integrated into this version of the review.

### Selection criteria

Eligible studies were randomised controlled trials (RCTs) or quasi-RCTs with between-subjects (parallel group) or within-subjects (cross-over) designs, interrupted time series studies, or controlled before-after studies comparing calorie labelling with no calorie labelling, applied to food (including non-alcoholic drinks) or alcoholic drinks. Eligible studies also needed to objectively measure participants' selection (with or without purchasing) or consumption, in real-world, naturalistic laboratory, or laboratory settings.

**Data collection and analysis**

Two review authors independently selected studies for inclusion and extracted study data. We applied the Cochrane RoB 2 tool and ROBINS-I to assess risk of bias in included studies. Where possible, we used (random-effects) meta-analyses to estimate summary effect sizes as standardised mean differences (SMDs) with 95% confidence intervals (CIs), and subgroup analyses to investigate potential effect modifiers, including study, intervention, and participant characteristics. We synthesised data from other studies in a narrative summary. We rated the certainty of evidence using GRADE.

**Main results**

We included 25 studies (23 food, 2 alcohol and food), comprising 18 RCTs, one quasi-RCT, two interrupted time series studies, and four controlled before-after studies. Most studies were conducted in real-world field settings (16/25, with 13 of these in restaurants or cafeterias and three in supermarkets); six studies were conducted in naturalistic laboratories that attempted to mimic a real-world setting; and three studies were conducted in laboratory settings. Most studies assessed the impact of calorie labelling on menus or menu boards (18/25); six studies assessed the impact of calorie labelling directly on, or placed adjacent to, products or their packaging; and one study assessed labels on both menus and on product packaging. The most frequently assessed labelling type was simple calorie labelling (20/25), with other studies assessing calorie labelling with information about at least one other nutrient, or calories with physical activity calorie equivalent (PACE) labelling (or both). Twenty-four studies were conducted in high-income countries, with 15 in the USA, six in the UK, one in Ireland, one in France, and one in Canada. Most studies (18/25) were conducted in high socioeconomic status populations, while six studies included both low and high socioeconomic groups, and one study included only participants from low socioeconomic groups. Twenty-four studies included a measure of selection of food (with or without purchasing), most of which measured selection with purchasing (17/24), and eight studies included a measure of consumption of food.

Calorie labelling of food led to a small reduction in energy selected (SMD −0.06, 95% CI −0.08 to −0.03; 16 randomised studies, 19 comparisons, 9850 participants; high-certainty evidence), with near-identical effects when including only studies at low risk of bias, and when including only studies of selection with purchasing. There may be a larger reduction in consumption (SMD −0.19, 95% CI −0.33 to −0.05; 8 randomised studies, 10 comparisons, 2134 participants; low-certainty evidence). These effect sizes suggest that, for an average meal of 600 kcal, adults exposed to calorie labelling would select 11 kcal less (equivalent to a 1.8% reduction), and consume 35 kcal less (equivalent to a 5.9% reduction). The direction of effect observed in the six non-randomised studies was broadly consistent with that observed in the 16 randomised studies.

Only two studies focused on alcoholic drinks, and these studies also included a measure of selection of food (including non-alcoholic drinks). Their results were inconclusive, with inconsistent effects and wide 95% CIs encompassing both harm and benefit, and the evidence was of very low certainty.

**Authors' conclusions**

Current evidence suggests that calorie labelling of food (including non-alcoholic drinks) on menus, products, and packaging leads to small reductions in energy selected and purchased, with potentially meaningful impacts on population health when applied at scale. The evidence assessing the impact of calorie labelling of food on consumption suggests a similar effect to that observed for selection and purchasing, although there is less evidence and it is of lower certainty. There is insufficient evidence to estimate the effect of calorie labelling of alcoholic drinks, and more high-quality studies are needed. Further research is needed to assess potential moderators of the intervention effect observed for food, particularly socioeconomic status. Wider potential effects of implementation that are not assessed by this review also merit further examination, including systemic impacts of calorie labelling on industry actions, and potential individual harms and benefits.

## PLAIN LANGUAGE SUMMARY

### Can calorie (energy) labelling change people's selection and consumption of food or alcohol?

**Key messages**

– Current evidence suggests that calorie (energy) labelling on menus, and on or next to products, leads to reductions in calories selected and bought from food and non-alcoholic drinks. The evidence for consumption (eating) suggests a similar effect, but there is less evidence and it is of lower quality.

– There is insufficient evidence to estimate the effect of calorie labelling for alcoholic drinks.

– Calorie labelling of food could lead to potentially meaningful impacts on population health when applied at scale, but we need more high-quality studies for consumption and for alcohol products.

**Why put calorie labels on products?**

Overconsumption of food and consumption of any alcohol products are important causes of poor health. Labelling menus and the packaging of products to show how much energy they contain ('calories', which is measured in kilocalories), may reduce the amount that people buy and consume, and help them choose healthier options.

**What did we want to find out?**

We investigated whether adding calorie labelling to food (including non-alcoholic drinks) and alcoholic drinks changes people's selection and consumption of those products.

**What did we do?**

We searched for studies comparing the effects of labelling products versus not labelling products on the selection and consumption of food and drinks in people of any age. We compared and summarised the results, and rated our confidence in the evidence, based on factors such as study methods and sizes.

**What did we find?**

We found 25 studies, all of which were conducted in high-income countries. Twenty-three studies involved food, while two studies involved food and alcohol products. Most of these studies were conducted in real-world settings such as restaurants or supermarkets.

**Main results**

We found that adding calorie labelling to food reduced the amount of calories selected by a small amount (16 studies, 9850 people). For example, if there was no labelling, people would select a meal that had 600 kilocalories but, when there was labelling, they would select a meal that had 589 kilocalories (11 kilocalories fewer).

Calorie labelling on food may reduce energy consumed (8 studies, 2134 people). For example, if there was no labelling, people would eat a meal that had 600 kilocalories but, when there was labelling, they would eat a meal that had 565 kilocalories (35 kilocalories fewer).

There was insufficient evidence to assess the effects of calorie labelling on alcohol products (selection of calories: 2 studies, 5756 people; selection of alcohol: 1 study, 205 people).

**What are the limitations of the evidence?**

We are confident in the results concerning calorie labelling on selection and purchasing of food (including non-alcoholic drinks). In contrast, we have little confidence in the results concerning calorie labelling on consumption of food (including non-alcoholic drinks) because most studies were conducted in laboratory settings for short periods and not all the studies provided enough information about how they were conducted. We are not confident in our estimates of the effects of calorie labelling on alcohol products because there were not enough studies.

**How up to date is this evidence?**

This review is up to date to 2 August 2021.

## SUMMARY OF FINDINGS

**Summary of findings 1. Calorie (energy) labelling for selection and consumption of food and non-alcoholic drinks**

**Calorie (energy) labelling for selection and consumption of food and non-alcoholic drinks**

**Population:** adults[a]

**Settings:** real-world (restaurant, store), naturalistic laboratory, and laboratory settings[b]

**Intervention:** calorie (energy) labelling

**Comparator:** no calorie (energy) labelling

| Outcomes | Estimated absolute effects* (95% CI) | | Relative effect** (95% CI) | No. of participants (no. of studies/ comparisons) | Certainty of the evidence (GRADE) | Comments |
|---|---|---|---|---|---|---|
| | Without calorie labelling | With calorie labelling | | | | |
| **Selection of energy from food and non-alcoholic drinks with or without purchasing (kcal)**<br><br>With exposure to intervention of between < 1 day and multiple weeks (2–13 weeks) | Mean energy selected for an average meal would be 600 (standard deviation 185) kcal | Mean energy selected for an average meal would be 11 kcal or 1.8% less with calorie labelling (15 kcal less to 6 kcal less; 2.5% less to 1% less) | Mean selection with calorie labelling was **0.06** standard deviations lower (0.08 lower to 0.03 lower) | 9850<br><br>(16 RCTs; 19 comparisons) | ⊕⊕⊕⊕<br>**High**[c] | Calorie labels decrease the amount of energy selected from food.<br><br>Our extrapolation of this summary effect (SMD) of −0.06 suggests that, if calorie labelling was implemented for an assumed average meal of 600 (standard deviation 185) kcal, adults would select 11 kcal less (15 kcal less to 6 kcal less), reducing energy purchased by 1.8% (2.5% less to 1% less). |
| **Consumption of energy from food and non-alcoholic drinks with or without purchasing (kcal)**<br><br>With exposure to intervention of <1 day | Mean energy consumed for an average meal would be 600 (standard deviation 185) kcal | Mean energy consumed for an average meal would be 35 kcal or 5.9% less with calorie labelling (61 kcal less to 9 kcal less; 10.2% less to 1.5% less) | Mean consumption with calorie labelling was **0.19** standard deviations lower (0.33 lower to 0.05 lower) | 2134<br><br>(8 RCTs; 10 comparisons) | ⊕⊕⊝⊝<br>**Low**[d] | Calorie labels may decrease the amount of energy consumed from food.<br><br>Our extrapolation of this summary effect (SMD) of −0.19 suggests that, if calorie labelling was implemented for an assumed average meal of 600 calories (standard deviation 185 kcal), adults would consume 35 kcal less (61 kcal less to 9 kcal less), reducing energy consumed by 5.9% (10.2% less to 1.5% less). |
| **Additional evidence from NRSI included in the review but not quantitatively synthesised in meta-analyses** | | | | | | |

| | | | | |
|---|---|---|---|---|
| **Selection of energy from food and non-alcoholic drinks with or without purchasing (kcal)**<br><br>With exposure to intervention of between 2 weeks to 5 years | Evidence from the 6 NRSI was broadly consistent with the direction of effect observed in randomised studies: 5/6 studies observed a numeric decrease in energy purchased, with a numeric increase in energy purchased in 1 study[e] | 6 studies (2 ITS, 4 CBA) | ⊕⊝⊝⊝<br>**Very low**[f] | — |

* The **estimated absolute effect** in the intervention group is based on an assumed value for the comparison group and the **relative effect** of the intervention. Assumed values are derived from mean and standard deviation values for population energy consumption from food in a representative sample of UK adults from the UK National Diet and Nutrition Survey Years 9-11 (Public Health England 2020); for selection outcomes it is assumed that all energy selected is consumed. The relative effect is derived from the primary random-effects meta-analysis for the outcome. See Measures of treatment effect and Effects of interventions for full details.

** Our interpretation is focused on the estimated absolute effect, but a Cohen's d effect size of ≤ 0.2 (i.e. 0.2 standard deviations) is commonly considered a small effect.

**CBA:** controlled before-after study; **CI:** confidence interval; **ITS:** interrupted time series; **NRSI:** non-randomised studies of interventions; **kcal:** kilocalorie (i.e. calorie); **RCT:** randomised controlled trial.

GRADE Working Group grades of evidence
**High certainty:** further research is very unlikely to change our confidence in the estimate of effect.
**Moderate certainty:** further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low certainty:** further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low certainty:** we are very uncertain about the estimate.

[a] Included studies focused primarily on adults, mostly from the general populations. One study included adults as well as adolescents over the age of 16 years (Harnack 2008); one study included adults as well as adolescents and children (Petimar 2019).
[b] Included studies focused on restaurants, stores, laboratories (selection); laboratories only (consumption).
[c] Certainty of evidence was not downgraded for any domain. See Effects of interventions for full details.
[d] Downgraded two levels: one level for risk of bias (all but one study in the meta-analysis were judged to have significant concerns regarding risk of bias), and by one level for indirectness (studies were conducted in laboratory settings with short-term exposures not representative of real-world implementation). See Effects of interventions for full details.
[e] We did not assume a meaningful quantitative summary estimate representing this body of six studies, as we considered their estimates to be insufficiently comparable to one another, and only four of six studies to have reported reasonably comparable outcome data (for mean selection per transaction, which ranged from 73 kcal lower to 18 kcal higher with calorie labelling). We have not made strong inferences from these studies, but have conducted a narrative synthesis (see Effects of interventions).
[f] Downgraded two levels (starting at 'low certainty' for NRSIs): one level for risk of bias (all studies were at moderate risk of bias) and one level for imprecision (because we were unable to generate and assess a meaningful summary effect size estimate with CIs). See Effects of interventions for full details.

### Summary of findings 2. Calorie (energy) labelling for selection and consumption of alcoholic drinks

**Calorie (energy) labelling for selection and consumption of alcoholic drinks**

**Population:** adults[a]

**Settings:** real-world (restaurant) and naturalistic laboratory (simulated online supermarket) settings

**Intervention:** calorie (energy) labelling

**Comparison:** no calorie (energy) labelling

| **Outcomes** | **Estimated absolute effects* (95% CI)** | | **Relative effect** (95% CI)** | **No. of participants (no. of studies/comparisons)** | **Certainty of the evidence (GRADE)** | **Comments** |
|---|---|---|---|---|---|---|
| | **Without calorie labelling** | **With calorie labelling** | | | | |
| **Selection of energy from alcoholic drinks (with or without purchasing) (kcal)**<br><br>With exposure to intervention of < 1 day | — | — | Mean selection with calorie<br><br>labelling was **0.05 standard deviations lower** (0.25 lower to 0.16 higher) | 5756<br><br>(2 RCTs; 2 comparisons) | ⊕⊝⊝⊝<br>**Very low**[b] | Calorie labels may have an effect on the amount of energy selected from alcohol, but the direction and size of this effect are very uncertain. |
| **Selection of alcohol from**<br><br>**alcoholic drinks (with or without purchasing) (units of alcohol)**<br><br>With exposure to intervention of < 1 day | — | — | Mean selection with calorie<br><br>labelling was **0.21 standard deviations lower** (0.49 lower to 0.06 higher) | 205<br><br>(1 RCT; 1 comparison) | ⊕⊝⊝⊝<br>**Very low**[c] | Calorie labels may have an effect on the amount of alcohol selected (in terms of alcohol units), but the direction and size of this effect are very uncertain. |
| **Consumption of energy from alcoholic drinks (with or without purchasing) (kcal)** | — | — | — | — | — | Not reported |
| **Consumption of alcohol from**<br><br>**alcoholic drinks (with or without purchasing) (units of alcohol)** | — | — | — | — | — | Not reported |

* We did not extrapolate the **estimated absolute effect** for alcoholic drinks outcomes, which would have been based on an assumed value for the comparison group and the **relative effect** of the intervention. The relative effect is derived from the primary random-effects meta-analysis for the outcome. See Measures of treatment effect and Effects of interventions for full details.

** Our interpretation of review findings is focused on the estimated absolute effect but a Cohen's d effect size of ≤ 0.2 (i.e. 0.2 standard deviations) is commonly considered a small effect.

**CI:** confidence interval; **kcal:** kilocalorie (i.e. calorie); **RCT:** randomised controlled trial

GRADE Working Group grades of evidence

**High certainty:** further research is very unlikely to change our confidence in the estimate of effect.
**Moderate certainty:** further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low certainty:** further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low certainty:** we are very uncertain about the estimate.

---

[a] Included studies focused on adults from general populations.

[b] Downgraded three levels: one level for indirectness (exposures to the intervention were short-term, and the intervention would feasibly and likely be implemented at a substantially larger scale and in a greater variety of real-world settings), one level for inconsistency (with evidence of substantial heterogeneity from tests of statistical inconsistency and homogeneity), and one level for imprecision (large sample size exceeding the optimal information size, but with wide CIs encompassing both appreciable benefit and appreciable harm). See Effects of interventions for full details.

[c] Downgraded three levels: one level for indirectness (exposure to the intervention was short-term, and the intervention would feasibly and likely be implemented at a substantially larger scale and in a variety of real-world settings not reflected in this single study), and two levels for imprecision (a single study with a sample size substantially lower than the optimal information size and with wide CIs encompassing both appreciable benefit and appreciable harm). See Effects of interventions for full details.

# BACKGROUND

## Description of the condition

Poor diets – including overconsumption of saturated fats, free sugars, and salt – and the consumption of alcohol, contribute to the global prevalence of obesity and the burden of non-communicable diseases, such as cardiovascular disease, diabetes, and many cancers (NCD Countdown 2020; Rehm 2018; Sheron 2016). Worldwide, 2.8 million people are estimated to die annually as a result of diseases caused by obesity (World Health Organization 2020), with higher prevalence in the most disadvantaged groups in high-income countries (Bann 2017; Health Survey for England 2018; NHS Digital 2019). There are also considerable economic costs associated with obesity (Tremmel 2017). For example, in the UK, attributable National Health Service (NHS) costs are expected to reach an estimated 9.7 billion pounds sterling (GBP) by 2050 (Department of Health and Social Care 2022). However, changing diet- and alcohol-related behaviours to halt and reverse rises in these potentially preventable diseases is difficult. While many people want to engage in behaviours that promote good health, most find it difficult to implement and maintain them (Sheeran 2016). This is in part due to physical environments that can exert considerable influence on routine and habitual health-related behaviours (Gardner 2022; Hollands 2016). Altering these environments appears a promising approach to achieving sustained behaviour change (Das 2012; Marteau 2012), including interventions that alter the small-scale physical environments in which health-related behaviours are performed (Hollands 2017). Calorie labelling (sometimes referred to as energy labelling) and broader nutritional labelling schemes are two key examples.

## Description of the intervention

Nutritional labels provide information about the nutritional content of a food or drink (e.g. the energy content or amount of fat, sugars, or salt). The type of information provided varies across countries; for example, with regard to which nutrients are presented and the format in which this information is communicated. The current review focused specifically on one type of these labels, calorie (or energy) labelling, which is the most frequent focus of both research and policy implementation (Crockett 2018; Department of Health and Social Care 2018a; Department of Health and Social Care 2020; Polden 2023; Robinson 2019; Zlatevska 2018). We considered calorie (or energy) labelling as it applies to both food products (including non-alcoholic drinks) and alcohol products. We defined calorie labelling as a label that explicitly quantified the calorie or energy value of a product, in kilocalories or other equivalent metric (e.g. kilojoules).

### Calorie labelling on food product packaging

Preprepared, often prepackaged, foods form a substantial part of dietary intake in many parts of the world. These food products are often complex items, with a mix of ingredients that make it difficult for consumers to know their nutritional or energy content. Many countries have implemented mandatory nutritional labelling in some parts of the food system, including the UK, the USA, Australia, New Zealand, Canada, Mexico, Argentina, Brazil, Chile, Colombia, Ecuador, Paraguay, Uruguay, Israel, Japan, India, China, Hong Kong, South Korea, Malaysia, Taiwan, Thailand, and Gulf Cooperation Council members (European Food Information Council 2018). In the European Union, Food Information Regulations made ingredient and nutrition declarations mandatory on the back of packaging for most prepackaged foods from December 2016 (European Union 2011). These regulations stipulated that manufacturers must provide nutritional information in a consistent format for most prepackaged foods, including information on energy content as well as on fat, saturated fat, carbohydrate, total sugars, protein, and salt (expressed per 100 g or per 100 mL of the food product). Additionally, manufacturers are able to repeat information in 'the principal field of vision'; in other words, on the front of the pack (Food and Drug Administration 2016). This is purely voluntary but, where provided, only information on energy or energy plus fat, saturated fats, sugars, and salt can be given, either per 100 g or 100 mL, or per portion, or both. These front-of-pack nutritional labelling schemes have usually been designed to guide consumer choice and sometimes include an interpretative component, such as reference to daily intake guidelines or colour coding to indicate relative healthiness. These can supplement, but not replace, the mandatory, back-of-pack nutrition declarations. In the UK, for example, a voluntary front-of-pack scheme using red, amber, and green colour coding according to nutrient content is widely used (European Food Information Council 2018).

### Calorie labelling on alcohol product packaging

Alcohol is energy dense (7.1 kcal/g) and the consumption of alcoholic drinks accounts for a substantial proportion (7.4%) of the total energy intake of adult drinkers aged 19 to 64 years in the UK (Public Health England 2020), estimated to be about 60% of the population (Statistica 2019). Most countries do not require the energy content to be displayed on alcohol products (World Health Organization 2018). In the European Union, most alcoholic drinks above 1.2% alcohol by volume are exempt from mandatory nutrition labelling, but energy declarations can be made on a voluntary basis. In the US, regulations also stipulate that nutrition content labelling is not required for alcoholic drinks (Alcohol and Tobacco Trade and Tax Bureau 2021). Therefore, most products do not display this information, and drinkers tend to underestimate the energy content of their drinks (Royal Society for Public Health 2014). Health advocacy organisations have called for the inclusion of calorie labelling for alcoholic drinks (Royal Society for Public Health 2018; World Health Organization 2017), and England's obesity strategy includes plans to consult on whether to make companies provide this information (Department of Health and Social Care 2020). The consultation on this policy that was announced in July 2020 has yet to take place at the time of writing in 2024.

### Calorie labelling on menus

In addition to labelling packaged foods and drinks, some countries have introduced labelling on menus. Mandatory calorie labelling in restaurants was first introduced in the state of New York (USA) in 2008 (Dumanovsky 2011). In 2016, the US Food and Drug Administration's (FDA) 'final rule' for all states became effective, requiring that calorie information be listed on menus and menu boards in chain restaurants with 20 or more locations, as well as in all vending machines (Food and Drug Administration 2016). Similarly, the Healthy Menu Choices Act 2015 came into force in 2017 in Ontario (Canada). Since 2011, some states in Australia have also implemented a labelling policy, requiring mandatory calorie labelling on menus in fast-food chains and in vending machines (Obesity Evidence Hub 2021). In England, out-of-home calorie labelling was mandated in April 2022, meaning large businesses (with more than 250 employees) are required to display calorie information on their menus, websites, and delivery platforms (Department of Health and Social Care 2021). In terms of alcohol calorie labelling on menus, this varies by country, but where calorie labelling has been implemented, including in the UK, the US, and Canada, alcohol is typically exempt from needing to display specific calorie content (e.g. Department of Health and Social Care 2021; Food and Drug Administration 2014; Government of Ontario 2023).

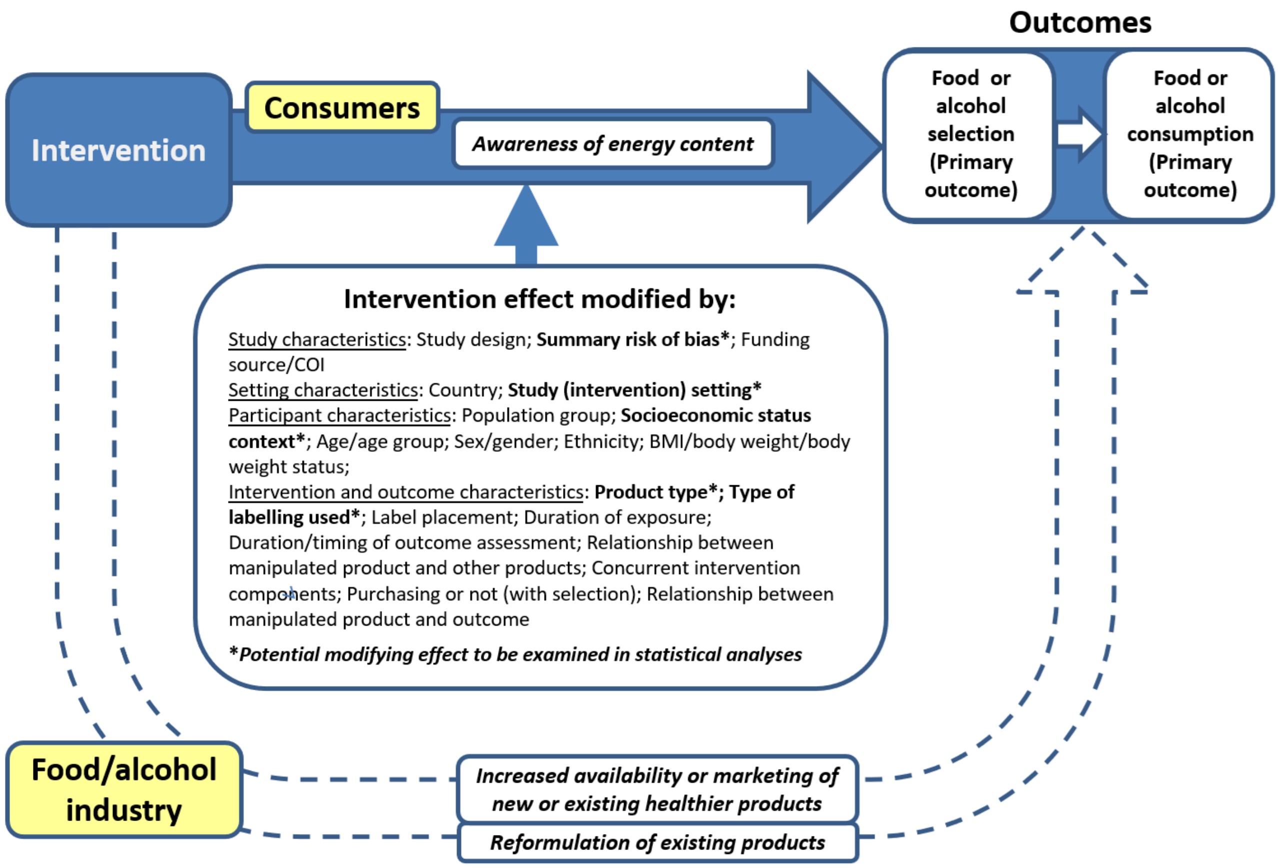
In the absence of international agreements, there has been, and continues to be, considerable variation in both the information provided and the presentation format for calorie and nutritional labelling. In terms of possible formats for calorie labels specifically, labels may include numerical information on energy content, energy as a proportion of daily guidelines, colours (e.g. traffic light labelling) to indicate a product's content relative to national guidance, or they may communicate energy in an alternative format, such as physical activity equivalents.

### How the intervention might work

Calorie labelling may impact on population health by changing the selection of food and drinks, with selection being a proximal determinant of healthier consumption (Cowburn 2005; Dobbs 2014; World Health Organization 2004; World Health Organization 2020).

A reduction in energy intake has the potential to meaningfully benefit population health, through attenuation of weight gain or reductions in bodyweight (Benton 2017; Robinson 2023a; Shelton 2014). Figure 1 presents a logic model of processes by which calorie labelling may impact upon food and alcohol selection and consumption, which are determinants of health outcomes. Concerning direct effects of viewing calorie labelling, a suggested key underlying mechanism is via increased knowledge of the energy content, related appraisals of the healthiness or other values of food and drink products, and subsequent deliberative and intentional actions to select or not select a particular food or drink (Muller 2016). Labelling interventions may also directly influence behaviour via less-conscious routes (Hollands 2016; Ventsel 2022; Zahedi 2023), especially where interpretative elements such as colours or symbols are included (Muller 2016). Possible key modifiers of any intervention effect include the setting in which people purchase the targeted products, such as a store or a restaurant, the type of label used, and the type of product that is labelled (e.g. main meals, snacks, or drinks). Other potential modifiers include the country in which the labelling is implemented, and socioeconomic status. Obesity in high-income countries is patterned by an individual's material, psychological, and social resources, reflected in levels of education, occupational status, family income, and area of residence (NHS Digital 2019; Winkleby 1992).

**Figure 1. Logic model. BMI: body mass index; COI: conflicts of interest.**




This review includes three planned subgroup analyses to investigate the type of label, setting, and socioeconomic status as potential effect modifiers (see Subgroup analysis and investigation of heterogeneity). We prioritised these for three reasons. First, it is widely acknowledged that labels differ in their presentation, and that this may lead to differential effects (e.g. see Daley 2020), and have implications for developing and implementing labelling interventions. Second, previous reviews have also shown differential effects of labelling interventions by study setting (e.g. Clarke 2021a), and this could have important implications for the implementation of labels in different real-world contexts. Third, socioeconomic status could modify any impact of calorie labelling, potentially via lower levels of understanding in more disadvantaged populations (Sarink 2016). It is important to know whether an intervention could exacerbate existing inequalities by having a lesser (or potentially even harmful) effect in people of low socioeconomic status, particularly given obesity rates are already patterned by socioeconomic status. More consistent reporting of differential intervention effectiveness will help to build the evidence base to test for and mitigate so-called "Intervention-Generated-Inequalities" (IGI) (Lorenc 2013). Participants in a study are sometimes categorised in terms of material and social deprivation based on individual-level characteristics or characteristics of where they reside (NHS Digital 2019; Winkleby 1992). Previous evaluations of the differential effects of labelling by socioeconomic status concluded that the current evidence was limited in both quantity and quality, and that further evaluation of the differential effectiveness of labelling by socioeconomic status was required (Sarink 2016). In accordance with methodological guidance on the importance of considering inequalities, incorporating study-level data on socioeconomic status of participants is intended to enable us to interpret any differential effects through a health equity lens (Hollands 2024; Welch 2012; Welch 2016).

This review focussed specifically on the direct effects of calorie labelling on people who select or consume the targeted food and alcohol products. However, as illustrated in Figure 1, there may also be indirect effects of implementing such interventions on people, because industry may modify its actions or practices. Calorie labelling could result in industry actions (e.g. changes in marketing) to dilute the impact of the intervention and related policies (Capewell 2018), or that could enhance its effects should the intervention lead to reformulation of high-calorie products or renovations of the product portfolio (Grummon 2021; Robinson 2021a; Shangguan 2019).

### Why it is important to do this review

The large and increasing burden of diet- and alcohol-related disease worldwide requires population-level interventions to promote sustained changes in behaviour. Although calorie labelling has been unsystematically implemented in North America, Europe, and Australasia, there is no consensus on the extent to which it might be effective in influencing selection and consumption of food and alcoholic drinks.

This uncertainty about the potential impacts of calorie labelling was not fully resolved by the previous version of this Cochrane review, as it identified only a small body of low- or very low-certainty evidence for all review outcomes (Crockett 2018). The previous version included only three randomised controlled trials (RCTs) examining selection with purchasing in real-world field settings. However, we were aware of several new potentially eligible studies published since the original review, including RCTs in real-world settings (Cawley 2020; Dubois 2021; Vasiljevic 2018; Vasiljevic 2019), and research investigating a relatively new form of calorie label, known as PACE (physical activity calorie equivalent) labels (Daley 2020; Reynolds 2022).

Now that a larger body of evidence has accumulated, it is important to generate up-to-date estimates of effects, incorporating new data from more-recent studies, with the potential for increased levels of certainty of evidence. This could help to fulfil an ongoing need for robust evidence to support decisions regarding the implementation of calorie labelling and the development and continuation of food and alcohol policy and programmes in the UK and globally. A range of UK Government reports (e.g. "*Calorie reduction: the scope and ambition for action*" (Public Health England 2018) and Chapter 2 of "*Childhood obesity: a plan for action*" (Department of Health and Social Care 2018b)) cite the previous version of this Cochrane review (Crockett 2018), and the meta-analytic effect size estimated in the previous version was used in calculations for the Department of Health's initial impact assessment for its consultation on mandating out-of-home calorie labelling (Department of Health and Social Care 2018a), a policy since implemented in England in April 2022 (Department of Health and Social Care 2021). The World Cancer Research Fund's (WCRF) global 'Nourishing' framework also includes a range of policy options to promote healthier eating and prevent obesity, including nutrition label standards and informing people about food and nutrition (World Cancer Research Fund 2021).

The previous version of this review focused on nutritional labelling applied to food (including non-alcoholic drinks) (Crockett 2018). For this updated review, we have narrowed our intervention focus to calorie labelling, as most studies in Crockett 2018 included calorie labelling as the sole component, or as part of the broader nutritional labelling scheme, and this form of labelling remains the principal focus of continued policy and research interest (Department of Health and Social Care 2020). Concurrently, we have broadened the range of target products of the intervention to include alcoholic drink products, reflecting ongoing policy interest (Royal Society for Public Health 2018; Department of Health and Social Care 2020).

## OBJECTIVES

- To estimate the effect of calorie labelling for food (including non-alcoholic drinks) and alcoholic drinks on selection (with or without purchasing) and consumption.
- To assess possible modifiers – label type, setting, and socioeconomic status – of the effect of calorie labelling on selection (with or without purchasing) and consumption of food and alcohol.

## METHODS

### Criteria for considering studies for this review

#### Types of studies

Similar to the approach in Crockett 2018 and as reflected in our published protocol (Clarke 2021b), we considered both RCTs and non-randomised studies of interventions.

We included RCTs or quasi-RCTs with between-subjects (parallel group) or within-subjects (cross-over) designs that compared a calorie labelling intervention with a no-label control (or equivalent, see Types of interventions). We included quasi-RCTs, in which the randomisation sequence was not truly random because of the difficulty of implementing true randomisation at an aggregate population level (Reeves 2020). We also included cluster-RCTs, when randomisation was by site (e.g. by restaurant), provided the study included at least two intervention sites and two control sites.

We included interrupted time series studies that compared selection or consumption before and after the implementation of calorie labelling. In line with the Cochrane Effective Practice and Organisation of Care (EPOC) group recommendations, and given the potential instability of real-world food and alcohol data, we only included interrupted time series studies if they had a clearly defined time point at which the intervention occurred and at least three observation periods both before and after the intervention with each observation period being of at least one month in duration (EPOC 2017). Authors were required to present these data within a graph or analyse them using regression analysis, preferably using segmented regression, or both. Based on Cochrane recommendations, we excluded studies that reported only a simple pre- and post-intervention comparison (EPOC 2017). We also included controlled before-after studies that measured selection or consumption before and after implementation of an intervention in non-randomised intervention and control groups. Such studies were required to have at least two intervention sites and two control sites, and the characteristics of the different groups had to be similar. If we identified controlled before-after designs that were eligible, members of the review team not involved in extracting data (SAJ, TMM, MP), blinded to study results, agreed and specified criteria for comparability dependent on the specific context of that study. At the outset, we expected intervention and control sites to be well-matched for at least type of site, geographic location, population size, and population demographic characteristics.

### Types of participants

Adults or children selecting (with or without purchasing) or consuming food or drink were eligible for inclusion. Selections included those made by an individual for their personal consumption, or for consumption by a small group to which the individual belonged, such as their family. Food or drink selections included those from any retail outlet, including shops or supermarkets, vending machines, bars, pubs, cafeterias, and both fast-food and other types of restaurants.

### Types of interventions

Eligible interventions included calorie labelling of a food (including non-alcoholic drinks) or an alcoholic drink product. Eligible labels were required to possess two sets of characteristics.

***1. Information about energy content.*** The label was required to provide information about energy contained in the product. Information was given specifying the absolute amount of energy contained in the product or in a serving size. The calorie label was required to display numeric information about the energy content; for example, calories in a meal/pack/serving/drink.

Ineligible labels were those that only used a relative or categorical descriptor; for example, low(er)/high(er) energy, or an indication that it met a certain threshold rather than an exact amount. Warning labels about the health implications of a product's energy content were deemed ineligible labels (e.g. Clarke 2021a). We also excluded logos or general health claims providing a summary assessment of the healthiness or content of a product. We excluded studies that purposefully mislabelled the energy content of products to mislead or misinform participants, meaning that products in included studies had to have their content accurately described.

***2. Visibility.*** The calorie labels were required to be visible at the point of selection or consumption. In some cases, the label was placed on the front-of-product packages or containers. In other cases, the calorie label did not appear on, but rather alongside, the product. Examples included labels on a shelf where the products were being displayed in a shop, on the exterior of a vending machine selling snacks, on the counter from which the food was being served in a cafeteria, or on a restaurant menu from which food or drinks were being selected. Studies were ineligible if information was provided separately, for example, on a company website.

As noted in Types of studies, the intervention labelling group had to be compared with a no-labelling control group (or equivalent, including the presence of labelling that was not health-, nutrition-, or energy content-related, or of back-of-pack nutritional labelling providing this was also present in the intervention group).

We only included interventions that combined a calorie label with other substantive discrete intervention components – either other types of (non-calorie-related) labelling, or other interventions unrelated to labelling – if we could isolate the effect of the calorie label. We excluded studies that assessed multiple intervention components that included calorie labels but did not allow the effect of the latter to be isolated. For example, if an intervention group combined a calorie label intervention and a pricing intervention, and this was compared to a control condition that did not include a pricing intervention, the specific effect of the calorie label could not be estimated.

Intervention details were ascertained through study reports – either text descriptions or images of the calorie label, or both. We contacted study authors if there were insufficient details on the intervention or comparator groups to assess the study's eligibility.

### Types of outcome measures

Eligible studies were required to assess an objectively measured and unconstrained behavioural outcome of food or drinks selected (with or without purchasing) or consumed (see more specific parameters below).

#### *Primary outcomes*

- Food and non-alcoholic drinks selected (with or without purchasing)
- Alcoholic drinks selected (with or without purchasing)
- Food and non-alcoholic drinks consumed
- Alcoholic drinks consumed

These are described in more detail below.

##### Selection (with or without purchasing) of food (including non-alcoholic drinks) and alcoholic drinks

Depending on how this was assessed, it could have reflected either total energy selected or the number of specified products (e.g. higher energy items) selected. Where total energy selected was available, we prioritised this. For alcoholic drinks, this outcome could have reflected either total energy selected or the number of specified products selected, or the volume or units of alcohol selected, or both. In our analyses, we considered a reduction in such selection to represent a beneficial effect of the intervention.

Studies were required to assess selection with or without purchasing either at the individual or population group level. An individual-level selection outcome measure required direct (and not self-reported) measurement of what was selected, whether assessed as energy or items selected. At a population level, selection with purchasing data had to be derived from sales data (e.g. supplied by the retailer from till receipts). Such data could be presented as selection of specified products or as total energy selected, calculated from the sales data presented.

We excluded studies that evaluated only intention or motivation to select or purchase, or hypothetical selection; studies without a measure of actual behaviour using real products that were received, or purchased; and studies that measured constrained selection, whereby the behaviour of participants is strictly regulated by either explicit instructions or some other action of the researcher. For example, we excluded studies where participants were unable to purchase or receive all the products that they had selected, as well as studies that required participants to complete additional measures during the intervention or outcome measurement period.

#### Consumption of food (including non-alcoholic drinks) and alcoholic drinks

Depending on how this was assessed, it could reflect either total energy consumed, or the number of specified products (e.g. higher energy items) consumed. Where total energy consumed was available, we prioritised this. For alcoholic drinks, this outcome could reflect either total energy consumed, or the number of specified products consumed, or the volume or units of alcohol consumed, or both. In our analyses, we considered a reduction in such consumption to represent a beneficial effect of the intervention.

Studies were required to assess consumption by an objective measure, calculating the amount of a snack, meal or drink consumed by subtracting the amount of food or drink remaining after consumption from the amount served. This was required to be specified as either the amount of food or drink product(s) consumed or total energy consumed calculated from the amount.

We excluded studies that evaluated only intention or motivation to consume, or hypothetical consumption; studies without a measure of actual behaviour using real products that were consumed; and studies that measured constrained consumption, whereby the behaviour of participants is strictly regulated by either explicit instructions or some other action of the researcher. For example, we excluded studies that presented a set of products to individual participants with an instruction to consume a given quantity, as well as studies that required participants to complete additional measures during the intervention or outcome measurement period.

As per our published protocol (Clarke 2021b), we did not expect any significant health-related harms to be consistently assessed or reported beyond those related to a change in behaviour in the unintended direction, that is, calorie labels causing higher energy consumption. This potential unintended consequence of energy labelling is an outcome that would have been captured by this review, and we also planned to record any adverse events or harms reported in the primary studies, but there were none reported.

For all primary outcomes, we had no time point restrictions and extracted data on the outcomes reported that were furthest away in time from the intervention.

### Search methods for identification of studies

In July 2021, we conducted the main electronic database search for studies of food and non-alcoholic drinks labelling from the search date (April 2017) of the previous Cochrane review on nutritional labelling (Crockett 2018) (noting that this previous review had a broader, more inclusive scope). Concurrently, we also conducted a separate search for studies on calorie labelling of alcoholic drinks – newly included for this update – from database inception. We integrated eligible studies from these searches into this review. This was supplemented by comprehensive Microsoft Academic Graph (MAG) searches conducted on 1 September 2021, using the latest MAG dataset that included records up to 2 August 2021. Because these comprehensive database searches were only fully integrated to this date, we therefore consider the evidence included in this review to be current to 2 August 2021.

Following subsequent updated 'top-up' searches in September 2023 (see Electronic searches for details), we screened, and considered for inclusion, evidence up to September 2023, but did not fully integrate the results of new studies identified at this point. These updated 'top-up' searches, covering the period up to 25 August 2023, identified one eligible study: Petimar 2022, which, while being provisionally accepted into the review, is yet to be fully integrated until the next review update. This provisionally accepted study is retained in Studies awaiting classification table, along with four additional studies for which we could not confidently determine their eligibility from available material. The decision not to fully integrate further results of newly included studies at this stage was a pragmatic one, based on a balance of the likelihood of this study changing results and conclusions, relative to the potential disadvantages of delaying publication further. In this instance, Petimar 2022 was a non-randomised study that would not be included in the meta-analysis for its outcome (and for which evidence was already judged at high certainty), and the data appeared to be consistent with the data for randomised and non-randomised studies. For further details, see Appendix 1.

#### Electronic searches

Appendix 2 reproduces full details of our electronic search strategies. Initially, we conducted conventional electronic searches of the following literature databases.

- Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library (2021, Issue 7)
- MEDLINE (OvidSP)
- Embase (OvidSP)
- PsycINFO (OvidSP)
- Applied Social Sciences Index and Abstracts (ASSIA) from Cambridge Scientific Abstracts (CSA) (ProQuest)
- Science Citation Index (Web of Science Core Collection, Thomson Reuters)

- Social Science Citation Index (Web of Science Core Collection, Thomson Reuters)
- Conference Proceedings Citation Index – Science (Web of Science)
- Conference Proceedings Citation Index – Social Science & Humanities (Web of Science)

The last two databases listed above are grey literature databases that specialise in indexing conference proceedings (Embase also includes indexing of conference proceedings). A first set of conventional searches of the electronic databases listed above aimed to retrieve records of eligible studies of calorie labelling applied to food and non-alcoholic drink products. This set of searches used a modified version of the search strategy used for the previous version of this review (Crockett 2018), but excluded terms relating to labels for fat, sugar, or salt content (no longer an eligible intervention), and covered a date range from 25 April 2017 up to 9 July 2021, thereby updating the (broader) electronic searches conducted for Crockett 2018. A second set of conventional electronic database searches aimed to retrieve records of eligible studies of calorie labelling applied to alcoholic drink products. We developed a new search strategy for the latter set of searches, which covered a date range from database inception up to 9 July 2021.

Alongside the conventional electronic database searches described above, we conducted two types of automated searches of the MAG dataset via EPPI-Reviewer (Thomas 2020) – a network graph search, and a custom search. In 2021, the MAG dataset was a regularly updated, open-access dataset, maintained by Microsoft, which comprised about 240 million bibliographic records of research articles on all topics across science, connected in a large network graph of conceptual, citation, and other relationships. At the same time, EPPI-Reviewer hosted a suite of tools, known as MAG Browser tools, which we had developed to enable automated searching of the MAG dataset.

Both types of MAG searches were conducted on 1 September 2021, using the latest MAG dataset that included records up to 2 August 2021. The network graph search automatically retrieved all those MAG records that were (on the date of search) connected in the MAG network graph, via specified relationships, to a set of 'seed' records (reports). The 'seed' records in this case were MAG records of 26 reports of all 28 studies included in the previous version of this review. The specified MAG network graph relationships were 'bi-directional citation' (i.e. MAG records in the bibliographies of, or cited by, the 26 'seed' records/ reports) and 'bi-directional recommendations' (i.e. MAG records closely related to the 26 'seed' records/ reports, based on a composite metric that quantified all MAG network graph relationships, and thereby either 'recommending' or 'recommended by' those 'seed' records). The second type of MAG search we conducted was a custom search. This search automatically retrieved all MAG records that were (on the date of search) tagged with one or more relevant MAG 'fields of study' specified in our custom search strategy (see Appendix 2 for full details). MAG 'fields of study' were topic terms, conceptually equivalent to indexing terms in conventional electronic databases (e.g. MeSH terms in MEDLINE and PubMed).

#### Updated 'top-up' searches (September 2023)

At the end of December 2021, the MAG dataset was replaced and superseded by the OpenAlex dataset. The OpenAlex dataset therefore now incorporates and maintains the MAG dataset and, correspondingly, EPPI-Reviewer has transitioned to incorporating re-engineered OpenAlex tools (replacing MAG Browser tools). OpenAlex tools enable network graph searches and custom searches of the OpenAlex dataset, as described above for MAG. A third type of automated search enabled by OpenAlex tools is known as the auto-update search (initially developed for the MAG dataset). Like network graph searches, auto-update searches are 'seeded' by included study reports and their corresponding OpenAlex records. These 'seed' records are subscribed to a novel machine learning recommender model (known as the 'auto-update model'), which automatically scores all 'new' records prospectively added to each sequential update of OpenAlex dataset (updated approximately monthly), and recommends (retrieves) 'new' OpenAlex records that are most likely to be eligible for inclusion in the subscribed systematic review. We subscribed study reports included on full text in this review to the auto-update model in March 2022 (i.e. once the initial, main tranche of full-text screening for the current update had been completed).

In September 2023, we retrieved the top 100 scoring records from multiple auto-update searches of 12 consecutive updates of the OpenAlex dataset, from 11 March 2022 up to 20 August 2023. We also conducted a further network graph search of the OpenAlex dataset from 1 January 2017 up to 25 August 2023, 'seeded' by records (reports) of included studies identified up to the latter date, which retrieved records either in the bibliographies, or cited by, that updated set of 'seed' records (i.e. a 'bi-directional citation' search, which is equivalent to a one-step forwards and backwards citation search). Finally, we conducted an updated custom search of the OpenAlex dataset from 1 January 2017 up to 25 August 2023, which retrieved further OpenAlex records tagged with the same topic terms (OpenAlex 'concepts', which superseded MAG 'fields of study') that we used in the initial MAG custom search (see Appendix 2 for full details).

In all cases, we uploaded retrieved bibliographic records to EPPI-Reviewer (Thomas 2020) for deduplication before screening (see Selection of studies). We did not integrate the results of new studies identified via these updated 'top-up' searches in September 2023 into the review (see Appendix 1 for details).

#### Searching other resources

Concurrent with the MAG searches conducted in September 2021, and again in July 2023, we also searched the following websites of key organisations in health and nutrition.

- Department of Health and Social Care, England (www.gov.uk/government/organisations/department-of-health)
- Department of Health and Social Care, Scotland (www.gov.scot/Topics/Health)
- Departments of Health and Social Care, Wales (gov.wales/topics/health/?lang=en)
- Department of Health, Northern Ireland (www.health-ni.gov.uk)
- European Commission (ec.europa.eu/commission/index en)
- Centers for Disease Control and Prevention (www.cdc.gov)
- World Health Organization (WHO) (who.int/en)
- National Institutes for Health Office of Disease Prevention (prevention.nih.gov)
- World Obesity Federation (www.worldobesity.org)
- Institute of Alcohol Studies (www.ias.org.uk/Home.aspx)

- National Institute on Alcohol Abuse and Alcoholism (www.niaaa.nih.gov/)
- Alcohol Change UK (alcoholchange.org.uk/)

We also searched trial registries for potentially relevant studies that were completed or in progress, using the WHO International Clinical Trials Registry Platform (ICTRP; apps.who.int/trialsearch/) and ClinicalTrials.gov (clinicaltrials.gov/) via CENTRAL, and the EU Clinical Trials Register (www.clinicaltrialsregister.eu/). In September 2021, we identified trial registrations for two studies that were assessed and included in the review (Clarke 2023a; Reynolds 2022).

We checked whether any included or eligible studies had postpublication amendments, including any retractions or errata, following guidance from the *Cochrane Handbook for Systematic Reviews of Interventions* (Lefebvre 2023). We checked the article webpage as enlisted on the journal website for any relevant retraction statements and errata as well as the Retraction Watch Database (retractiondatabase.org/). In addition, for studies that were identified from MEDLINE or Embase, we also checked the complete version of its most recent citation in the corresponding database.

## Data collection and analysis

### Selection of studies

The study selection process was managed using EPPI-Reviewer (Thomas 2020). First, title-abstract records retrieved by the electronic searches described above were deduplicated using a three-stage, semi-automated procedure: 1. records were automatically scored based on the similarity of their text features, and those with similarity scores $^{3}$0.70 threshold were organised into duplicate groups, with one record marked as the master in each group; 2. records with a similarity score $^{3}$0.85 threshold, compared with the corresponding master record, were automatically marked as duplicates and discarded; 3. records with a similarity score in the 0.70 to 0.849 range, compared with a designated master record, were manually checked and either marked as duplicates and discarded, or else marked as not duplicates and retained for potential screening.

Screening was conducted in two stages. In the first stage of screening, the (semi-automated) title-abstract screening stage, retained (deduplicated) title-abstract records were manually screened in a prioritised order determined by active learning (i.e. using 'priority screening mode' in EPPI-Reviewer). Specifically, a rank-ordered list of those records yet to be screened was continually reprioritised by a machine learning algorithm, which progressively learned to distinguish between potentially eligible and clearly ineligible records, initially 'seeded' by records screened on title-abstract when producing the previous version of this review. Two review authors (from NC, EP, GJH, IS) independently assessed each prioritised title-abstract record against the eligibility criteria, with disagreements resolved by discussion and consensus, or else in discussion with a third review author acting as arbiter. To ensure that no authors who were involved in the primary research studies were involved in screening, we involved additional authors where necessary. For the initial, main tranche of title-abstract screening, we truncated screening in priority screening mode once we had not encountered a potentially eligible record for more than 700 records, temporarily setting aside the remaining unscreened records. For the second tranche of title-abstract screening, we added further retained (deduplicated) records from the updated searches (September 2023) to the pool of unscreened records from the main tranche, and continued screening in priority screening mode, until we had not encountered a potentially eligible record for more than 700 records, at which point we truncated title-abstract screening for the current update of this review and set aside the remaining unscreened records. In the second stage of screening, we retrieved the corresponding full-text reports of potentially eligible studies identified during the title-abstract screening phase, and manually assessed the study reports for inclusion using the same eligibility criteria and procedures described above. We also retrieved and screened potentially eligible reports identified by searching other resources.

Two review authors (NC, EP) independently assessed each report against the eligibility criteria described above, with disagreements resolved by discussion and consensus, or else in discussion with a third review author (GJH) acting as arbiter. Finally, we linked multiple reports of the same study.

### Data extraction and management

We developed a data extraction form based on the one used for the previous version of this review. Two review authors (NC, EP) independently piloted the draft to ensure that it enabled reliable and accurate extraction of appropriate data. Two review authors then independently extracted all data on study characteristics along with results (NC, EP). If a review author was also an author of an included study, another review author was involved in the data extraction process. Once the first phase of data extraction was complete, the first author (NC) reconciled the two sets of data extraction forms. Where there were inconsistencies, the two data extractors met to discuss and reach a consensus (NC, EP). Where outcome data were missing or unclear, we contacted study authors. Finally, one author (NC) entered the data into Review Manager (RevMan 2024), and a second author checked the data entry (GJH). When multiple articles reported data from the same study, we treated the articles as one study, selecting the principal results article as the study's primary reference. For each eligible study, we collected the data summarised below, representing the core dataset we considered was required based on the review's eligibility criteria and its logic model:

#### Study characteristics

- Study design
- Summary risk of bias assessments
- Information on funding source and potential conflicts of interest from funding

#### Setting characteristics

- Geographical setting: country (and country classification by income)
- Study (intervention) setting: restaurant (field); store (field); naturalistic laboratory; laboratory

#### Participant characteristics

- Population group (e.g. general population)
- Socioeconomic status (e.g. by reference to education, occupation, income, location, a combination of these)
- Age/age group

- Sex/gender (e.g. male, female)
- Ethnicity
- Body mass index/bodyweight/bodyweight status

**Intervention and outcome characteristics**

- Product type: food; alcohol
- Type of labelling used: simple calorie labelling; calories with PACE labelling; calorie labelling with information about at least one other nutrient
- Label placement
- Duration of exposure
- Duration/timing of outcome assessment
- Relationship between manipulated product and other available products
- Concurrent intervention component(s) (in factorial design/ confounded with comparison of interest)
- Target selection with purchasing; selection without purchasing; consumption
- Relationship between manipulated product and outcome

**Outcome data**

Where studies included more than one eligible measure of selection or consumption, we used the measure of selection or consumption (pre)specified by the study authors as the primary outcome. If study authors specified no primary outcome, we used the measure of selection or consumption that accounted for the largest proportion of the overall diet. For example, if a study reported consumption measures of both total energy intake from a meal as well as energy intake from a specific food only (e.g. a chocolate cake), we used total energy intake from a meal.

We also aimed to use measures that related to energy or alcohol content rather than, for example, quantity, mass, or volume. If studies reported comparable outcomes using a metric that was not a measure of energy (such as grams or millilitres), in consultation with the public health nutritionist on the review team (SAJ), we determined if it was possible and appropriate to convert these to calories (e.g. by using the formula presented in DeGroot 2012). For alcoholic drinks, if necessary, we converted alcohol by volume to alcohol content in units of alcohol.

**Assessment of risk of bias in included studies**

Two review authors (NC, EP) independently assessed the risk of bias for the primary outcomes for each included study (see Primary outcomes). We resolved any disagreements through discussion, involving another review author as necessary (GJH). To ensure that no authors who were involved in the primary research studies were involved in risk of bias assessment, we involved additional authors where necessary. We assessed risk of bias for RCTs using the RoB 2 tool (Sterne 2019), using the accompanying Excel tool to manage the risk of bias assessments, and included a risk of bias table. We assessed the effect of the assignment to the intervention for the following domains: bias arising from the randomisation process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in measurement of the outcome, and bias in selection of the reported result.

For cross-over trials, we followed the supplementary RoB 2 guidance for cross-over trials. This revised tool addresses bias arising from period and carryover effects (Higgins 2023; Higgins 2021b).

For cluster-RCTs, we followed the supplementary RoB 2 guidance for assessing risk of bias of these designs (Eldridge 2020). This revised tool includes assessing bias arising from the identification or recruitment of participants into clusters, and bias arising from the timing of identification and recruitment of individual participants in relation to timing of randomisation.

For each study, we judged each included outcome as 'low risk of bias', 'some concerns', or 'high risk of bias', according to these criteria:

- low risk of bias: when all domains were classified at low risk;
- some concerns: when one or more domains were classified as raising some concerns, but no domain was classified at high risk of bias;
- high risk of bias: when one or more domains were classified at high risk of bias, or multiple domains were classified as raising some concerns to the extent that these reduce confidence in the results.

We assessed the risk of bias of controlled before-after and interrupted time-series studies using the ROBINS-I tool (Sterne 2016a), and appropriate guidance (Sterne 2016b; Sterne 2020). This tool focusses on a specific result, uses a fixed set of domains of bias, and leads to an overall risk of bias judgement. For each study, we specified a target trial (a hypothetical randomised trial whose results should be the same as the non-randomised trial under consideration). The effect of interest was the effect of assignment to the intervention at baseline regardless of the extent to which the interventions were received (i.e. intention-to-treat). Before the tool was completed, we specified whether the predefined critical confounders and co-interventions were identified in the study. Based on findings from the previous iteration of the review (Crockett 2018), and the research team's knowledge of the literature, at the outset we considered the potential confounding domains to be: key participant demographic characteristics (age, sex, and socioeconomic status distributions), and key study setting (intervention setting) characteristics (major events, different geographic areas, different retail environments). We considered possible co-interventions to be changes to labelling, other changes in product packaging or presentation, or any other interventions within the same physical environment (as classified by the TIPPME typology of these interventions (Hollands 2017)), or economic environment, such as price changes. We assessed the following domains: bias due to confounding; bias in selection of participants into the study; bias in classification of intervention; bias due to deviations from intended interventions; bias due to missing data; bias in measurement of outcomes; and bias in selection of reported result. This tool leads to judgements for each risk of bias domain, and for overall risk of bias, which can be 'low', 'moderate', 'serious', or 'critical'.

**Measures of treatment effect**

Selection data could be either dichotomous (e.g. a selection with more versus less energy) or continuous (e.g. total or mean amount of energy purchased), while we anticipated that consumption would typically be assessed using continuous data only (e.g. total or mean energy consumed) on the same measurement scale (i.e. kilocalories or data we could transform to kilocalories). We intended to calculate a mean difference (MD) with 95% confidence intervals (CIs) for each study when possible. In practice, all included outcomes were represented by continuous data, and it was only possible to calculate standardised mean differences (SMD) with 95% CIs because different measures were used, or the study investigators did not report the necessary data. We intended to report a risk ratio (RR) as the effect size (Mantel-Haenszel method) had we included dichotomous data.

In order to re-express SMD effect sizes using a more familiar metric, for estimating absolute effects for analyses of food and non-alcoholic drinks, we calculated the percentage change in calories consumed over a typical meal (given that most of the included studies manipulated labelling over the period of a meal), using a mean average of 600 (standard deviation (SD) 185) kcal as a baseline value for the comparison group. This amount was based on mean daily energy intake from food in UK adults (aged 19 to 64 years) of 1752 (SD 540) kcal, using data from the most recent available years (9 to 11) of the UK National Diet and Nutrition Survey (Public Health England 2020). We assumed a consistent relationship between the mean average and the variance for different values for energy intake, and that for selection outcomes, all energy selected is consumed. For alcoholic drinks, we had intended to apply a similar approach using data from the most recent available years (9 to 11) of the UK National Diet and Nutrition Survey (Public Health England 2020). This would have involved using a mean of 145.6 (SD 135.8) kcal as a baseline for daily energy consumption from alcohol, based on daily amount of alcohol consumed among UK adults (aged 19 to 64 years) who consume alcohol estimated at 20.8 (SD 19.4) g of alcohol (with 1 g of alcohol containing 7 kcal) by the most recent available years (9 to 11) (Public Health England 2020). We did not conduct this extrapolation as we judged it inappropriate given the very low-certainty evidence with wide CIs for the outcome of selection of alcoholic drinks.

There are important limitations of such translations or extrapolations (Hollands 2019). First, these re-expressed values relate to UK populations (although readers could translate the results using similar methods using representative survey data from other countries). Second, because such translations necessarily extrapolate beyond the scope of the included studies, they are intended only to be illustrative for guiding interpretation of the meta-analyses. In this particular case, there is additional extrapolation in assuming that the variance associated with total daily energy intake will be proportionate for lower levels of energy intake from a meal. The survey data used are also based on self-report, and there is some evidence to suggest self-report data may be underestimates (Cook 2000). Finally, for alcoholic drinks there is an additional issue that data such as that described above does not account for additional sources of energy intake in drinks that are consumed, other than the alcohol content itself.

In all the interrupted time series studies, we presented the results as described by the study authors, typically as regression analyses. We did not attempt any further re-analysis using segmented time series regression techniques as in all cases the review's statistician judged that the data were already appropriately analysed by the study authors or we did not consider the study to be of sufficient quality to warrant re-analysis, or both.

#### Unit of analysis issues

For cluster-RCTs, we estimated the effect accounting for the clustered study design, using reported test statistics (t statistics, F statistics, or P values) to calculate standard errors where necessary, and contacting authors for any required data that were missing.

We handled unit of analysis errors from studies with multiple intervention groups in line with the relevant Cochrane guidance (Higgins 2023). We combined groups that used similar types of calorie label and differed only in their presentational format, but included interventions as separate comparisons if treatment groups comprised multiple eligible calorie label types as per our prespecified criteria. For studies contributing multiple comparisons, we adjusted the study weights to account approximately for the statistical dependencies between comparisons by dividing the sample size of the common intervention group as evenly as possible between the comparisons. For handling cross-over trials that did not report data necessary for inclusion in the meta-analysis, we had planned to approximate paired analysis by imputing missing SDs where possible, following the methods outlined in Chapter 23 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2023). However, we identified no cross-over trials for inclusion.

#### Dealing with missing data

We included all data in the review using an intention-to-treat approach where possible. For studies reporting dropouts or withdrawals, we extracted relevant information on the extent and reasons for missing data using the Cochrane RoB 2 tool. We attempted to obtain missing outcome data where possible by contacting authors. For three studies that did not report the results in a format required for analysis, we imputed values following contact with the authors (Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019). It was determined in discussion with the statistician who conducted the analysis for these studies (also the statistician in our review team) that a meaningful SMD could not be extracted from the model used and reported by the authors. The modelling (generalised additive (mixed) models for location scale and shape) utilised for these studies used splines to perform smoothing, but smoothed models are as yet not supported in available packages for calculating effect sizes from such data. As such, we used available data to extrapolate statistical effect sizes. The studies reported intervention effect sizes as percentage changes (to one decimal place) in calories selected and their associated 95% CIs, while in this review we extrapolated statistical effect sizes to percentage changes in calorie selection and consumption based on mean average and variance values from UK population estimates from national survey data (see Measures of treatment effect). Therefore, for these three studies for which we had known values for percentage changes in calorie selection (to one decimal place), we imputed the SMD and standard error values that would be extrapolated to the observed percentage change. In doing so, we erred towards using conservative point estimate and variance values that produced (i.e. would be extrapolated to) equivalent percent change values to one decimal place, and ensured CIs were evenly spaced from the point estimate. The resulting imputed standard error values were also similar to the standard errors extracted for the two largest randomised studies included in the meta-analysis. This approach has similar caveats to that which apply more generally to such translations (see Measures of treatment effect). As effect size data from these studies were imputed, we conducted a sensitivity analysis (see Effects of interventions) removing these three studies (Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019), and which did not affect the results. We had also previously considered imputation of the effect relative to available baseline data from the most precisely estimated comparable study (Vasiljevic 2019), but this was less conservative, suggesting larger effects.

**Assessment of heterogeneity**

In order to deal with inevitable methodological variability among studies that evaluated food or drinks consumed in real-world or laboratory settings, we considered studies that evaluated calorie labelling as their interventions, and calorie selection or consumption as their outcomes, to be similar enough to be meaningfully combined in respective meta-analyses. We assessed clinical heterogeneity in participants, interventions, and outcomes by comparing study characteristics across studies and applying prespecified subgroup analyses to disentangle key characteristics (i.e. label type, study setting, socioeconomic status) that could contribute to observed heterogeneity. We assessed statistical heterogeneity by visually examining the extent to which CIs overlapped and by formal statistical tests of homogeneity ($Chi^2$) and measures of inconsistency ($I^2$ statistic) and heterogeneity ($Tau^2$). We interpreted the levels of heterogeneity made based on the recommendations of Deeks 2023.

**Assessment of reporting biases**

We used funnel plots to identify small-study effects, which in turn, indicate publication bias. We only used funnel plots if the meta-analyses included at least 10 studies, based on the recommendations of Sterne 2019.

**Data synthesis**

Where studies reported a number of different types of interventions or outcome measures, we followed the procedures described below.

- If treatment groups comprised multiple eligible calorie label types as per our prespecified criteria (e.g. calorie label with PACE compared with a simple calorie label comprising energy content only), then we included these interventions as separate comparisons in the analysis. We combined treatment groups into the same comparison where they had used similar types of calorie label and differed only in their presentational format (e.g. PACE labels that displayed calories as walking and PACE labels that displayed calories as running).
- For studies that used factorial designs to investigate the effects of calorie labelling as one of multiple intervention components, we combined outcome data across groups to capture the main effect attributable to calorie labelling. As previously stated, we excluded studies of only interventions with substantive concurrent components that were unrelated to but intrinsically confounded with calorie labelling.
- Where studies assessed the impact of calorie labelling relating to a range of products and it was not possible to extract a summary effect adequately representing the range of products, we included data representing the most complete range or the largest proportion of overall dietary energy; for example, sales of main meals/entrées (as opposed to sales of a side dish) (e.g. Dubbert 1984).
- Where studies reported several eligible selection or consumption outcomes, we used the primary outcomes specified by the study authors. If an outcome was not specified as the primary outcome, we prioritised the measure that accounted for the largest proportion of the overall diet. For example, if a study reported consumption measures of both total energy intake from a meal and the energy intake of a specific food (e.g. chocolate cake), we selected total energy intake from a meal. If outcomes were reported that related to increased consumption of lower energy foods and decreased consumption of higher energy foods, we prioritised the latter.

We analysed food (including non-alcoholic drinks) and alcohol studies separately. We also analysed selection and consumption outcomes separately, including separate meta-analyses. We expected that we would mainly encounter studies where the public health goal was, or aligned with, a decrease in the energy consumed from the labelled food or drinks. We did not encounter any studies where the intervention specifically aimed to increase consumption of a labelled food in order to improve population health – such as foods beneficial to health including fruit and vegetables.

We used Review Manager to perform random-effects meta-analyses, using the inverse variance weighting approach for continuous data (RevMan 2024). We conducted meta-analyses only for the results of RCTs, cluster-RCTs and quasi-RCTs. Where meta-analyses included studies for which we did not have necessary summary data for each intervention group, we instead included overall estimates of effect – standardised mean differences and standard errors – for all comparisons in that meta-analysis. Where possible, we computed such overall estimates of effect for such studies using already available summary data (means and SDs) for each intervention group entered separately as continuous outcome type data. We combined data for overall estimates of effect in meta-analyses using the generic inverse variance method with effect sizes reported as standardised mean differences. For studies that contributed multiple pairwise comparisons to a meta-analysis, we included each pairwise comparison separately, but adjusted the study weights to account approximately for the statistical dependencies between comparisons by dividing the sample size of the common group as evenly as possible between the comparisons.

Because of the increased risk of bias, we intended to summarise data from controlled before-after studies and interrupted time series studies in a narrative synthesis. In these and other cases where studies provided data that could not be included in meta-analysis, we used an acceptable narrative synthesis method as described in the *Cochrane Handbook for Systematic Reviews of Interventions* by using an additional table to present results in a systematic format (McKenzie 2023). We included the following variables: key participant, intervention, and study characteristics, potential key modifiers (as specified by our planned subgroup analyses), comparison group, primary outcome, and results (summary effect estimates) as reported by study authors.

We did not encounter any randomised cross-over trials, but we planned to combine cross-over trials that used an appropriate method of analysis (e.g. a paired analysis) with parallel design trials in the data synthesis and include them in the meta-analysis. For cross-over trials not appropriately reported, we planned to approximate a paired analysis by imputing SDs where possible, and not restrict analysis to the first period only. If dichotomous data were reported, we planned to discuss the most suitable approach for inclusion, guided by our statistician (MP). If we had identified and included cross-over trials and had found that pooled effect estimates from those studies suggested a systematic difference in effect size from parallel group studies, then we planned to exclude them in a sensitivity analysis (e.g. if there was considerable heterogeneity defined in the *Cochrane Handbook for Systematic Reviews of Interventions* as an $I^2$ statistic of 75% to 100% due to a difference between cross-over and parallel groups; Deeks 2023). We also planned to identify any concerns with specific cross-over trials (e.g. issues such as carryover effects) when assessing risk of bias for these types of studies using the modified RoB 2 tool (Higgins 2023). As such, interpretation of analyses including these study designs would have taken account of risk of bias considerations, including within formal assessments of evidence certainty for any given outcome.

### Subgroup analysis and investigation of heterogeneity

We conducted the following subgroup analyses using a test of subgroup differences ($Chi^2$) to determine the strength of evidence for possible effect modifiers and explore them as sources of heterogeneity:

- Environmental setting (restaurants; stores; naturalistic laboratories; laboratories). We classified naturalistic laboratories as those that attempted to mimic the real-world experience of purchasing food or drink.
- Type of calorie label used (simple calorie labelling; calories with PACE labelling; calorie labelling with information about at least one other nutrient (e.g. Food Standards Agency multiple traffic light labelling)).
- Socioeconomic status of participants (at the study level). Based on study authors' explicit descriptions of the study sample or setting indicating that they sampled a population with specific socioeconomic characteristics indicative of relative deprivation (e.g. low socioeconomic status, or a combination of low and high socioeconomic status) in terms of education, occupation, income, location, or a combination. Where this information was not provided, we assumed that a high socioeconomic status population was targeted. As in the previous version of the review (Hollands 2019), our planned subgroups were: low; high; and both low and high.

### Sensitivity analysis

We repeated meta-analyses including only studies judged to be at low risk of bias. We also re-ran any relevant analysis to exclude studies with imputed data.

### Summary of findings and assessment of the certainty of the evidence

We prepared two summary of findings tables for the primary outcomes, further organised according to the two product types: food and non-alcoholic drinks, and alcoholic drinks.

- Selection (with or without purchasing) and consumption of food and non-alcoholic drinks
- Selection (with or without purchasing) and consumption of alcoholic drinks

As specified in the *Cochrane Handbook for Systematic Reviews of Interventions* (Schünemann 2021), we included the following sections in the summary of findings tables.

- Population and setting addressed by the available evidence
- Intervention and comparison interventions (calorie labelling (intervention) versus no calorie labelling (comparison))
- Health outcomes
- Illustrative risk (risk with no labelling and risk with calorie labelling), and the absolute and relative effects for each outcome
- Number of participants and studies contributing to the analysis of each outcome
- GRADE assessment for the certainty of the evidence for each outcome
- Comments and explanations

Using the GRADE framework (Guyatt 2011), we assessed the certainty of each body of evidence relating to primary outcomes that are incorporated into summary estimates of effect. Two review authors (from NC, GJH, EP, IS) independently assessed the certainty of each body of evidence using the GRADE framework. We resolved any disagreements through discussion or by involving another review author. To ensure that no authors who were involved in the primary research studies were involved in GRADE assessment for those studies, we involved additional authors where necessary.

We considered risk of bias (study limitations), imprecision, indirectness, inconsistency, and publication bias in the GRADE assessments. Our risk-of-bias assessments using the RoB 2 tool for randomised studies and the ROBINS-I tool for non-randomised studies supported our GRADE assessment process (Schünemann 2021). Each body of evidence was given a GRADE rating. There are four standard GRADE levels of certainty: high, moderate, low, and very low. We assigned to the evidence from RCTs and non-randomised studies an initial certainty rating and levels were downgraded accordingly. We reported the degree of certainty assigned for each outcome, along with justification for the decisions in Effects of interventions. Justifications underpinning GRADE assessments were also included in the individual summary of findings tables.

# RESULTS

## Description of studies

### Results of the search

Our study selection process is shown in a PRISMA flow diagram (Figure 2; Page 2021). Searches were first conducted in July 2021 and updated in September 2023, and retrieved a total of 31,024 study records. Following the removal of 15,302 duplicated records, 15,722 records were processed in accordance with the semi-automated screening workflow described in Selection of studies. This semi-automated workflow excluded 9335 records, and a further 6262 records were manually screened and excluded by two review authors. We retrieved 125 full-text reports for eligibility assessment. From these, we included 14 new studies (from 34 full-text reports). These were added to 11 studies (15 reports) assessed as still eligible that had been included in the previous version of this review (Crockett 2018). In total, we included 25 studies in the review, described in 49 full-text reports. See Characteristics of included studies table.

**Figure 2. PRISMA flow diagram. [1] The review question addressed by the previous version of this Cochrane review (Crockett 2018) was amended for this current version and is therefore being addressed according to a new protocol, published 11 June 2021 (Clarke 2021b). The previous version of this review included a total of 32 reports (of 28 included studies), of which we included 15. Of the remaining 17 reports of studies that were not included, 16 were judged to be no longer eligible and were excluded from the current version, and 1 was categorised as awaiting classification, because the data available were insufficient to determine eligibility and authors could not be contacted. [2] A single report of a potentially eligible study of alcoholic drinks was identified by non-systematic scoping searches in advance of the searches conducted for this updated review. Two additional potentially eligible trial registrations were identified. [3] These records – which were not prioritised for title-abstract screening by active learning have been retained and set aside, pending the next major update of this review. During the next major update of this review, these records will be incorporated into the pool of candidate records to be prioritised (using**

**active learning) for title-abstract screening along with further records retrieved by the next tranche of updating searches. Ref.: Page 2021. n: number of records/studies.**

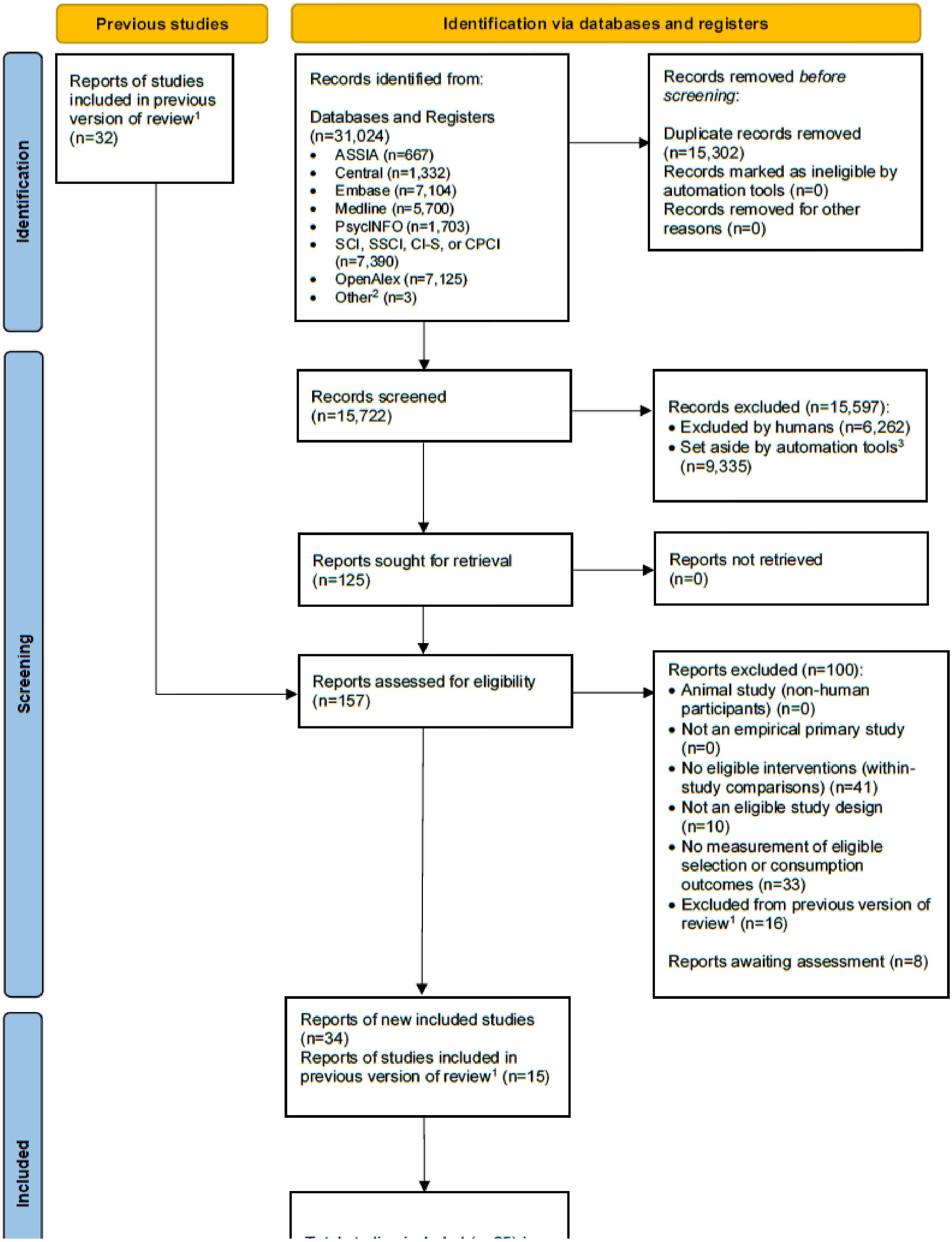
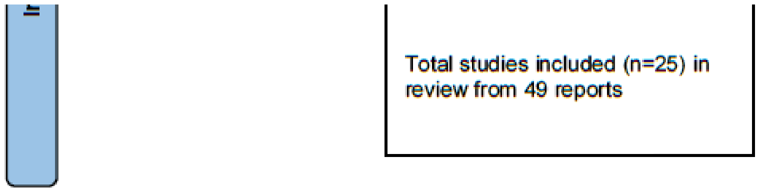


**Figure 2. (Continued)**




We categorised eight studies as awaiting classification (see Characteristics of studies awaiting classification table), and which will be considered for inclusion in the next update of this review.

We excluded 84 reports of studies at the full-text screening stage. We also excluded 16 reports of studies that had been included in a previous version of this review that had different eligibility criteria (Crockett 2018). See Characteristics of excluded studies table for details.

### Study characteristics

Twenty-five studies met the inclusion criteria for the review, of which 18 were RCTs. Of these RCTs, nine were randomised by individual (Clarke 2023a; Harnack 2008; James 2015; Platkin 2014; Roberto 2010; Roberto 2012; Robertson 2020; Temple 2010; VanEpps 2016), and nine were randomised by cluster, with either table or site being the unit of randomisation. Six cluster-RCTs were new to this updated review (Cawley 2020; Dubois 2021; Oliveira 2018; Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019); a further three RCTs randomised by cluster were included from the previous review (Ellison 2013; Ellison 2014; Hammond 2013). There was one quasi-RCT (Berry 2019), and six non-randomised studies, which were interrupted time series studies (Elshiewy 2018; Petimar 2021), or controlled before-after studies (Bollinger 2011; Elbel 2009; Fichera 2020; Petimar 2019).

### Settings and participants

Twenty-four studies were conducted in high-income countries, of which 15 were in the USA (Berry 2019; Bollinger 2011; Cawley 2020; Elbel 2009; Ellison 2013; Ellison 2014; Harnack 2008; James 2015; Petimar 2019; Petimar 2021; Platkin 2014; Roberto 2010; Roberto 2012; Temple 2010; VanEpps 2016), six in the UK (Clarke 2023a; Elshiewy 2018; Fichera 2020; Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019), one in Ireland (Robertson 2020), one in France (Dubois 2021), and one in Canada (Hammond 2013). One study was conducted in Brazil, which is classed as a low- or middle-income country (Oliveira 2018).

Sixteen studies were conducted in real-world field settings, of which 10 were in restaurants (Berry 2019; Bollinger 2011; Cawley 2020; Elbel 2009; Ellison 2013; Ellison 2014; Oliveira 2018; Petimar 2019; Petimar 2021; VanEpps 2016), three in worksite cafeterias (Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019), and three in supermarkets (Dubois 2021; Elshiewy 2018; Fichera 2020). The remaining nine studies were conducted in laboratory settings, of which six took place in naturalistic laboratory settings that attempted to mimic selection and purchasing in a real-world setting (Clarke 2023a; Hammond 2013; Harnack 2008; James 2015; Platkin 2014; Roberto 2012), and three were in laboratory settings that did not attempt to wholly mimic a real-world setting (Roberto 2010; Robertson 2020; Temple 2010).

Fourteen studies recruited a general population sample (Berry 2019; Bollinger 2011; Cawley 2020; Clarke 2023a; Dubois 2021; Elbel 2009; Elshiewy 2018; Fichera 2020; Harnack 2008; Petimar 2019; Petimar 2021; Roberto 2010; Roberto 2012; Robertson 2020), four studies recruited worksite employees from the general population (Reynolds 2022; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019), six studies recruited mainly student populations (Ellison 2013; Ellison 2014; Hammond 2013; James 2015; Oliveira 2018; Temple 2010), and one study recruited a specific sample of females, who were classified as overweight or obese (Platkin 2014).

In terms of socioeconomic status of the study samples, six studies were conducted purposefully in both high and low socioeconomic status samples (Dubois 2021; Harnack 2008; Petimar 2019; Petimar 2021; Vasiljevic 2018; Vasiljevic 2019), one was conducted in a low socioeconomic context (Elbel 2009); and the remaining 18 studies were conducted in high socioeconomic contexts (Berry 2019; Bollinger 2011; Cawley 2020; Clarke 2023a; Ellison 2013; Ellison 2014; Elshiewy 2018; Fichera 2020; Hammond 2013; James 2015; Oliveira 2018; Platkin 2014; Reynolds 2022; Roberto 2010; Roberto 2012; Robertson 2020; Temple 2010; VanEpps 2016).

The studies did not consistently describe full details of study participants (including age, gender/sex, and ethnicity) as this information was not always available for studies that randomised by site (Dubois 2021; Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019) or recorded supermarket transactions (Bollinger 2011; Elshiewy 2018; Petimar 2021). Of the studies that did include these details:

- 16 studies recruited participants aged 18 years or over (Berry 2019; Cawley 2020; Clarke 2023a; Elbel 2009; Ellison 2013; Ellison 2014; Hammond 2013; James 2015; Platkin 2014; Roberto 2010; Roberto 2012; Robertson 2020; Temple 2010; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019); one study recruited participants aged 20 years or over (Oliveira 2018), one study included adolescents over 16 years as well as adults (Harnack 2008), and one study recruited children, adolescents, and adults (Petimar 2021).
- 14 studies that included a mean age (Berry 2019; Cawley 2020; Clarke 2023a; Elbel 2009; Fichera 2020; James 2015; Platkin 2014; Reynolds 2022; Roberto 2010; Roberto 2012; Robertson 2020; Temple 2010; Petimar 2019; VanEpps 2016), this ranged from 21.9 years (James 2015) to 48.6 years (Fichera 2020).
- 21 studies reported the sex/gender of participants (Berry 2019; Cawley 2020; Clarke 2023a; Elbel 2009; Ellison 2013; Ellison 2014; Fichera 2020; Hammond 2013; Harnack 2008; James 2015; Oliveira 2018; Petimar 2019; Platkin 2014; Reynolds 2022; Roberto 2010; Roberto 2012; Robertson 2020; Temple 2010;

VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019) ranging from 28% female (Reynolds 2022), to 100% female (Platkin 2014). Four studies did not report or were unable to assess this (Bollinger 2011; Dubois 2021; Elshiewy 2018; Petimar 2021).

- 12 studies reported some element of ethnicity (Cawley 2020; Clarke 2023a; Elbel 2009; Hammond 2013; Harnack 2008; James 2015; Petimar 2019; Platkin 2014; Roberto 2010; Roberto 2012; Temple 2010; VanEpps 2016). Eight studies reported the percentage of white participants (Cawley 2020; Clarke 2023a; Hammond 2013; James 2015; Petimar 2019; Roberto 2010; Roberto 2012; VanEpps 2016), which ranged between 25% (Petimar 2019) and 88% (James 2015). Eight studies reported the percentage of Black participants (Clarke 2023a; Elbel 2009; James 2015; Petimar 2019; Platkin 2014; Roberto 2010; Roberto 2012; VanEpps 2016), which ranged between 4% (James 2015) and 66% (Elbel 2009). Nine studies reported the percentage of Hispanic participants (Cawley 2020; Cawley 2020; Clarke 2023a; Elbel 2009; James 2015; Petimar 2019; Roberto 2010; Roberto 2012; VanEpps 2016), which ranged between 1% (Cawley 2020; VanEpps 2016) and 45% (Platkin 2014). Seven studies reported the percentage of Asian participants (Cawley 2020; Harnack 2008; James 2015; Petimar 2019; Platkin 2014; Roberto 2010; Roberto 2012; VanEpps 2016), which ranged between 5% (James 2015; Petimar 2019) and 24% (Cawley 2020). Nine studies reported ethnicity in an 'other' category (Cawley 2020; Hammond 2013; Harnack 2008; James 2015; Petimar 2019; Platkin 2014; Roberto 2010; Roberto 2012; VanEpps 2016), which ranged between 1% (James 2015) and 97% (Harnack 2008). Temple 2010 reported a 43% ethnic minority sample.
- 11 studies reported weight status or BMI in some form (Cawley 2020; Clarke 2023a; Hammond 2013; Harnack 2008; James 2015; Petimar 2019; Platkin 2014; Roberto 2010; Roberto 2012; Temple 2010; VanEpps 2016). Seven of these studies provided mean BMIs, of which the mean BMI was less than 25 kg/m$^2$ in one study (James 2015) and between 25 kg/m$^2$ and 30 kg/m$^2$ in six studies (Clarke 2023a; Petimar 2019; Platkin 2014; Roberto 2010; Temple 2010; VanEpps 2016). Harnack 2008 reported participants as being 42.6% normal weight range, 27.9% overweight, and 29.6% obese. Hammond 2013 reported participants as 1.6% underweight, 42.5% normal weight, 31.4% overweight, and 24.5% obese. Cawley 2020 reported that 7% of the sample had a BMI over 30 kg/m$^2$, and in Roberto 2012, 22% of the sample was classified as overweight or obese. Overall, six of the studies that reported weight status had a sample that had an average BMI score within the overweight category (Clarke 2023a; Petimar 2019; Platkin 2014; Roberto 2010; Temple 2010; VanEpps 2016), or indicated a majority of the sample as overweight or obese (Hammond 2013; Harnack 2008).

**Interventions and comparisons**

All 25 studies involved manipulations as applied to food products (including non-alcoholic drinks); two studies also included alcoholic drinks (Cawley 2020; Clarke 2023a).

Twenty studies included a simple calorie label versus no calorie label comparison (Berry 2019; Bollinger 2011; Cawley 2020; Clarke 2023a; Elbel 2009; Ellison 2013; Ellison 2014; Hammond 2013; Harnack 2008; James 2015; Petimar 2019; Petimar 2021; Platkin 2014; Roberto 2010; Roberto 2012; Robertson 2020; Temple 2010; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019). Two of these studies also included an additional labelling arm that was eligible for inclusion in the review: one study also assessed calorie labelling with information about at least one other nutrient (Hammond 2013), and one study also included calories with PACE labelling (Platkin 2014). Four studies compared only calorie labelling with information about at least one other nutrient to no label (Dubois 2021; Elshiewy 2018; Fichera 2020; Oliveira 2018), while one study only compared calories with PACE labelling to no label (Reynolds 2022). Seven studies evaluated a second or third treatment arm that contained information about the same product characteristic presented in multiple ways (e.g. simple calorie labelling; calorie labelling with additional colour coding), so we combined these arms as a single calorie labelling intervention (Ellison 2013; Ellison 2014; Hammond 2013; Platkin 2014; Roberto 2010; Roberto 2012; VanEpps 2016).

Eighteen studies investigated calorie labels on menus (Berry 2019; Bollinger 2011; Cawley 2020; Elbel 2009; Ellison 2013; Ellison 2014; Hammond 2013; Harnack 2008; James 2015; Oliveira 2018; Petimar 2019; Petimar 2021; Platkin 2014; Roberto 2010; Robertson 2020; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019). Five studies investigated calorie labels on products, with labels either on product packaging (Elshiewy 2018; Fichera 2020; Roberto 2012), or adjacent to the product (Clarke 2023a; Temple 2010); one study displayed calorie labels on menus, adjacent to products, and on packaging (Reynolds 2022); and one study assessed calorie labels on products, leaflets, and aisles in supermarkets (Dubois 2021). Most studies displayed calorie information on all menu items (Berry 2019; Bollinger 2011; Cawley 2020; Clarke 2023a; Elbel 2009; Ellison 2014; Hammond 2013; Harnack 2008; James 2015; Oliveira 2018; Petimar 2019; Petimar 2021; Platkin 2014; Roberto 2010; Robertson 2020; VanEpps 2016). One study that displayed calorie information on menus did not give calorie information for all items (specifically dessert specials and drinks) (Ellison 2013). Another study investigating calorie labels on cereal packaging provided calorie information for cereal but not milk (Roberto 2012). It was unclear for one study if soft drinks were labelled, but this was assumed as all other food products were labelled and energy was calculated for all products including drinks (Temple 2010). Three studies in supermarket settings displayed FOP calorie labels on only certain products, for example, store-branded products (Dubois 2021; Elshiewy 2018; Fichera 2020). Three studies in worksite cafeterias could not provide calorie information for all items, but most items were labelled in each study (Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019). For example, in Vasiljevic 2019, 87% of products across all sites were labelled.

Twenty-one studies compared a calorie labelling intervention to no label (Berry 2019; Bollinger 2011; Cawley 2020; Clarke 2023a; Elbel 2009; Ellison 2013; Ellison 2014; Hammond 2013; Harnack 2008; James 2015; Oliveira 2018; Petimar 2019; Petimar 2021; Platkin 2014; Reynolds 2022; Roberto 2010; Robertson 2020; Temple 2010; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019). Four studies compared additional front-of-pack calorie labelling to solely current labelling on products, which was on the back of packaging (Elshiewy 2018; Fichera 2020; Dubois 2021; Roberto 2012).

Nine studies included additional arms that were not eligible for inclusion in the review. One study assessed health warning labels (Clarke 2023a); one study assessed serving size information (Roberto 2012); two studies assessed recommended daily allowance information (Roberto 2010; Robertson 2020); three studies assessed other nutritional labelling that did not include energy information, such as colour coding or traffic light symbols (Dubois 2021; Ellison 2013; VanEpps 2016); and two studies assessed price manipulations (Ellison 2014; Harnack 2008). Where possible, and unless there was clear evidence for an interaction, groups that included calorie information in a factorial design were combined.

Twenty-four studies measured selection (Berry 2019; Bollinger 2011; Cawley 2020; Clarke 2023a; Dubois 2021; Elbel 2009; Ellison 2013; Ellison 2014; Elshiewy 2018; Fichera 2020; Hammond 2013; Harnack 2008; James 2015; Oliveira 2018; Petimar 2019; Petimar 2021; Platkin 2014; Reynolds 2022; Roberto 2010; Roberto 2012; Robertson 2020; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019). For studies that measured selection, most measured selection with purchasing (Berry 2019; Bollinger 2011; Cawley 2020; Dubois 2021; Elbel 2009; Ellison 2013; Ellison 2014; Elshiewy 2018; Fichera 2020; Harnack 2008; Petimar 2019; Petimar 2021; Reynolds 2022; Robertson 2020; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019), with six studies measuring selection without purchasing (Hammond 2013; James 2015; Oliveira 2018; Platkin 2014; Roberto 2010; Roberto 2012), and one study measuring selection with the intention to purchase from a subsequent online shop (with actual purchasing also then being assessed) (Clarke 2023a). Most studies measured selection by calculating the amount of energy or number of healthy items selected from menu items (Berry 2019; Cawley 2020; Ellison 2013; Hammond 2013; Harnack 2008; James 2015; Oliveira 2018; Platkin 2014; Roberto 2010) or receipts (Clarke 2023a; Elbel 2009; Ellison 2014; Petimar 2019; Robertson 2020; VanEpps 2016), with eight studies measuring selection from transaction data (Bollinger 2011; Dubois 2021; Elshiewy 2018; Fichera 2020; Petimar 2021; Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019). One study assessed selection by measuring the amount of cereal poured (Roberto 2012). Seven studies that measured selection also measured consumption (Hammond 2013; Harnack 2008; James 2015; Platkin 2014; Roberto 2010; Roberto 2012; Robertson 2020), with one study only measuring consumption (Temple 2010). All studies assessed consumption by subtracting the weight of items consumed from total items ordered or available and converting this to energy. All of the studies, apart from four, measured energy selected or consumed per transaction or meal: one study measured energy selected or consumed per individual item within a transaction (Dubois 2021), and three studies measured total energy selected or consumed per day (Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019).

Eighteen studies assessed energy selected or consumed from all products that participants had access to or were presented with, irrespective of whether it was labelled (Berry 2019; Bollinger 2011; Cawley 2020; Clarke 2023a; Elbel 2009; Ellison 2014; Hammond 2013; Harnack 2008; James 2015; Oliveira 2018; Petimar 2019; Petimar 2021; Platkin 2014; Roberto 2010; Roberto 2012; Robertson 2020; Temple 2010; VanEpps 2016). For the only study conducted in supermarket settings that was included in the meta-analysis, energy from labelled items was the outcome extracted for inclusion as those were the data provided by the authors (Dubois 2021). Non-randomised studies reported outcomes relating to both labelled categories only and total energy (Elshiewy 2018; Fichera 2020). For the studies conducted in worksite cafeterias, we extracted the primary outcome, which was energy purchased from labelled items only (labelling being applied to most products) (Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019).

Duration of exposure to, and assessment of, the intervention ranged between less than one day (Berry 2019; Cawley 2020; Clarke 2023a; Hammond 2013; Harnack 2008; James 2015; Oliveira 2018; Platkin 2014; Roberto 2010; Roberto 2012; Robertson 2020; Temple 2010) and multiple weeks (2 to 13 weeks) (Dubois 2021; Ellison 2013; Ellison 2014; Reynolds 2022; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019) in randomised/quasi-randomised studies, and two weeks to five years in non-randomised studies (Bollinger 2011; Elbel 2009; Elshiewy 2018; Fichera 2020; Petimar 2019; Petimar 2021).

#### Excluded studies

We excluded 100 study reports at the full-text screening stage; 41 of these were ineligible interventions and 33 had a lack of measurement of eligible selection/consumption outcomes. See Characteristics of excluded studies table for further details.

#### Studies awaiting classification

We categorised eight studies as awaiting classification, which will be considered for inclusion in the next update of this review (Clements 2016; Dos Santos 2015; Elshiewy 2022; Girz ongoing; Lee 2018; Petimar 2022; van Doorn 2023; Zhu 2023). See Characteristics of studies awaiting classification table for further details.

### Risk of bias in included studies

#### Randomised studies

We used the RoB 2 tool to assess risk of bias for each of the included studies, with a summary provided in Figure 3 (randomised studies) and Figure 4 (cluster-RCTs). There were at least some concerns about overall risk of bias for 12/19 randomised studies, with one of these assessed at high risk of bias (Ellison 2014). The other seven studies were at low risk of bias (Cawley 2020; Clarke 2023a; Dubois 2021; Reynolds 2022; Robertson 2020; Vasiljevic 2018; Vasiljevic 2019). A summary for each domain by outcome is presented below.

**Figure 3. Risk of bias assessment for randomised studies (figure generated using the robvis tool (McGuinness 2021)). Full information on the judgements for each domain available at: https://osf.io/qya8u/**

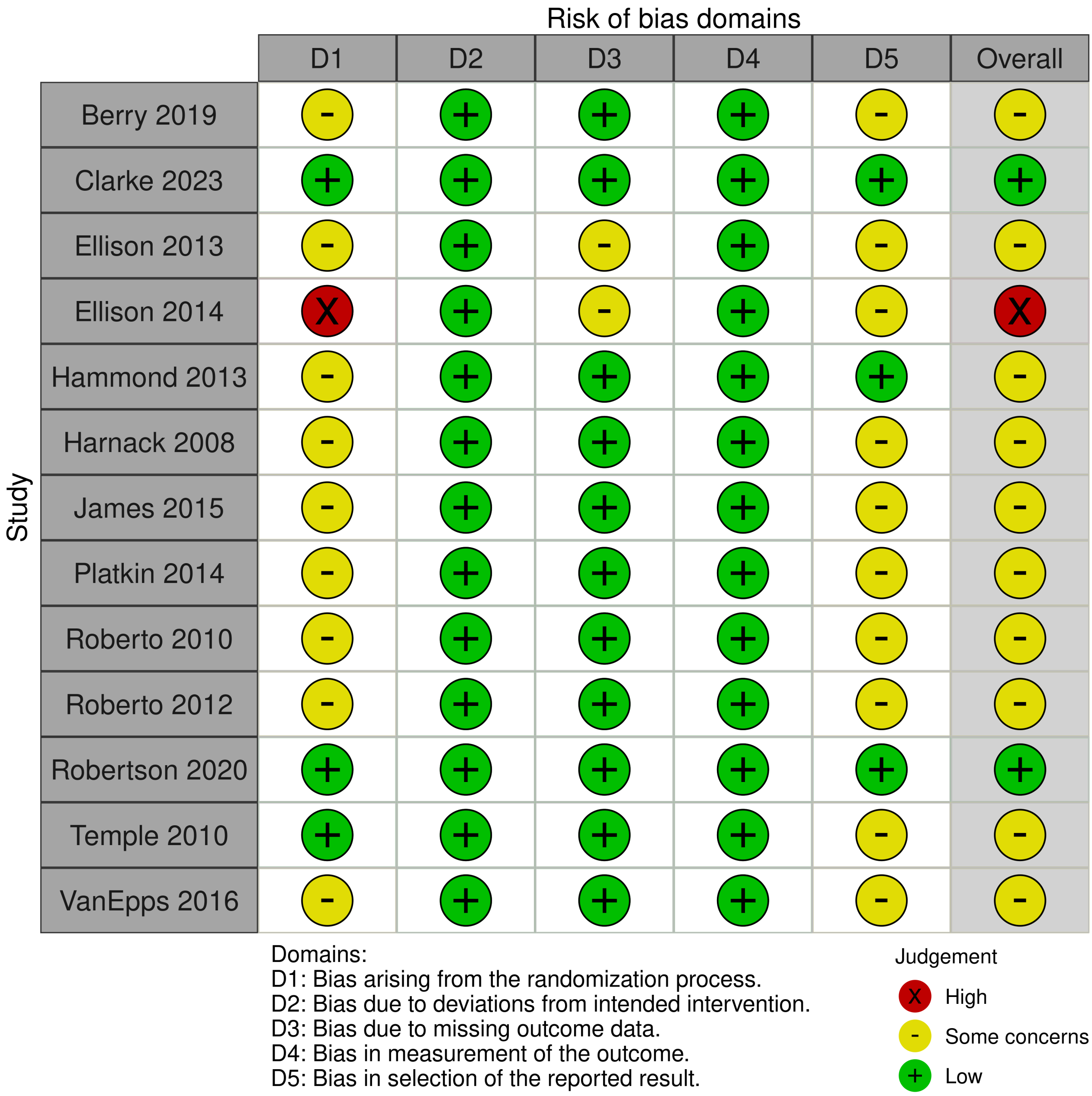
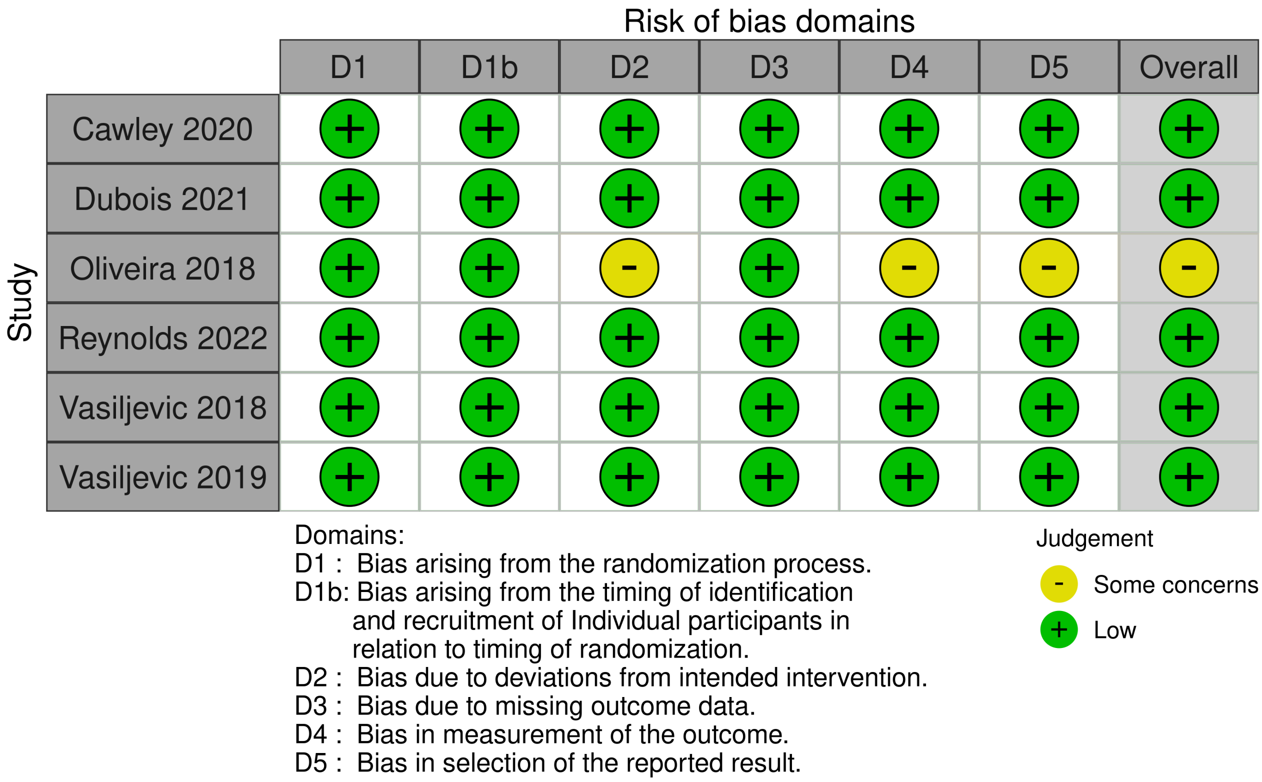


**Figure 4. Risk of bias assessment for cluster-randomised studies (figure generated using the robvis tool (McGuinness 2021)). Full information on the judgements for each domain available at: https://osf.io/qya8u/**




***Selection (with or without purchasing)***

*Bias arising from the randomisation process.* There were some concerns for this component in 9/18 randomised studies with a selection (with or without purchasing) outcome, mainly due to an absence of detail in describing the randomisation and allocation concealment processes. One study was at high risk of bias as the randomisation process was unclear and there were substantial baseline imbalances which the authors acknowledged could be due to the intervention assignment (Ellison 2014).

*Bias arising from the timing of identification and recruitment of individual participants into clusters.* This component only applied to cluster-RCTs, all of which were at low risk of bias for this domain.

*Bias due to deviations from intended interventions.* Most studies (17/18) were at low risk for this domain. There were some concerns about one study, for which we judged there was insufficient information (including on the analysis) to determine that there would have been no impact of deviations from the intervention (Oliveira 2018).

*Bias due to missing outcome data.* Most studies (16/18) were at low risk for this domain. Two studies were judged as having some concerns as the number of randomised participants was not reported, meaning we were unable to assess the amount of missing outcome data (Ellison 2013; Ellison 2014).

*Bias in measurement of the outcome.* Most studies (17/18) were at low risk as they used an objective measure of selection or purchasing. One study was judged as having some concerns as there was insufficient detail on the outcome measurement (i.e. how foods were classed into healthy versus unhealthy categories, and energy details were not provided) (Oliveira 2018).

*Bias in selection of the reported result.* We judged eight studies at low risk of bias, as they reported preregistered analysis intentions (Cawley 2020; Clarke 2023a; Dubois 2021; Hammond 2013; Reynolds 2022; Robertson 2020; Vasiljevic 2018; Vasiljevic 2019). The remaining studies (10/18) had some concerns due to there not being preregistered publicly available analysis intentions (e.g. in a protocol, trial registration, or statistical analysis plan).

***Consumption***

*Bias arising from the randomisation process.* There were some concerns for this component in 6/8 studies, mainly due to an absence of detail in describing the randomisation and allocation concealment processes.

*Bias due to deviations from intended interventions.* All studies were at low risk for this domain.

*Bias due to missing outcome data.* All studies were at low risk for this domain.

*Bias in measurement of the outcome.* All studies were at low risk as they used an objective measure of consumption.

*Bias in selection of the reported result.* We judged two studies at low risk of bias, as they reported preregistered analysis intentions (Hammond 2013; Robertson 2020). Most studies (6/8) had some concerns due to there not being preregistered publicly available analysis intentions (e.g. in a protocol, trial registration, or statistical analysis plan).

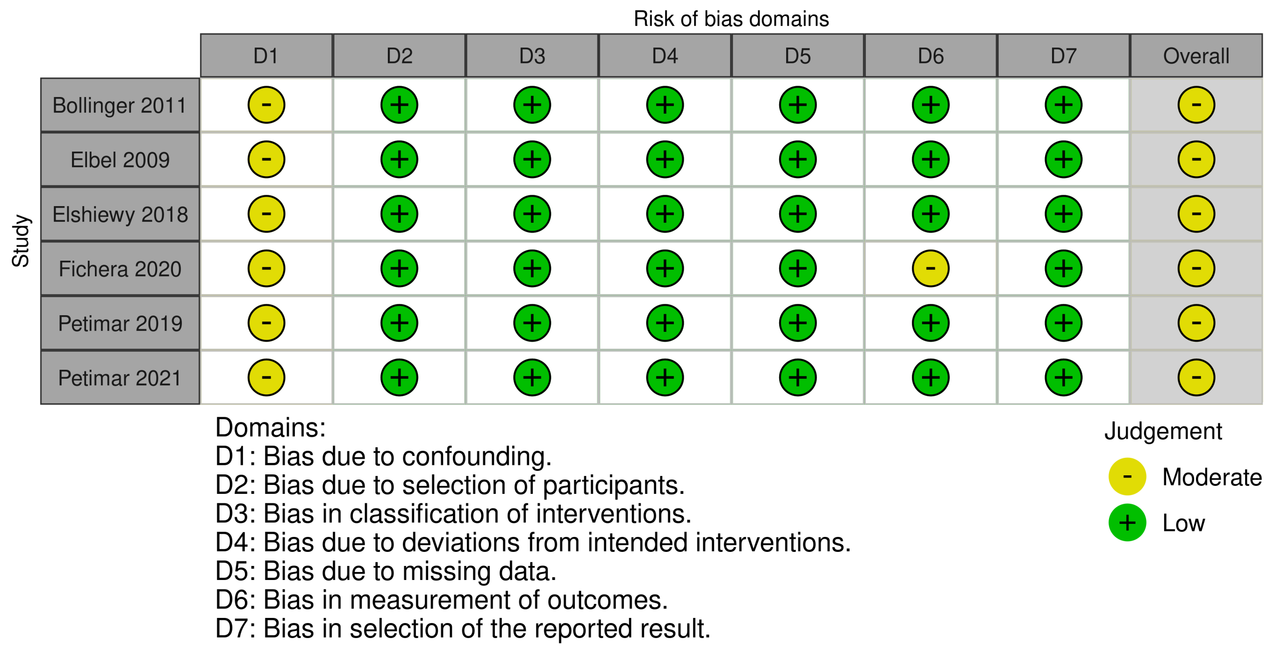
Full information on the judgements for each domain is available at https://osf.io/qya8u/.

**Non-randomised studies**

***Selection (with or without purchasing)***

We used the ROBINS-I tool to assess risk of bias for each of the non-randomised studies, with a summary provided in Figure 5. All six studies were assessed at moderate overall risk of bias (Bollinger 2011; Elbel 2009; Elshiewy 2018; Fichera 2020; Petimar 2021; Petimar 2019), principally due to all studies having the potential for confounding, even if there was an appropriate analysis to control for confounding domains. At the outset, we considered potential confounding domains to be key participant demographic characteristics (age, sex, and socioeconomic status distributions), and key study (intervention setting) characteristics (major events, different geographic areas, different retail environments). A moderate risk of bias rating suggests a study is sound for a non-randomised study but cannot be compared to a well-performed RCT (Sterne 2016b). A summary for each of the individual risk of bias domains is presented below.

**Figure 5. Risk of bias assessment (summary of ROBINS-I) for non-randomised studies (figure generated using the robvis tool (McGuinness 2021)). Full information on the judgements for each domain available at: https://osf.io/qya8u/**




*Bias due to confounding.* All studies were at moderate risk of bias for the potential for confounding, although we judged that confounding domains – including time varying confounding where applicable – were appropriately controlled for in the analysis in each study. For most studies there were no concerns in any study about validity or reliability of their measurement, although there was not enough information reported in one study to assess this (Elbel 2009).

*Bias in selection of participants into the study.* All studies were at low risk of bias for this domain. Most studies predefined selection of participants into the study by geographical area based on calorie labelling legislation and all participants who would have been eligible were included (i.e. transactions were recorded within a given time period). One study used a random sample of participants, but this was carried out before the intervention was implemented (Elshiewy 2018). Another study recruited participants from a pre-existing representative panel of participants, but the sample was recruited independently of the study via a research agency (Fichera 2020).

*Bias in classification of interventions.* All studies were at low risk of bias for this domain. The intervention status was well-defined for all studies and was based on the introduction of calorie labelling policies or legislation.

*Bias due to deviations from intended interventions.* All studies were at low risk of bias for this domain. Most studies (5/6) did not report any deviations from the intended intervention. One study highlighted that fines were given to fast-food restaurants that were not in compliance with regulations (Elbel 2009), which suggests that some restaurants did not fully comply with the labelling policy. However, study authors noted that it was likely that most sites implemented labelling as planned and therefore effects of this on the outcome would be minimal. No studies included co-interventions.

*Bias due to missing data.* All studies were at low risk of bias for this domain. Two studies reported some missing outcome data, but both included sensitivity analyses and reported that results were robust for any missing data (Fichera 2020; Petimar 2021).

*Bias in measurement of the outcome.* Most (5/6) studies were at low risk of bias for this domain as they used an objective (i.e. not self-report) measure of behaviour (i.e. selection measured using receipts or transactions). One study was at moderate risk of bias as it relied on participants recording their own purchases using a scanner in the home (Fichera 2020).

*Bias in selection of the reported result.* We judged all studies at low risk of bias as there was no indication that there was selection of the reported analysis from multiple analyses or subgroups for any study and the statistician on this review's author team confirmed that analyses for each study were appropriate.

Full information on the judgements for each domain is available at https://osf.io/qya8u/.

## Effects of interventions

See: **Summary of findings 1** Calorie (energy) labelling for selection and consumption of food and non-alcoholic drinks; **Summary of findings 2** Calorie (energy) labelling for selection and consumption of alcoholic drinks

### Effects of calorie labelling on selection of food (including non-alcoholic drinks)

See Summary of findings 1.

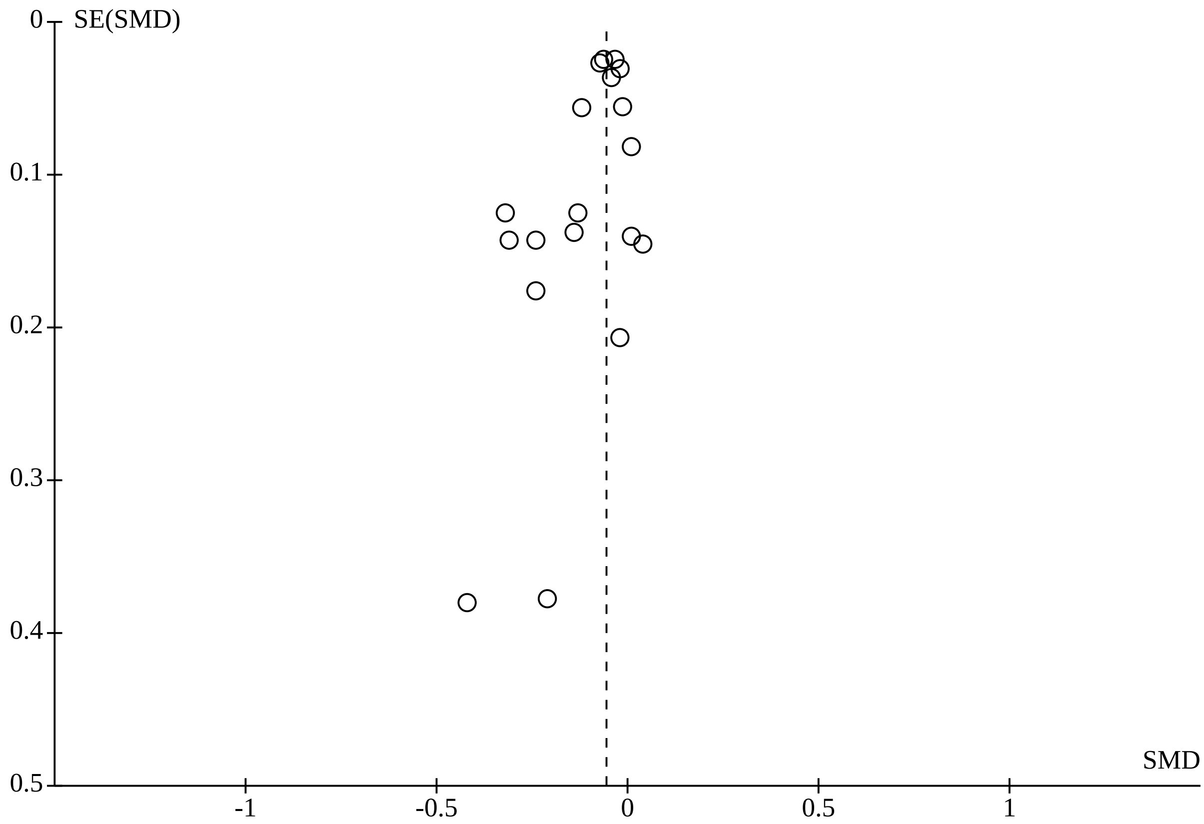
#### *Randomised studies*

For our planned primary analysis, outcome data were available for 19 comparisons from 16 studies, involving 9850 participants (including participating clusters) (Cawley 2020; Clarke 2023a; Dubois 2021; Ellison 2013; Ellison 2014; Hammond 2013; Harnack 2008; James 2015; Platkin 2014; Reynolds 2022; Roberto 2010; Roberto 2012; Robertson 2020; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019). All studies introduced calorie labelling on or adjacent to products or on menus. Most studies (14/16) investigated simple calorie labels compared to no labels (Cawley 2020; Clarke 2023a; Ellison 2013; Ellison 2014; Hammond 2013; Harnack 2008; James 2015; Platkin 2014; Roberto 2010; Roberto 2012; Robertson 2020; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019), of which one also compared calorie labelling with information about at least one other nutrient to no labels (Hammond 2013), and one also compared calories with PACE labelling to no labels (Platkin 2014). For the remaining two studies, one study only investigated calorie labels with information about at least one other nutrient to no labels (Dubois 2021), and one study only compared calories with PACE labelling to no labels (Reynolds 2022). Half of the studies (8/16) were conducted in real-world settings (seven in restaurants and one in a store), six in naturalistic laboratory settings (Clarke 2023a; Hammond 2013; Harnack 2008; James 2015; Platkin 2014; Roberto 2012), and two in laboratory settings (Roberto 2010; Robertson 2020). Most studies measured selection with purchasing (10/16; Cawley 2020; Dubois 2021; Ellison 2013; Ellison 2014; Harnack 2008; Reynolds 2022; Robertson 2020; VanEpps 2016; Vasiljevic 2018; Vasiljevic 2019), with five studies measuring selection without purchasing (Hammond 2013; James 2015; Platkin 2014; Roberto 2010; Roberto 2012), and one study measuring selection with the intention to purchase from a subsequent online shop (with actual purchasing also then being assessed) (Clarke 2023a).

Calorie labelling of food led to a small reduction in energy selected (SMD −0.06, 95% CI −0.08 to −0.03; $P < 0.0001$, $I^2 = 2\%$; 16 randomised studies, 19 comparisons, 9850 participants; high-certainty evidence; Analysis 1.1). The $I^2$ statistic indicated that almost none of the total variance in effect estimates was attributable to study heterogeneity. Our extrapolation of the size of this summary effect suggests that if calorie labelling was implemented for an assumed average meal of 600 calories (SD 185) kcal, adults would select 11 kcal less (15 kcal less to 6 kcal less), reducing energy purchased by 1.8% (2.5% less to 1% less).

GRADE assessment indicated that the evidence for this outcome was of high certainty, meaning that it provides a very good indication of the likely effect, with a low likelihood that the actual effect will be substantially different, and that further research is very unlikely to change our confidence in the estimate of effect. Our justification for not downgrading on any domains was as follows. We did not downgrade for risk of bias as most of the weight in the meta-analysis was derived from studies at low risk of bias, corroborated by a sensitivity analysis of only low risk of bias studies generating a similar result. We did not downgrade on inconsistency as there were low levels of statistical inconsistency and heterogeneity in the analysis. We did not downgrade for indirectness because most of the weight in the meta-analysis derived from real-world field studies in a range of settings and with the content of studied interventions comparable to what would likely be implemented. The intervention would likely be implemented at a substantially larger scale than was reflected in the included randomised studies, but while this confers uncertainty about how predictably estimates would apply to real-world implementation at scale (e.g. the degree to which intervention fidelity would be maintained, although particularly if mandated and enforced through regulation, it cannot be presumed this would be lesser than is observed in trials), we judged that it is likely impractical to conduct randomised trials of these interventions that use substantially increased samples and durations of assessment. We did not downgrade for imprecision, as there was a precise effect estimate with a large sample size exceeding the optimal information size and narrow CIs that only encompass benefit of the intervention. Finally, we did not downgrade on publication bias as this was not strongly suspected based on the funnel plot, with only a minimally sufficient number of plots with no clear asymmetry (Figure 6).

**Figure 6. Funnel plot**




We conducted three prespecified subgroup analyses concerning moderation of intervention effects by setting, label type, and socioeconomic status. Subgroup analyses suggested no evidence of a difference in effect by setting ($Chi^2$ = 5.68, df = 3, P = 0.13, $I^2$ = 47.2%; Analysis 1.2), although there was a notably larger summary mean effect size in laboratory settings (SMD −0.29, 95% CI −0.49 to −0.09) compared to restaurants (SMD −0.05, 95% CI −0.09 to −0.02), stores (SMD −0.05, 95% CI −0.08 to −0.01), and naturalistic laboratories (SMD −0.07, 95% CI −0.16 to 0.03). There was no evidence of a difference by label type ($Chi^2$ = 0.31, df = 2, P = 0.86, $I^2$ = 0%; Analysis 1.3): simple calorie labels (SMD −0.07, 95% CI −0.11 to −0.03); calories with PACE labelling (SMD −0.05, 95% CI −0.12 to 0.03); calorie labelling with information about at least one other nutrient (SMD −0.06, 95% CI −0.11 to −0.01). Finally, subgroup analyses suggested no evidence of a difference ($Chi^2$ = 3.92, df = 1, P = 0.05, $I^2$ = 74.5%; Analysis 1.4) in studies with participants from a high socioeconomic status (SMD −0.08, 95% CI −0.12 to −0.05) compared to those with participants of both low and high socioeconomic status (SMD −0.04, 95% CI −0.07 to −0.01). A further exploratory subgroup analysis found no evidence that calorie labels were differentially effective whether placed on menus, or adjacent to and on product packaging (Analysis 1.5).

We conducted a series of planned sensitivity analyses of the primary analysis of the selection outcome. First, we used a subgroup analysis concerning the seven studies that were at low risk of bias (Cawley 2020; Clarke 2023a; Dubois 2021; Reynolds 2022; Robertson 2020; Vasiljevic 2018; Vasiljevic 2019). The effect in the low risk of bias studies was comparable to the overall effect size (SMD −0.05, 95% CI −0.07 to −0.02). There was a larger summary mean effect size for studies judged to have some concerns or at high risk of bias (SMD −0.12, 95% CI −0.19 to −0.06) ($Chi^2$ = 4.40, df = 1, P = 0.04, $I^2$ = 77.3%; Analysis 1.6). Second, for three studies that did not report the results in the format required for our analysis, we imputed values (Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019) (see Broken link). We carried out a sensitivity analysis removing these three studies. This did not affect the interpretation of the primary analysis with a summary mean effect size that was slightly larger (SMD −0.07, 95% CI −0.10 to −0.04; P < 0.0001; Analysis 1.7). As the analysis of selection outcomes included studies with purchasing and studies without purchasing, we also carried out a sensitivity analysis to check whether this modified the result. As would be expected given most studies of selection included actual purchasing, excluding studies without purchasing did not change effect estimates (SMD −0.05, 95% CI −0.07 to −0.03; P < 0.0001).

Two randomised (or quasi-randomised) studies could not be included in the analysis as one did not report the number of participants in each group (Berry 2019), and another did not report energy, only the number of healthy products selected (Oliveira 2018) (see Table 1). Both were conducted in restaurant field settings and investigated calorie labelling on menus. Berry 2019 measured selection with purchasing and Oliveira 2018 selection only. In Berry 2019, calorie information on menus did not influence energy ordered. Similarly, in Oliveira 2018 there were no differences in the number of healthy food choice items selected.

***Non-randomised studies***

Six non-randomised studies met the inclusion criteria: two interrupted time series studies (Elshiewy 2018; Petimar 2021) and four CBA studies (Bollinger 2011;Elbel 2009; Fichera 2020; Petimar 2019). Table 1 provides a narrative synthesis of these studies. Four studies compared energy purchasing before and after the implementation of calorie labelling in restaurants in the USA: three in fast-food chains (Elbel 2009; Petimar 2019; Petimar 2021), and one in a coffee shop chain (Bollinger 2011). Two studies compared energy purchasing before and after the implementation of calorie labelling on products in stores in the UK (Elshiewy 2018; Fichera 2020).

Of the four studies that investigated calorie labelling in restaurants in the USA, two studies found that calorie labelling reduced energy purchased. Bollinger 2011 investigated the implementation of calorie labelling in 222 coffee shops in New York City compared to 94 control sites in Boston and Philadelphia. There was an average decrease in energy purchased of 14.4 kcal per transaction after implementation of calorie labelling, mainly driven by changes in food purchasing (P = 0.001), with no evidence of impact on energy purchased from drinks. Petimar 2021 measured transactions over a four-year period (two years after franchise labelling and one year after nationwide labelling) and reported a mean reduction of 73 kcal per transaction (95% CI −81 to −65), equivalent to a 4.7% reduction. Two studies reported no evidence for a reduction in energy purchased. Elbel 2009 reported no differences in purchasing behaviour in five control sites and 14 intervention sites, after calorie labelling was implemented, with an increase of 21 calories after labelling in one city (New York City) and 3 calories in another city that did not implement calorie labelling (Newark). Petimar 2019 reported that although there was a decrease in energy purchased, calorie labelling was not associated with a reduction in energy purchased in intervention compared to control sites over a four-year period in 5948 participants (adults: −19 calories, 95% CI −112 to 75; children: −13 calories, 95% CI −108 to 135; adolescents: −49 calories, 95% CI −126 to 38). Of the two studies that investigated calorie labelling on products in stores in the UK, both found a reduction in energy purchased from labelled products. Elshiewy 2018 evaluated changes before and after calorie front-of-pack labelling in UK supermarkets one year before and one year after labelling was implemented on three store brand categories: cookies, soft drinks, and breakfast cereals. In 4,131,570 purchase transactions from 188,062 participants, there was a reduction of 9.5% in energy purchased from labelled products. Fichera 2020 assessed the implementation of calorie labelling in nine large retailers where four introduced voluntary front-of-pack labelling on foods recommended for labelling by the Food Standards Agency, these included ready meals, burgers, pizzas, sandwiches, cereals, dairy foods, and cookies. Over a three-year period, in 360,921 observations from more than 20,707 households, there was a reduction in energy purchased from labelled food purchases of 588 kcal per month (P < 0.01). This study also found a reduction in energy purchased across all products (labelled and unlabelled).

We did not include quantitative summary estimates from non-randomised studies in Summary of findings 1. This is because we could not assume a meaningful summary estimate representing this body of studies, with estimates not being comparable to those from randomised studies or to one another. We considered only four of six studies to have reported reasonably comparable data (on mean selection per transaction, which ranged from 73 kcal lower to 18 kcal higher with calorie labelling). Non-randomised evidence may in principle complement randomised evidence when the reason to downgrade evidence from RCTs is due to indirectness (Cuello-Garcia 2022). In this case, while evidence for this outcome was already judged at high certainty without downgrading in any domain, the principal reason for not downgrading on indirectness was that we judged that it was impracticable to conduct RCTs at a larger scale. Therefore, evidence from non-randomised studies conducted at a larger scale in real-world contexts could remain additionally informative and complementary to randomised evidence, justifying its inclusion in the review as per our intention in this review's protocol (Clarke 2021b).

Although for reasons outlined above, we could not summarise the non-randomised data sufficiently well to make strong inferences to directly complement our analysis of randomised studies, including quantitative data concerning the magnitude of effects, findings were broadly consistent with the direction of effect observed in randomised studies: five of six non-randomised studies observed a numeric decrease in energy purchased, with a numeric increase in energy purchased in one study. We narratively synthesised this evidence, presented in Table 1, and also conducted a sign test. We deemed this method to be the most applicable in this case given the studies concerned were heterogeneous in their methods and design, reported different effect measures, and not all studies reported P values. Therefore, this choice is in line with both Cochrane guidance and reporting guidelines on synthesis without meta-analysis in systematic reviews (Campbell 2020). Although the sign test suggested findings were broadly consistent (5/6 studies favoured the intervention; 83%, 95% CI 44% to 97%; P = 0.22), due to the conservative nature of this test with such few studies, we could not establish that effects consistently favoured (or did not favour) the intervention.

GRADE assessment indicated that the non-randomised evidence for this outcome was of very low certainty, meaning that it does not provide a reliable indication of the likely effect, and the likelihood that the actual effect will be substantially different is very high. We reached this judgement through downgrading certainty of the evidence four times. It was initially downgraded twice because GRADE assigns a default starting rating of 'low certainty' for evidence from non-randomised studies. We further downgraded the evidence one level for risk of bias (because all the studies were judged to have significant concerns regarding risk of bias), and one level for imprecision (because we were unable to generate and assess a meaningful summary effect size estimate with CIs). We did not downgrade for indirectness (because the studies were all large-scale real-world field studies), inconsistency (the direction of effect was broadly consistent with five of six non-randomised studies observing a numeric decrease in energy purchased, although we could not convincingly demonstrate consistency), or publication bias (the studies were large scale and we had insufficient additional information to suspect publication bias).

**Effects of calorie labelling on consumption of food (including non-alcoholic drinks)**

See Summary of findings 1.

For our planned primary analysis, 10 comparisons from eight studies, involving 2134 participants, provided data (Hammond 2013; Harnack 2008; James 2015; Platkin 2014; Roberto 2010;

Roberto 2012; Robertson 2020; Temple 2010). All studies introduced calorie labelling on products or on menus. All studies compared simple calorie labels to no labels. One study also compared calorie labelling with information about at least one other nutrient to no labels (Hammond 2013), and one study also compared calories with PACE labelling to no labels (Platkin 2014). Five studies took place in naturalistic laboratory settings (Hammond 2013; Harnack 2008; James 2015; Platkin 2014; Roberto 2012), and three were in laboratory settings (Roberto 2010; Robertson 2020; Temple 2010).

Calorie labels may decrease the amount of energy consumed from food (SMD −0.19, 95% CI −0.33 to −0.05; P = 0.006; $I^2$ = 47%; 8 randomised studies, 10 comparisons, 2134 participants; low-certainty evidence; Analysis 1.8). The $I^2$ statistic indicates that a moderate amount of the total variance in study-level estimates of this effect was attributable to statistical heterogeneity. Our extrapolation of the size of this summary effect suggests that, if calorie labelling was implemented for an assumed average meal of 600 (SD 185) kcal, adults would consume 35 kcal less (95% CI −61 kcal to −9 kcal), reducing energy consumed by 5.9% (95% CI −10.2% to −1.5%).

GRADE assessment indicated that the evidence for this outcome was of low certainty, meaning that the current evidence does not provide a reliable indication of the likely effect, and that the likelihood that the actual effect will be substantially different is very high. We reached this judgement through consideration of the following criteria. We downgraded the evidence one level due to serious study limitations because all but one study in the meta-analysis was judged to have significant concerns regarding risk of bias. We further downgraded the evidence one level for indirectness because all studies were conducted in laboratory settings (albeit varying in how naturalistic they were designed to be) with short-term exposures not representative of real-world implementation. We did not downgrade the evidence for inconsistency, imprecision, or publication bias.

We conducted three prespecified subgroup analyses concerning setting, label type, and socioeconomic status. It is advised that subgroup analyses require at least 10 studies (Deeks 2023), but these analyses include only 10 data points from eight studies. Therefore, they are reported for completeness and we do not consider them reliable for interpretation; we therefore only report them here in this section and do not interpret them further. Subgroup analyses for the consumption outcome suggested a difference in effect by setting ($Chi^2$ = 4.34, df = 1, P = 0.04, $I^2$ = 76.9%; Analysis 1.9), with a larger summary mean effect size in laboratory settings (SMD −0.45, 95% CI −0.76 to −0.14) compared to naturalistic laboratories (SMD −0.09, 95% CI −0.21 to 0.02). There was no evidence of a difference by label type ($Chi^2$ = 0.30, df = 2, P = 0.86, $I^2$ = 0%; Analysis 1.10) (simple calorie labels: SMD −0.19, 95% CI −0.35 to −0.02; calories with PACE labelling: SMD −0.39, 95% CI −1.13 to 0.36; calorie labelling with information about at least one other nutrient: SMD −0.23, 95% CI −0.50 to 0.04). Subgroup analyses by socioeconomic status suggested a larger effect ($Chi^2$ = 5.50, df = 1, P = 0.02, $I^2$ = 81.8%; Analysis 1.11) in studies with participants of high socioeconomic status (SMD −0.23, 95% CI −0.37 to −0.10) compared to those studies with participants of both low and high socioeconomic status (SMD 0.02, 95% CI −0.14 to 0.18).

We carried out a planned sensitivity analysis concerning one study that was at low risk of bias (Robertson 2020). The effect in the low risk of bias study was larger than the overall effect size (SMD −0.68, 95% CI −1.13 to −0.24), compared to studies judged to have some concerns or a high risk of bias (SMD −0.14, 95% CI −0.26 to −0.02) ($Chi^2$ = 5.29, df = 1, P = 0.02, $I^2$ = 81.1%; Analysis 1.12). This suggests that the primary meta-analysis result may be a conservative estimate, albeit noting that there was only one low risk of bias study.

**Effects of calorie labelling on selection of alcoholic drinks**

See Summary of findings 2.

For our planned primary analysis, two comparisons from two studies, involving 5756 participants provided data (Cawley 2020; Clarke 2023a). One study introduced calorie labels adjacent to products in a naturalistic supermarket context, and one study introduced calorie labels on menus in a restaurant (Cawley 2020). Both studies investigated simple calorie labels compared to no labels, with one study measuring selection with purchasing, and one study measuring selection with the intention to purchase (with actual purchasing also being subsequently assessed).

Calorie labels may have an effect on the amount of energy selected from alcohol, but the direction and size of this effect are very uncertain (SMD −0.05, 95% CI −0.25 to 0.16; P = 0.66, $I^2$ = 59%; very low-certainty evidence; Analysis 2.1). The $I^2$ statistic indicates that a substantial amount of the total variance in study-level estimates of this effect was attributable to statistical heterogeneity, which is consistent with differences in characteristics between the two studies in this analysis, including differences in setting type (supermarket versus restaurant) and label characteristics (adjacent to products versus on menus).

GRADE assessment indicated that the evidence for this outcome is of very low certainty, meaning that current evidence does not provide a reliable indication of the likely effect, and the likelihood that the actual effect will be substantially different is very high. We reached this judgement through consideration of the following criteria. We downgraded the current evidence one level due to indirectness. Studies were conducted in a field setting and a naturalistic laboratory setting (a simulated online supermarket) and the content of studied interventions was comparable to what would likely be implemented. However, exposures to the intervention were short-term, and the nature of the intervention means that it would feasibly and likely be implemented at a substantially larger scale (e.g. at national or regional level, including via manufacturing or retail companies) and in a greater variety of real-world settings (e.g. supermarkets, pubs, and bars) than was reflected in the included randomised studies conducted at relatively small scale. Furthermore, we judge that RCTs of these interventions that involve larger samples and longer durations of assessment could feasibly be conducted that may be able to adequately fulfil this criterion. In sum, this gives uncertainty about how predictably estimates of intervention effects from this review would apply to real-world implementation at scale. This includes the degree to which intervention fidelity would be maintained (although particularly if mandated and enforced through regulation, it cannot be presumed this would be lesser than is observed in trials) as well how much other actors in the system could dampen or increase potential effects (e.g. activities of the alcohol industry). Certainty was also downgraded for inconsistency because there was evidence of substantial heterogeneity from tests of statistical inconsistency and homogeneity, and imprecision, because although there was a large sample size exceeding the optimal information size, there were wide CIs encompassing both benefit and harm. We did not downgrade the certainty of evidence on the remaining GRADE domains (risk of bias because the meta-analysis derived solely from low risk of bias studies, or publication bias because with only two studies there was an insufficient number of data points to adequately inform our judgement).

One study additionally reported the amount of alcohol that was selected (in terms of alcohol units), finding no evidence of an effect on this outcome (19.6 (SD 20.9) alcohol units selected in the calorie labelling group versus 23.7 (SD 17.5) alcohol units selected in the no-label group) (Clarke 2023a). In line with other outcomes, we calculated an effect size as a SMD of −0.21 (95% CI −0.49 to 0.06; Analysis 2.2). GRADE assessment indicated that the evidence for this outcome was of very low certainty, meaning that current evidence does not provide a reliable indication of the likely effect, and the likelihood that the actual effect will be substantially different is very high. We reached this judgement through consideration of the following criteria. We downgraded the current evidence one level due to indirectness, as the study was conducted in a simulated online supermarket with short-term exposure to the intervention, and the nature of the intervention means that it would feasibly and likely be implemented at a substantially larger scale (e.g. at national or regional level, including via manufacturing or retail companies) and in a variety of real-world settings (e.g. supermarkets, pubs, and bars) not reflected in this single study. Furthermore, we judged that RCTs that involve larger samples and longer durations of assessment could feasibly be conducted that may be able to adequately fulfil this criterion. In sum, this gives uncertainty about how predictably this estimate of the intervention's effect would apply to real-world implementation at scale. Certainty was further downgraded by two levels for imprecision, because the effect estimate was derived from a single study with a sample size substantially lower than the optimal information size and with wide CIs encompassing both appreciable benefit and appreciable harm. We did not downgrade the certainty of evidence (although it was already judged at very low certainty) the remaining GRADE domains (risk of bias, inconsistency, publication bias) as this was only a single study, and was assessed as being at low risk of bias.

### Effects of calorie labelling on consumption of alcoholic drinks

The search identified no studies.

## DISCUSSION

### Summary of main results

The evidence in this review suggests that people select and purchase less energy from food (including non-alcoholic drinks) when calorie labels are displayed on or adjacent to food products and packaging. It also suggests that people consume less energy from food when food products display calorie labels, although the evidence for this effect was less certain. For alcohol, there was insufficient evidence to estimate whether people select and purchase less energy from alcohol when alcoholic drinks display calorie labels, and evidence was entirely absent for consumption of alcohol. Therefore, the remainder of this section concerns food products (including non-alcoholic drinks) only.

The summary effect sizes derived from meta-analyses for food were small but potentially important, suggesting impacts of 1.8% on energy selected and purchased, and of 5.9% on energy consumed. Although these estimates are only illustrative due to limitations of the underlying data and the nature of extrapolation (see Measures of treatment effect), if effects of such magnitudes were realised, extended to food selection and consumption over a day, and sustained over time, this could result in meaningful reductions in energy intake. For example, 10-year weight gain between 1999 and 2009 among adults in England (i.e. 9 kg at the 90th percentile) has been estimated to be equivalent to extra energy intake of 24 kcal per day over the same period (approximately 1.4% of average daily energy intake from food) (Department of Health 2011). Sustained reductions in energy intake exceeding this level are therefore potentially valuable in preventing further weight gain. The relatively small effects observed are not unexpected given evidence of the limited ability of information provision to change health-related behaviour (Hollands 2012; Marteau 2023a). It has also been observed that calorie labelling is often not noticed and may only be used by a small proportion of motivated consumers (Polden 2023). The large sample size for the meta-analysis of the selection outcome allows for the estimation of a relatively precise effect of such small magnitude (1.8%, 95% CI 1.0% to 2.5%), where a smaller sample size would not allow one to distinguish between a small effect and no effect.

The certainty of the evidence in the meta-analysis for the selection outcome was high, which means that we can be confident that the true effect is similar to the estimated effect. However, the certainty of the evidence in the meta-analysis for the consumption outcome was low, meaning that the true effect of calorie labelling might be markedly different from the estimated effect.

Concerning the second objective of the review – addressing the impact of potential modifiers – for selection of food, we found no evidence for an association between summary effect sizes and variations in study setting, labelling type, or the socioeconomic status of participants (noting there were no studies with only high deprivation samples). These findings should be interpreted with caution and would need to be confirmed by evidence from future studies, but they tentatively suggest the potential for calorie labelling to reduce selection of food among a variety of settings and a broad range of people. There were insufficient data to enable similar assessments of potential modifiers for the consumption of food.

### Overall completeness and applicability of evidence

The evidence for the effect of energy labelling for food (including non-alcoholic drinks) was relatively complete and applicable. For the selection outcome, the synthesised evidence was collected from 16 randomised studies and six non-randomised studies, encompassing a large sample of participants. The studies were predominantly from field and naturalistic laboratory settings. Most field studies were conducted in restaurants with only two studies using supermarket settings and one of these could not be included in the meta-analysis. Although we identified two large non-randomised studies in supermarket settings which were broadly in accordance with the results from the meta-analysis, the lack of randomised studies in these settings could limit the applicability of the evidence. Additionally, while most of the non-randomised studies were long-term (i.e. over periods of up to four years), the randomised studies typically assessed outcomes over relatively short periods. However, for the selection outcome, findings were broadly consistent between included randomised and non-randomised studies in terms of the direction of the effect.

Another factor that may limit the applicability of the results to a broader range of contexts is that most of the studies were from low social and material deprivation contexts, with only four studies clearly including a mix of high and low deprivation participants, and no studies with a high deprivation sample. Given evidence of differences in nutrition literacy between different socioeconomic status groups (Bhawra 2022), this may limit the broad applicability of the findings. The scarcity of evidence in high deprivation contexts also means we were unable to formally assess differences between subgroups of differing socioeconomic status. Concerns in this regard are somewhat assuaged by recent analyses that suggest impacts of calorie labelling are unlikely to be patterned by socioeconomic status, although evidence from real-world settings remains scarce (Robinson 2023b). Additionally, we identified very limited evidence (one study) from low- and middle-income countries. It is possible that the effectiveness of calorie labelling is similar across countries, although it could differ given that exposure to prepackaged products, and patterns of obesity, may vary (Ameye 2019). In terms of its implementation, calorie labelling appears to be similarly feasible across countries, due to, for example, its likely relatively low costs due to being delivered at scale through small changes to menus or packaging, but this remains to be comprehensively assessed.

The evidence for the effect of calorie labelling on the consumption of food (including non-alcoholic drinks) was less complete, with most of these studies conducted in laboratory settings. Even though some of these were naturalistic laboratories, these are not equivalent to uncontrolled real-world settings. These studies usually also involved a single exposure to the calorie labelling intervention over a relatively short time period, with short-term outcomes. This limits the degree to which such effects of calorie labelling can be generalised to complex real-world food environments.

The completeness and applicability of evidence for alcoholic drinks was very limited, with only two included studies investigating effects of calorie labelling on selection, and an absence of evidence for consumption. The two included studies for alcohol selection were conducted in a field setting and a naturalistic laboratory setting (replicating an online supermarket). The very small number of studies on alcohol compared with food products is in line with other reviews of interventions in physical microenvironments (Hollands 2013), and with other reviews highlighting a lack of real-world evidence including objective purchasing and consumption outcomes for alcohol calorie labelling (Robinson 2021b). Although there is increasing research and policy interest in alcohol calorie labelling interventions, research is in its infancy compared to food calorie labelling. This may be because, typically, alcohol labelling interventions have focused on other types of labels, such as unit or health warning labels (e.g. Blackwell 2018; Clarke 2021b; Clarke 2023a). In addition, alcohol consumption can be difficult to test in laboratory and field settings, for example, due to licencing or ethical issues.

We found no evidence to suggest associations between the effects of calorie labelling on food selection, and setting type, labelling type, or socioeconomic status (although this subgroup analysis only compared low deprivation samples to samples that had a mix of both low and high deprivation), suggesting that results are reasonably consistent across at least different setting types and labelling types. We were unable to assess the statistical associations between such study characteristics and the effects of calorie labelling on food consumption, or alcohol selection or consumption.

A final limitation to note concerning completeness, is that the evidence in this review is only current to August 2021, being when the full search was performed with its results fully incorporated. Although updated searches were conducted in September 2023, their results are not fully integrated into this version of the review, and would instead be incorporated in a future update.

## Quality of the evidence

Most individual studies included in the meta-analysis for the effect of calorie labelling on selection of food were subject to concerns about risk of bias, although most of the weight in the analysis was derived from studies judged to be at low risk of bias. A sensitivity analysis including only these low risk of bias studies showed the effect remained similar, albeit slightly smaller. For non-randomised studies for this outcome, all studies were at moderate risk of bias, principally due to all studies having the potential for confounding, even if there was an appropriate analysis to control for confounding domains. For the effect of calorie labelling on consumption of food, only one study was at low risk of bias, with the other studies subject to some concerns, commonly because they did not have preregistered protocols or analysis intentions.

Only two studies were included in the meta-analysis for the effect of calorie labelling on selection of alcoholic drinks and both were assessed at low risk of bias.

At the level of the evidence available for each outcome, considering the full set of GRADE criteria (detailed in Effects of interventions), we judged the evidence base to be of high certainty for the effect of calorie labelling on selection of food and non-alcoholic drinks. This suggests we can be confident that the true effect is similar to the estimated effect. We judged the evidence to be of low certainty for the effect of calorie labelling on consumption of food and non-alcoholic drinks, and very low certainty for the effect of calorie labelling on selection of alcoholic drinks, suggesting limited confidence in these effects and necessary due caution in their interpretation.

## Potential biases in the review process

At least two review authors independently performed selection of studies, data extraction, and quality assessment processes, thus reducing the potential for error and bias. However, there are potential biases that may have occurred at each stage of the review, which are outlined below.

Comprehensive searches of electronic databases designed and conducted by an information specialist, as well as supplementary citation searches and searches of other resources, were likely to have identified all relevant publications of randomised and non-randomised studies investigating calorie labelling. However, there were some additional databases included in the previous version of the review (e.g. CINAHL, TRoPHI), which we did not include. Therefore, it is possible, although unlikely given our extensive search strategy, that additional studies may have been identified if additional databases had been included. In our updated 'top-up' searches, we also identified one additional eligible study and four potentially eligible studies that we did not fully integrate into the review at this point (see Appendix 1). It is possible that including any of these studies could have changed our results or interpretation, although we consider this very unlikely given the high certainty of the current evidence for the effect of calorie labelling on selection of food and non-alcoholic drinks.

In terms of the synthesis of study results, we imputed some study results as they were not reported or available in a useable format for analysis. While these studies could have changed the effect estimates, we conducted a sensitivity analysis which removed them and it did not affect our results.

We also considered whether to downgrade for indirectness for the GRADE assessment for the selection outcome. If we had decided to downgrade for indirectness, then the GRADE rating would have been moderate, implying less certainty for this outcome. The decision not to downgrade is explained in Effects of interventions.

Finally, our interpretation of the findings could have been weighted differently if we had included a comprehensive consideration of the potential harms of calorie labelling, including attempting to capture studies designed to assess this, which future review updates should consider.

## Agreements and disagreements with other studies or reviews

The findings are broadly in line with the previous version of this Cochrane review, which concluded from small bodies of low-quality evidence that calorie labelling may reduce energy purchased and consumed from food (Crockett 2018). Although Crockett 2018 had a broader focus on nutritional labelling schemes (see Why it is important to do this review), and differed in its inclusion criteria and methods, the most common labelling type in its included studies was simple calorie labelling (12/28 studies) with a further six studies that investigated calorie labelling with various forms of additional information. For its outcome of selection with purchasing, meta-analysis of three RCTs of calorie labelling found a reduction in energy purchased equivalent to 7.8% (95% CI 2.5% to 13.1%) for a meal of 600 kcal. For consumption from a range of foods, meta-analysis of eight RCTs largely of calorie labelling (7/8 studies) found a similarly sized reduction in energy consumed. The effect size estimates in the current review, while in the same direction, are different (smaller), because the sets of studies are substantively different. In particular, for the selection outcome for food, the vast majority of the weight in the current meta-analysis came from studies that were newly included. The new studies comprised 16/19 comparisons and contributed about 95% of the weight in the current meta-analysis. For consumption, there were also differences in the sets of studies between these two reviews, although this was less marked with only two new studies added, and a further two no longer being eligible for inclusion. While the certainty of evidence for consumption was low for both the current and the 2018 Cochrane review, the evidence base is now considerably stronger for the selection outcome, with a large number of informative real-world studies newly included in the current review. Reflecting this, the certainty of evidence for this outcome was low in the 2018 review but is now high.

Beyond Cochrane reviews, there are numerous systematic reviews published since 2018 that concern the impacts of nutritional or calorie (or both) labelling schemes with varying degrees of overlap with the current review (see also Crockett 2018). We compared our findings to those from five recent systematic reviews published since Crockett 2018, and which assessed at least some studies of calorie labelling interventions and included at least one type of similar behavioural outcome in consumers (i.e. selection, purchasing, or consumption) (Croker 2020; Daley 2020; Feteira-Santos 2020; Song 2021; Robinson 2021b).

Croker 2020 included a wide range of front-of-pack labelling schemes with no specific focus on calorie labelling but did include relevant guideline daily amount and traffic light labels that commonly include specific energy information. They also limited inclusion to studies using objectively measured outcomes. They found that front-of-pack labelling in general resulted in healthier purchasing in terms of nutrient and energy content, and in terms of specific label types, traffic light labelling led to lower amounts of sodium being purchased relative to no labelling. The authors also concluded that labelling resulted in healthier purchasing in interrupted time series studies. Regarding consumption, evidence was limited and inconsistent.

Daley 2020 found that PACE labels reduced calories selected and consumed relative to no labels. However, of the 15 included studies, most were at unclear risk of bias, the majority concerned hypothetical food selection, and only one was conducted in a naturalistic setting.

Feteira-Santos 2020, in a systematic review with narrative synthesis, included a wide range of interpretative front-of-pack nutritional labels in their review, with no specific focus on calorie labelling but including relevant guideline daily amount and traffic light labels. The authors concluded that, in general, labelling had beneficial impacts on a range of outcomes including selection, and purchase intention, of products with better nutritional quality, as well as enhanced understanding of the nutritional content and perception of the healthiness of products. They could not determine superiority of any particular labelling scheme among the many types assessed.

Song 2021, in a systematic review with network meta-analysis, included a wide range of food labelling schemes including relevant traffic light labels with 'detailed' versions that provide information on specific nutrients (including energy). They found that traffic light labelling was associated with an increased probability of selecting healthier products, as well as at the level of specific nutrient reductions in energy, sodium, fat, and saturated fat content of purchases.

Robinson 2021b, in a rapid review of calorie labelling for alcoholic drinks, identified no studies assessing its impact on selection and purchasing behaviours. Only one study was identified that assessed actual consumption. This was conducted in an artificial laboratory setting and so was not included in our review, and found no significant effect of calorie labelling on consumption (Maynard 2018). Similar to our review, the review authors concluded that, for behavioural responses to calorie labelling for alcoholic drinks, there was a very low level of evidence with a high degree of uncertainty.

Across all of these systematic reviews, there was considerable variation in the focus, methods, inclusion and exclusion criteria, and measures of effect size, meaning that direct comparison of numeric results is not possible or otherwise meaningful. For example, these reviews typically include evidence from contexts that we purposefully excluded from the current review since we considered them less informative of real-world behaviour, such as artificial laboratory studies using hypothetical, motivational, constrained, or regulated outcomes. In spite of this, overall, their findings are broadly consistent with the current review.

At the most general level, all the reviews of labelling applied to food concluded that there was likely a beneficial effect of nutritional or calorie labelling schemes, including for guideline daily amount and traffic light labels that are particularly relevant to the focus of the current review. Of the two reviews that specifically focused on types of calorie labelling (Daley 2020; Robinson 2021b), the findings of Daley 2020 were consistent with ours in terms of direction of effect for PACE labelling, albeit based on a different corpus of studies, while as mentioned, Robinson 2021b concluded similarly to us that current evidence for calorie labelling for alcoholic drinks is not yet sufficiently informative (despite some more-recent evidence being available to be included in the current review).

## AUTHORS' CONCLUSIONS

### Implications for practice

The key implications of this review for public health policy and practice concern food products (including non-alcoholic drinks), since there is insufficient evidence concerning calorie labelling applied to alcohol products to credibly determine likely impacts.

This review focused specifically on the effects of calorie labelling on consumer behaviour, and the evidence relates principally to such labelling being applied across a wide range of available foods in a given setting. The current evidence suggests with high certainty that calorie labelling for food and non-alcoholic drinks added to menus, or to the front of (or adjacent to) packaged food, will lead to small reductions in energy selected and purchased in food retail settings including shops, supermarkets, and restaurants. The included studies that directly assess the impact on consumption suggest calorie labels may decrease the amount of energy consumed from food, although the evidence is of low certainty. Taken together, we consider that if estimated effects were realised at scale, then calorie labelling would have the potential to contribute to reductions in the energy intake from food, and therefore to lead to small but meaningful benefits to population health, through attenuation of weight gain or reductions in bodyweight.

It is demonstrably feasible to implement calorie labelling at a national scale via legislation and regulation, such as the policy in England introduced in April 2022 (Polden 2023). Furthermore, detailed implementation guidance is available for existing calorie labelling policies, for example, in England (Department of Health and Social Care 2021) and Northern Ireland (Food Standards Agency 2021), which could inform some of the more detailed considerations necessary when calorie labelling interventions are designed and implemented. Any such implementation should be supported and assessed via comprehensive long-term linked evaluations concurrent with the intervention and that include monitoring of progress so that adjustments can be made if the expected outcomes are not realised (Marteau 2023b). At present, the evidence base is limited in terms of assessment of long-term intervention exposures and the sustainability of effects, as well of implementation at scale, although a relatively small body of eligible non-randomised studies is able to at least partly address these limitations.

Given the estimated small effects of calorie labelling, wider systemic approaches – which could include labelling as one of an array of different interventions targeting varying levels, scales, and points of influence – may elicit greater and more durable effects (Fanzo 2021; Marteau 2023a; Waterlander 2018). Such multi-faceted approaches would include addressing systemic structural factors and counter criticisms that an overt focus on more agentic interventions might propagate and facilitate industry actors in promoting a narrative of individual responsibility for behaviour (Lacy-Nichols 2022). Unlike presumed mechanisms underlying consumer-focused labelling interventions, a wider set of policy actions could place less overall reliance on individual conscious engagement, cognitive resources, and agency, all of which may be patterned by socioeconomic position (Adams 2016; Hollands 2016; Lawson 2018). This could include complementary scalable interventions to alter economic and physical micro- and macro-environments (Glanz 2005; Hollands 2017; Shaw 2023; Swinburn 1999), such as taxes and subsidies (Anderson 2021; Shemilt 2013), measures to limit exposure to marketing (Mytton 2020; Yau 2022), product reformulation (Basto-Abreu 2018), and changing ranges of available healthier and less healthy products (Brimblecombe 2020; Clarke 2023b; De-loyde 2024; Marteau 2022; Reynolds 2021).

### Implications for research

This review reveals six important implications for research.

First, future studies should focus on alcohol, for which there was very limited evidence for selection, and a complete absence of evidence for consumption (see Overall completeness and applicability of evidence). This is particularly important given the global health burden of excess alcohol consumption and related ongoing policy interest in alcohol calorie labelling (Department of Health and Social Care 2020).

Second, while many of the included studies were conducted in real-world settings, most of these were in restaurant settings, with few studies conducted in supermarket settings. The studies in supermarkets that were included in this review were mainly high-quality non-randomised studies, with only one randomised controlled trial in a supermarket setting included in the meta-analysis.

Third, most studies investigated simple calorie labels on menus (i.e. with few studies investigating the impact of PACE (physical activity calorie equivalent) or additional nutrient information). Studies investigating different label types could also directly compare calorie labelling to alternative or complementary labelling interventions outside the scope of this review, for example, interpretative nutritional labelling schemes that include single summary indicators, such as Nutri-Score labels (Aguenaou 2021; Dubois 2021), or health warning labels (Clarke 2021c; Grummon 2020). More granular elements of calorie labelling, such as label design and formatting (e.g. size, typography, colour use), could also be considered.

Fourth, all studies examining the effect on consumption were conducted either in naturalistic laboratory or artificial laboratory settings, with no studies conducted in field settings. High-quality studies of food consumption in such contexts would be valuable, although we acknowledge that while selection and purchasing outcomes are feasible to collect in real-world settings, this is considerably more challenging for consumption outcomes.

Fifth, more food and alcohol studies are needed in lower- and middle-income countries, as well as studies that focus on only high-deprivation samples or representative samples encompassing a range of deprivation (see Overall completeness and applicability of evidence).

The above implications for research relate to the continued assessment of the likely impact of calorie labels on food and alcohol selection and consumption. However, particularly given this review suggests calorie labelling is likely to be effective in changing at least some of these outcomes, a sixth and final priority area for further research should be the comprehensive evaluation of calorie labelling policies that are and will continue to be implemented. Examples include the potential for harms or unintended consequences to consumers of calorie labelling. For example, there are concerns that calorie labelling could have detrimental effects on people with lived experience or at risk of disordered eating (Duffy 2023; Putra 2023). Additionally, exposure to calorie labels on alcoholic drinks could lead to increased alcohol use if consumers had previously overestimated the calories in these products, and are therefore pleasantly surprised by, or intend to compensate for, this (Atkinson 2024). While our review found no evidence for a risk of potential adverse effects of the intervention in leading to increases in energy or alcohol selected and consumed, there was an absence of reporting of other possible harms, such as mental health outcomes. We recommend that future updates to this review should include potential harms by assessing and synthesising any adverse effects reported in studies in line with appropriate guidance on including harms in systematic reviews (Qureshi 2022). The lack of evidence on harms in the included studies in the current review could be due to the nature of included studies focusing on short-term behavioural outcomes and often not involving monitoring of individual participants. Ultimately, any implementation of calorie labelling will necessarily have to involve decisions by policymakers and other stakeholders as to how to balance many potential impacts in opting to implement or not implement a policy, or designing its characteristics and those of any wider package of actions to minimise anticipated negative impacts.

Research should also examine factors related to industry actions. These could include pre- and post-implementation considerations and responses of industry actors (Kerins 2020), which could, for example, result in actions (e.g. changes in marketing) to dilute the impacts of the intervention and related policies (Capewell 2018), or that could enhance its effects should the intervention lead to reformulation of high-calorie products or renovations of the product portfolio (Grummon 2021; Robinson 2021a; Shangguan 2019). They also include associated costs, such as known financial costs to businesses of implementation (Kerins 2020), but also potentially broader sociopolitical costs (and benefits). Developing and implementing policies, such as for calorie labelling, likely has associated opportunity costs in reducing focus on, and opportunities for, alternative policies that may be more effective, and relatedly, depleting the accrued political will to act (Burris 2018). Sustainability should also be considered, with requirements for calorie labels potentially making it more difficult to rapidly update or change menu options to use leftover food in the out-of-home sector.

Sixth, it also merits noting that because the implementation and evaluation of these interventions fundamentally relies on the accurate measurement of calories, important inaccuracies in the estimated calorie content of products need to be addressed (Urban 2011).

## ACKNOWLEDGEMENTS

Production of this review was funded in part by a Collaborative Award in Science from the Wellcome Trust (ref: 206853/Z/17/Z). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Cochrane Public Health supported the authors in the development of this review. The following people conducted the editorial process for this review.

- Sign-off Editor (final editorial decision): Solange Durao, South African Medical Research Council
- Managing Editors (conducted editorial policy checks, selected peer reviewers, collated peer-reviewer comments, provided editorial comments/guidance to authors, edited the article): Hannah Payne and Joey Kwong, Cochrane Central Editorial Service
- Copy Editor (copy-editing and production): Anne Lawson, Cochrane Central Production Service
- Peer-reviewers (provided comments and recommended an editorial decision): Amanda Daley, Loughborough University (clinical/content review); Erin Hobin, Public Health Ontario (clinical/content review); Radheshyam Meher, AIIMS, New Delhi (consumer review); Nuala Livingstone, Cochrane Evidence Production and Methods Directorate (methods review); Ina Monsef, Cochrane Haematology, Institute of Public Health, Faculty of Medicine and University Hospital Cologne, University of Cologne, Germany (search review).

# REFERENCES

## References to studies included in this review


**Berry 2019** *{published data only}*

* Berry C, Burton S, Howlett E, Newman CL. Understanding the calorie labeling paradox in chain restaurants: why menu calorie labeling alone may not affect average calories ordered. *Journal of Public Policy & Marketing* 2019;**38**:192-213.

Berry C. The Effects of Calorie Information Disclosures on Food Choice and Retailer-related Outcomes [Dissertation]. Fayetteville (AR): University of Arkansas, 2017.

**Bollinger 2011** *{published data only}*

Bollinger B, Leslie P, Sorensen A. Calorie posting in chain restaurants. *American Economic Journal-Economic Policy* 2011;**3**(1):91-128.

**Cawley 2020** *{published data only}*

AEARCTR-0000940. The impact of restaurant menu calorie labeling on food choice. www.socialscienceregistry.org/trials/940 (first received 2 August 2018). [REGISTRATION: https://www.socialscienceregistry.org/trials/940]

* Cawley J, Susskind A, Willage B. The impact of information disclosure on consumer behavior: evidence from a randomized field experiment of calorie labels on restaurant menus. *Journal of Policy Analysis and Management* 2020;**39**:1020-42.

Cawley J, Susskind A, Willage B. The impact of information disclosure on consumer behavior: evidence from a randomized field experiment of calorie labels on restaurant menus. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=3226841 (accessed prior to 5 December 2024).

Cawley J, Susskind AM, Barton W. Does information disclosure improve consumer knowledge? *American Journal of Health Economics* 2021;**7**(4):427-56. [DOI: 10.1086/714987]

Cawley J, Susskind AM, Willage B. Does information disclosure improve consumer knowledge? Evidence from a randomized experiment of restaurant menu calorie labels. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=3597858 (accessed prior to 5 December 2024).

Cawley J. The impact of restaurant menu calorie labels on food choice: results of a randomized field experiment. 7th Annual Conference of the American Society of Health Economists; 2018 Jun 10-13; Atlanta (GA).

**Clarke 2023a** *{published data only}*

* Clarke N, Ferrar J, Pechey E, Ventsel M, Pilling MA, Munafo MR, et al. Impact of health warnings and calorie labels on selection and purchasing of alcoholic and non-alcoholic drinks: a randomised controlled trial. *Addiction (Abingdon, England)* 2023;**118**(12):2327-41.

ISRCTN10313219. Impact of health warning labels and calorie information on the selection and purchasing of alcohol. www.isrctn.com/ISRCTN10313219 (first received 17 August 2021). [DOI: 10.1186/ISRCTN10313219]

**Dubois 2021** *{published data only}*

* Dubois P, Albuquerque P, Allais O, Bonnet C, Bertail P, Combris P, et al. Effects of front-of-pack labels on the nutritional quality of supermarket food purchases: evidence from a large-scale randomized controlled trial. *Journal of the Academy of Marketing Science* 2021;**49**:119-38.

Dubois P, Allais O, Albuquerque P, Bertail P, Bonnet C, Chandon P, et al. Impact of analytic and synthetic front-of-pack graphical nutrition labels on the nutritional quality of supermarket food purchases: evidence from a French randomized control experiment. Research Papers In Economics 2018.

Effects of front-of-pack nutrition labels on food purchases: a randomised controlled trial. www.isrctn.com/ISRCTN58212763 (first received 7 November 2017). [DOI: 10.1186/isrctn58212763]

**Elbel 2009** *{published data only}*

Elbel B, Kersh R, Brescoll VL, Dixon LB. Calorie labeling and food choices: a first look at the effects on low-income people in New York City. *Health Affairs* 2009;**28**:w1110-w1121.

**Ellison 2013** *{published data only}*

Ellison B, Lusk J, Davis D. Looking at the label and beyond: the effects of calorie label, health consciousness and demographics on calorie intake in restaurants. *International Journal of Behavioural Nutrition and Physical Activity* 2013;**10**(21):1-9.

**Ellison 2014** *{published data only}*

Ellison B, Lusk J, Davis D. The impact of restaurant calorie labels on food choice: results from a field experiment. *Economic Inquiry* 2014;**52**(2):666-81.

**Elshiewy 2018** *{published data only}*

Elshiewy O, Boztug Y. When back of pack meets front of pack: how salient and simplified nutrition labels affect food sales in supermarkets. *Journal of Public Policy & Marketing* 2018;**37**:55-67.

**Fichera 2020** *{published data only}*

Fichera E, von Hinke S. The response to nutritional labels: evidence from a quasi-experiment. *Journal of Health Economics* 2020;**72**:1-17.

**Hammond 2013** *{published data only}*

* Hammond D, Goodman S, Hanning R, Danel S. A randomised trial of calorie labeling on menus. *Preventive Medicine* 2013;**57**(6):860-6.

NCT01948752. Clinical trial of menu labeling. clinicaltrials.gov/ct2/show/NCT01948752 (first received 24 September 2013).

**Harnack 2008** *{published data only}*

* Harnack LJ, French SA, Oakes JM, Story MT, Jeffery RW, Rydell SA. Effects of calorie labeling and value size pricing on fast food meal choices: results from an experimental trial. International Journal of Behavioral Nutrition and Physical Activity 2008;**5**(63):13.

NCT00127660. Effect of nutrition labeling on fast food choices. clinicaltrials.gov/ct2/show/NCT00127660 (first received 8 August 2005).

**James 2015** *{published data only}*

James A, Adams-Huet B, Crisp K, Mitchell J, Dart L, Turner M, et al. The effect of menu labels, displaying minutes of brisk walking needed to burn food calories, on calories ordered and consumed in young adults. FASEB Journal 2013. [DOI: 10.1096/fasebj.27.1_supplement.367.2]

* James A, Adams-Huet B, Shah M. Menu labels displaying the kilocaloric content or the exercise equivalent: effects on energy ordered and consumed in young adults. *American Journal of Health Promotion* 2015;**29**(5):294-302.

**Oliveira 2018** *{published data only}*

Oliveira RC, Fernandes AC, Proenca RP, Hartwell H, Rodrigues VM, Colussi CF, et al. Menu labelling and healthy food choices: a randomised controlled trial. *British Food Journal* 2018;**120**:788-803.

**Petimar 2019** *{published data only}*

Petimar J, Zhang F, Cleveland LP, Simon D, Gortmaker SL, Polacsek M, et al. Estimating the effect of calorie menu labeling on calories purchased in a large restaurant franchise in the southern United States: quasi-experimental study. *BMJ (Clinical Research Ed.)* 2019;**367**:l5837.

**Petimar 2021** *{published data only}*

Petimar J, Moran AJ, Ramirez M, Block JP. A natural experiment to evaluate the nutritional content of restaurant meal purchases after calorie labeling. *Journal of the Academy of Nutrition and Dietetics* 2020;**120**(12):2039-46.

Petimar J, Ramirez M, Rifas-Shiman SL, Linakis S, Mullen J, Roberto CA, et al. Evaluation of the impact of calorie labeling on McDonald's restaurant menus: a natural experiment. *International Journal of Behavioral Nutrition and Physical Activity* 2019;**16**:99.

* Petimar J, Zhang F, Rimm EB, Simon D, Cleveland LP, Gortmaker SL, et al. Changes in the calorie and nutrient content of purchased fast food meals after calorie menu labeling: a natural experiment. *PLOS Medicine* 2021;**18**:e1003714.

Petimar J, Zhang F, Simon D, Cleveland L, Gortmaker S, Bleich S, et al. Changes in nutrient content of meal purchases after calorie menu labeling. *Obesity (Silver Spring, Md.)* 2020;**28**:Poster 083.

**Platkin 2014** *{published data only}*

* Platkin C, Yeh M-C, Hirsch K, Weiss Wiewel E, Lin C-Y, Tung H-J, et al. The effect of menu labeling with calories and exercise equivalents on food selection and consumption. *BMC Obesity* 2014;**1**:21.

Platkin C. Think before you eat: calories and exercise equivalents presented on menus at point-of-choice. Dissertation 2017.

**Reynolds 2022** *{published data only}*

ISRCTN31315776. Effect of physical activity calorie equivalent (PACE) labels on food purchased in worksite cafeterias. www.isrctn.com/ISRCTN31315776 (first received 22 March 2021). [REGISTRATION: https://www.isrctn.com/ISRCTN31315776]

* Reynolds JP, Ventsel M, Hobson A, Pilling M, Pechey R, Jebb SA, et al. Evaluation of physical activity calorie equivalent (PACE) labels' impact on energy purchased in cafeterias: a stepped-wedge randomised controlled trial. *PLOS Medicine* 2022;**19**(11):e1004116.

**Roberto 2010** *{published data only}*

NCT01626729. Facts-up-front versus traffic-light food labels. clinicaltrials.gov/ct2/show/NCT01626729 (first received 25 June 2012).

* Roberto CA, Larsen PD, Agnew H, Baik J, Brownell KD. Evaluating the impact of menu labeling on food choices and intake. *American Journal of Public Health* 2010;**100**(2):312-8.

**Roberto 2012** *{published data only}*

Roberto CA, Shivaram M, Martinez O, Boles C, Harris JL, Brownell KD. The Smart Choices front-of-package nutrition label. Influence on perceptions and intake of cereal. *Appetite* 2012;**58**:651-7.

**Robertson 2020** *{published data only}*

Lunn P, Robertson D. The position of calories on menus influences how much people eat. Research Papers In Economics 2019.

Robertson DA, Lunn P. The effect of spatial position of calorie information on choice, consumption and attention. Economic and Social Research Institute (ESRI).

* Robertson DA, Lunn PD. The effect of spatial location of calorie information on choice, consumption and eye movements. *Appetite* 2020;**144**:104446.

**Temple 2010** *{published data only}*

Temple JL, Johnson K, Recupero K, Suders H. Nutrition labels decrease energy intake in adults consuming lunch in the laboratory. *Journal of the American Dietetic Association* 2010;**110**:1094-7.

**VanEpps 2016** *{published data only}*

NCT01473225. Can calorie labels increase caloric intake? clinicaltrials.gov/ct2/show/NCT01473225 (first received 17 November 2011).

* VanEpps EM, Downs JS, Loewenstein G. Calorie label formats: using numeric and traffic light calorie labels to reduce lunch calories. *Journal of Public Policy & Marketing* 2016;**35**(1):26-36.

**Vasiljevic 2018** *{published data only}*

Vasiljevic M, Cartwright E, Pechey R, Hollands GJ, Couturier DL, Jebb SA, et al. Physical micro-environment interventions for healthier eating in the workplace: protocol for a stepped wedge randomised controlled pilot trial. Pilot and Feasibility Studies 2017;**3**:27.

Vasiljevic M, Cartwright E, Pilling M, Lee MM, Bignardi G, Pechey R, et al. Impact of calorie labelling in worksite cafeterias: a stepped wedge randomised controlled pilot trial. *International Journal of Behavioral Medicine* 2018;**15**(1):41.

* Vasiljevic M, Cartwright E, Pilling M, Lee MM, Bignardi G, Pechey R, et al. Impact of calorie labelling in worksite cafeterias: a stepped wedge randomised controlled pilot trial. *International Journal of Behavioral Nutrition and Physical Activity* 2018;**15**:41.

**Vasiljevic 2019** *{published data only}*

ISRCTN20474205. Enhanced calorie labelling to improve employees' diets in workplace cafeterias. www.isrctn.com/ISRCTN20474205 (first received 21 February 2018).

* Vasiljevic M, Fuller G, Pilling M, Hollands GJ, Pechey R, Jebb SA, et al. What is the impact of increasing the prominence of calorie labelling? A stepped wedge randomised controlled pilot trial in worksite cafeterias. *Appetite* 2019;**141**:104304.

## References to studies excluded from this review

**Acton 2019** *{published data only}*

* Acton RB, Jones AC, Kirkpatrick SI, Roberto CA, Hammond D. Taxes and front-of-package labels improve the healthiness of beverage and snack purchases: a randomized experimental marketplace. *International Journal of Behavioral Nutrition and Physical Activity* 2019;**16**(1):46.

**Acton 2020a** *{published data only}*

* Acton RB, Hammond D. Impact of sugar taxes and front-of-package nutrition labels on purchases of protein, calcium and fibre. *Preventive Medicine* 2020;**136**:106091.

**Acton 2020b** *{published data only}*

* Acton R. The Influence of Sugar Taxes and Front-of-Package Nutrition Labels on Consumer Purchasing Behaviours: a Randomized Experimental Marketplace [thesis]. Ontario (Canada): University of Waterloo, 2020.

**Acton 2021** *{published data only}*

* Acton RB, Kirkpatrick SI, Hammond D. Exploring the main and moderating effects of individual-level characteristics on consumer responses to sugar taxes and front-of-pack nutrition labels in an experimental marketplace. *Canadian Journal of Public Health. Revue Canadienne de Sante Publique* 2021;**112**(4):647-62.

**Allan 2015** *{published data only}*

Allan JL, Johnston M, Campbell N. Snack purchasing is healthier when the cognitive demands of choice are reduced: a randomized controlled trial. *Health Psychology* 2015;**34**(7):750-5.

**Allan 2020** *{published data only}*

Allan JL, Powell DJ. Prompting consumers to make healthier food choices in hospitals: a cluster randomised controlled trial. *International Journal of Behavioral Nutrition and Physical Activity* 2020;**17**(1):86.

**Al-Otaibi 2021** *{published data only}*

Al-Otaibi H, Al-Sandal T, Elkatr HO. Is calorie labeling on menus related to weight disturbances among females in Saudi Arabia? *Journal of Nutrition and Metabolism* 2021;**2021**:4041451. [DOI: 10.1155/2021/4041451]

**Bailey 2022** *{published data only}*

Bailey RL, Kwon K, Garcia CJ, Wang P. Fast food menu calorie labeling contexts as complex contributing factors to overeating. *Appetite* 2022;**173**:105992-2. [DOI: 10.1016/j.appet.2022.105992]

**Balasubramanian 2002** *{published data only}*

Balasubramanian SK, Cole C. Consumers' search and use of nutrition information: the challenge and promise of the Nutrition Labeling and Education Act. *Journal of Marketing* 2002;**66**(3):112-27. [Study 3 in published paper]

**Baum 2017** *{published data only}*

* Baum JI, Gaines BL, Kubas GC, Mitchell CF, Russell SL. Educational nutrition messaging at breakfast reduces snack intake and influences snack preferences in adult men and women. *Appetite* 2017;**117**:67-73.

**Bergen 2006** *{published data only}*

Bergen D, Yeh M-C. Effects of energy-content labels and motivational posters on sales of sugar-sweetened beverages: stimulating sales of diet drinks among adults study. *Journal of the American Dietetic Association* 2006;**106**(11):1866-9.

**Bergman 2021** *{published data only}*

* Bergman C, Tian Y, Moreo A, Raab C. Menu engineering and dietary behavior impact on young adults' kilocalorie choice. *Nutrients* 2021;**13**(7):2329.

**Byrd 2017** *{published data only}*

* Byrd K. Sodium and the restaurant consumer: associations with dining out frequency and menu nutrition information. Dissertation Abstracts International: Section B: the Sciences and Engineering. https://docs.lib.purdue.edu/dissertations/AAI10270891/ (accessed 6 January 2025).

**Byrd 2021** *{published data only}*

* Byrd K, Almanza B. Restaurant menu labeling for calories and sodium: effect of consumer mindset of immediate versus future consequences. *Journal of Food Service Business Research* 2021;**24**(3):310-47.

**Carter 2018** *{published data only}*

* Carter KA, González-Vallejo C. Nutrient-specific system versus full fact panel: testing the benefits of nutrient-specific front-of-package labels in a student sample. *Appetite* 2018;**125**:512-26.

**Cavanagh 2014** *{published data only}*

Cavanagh KV, Kruja B, Forestall CA. The effect of brand and caloric information on flavor perception and food consumption in restrained and unrestrained eaters. *Appetite* 2014;**82**:1-7.

**Chu 2009a** *{published data only}*

* Chu YH, Frongillo EA, Jones SJ, Kaye GL. Improving patrons' meal selections through the use of point-of-selection nutrition labels. *American Journal of Public Health* 2009;**99**(11):2001-5.

**Chu 2009b** *{published data only}*

* Chu YH, Frongillo EA, Jones SJ, Kaye GL. Nutrition labels at point-of-selection in a food-service operation improves patrons' meal selection. *FASEB Journal* 2009;**23**(1):336.5.

**Cioffi 2015** *{published and unpublished data}*

* Cioffi CE, Levitsky D, Pacanowski CR, Bertz F. A nudge in a healthy direction. The effect of nutrition labels on food purchasing behaviors in university dining facilities. *Appetite* 2015;**92**:7-14.

Cioffi CE, Levitsky D. An Examination of the Effect of Nutrition Labels on Food Purchasing Behavior in Cornell Dining Units [Honors Degree Research Paper]. Cornell (USA): University of Cornell, 2012.

**Cornil 2017** *{published data only}*

* Cornil Y, Chandon P. Sensory or nutrition menu labeling? A field experiment in aligning public health and restaurant business goals. *Advances in Consumer Research* 2017;**45**:574-5. [http://www.acrwebsite.org/volumes/1024733/volumes/v45/NA-45]

**Crockett 2014** *{published and unpublished data}*

* Crockett RA, Jebb SA, Hankins M, Marteau TM. The impact of nutritional labels and socioeconomic status on energy intake: an experimental field study. *Appetite* 2014;**81**:12-9.

**Crosetto 2017** *{published data only}*

* Crosetto P, Lacroix A, Muller L, Ruffieux B. Modifications of food purchases in response to five nutrition simplified labelling. *Cahiers de Nutrition et de Dietetique* 2017;**52**(3):129-33.

**Crosetto 2018** *{published data only}*

* Crosetto P, Lacroix A, Muller L, Ruffieux B. Nutritional and economic impact of 5 alternative front-of-pack nutritional labels: experimental evidence. Research Papers In Economics.

**Crosetto 2020** *{published data only}*

* Crosetto P, Lacroix A, Muller L, Ruffieux B. Nutritional and economic impact of five alternative front-of-pack nutritional labels: experimental evidence. *European Review of Agricultural Economics* 2020;**47**(2):785-818.

**Dallas 2017** *{published data only}*

* Dallas SK, Liu P, Ubel PA. Don't count calorie labeling out: calorie counts on the left side of menu items lead to lower calorie food choices. *Advances in Consumer Research* 2017;**45**:576-7. [https://www.acrwebsite.org/volumes/1023629/volumes/v45/NA-45]

**Dallas 2019** *{published data only}*

* Dallas SK, Liu PJ, Ubel PA. Don't count calorie labeling out: calorie counts on the left side of menu items lead to lower calorie food choices. *Journal of Consumer Psychology* 2019;**29**(1):60-9.

**Defago 2017** *{published data only}*

* Defago D, Geng JF, Molina O, Maria DS. Digestible information: the impact of multiple traffic light nutritional labeling in a developing country. University Library of Munich, Munich (Germany) 2017.

**Dubbert 1984** *{published data only}*

* Dubbert PM, Johnson WG, Schlundt DG, Montague NW. The influence of caloric information on cafeteria food choices. *Journal of Applied Behavior Analysis* 1984;**17**(1):85-92.

**Ebneter 2013** *{published data only}*

* Ebneter D, Latner J, Nigg C. Is less always more? The effects of low fat labeling and caloric information on food intake, calorie estimates, taste preference, and health attributions. *Appetite* 2013;**68**:92-7.

**Egnell 2019** *{published data only}*

* Egnell M, Boutron I, Péneau S, Ducrot P, Touvier M, Galan P, et al. Front-of-pack labeling and the nutritional quality of students' food purchases: a 3-arm randomized controlled trial. *American Journal of Public Health* 2019;**109**(8):1122-9.

**Erdem 2021** *{published data only}*

* Erdem S. Investigating the effect of restaurant menu labelling on consumer food choices using a field experiment. *British Food Journal* 2021;**124**(11):3447-67. [DOI: 10.1108/bfj-04-2021-0432]

**Fang 2019** *{published data only}*

* Fang D, Nayga RM, Snell HA, West GH, Bazzani C. Evaluating USA's new nutrition and supplement facts label: evidence from a non-hypothetical choice experiment. *Journal of Consumer Policy* 2019;**42**(4):545-62.

**Finkelstein 2019a** *{published data only}*

* Finkelstein EA, Ang FJ, Doble B, Wong WH, van Dam RM. A randomized controlled trial evaluating the relative effectiveness of the multiple traffic light and Nutri-Score front of package nutrition labels. *Nutrients* 2019;**11**(9):2236.

**Finkelstein 2019b** *{published data only}*

* Finkelstein E, Doble B, Ang FJ, Wong MW, van Dam RM. A randomized controlled trial testing the effects of Singapore's front-of-pack healthier choice symbol label with or without a physical activity equivalent label on food purchases and measures of diet quality. Research Square 2019.

**Finkelstein 2020** *{published data only}*

* Finkelstein EA, Ang FJ, Doble B. Randomized trial evaluating the effectiveness of within versus across-category front-of-package lower-calorie labelling on food demand. *BMC Public Health* 2020;**20**(1):312.

**Finkelstein 2021** *{published data only}*

* Finkelstein EA, Doble B, Ang FJ, Wong WH, van Dam RM. A randomized controlled trial testing the effects of a positive front-of-pack label with or without a physical activity equivalent label on food purchases. *Appetite* 2021;**158**:104997.

**Girz 2012** *{published data only}*

* Girz L, Polivy J, Herman CP, Lee H. The effects of calorie information on food selection and intake. *International Journal of Obesity* 2012;**36**(10):1340-5.

**Goodman 2018** *{published data only}*

* 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.

**Guha 2018** *{published data only}*

* Guha A, Biswas A, Grewal D, Bhowmick S, Nordfalt, J. An empirical analysis of the joint effects of shoppers' goals and attribute display on shoppers' evaluations. *Journal of Marketing* 2018;**82**(3):142-56.

**Gustafson 2018** *{published data only}*

* Gustafson CR, Zeballos E. The effect of ingredient-specific calorie information on calories ordered. *Preventive Medicine Reports* 2018;**12**:186-90. [DOI: 10.1016/j.pmedr.2018.09.013]

**Hartley 2019a** *{published data only}*

* Hartley C, Keast RS, Liem DG. The response of more health focused and less health focused people to a physical activity calorie equivalent label on discretionary snack foods. *Nutrients* 2019;**11**(3):525.

**Hartley 2019b** *{published data only}*

* Hartley IE, Keast RS, Liem DG. Physical activity-equivalent label reduces consumption of discretionary snack foods. *Public Health Nutrition* 2019;**21**(8):1435-43.

**Holmes 2013** *{published data only}*

Holmes AS, Serrano EL, Machin JE, Duetsch T, Davis GC. Effect of different children's menu labeling designs on family purchases. *Appetite* 2013;**62**:198-202.

**Huseynov 2021** *{published data only}*

* Huseynov S, Palma MA, Ahmad G. Does the magnitude of relative calorie distance affect food consumption. *Journal of Economic Behavior & Organization* 2021;**188**:530-51.

**IRCT20181002041201N1** *{published data only}*

IRCT20181002041201N1. The effect of physical activity labeling which indicates the amount of physical activity needed to burn energy of chosen food, and traffic light labeling on food choice. irct.behdasht.gov.ir/trial/34182 (first received 25 October 2018).

**ISRCTN90365793** *{published data only}*

* ISRCTN90365793. Prompting healthier snack choices in Scottish hospitals. trialsearch.who.int/Trial2.aspx?TrialID=ISRCTN90365793 (first received 8 October 2019). [REGISTRATION: https://trialsearch.who.int/Trial2.aspx?TrialID=ISRCTN90365793]

**Jin 2020** *{published data only}*

* Jin HZ, Li YN, Li DJ, Zheng J. The effects of physical activity calorie equivalent labeling on dieters' food consumption and post-consumption physical activity. *Journal of Consumer Affairs* 2020;**54**(2):723-41.

**Jue 2012** *{published data only}*

* Jue JJ, Press MJ, McDonald D, Volpp KG, Asch DA, Mitra N, et al. The impact of price discounts and calorie messaging on beverage consumption: a multi-site field study. *Preventive Medicine* 2012;**55**(6):629-33.

**Julia 2021** *{published data only}*

* Julia C, Arnault N, Agaësse C, Fialon M, Deschasaux-Tanguy M, Andreeva VA, et al. Impact of the front-of-pack label Nutri-Score on the nutritional quality of food choices in a quasi-experimental trial in catering. *Nutrients* 2021;**13**(12):4530. [DOI: 10.3390/nu13124530]

**Karnik 2018** *{published data only}*

* Karnik H, Peterson CH. Impact of point of sales nutritional labels on food purchase: evidence from the rural Midwest. 2018 Agricultural & Applied Economics Association Annual Meeting; 2018 Aug 5-7; Washington (DC).

**Kral 2002** *{published data only}*

* Kral TV, Roe LS, Rolls BJ. Does nutrition information about the energy density of meals affect food intake in normal-weight women? *Appetite* 2002;**39**:137-45.

**Lee 2019** *{published data only}*

* Lee KM, Hammond D, Hobin E, Price M, Olstad DL, Minaker LM, et al. Examining the impact of numeric versus traffic light calorie labelling at the point-of-purchase on young adults' food and beverage purchases. Dietitians of Canada 2019 National Conference; 2019 Jun 5-8; Ottawa (Canada). [URL: https://www.morressier.com/o/event/5cdc2526bf4ffc620542101b/article/5cdc25c8bf4ffc62054210c0]

**Machin 2018** *{published data only}*

* Machin L, Aschemann-Witzel J, Curutchet MR, Gimenez A, Ares G. Does front-of-pack nutrition information improve consumer ability to make healthful choices? Performance of warnings and the traffic light system in a simulated shopping experiment. *Appetite* 2018;**121**:55-62.

**Machin 2019** *{published data only}*

* Machin L, Curutchet MR, Gimenez A, Aschemann-Witzel J, Ares G. Do nutritional warnings do their work? Results from a choice experiment involving snack products. *Food Quality and Preference* 2019;**77**:159-65.

**Mantzari 2020** *{published data only}*

* Mantzari E, Pechey R, Codling S, Sexton O, Hollands GJ, Marteau TM. The impact of 'on-pack' pictorial health warning labels and calorie information labels on drink choice: a laboratory experiment. *Appetite* 2020;**145**:104484.

**Maynard 2018** *{published data only}*

* Maynard OM, Langfield T, Attwood AS, Allen E, Drew I, Votier A, et al. No impact of calorie or unit information on ad libitum alcohol consumption. *Alcohol and Alcoholism* 2018;**53**(1):12-9.

**Mazza 2018** *{published data only}*

* Mazza MC, Dynan L, Siegel RM, Tucker AL. Nudging healthier choices in a hospital cafeteria: results from a field study. *Health Promotion Practice* 2018;**19**(6):925-34.

**McCrickerd 2020** *{published data only}*

* McCrickerd K, Tang CS, Forde CG. The independent and combined impact of front-of-pack labelling and sensory quality on calorie estimations and portion selection of commercial food products. *Food Quality and Preference* 2020;**79**:103766.

**McElroy 2016** *{published data only}*

* McElroy K, Clements M, Feehan O, Cronin BE, Griffin B, Douglas FE, et al. Calories on menus: does impact diminish with duration or flourish after festivities? *Proceedings of the Nutrition Society* 2016;**75**:E87.

**McInerney 2017** *{published data only}*

* McInerney M, Hutchins M, Peterson Y, Honselman C. Menu labels decrease the number of calories selected from fast-food restaurants among college students: a pilot study. *Journal of the Academy of Nutrition and Dietetics* 2017;**117**(9):A90.

**Montford 2017** *{published data only}*

* Montford WJ. Informing consumer decision-making: two empirical studies. Dissertation Abstracts International Section A: Humanities and Social Sciences 2017;**77**:11-A(E).

**Muller 2020a** *{published data only}*

* Muller L, Ruffieux B. What makes a front-of-pack nutritional labelling system effective: the impact of key design components on food purchases. *Nutrients* 2020;**12**(9):1-25.

**Muller 2020b** *{published data only}*

Muller L, Ruffieux B. What makes a front-of-pack nutritional labelling system effective: the impact of key design components on food purchases. Research Papers In Economics 2020.

**Nassab 2017** *{published data only}*

* Nassab K. The effect of caloric labeling on consumers calorie intake. https://scholarexchange.furman.edu/scjas/2017/all/86 (accessed prior to 5 December 2024).

**NCT03553043** *{published data only}*

NCT03553043. Energy labelling for alcohol drinks in New Zealand: consumers perceptions and impacts on purchase behaviour. clinicaltrials.gov/study/NCT03553043 (first received 20 May 2018).

**NCT03761342** *{published data only}*

NCT03761342. Evaluating the relative effectiveness of two front-of-pack nutrition labels. clinicaltrials.gov/study/NCT03761342 (first received 28 November 2018). [REGISTRATION: https://clinicaltrials.gov/show/NCT03761342 2018.]

**NCT04172337** *{published data only}*

NCT04172337. Testing the effects of Singapore's front-of-pack healthier choice symbol label with or without a physical activity equivalent label. clinicaltrials.gov/study/NCT04172337 (first received 18 November 2019). [REGISTRATION: https://clinicaltrials.gov/show/NCT04172337]

**NCT04252898** *{published data only}*

NCT04252898. Impact of a front-of-pack nutrition labelling on food purchases in contract catering. clinicaltrials.gov/study/NCT04252898 (first received 24 January 2020). [https://explore.openalex.org/works/W4241218153]

**Neal 2017** *{published data only}*

ACTRN12614000964617. The effects of four front-of-pack labelling schemes compared to standard Nutrition Information Panel, on the healthiness of food purchases among adult Australian consumers [The effects of four front-of-pack labelling schemes compared to standard Nutrition Information Panel, on mean nutrient profiling score of food purchases among adult Australian consumers: a randomised trial]. www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=366945 (first received 21 August 2014). [www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=366945 (first received 21 August 2014). [CRS ID: 7222224]]

* Neal B, Crino M, Dunford E, Gao A, Greenland R, Li N, et al. Effects of different types of front-of-pack labelling information on the healthiness of food purchases – a randomised controlled trial. *Nutrients* 2017;**9**(112):1284.

**Neuhofer 2020** *{published data only}*

* Neuhofer Z, McFadden BR, Rihn A, Wei X, Khachatryan H, House L. Can the updated nutrition facts label decrease sugar-sweetened beverage consumption? *Economics and Human Biology* 2020;**37**:100867.

**Ni Mhurchu 2017** *{published data only}*

* Ni Mhurchu N, Volkova E, Jiang Y, Eyles H, Michie J, Neal B, et al. Effects of interpretive nutrition labels on consumer food purchases: the Starlight randomized controlled trial. *American Journal of Clinical Nutrition* 2017;**105**(3):695-704.

**Nordström 2020** *{published data only}*

* Nordström J, Thunström L, Van't Veld K, Shogren JF, Ehmke M. Strategic ignorance of health risk: its causes and policy consequences. *Behavioural Public Policy* 2020;**7**:83-114.

**Oh 2020** *{published data only}*

* Oh GE, Huh YE, Mukhopadhyay A. Informed indulgence: the effects of nutrition information provision and dietary restraint on consecutive food consumption decisions. *Psychology & Health* 2020;**36**:1314-35.

**Otto 2020** *{published data only}*

* Otto AS, Davis B, Wakefield K, Clarkson JJ, Inman J. Consumer strategies to improve the efficacy of posted calorie information: how provincial norms nudge consumers to healthier consumption. *Journal of Consumer Affairs* 2020;**54**(1):311-41.

**Rayner 2017** *{published data only}*

* Rayner M, Harrington R, Scarborough P. The impact of the UK traffic-light labelling system on diets. *Annals of Nutrition & Metabolism* 2017;**71**:101-2.

**Reinhardt 2019** *{published data only}*

* Reinhardt D, Hurtienne J. Only one item left? Heuristic information trumps calorie count when supporting healthy snacking under low self-control. 2019 CHI Conference on Human Factors in Computing Systems; 2019 May 4-8; Glasgow (UK).

**Rising 2017** *{published data only}*

* Rising CJ, Bol N. Nudging our way to a healthier population: the effect of calorie labeling and self-control on menu choices of emerging adults. *Health Communication* 2017;**32**(8):1032-8.

**Sandoval 2017** *{published data only}*

* Sandoval LA, Carpio CE, Sanchez M, Borja I, Cabrera T. The effect of 'traffic-light' nutritional labelling on carbonated soft drink purchases in Ecuador. Agricultural & Applied Economics Association Annual Meeting; 2017 July 30-August 1; Chicago (IL).

**Scourboutakos 2016** *{published data only}*

* Scourboutakos MJ. Eating-out: a study of the nutritional quality of Canadian chain restaurant foods and interventions to promote healthy eating. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 2016;**78**:9-B(E).

**Seyedhamzeh 2020** *{published data only}*

* Seyedhamzeh S, Nedjat S, Hosseini H, Shakibazedeh E, Viera AJ, Dorosty M, et al. Potential effect of different nutritional labels on food choices among mothers: a study protocol. *BMC Public Health* 2020;**20**(1):294.

**Shin 2020** *{published data only}*

* Shin S, van Dam RM, Finkelstein EA. The effect of dynamic food labels with real-time feedback on diet quality: results from a randomized controlled trial. *Nutrients* 2020;**12**(7):2158.

**Shoychet 2023** *{published data only}*

* Shoychet G, Lowe CJ, Bodell LP. Does menu labelling influence food choice and consumption in female undergraduate students? *Canadian Journal Of Behavioural Science* 2023;**55**(3):220-7. [DOI: 10.1037/cbs0000336]

**Silva 2022** *{published data only}*

* Silva A, Mhurchu CN, Anastácio LR. Comparison of two front-of-pack nutrition labels for Brazilian consumers using a smartphone app in a real-world grocery store: a pilot randomized controlled study. *Frontiers in Nutrition* 2022;**9**:898021. [DOI: 10.3389/fnut.2022.898021]

**Tangari 2019** *{published data only}*

* Tangari AH, Bui M, Haws KL, Liu PJ. That's not so bad, I'll eat more! Backfire effects of calories-per-serving information on snack consumption. *Journal of Marketing* 2019;**83**(1):133-50.

**Tapper 2019** *{published data only}*

* Tapper K, Yarrow K, Farrar S, Mandeville K. The effect of calorie labelling, motivation and habits on the calorie content of items selected from a coffee shop menu. 43rd Meeting of the British Feeding and Drinking Group; 2019 Apr 16-17; Swansea (UK).

**Tapper 2021a** *{published data only}*

* Tapper K, Yarrow K, Farrar S, Mandeville K. The effect of calorie labeling, motivation and habits on the calorie content of items selected from a university coffee shop menu. *Obesity Facts* 2021;**14**:39.

**Thunstrom 2018** *{published data only}*

Thunstrom L. Incidence of an emotional tax: the case of calorie menu labeling. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=3272036 (accessed prior to 5 December 2024). [papers.ssrn.com/sol3/papers.cfm?abstract_id=3272036]

**Thunstrom 2019** *{published data only}*

* Thunstrom L. Welfare effects of nudges: the emotional tax of calorie menu labeling. *Judgment and Decision Making* 2019;**14**(1):11-25.

**Turnwald 2019a** *{published data only}*

* Turnwald BP, Bertoldo JD, Perry MA, Policastro P, Timmons M, Bosso C, et al. Increasing vegetable intake by emphasizing tasty and enjoyable attributes: a randomized controlled multisite intervention for taste-focused labeling. *Psychological Science* 2019;**30**(11):1603-15.

**Turnwald 2019b** *{published data only}*

* Turnwald BP, Crum AJ. Smart food policy for healthy food labeling: leading with taste, not healthiness, to shift consumption and enjoyment of healthy foods. *Preventive Medicine* 2019;**119**:7-13.

**Urminsky 2019** *{published data only}*

* Urminsky O, Goswami I. The 'mere reminder' effect of salient calorie labeling. https://ssrn.com/abstract=3448160 (accessed prior to 5 December 2024).

**VanEpps 2021** *{published data only}*

* VanEpps EM, Molnar A, Downs JS, Loewenstein G. Choosing the light meal: real-time aggregation of calorie information reduces meal calories. *Journal of Marketing Research* 2021;**58**(5):948-67. [DOI: 10.1177/00222437211022367]

**Vermeer 2011** *{published data only}*

* Vermeer WM, Steenhuis IH, Leeuwis FH, Bos AE, De Boer M, Seidell JC. View the label before you view the movie: a field experiment into the impact of portion size and guideline daily amounts labelling on soft drinks in cinemas. *BMC Public Health* 2011;**11**:438.

**Viera 2017** *{published data only}*

* Viera AJ, Tuttle L, Olsson E, Gras-Najjar J, Gizlice Z, Hales D, et al. Effects of physical activity calorie expenditure (PACE) labeling: study design and baseline sample characteristics. *BMC Public Health* 2017;**17**(1):702.

**Viera 2019** *{published data only}*

* Viera AJ, Gizlice Z, Tuttle L, Olsson E, Gras-Najjar J, Hales D, et al. Effect of calories-only vs physical activity calorie expenditure labeling on lunch calories purchased in worksite cafeterias. *BMC Public Health* 2019;**19**(1):107.

**Wansink 2006** *{published data only}*

* Wansink B, Chandon P. Can "Low-Fat" nutrition labels lead to obesity? *Journal of Marketing Research* 2006;**43**(4):605-17. [Study 1 of three presented in the paper]

**Whitt 2017** *{published data only}*

* Whitt O, Pitts S, Jilcott B, Rafferty AP, Payne CR, Ng SW. The effects of traffic light labelling versus cartoon labelling on food and beverage purchases in a children's hospital setting. *Pediatric Obesity* 2018;**13**:265-8. [DOI: 10.1111/ijpo.12232]

**Zhang 2020** *{published data only}*

* Zhang S, Zhou L, Zhao Y. How to present calorie information on the electronic menu to help people order more healthily. 22nd International Conference On Human-Computer Interaction; 2020 Jul 19-24; Copenhagen (Denmark).

### References to studies awaiting assessment

**Clements 2016** *{published data only}*

* Clements M, Feehan O, McElroy K, Griffin B, Cronin BE, Douglas FE, et al. Calories on menus and that 'Friday feeling'. *Proceedings of the Nutrition Society* 2016;**75 (OCE3)**:E89.

**Dos Santos 2015** *{published data only}*

* Dos Santos, De Souza B, Hansen G, Perez-Cueto FJ. The effects of nudging in coffee machines on two types of companies in Denmark. *Annals of Nutrition & Metabolism* 2015;**67**:516.

**Elshiewy 2022** *{published data only}*

* Elshiewy O. The impact of voluntary front-of-pack nutrition-label introduction on purchase behavior. Dissertation.

**Girz ongoing** *{unpublished data only}*

Girz L. The Impact of calorie labels and a confederate on ordering behaviour. In process.

**Lee 2018** *{published data only}*

* Lee S. Impacts of calorie information and BMI rate on consumers' decisions at the point of purchase in fast food restaurants. *Gwan'gwang Yeon'gu* 2018;**33**:433-46. [DOI: 10.21719/ijtms.33.8.23]

**Petimar 2022** *{published data only}*

* Petimar J, Grummon AH, Zhang F, Gortmaker SL, Moran AJ, Polacsek M, et al. Assessment of calories purchased after calorie labeling of prepared foods in a large supermarket chain. *JAMA Internal Medicine* 2022;**182**(9):965. [DOI: 10.1001/jamainternmed.2022.3065]

**van Doorn 2023** *{published data only}*

* van Doorn J, Luiting-Drijfhout M, van Ittersum K. Waste on impulse? Food ordering, calorie intake and waste in out-of-home consumption. *Journal of Business Research* 2023;**165**:114035-5. [DOI: 10.1016/j.jbusres.2023.114035]

**Zhu 2023** *{published data only}*

* Zhu C, Lopez RA, Zhen C, Zhao Q. Consumer responses to nutrition labels in China. *Journal of the Agricultural and Applied Economics Association* 2023;**2**:278-94. [DOI: 10.1002/jaa2.57]

### Additional references

**Adams 2016**

Adams J, Mytton O, White M, Monsivais P. Why are some population interventions for diet and obesity more equitable and effective than others? The role of individual agency. *PLOS Medicine* 2016;**13**(4):e1001990. [DOI: 10.1371/journal.pmed.1001990]

**Aguenaou 2021**

Aguenaou H, El Ammari L, Bigdeli M, El Hajjab A, Lahmam H, Labzizi S, et al. Comparison of appropriateness of Nutri-Score and other front-of-pack nutrition labels across a group of Moroccan consumers: awareness, understanding and food choices. *Archives of Public Health* 2021;**79**:71. [DOI: 10.1186/s13690-021-00595-3]

**Alcohol and Tobacco Trade and Tax Bureau 2021**

Alcohol and Tobacco Trade and Tax Bureau. Alcohol beverage labeling: statements related to calorie and nutrient content, serving facts, alcohol facts, and sugar content. https://www.ttb.gov/labeling-wine/alcohol-beverage-labeling (accessed prior to 5 December 2024).

**Ameye 2019**

Ameye H, Swinnen J. Obesity, income and gender: the changing global relationship. *Global Food Security* 2019;**23**:267-81. [DOI: 10.1016/j.gfs.2019.09.003]

**Anderson 2021**

Anderson P, O'Donnell A, Kaner E, Llopis EJ, Manthey J, Rehm J. Impact of minimum unit pricing on alcohol purchases in Scotland and Wales: controlled interrupted time series analyses. *Lancet. Public Health* 2021;**6**(8):E557-65. [DOI: 10.1016/S2468-2667(21)00052-9]

**Atkinson 2024**

Atkinson AM, Meadows BR, Hobin E, Vanderlee LM, Sumnall H. Qualitative analysis of UK women's attitudes to calorie-based alcohol marketing and alcohol calorie labelling. *Health Promotion International* 2024;**39**(1):daae006. [DOI: 10.1093/heapro/daae006]

**Bann 2017**

Bann D, Johnson W, Li L, Kuh D, Hardy R. Socioeconomic inequalities in body mass index across adulthood: coordinated analyses of individual participant data from three British birth cohort studies initiated in 1946, 1958 and 1970. *PLOS Medicine* 2017;**14**(1):e1002214. [DOI: 10.1371/journal.pmed.1002214]

**Basto-Abreu 2018**

Basto-Abreu A, Braverman-Bronstein A, Camacho-García-Formentí D, Zepeda-Tello R, Popkin BM, et al. Expected changes in obesity after reformulation to reduce added sugars in beverages: a modeling study. *PLOS Medicine* 2018;**15**(10):e1002664. [DOI: 10.1371/journal.pmed.1002664]

**Benton 2017**

Benton D, Young HA. Reducing calorie intake may not help you lose body weight. *Perspectives on Psychological Science* 2017;**12**(5):703-714. [DOI: 10.1177/1745691617690878]

**Bhawra 2022**

Bhawra J, Kirkpatrick SI, Hall MG, Vanderlee L, Thrasher JF, Hammond D. Correlates of self-reported and functional understanding of nutrition labels across 5 countries in the 2018 international food policy study. *Journal of Nutrition* 2022;**152**(Suppl 1):13S-24S. [DOI: 10.1093/jn/nxac018]

**Blackwell 2018**

Blackwell AK, Drax K, Attwood AS, Munafò MR, Maynard OM. Informing drinkers: can current UK alcohol labels be improved? *Drug and Alcohol Dependence* 2018;**192**:163-70. [DOI: 10.1016/j.drugalcdep.2018.07.032]

**Brimblecombe 2020**

Brimblecombe J, McMahon E, Ferguson M, De Silva K, Peeters A, Miles E, et al. Effect of restricted retail merchandising of discretionary food and beverages on population diet: a pragmatic randomised controlled trial. *Lancet Planetary Health* 2020;**4**(10):e463-e473. [DOI: 10.1016/S2542-5196(20)30202-3]

**Burris 2018**

Burris S. Taking opportunity costs seriously in public health law. *Public Health Reports* 2018;**133**(6):726-8. [DOI: 10.1177/0033354918805741]

**Campbell 2020**

Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ (Clinical Research Ed.)* 2020;**368**:l6890. [DOI: 10.1136/bmj.l6890]

**Capewell 2018**

Capewell S, Lloyd-Williams F. The role of the food industry in health: lessons from tobacco? *British Medical Bulletin* 2021;**125**(1):131-43. [DOI: 10.1093/bmb/ldy002]

**Clarke 2021a**

Clarke N, Pechey E, Kosīte D, König LM, Mantzari E, Blackwell AK, et al. Impact of health warning labels on selection and consumption of food and alcohol products: systematic review with meta-analysis. *Health Psychology Review* 2021;**15**(3):430-53. [DOI: 10.1080/17437199.2020.1780147]

**Clarke 2021c**

Clarke N, Blackwell AK, De-Loyde K, Pechey E, Hobson A, Pilling M, et al. Health warning labels and alcohol selection: a randomised controlled experiment in a naturalistic shopping laboratory. *Addiction (Abingdon, England)* 2021;**116**(12):3333-45. [DOI: 10.1111/add.15519]

**Clarke 2023b**

Clarke N, Blackwell AK, Ferrar J, De-loyde K, Pilling MA, Munafo MR, et al. Impact on alcohol selection and online purchasing of changing the proportion of available non-alcoholic versus alcoholic drinks: a randomised controlled trial. *PLOS Medicine* 2023;**20**(3):e1004193. [DOI: 10.1371/journal.pmed.1004193]

**Cook 2000**

Cook A, Pryer J, Prakash S. The problem of accuracy in dietary surveys. Analysis of the over 65 UK National Diet and Nutrition Survey. *Journal of Epidemiology and Community Health* 2000;**54**:611-6.

**Cowburn 2005**

Cowburn G, Stockley L. Consumer understanding and use of nutrition labelling: a systematic review. *Public Health Nutrition* 2005;**8**(1):21-8.

**Crockett 2018**

Crockett RA, King SE, Marteau TM, Prevost AT, Stubbs B, Bignardi G, et al. Nutritional labelling for healthier food or non-alcoholic drink purchasing and consumption. *Cochrane Database of Systematic Reviews* 2018, Issue 2. Art. No: CD009315. [DOI: 10.1002/14651858.CD009315.pub2]

**Croker 2020**

Croker H, Packer J, Russell S, Stansfield C, Viner RM. Front of pack nutritional labelling schemes: a systematic review and meta-analysis of recent evidence relating to objectively measured consumption and purchasing. *Journal of Human Nutrition and Dietetics* 2020;**33**:518-37. [DOI: 10.1111/jhn.12758]

**Cuello-Garcia 2022**

Cuello-Garcia CA, Santesso N, Morgan RL, Verbeek J, Thayer K, Ansari MT, et al. GRADE guidance 24 optimizing the integration of randomized and non-randomized studies of interventions in evidence syntheses and health guidelines. *Journal of Clinical Epidemiology* 2022;**142**:200-8. [DOI: 10.1016/j.jclinepi.2021.11.026]

**Daley 2020**

Daley AJ, McGee E, Bayliss S, Coombe A, Parretti HM. Effects of physical activity calorie equivalent food labelling to reduce food selection and consumption: systematic review and meta-analysis of randomised controlled studies. *Journal of Epidemiology and Community Health* 2020;**74**(3):269-75. [DOI: 10.1136/jech-2019-213216]

**Das 2012**

Das P, Horton R. Rethinking our approach to physical activity. *Lancet* 2012;**80**(9838):189-90. [DOI: 10.1016/S0140-6736(12)61024-1]

**De-loyde 2024**

De-loyde K, Ferrar J, Pilling MA, Hollands GJ, Clarke N, Matthews J, et al. The impact of introducing alcohol-free beer options in bars and public houses on alcohol sales and revenue: a randomised crossover field trial. *Addiction (Abingdon, England)* 2024;**119**(6):1071-9. [DOI: 10.1111/add.16449]

**Deeks 2023**

Deeks JJ, Higgins JP, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (updated August 2023). Cochrane, 2023. Available from training.cochrane.org/handbook.

**DeGroot 2012**

DeGroot MH, Schervish MJ. Probability and Statistics. 4th edition. Boston (MA): Addison-Wesley, 2012.

**Department of Health 2011**

Department of Health. Statement of the Calorie Reduction Expert Group (policy paper). http://www.gov.uk/government/publications/statement-of-the-calorie-reduction-expert-group. (accessed prior to 5 December 2024).

**Department of Health and Social Care 2018a**

Department of Health and Social Care. Mandating energy labelling of food and drink in out-of-home settings (Impact Assessment IA No: 13009 (Consultation stage)). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/751532/impact-assessment-for-consultation-on-calorie-labelling-outside-of-the-home.pdf (accessed prior to 5 December 2024).

**Department of Health and Social Care 2020**

Department of Health and Social Care. Tackling obesity: empowering adults and children to live healthier lives. http://www.gov.uk/government/publications/tackling-obesity-government-strategy/tackling-obesity-empowering-adults-and-children-to-live-healthier-lives (accessed prior to 8 May 2021).

**Department of Health and Social Care 2021**

Department of Health and Social Care. Calorie labelling in the out of home sector: implementation guidance. https://www.gov.uk/government/publications/calorie-labelling-in-the-out-of-home-sector/calorie-labelling-in-the-out-of-home-sector-implementation-guidance (accessed prior to 5 December 2024).

**Department of Health and Social Care 2022**

New obesity treatments and technology to save the NHS billions. https://www.gov.uk/government/news/new-obesity-treatments-and-technology-to-save-the-nhs-billions (accessed prior to 5 December 2024).

**Department of Health and Social Care 2018b**

Department of Health and Social Care. Childhood obesity: a plan for action, Chapter 2; June 2018. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/718903/childhood-obesity-a-plan-for-action-chapter-2.pdf (accessed prior to 5 December 2024).

**Dobbs 2014**

Dobbs R, Sawers C, Thompson F, Manyika J, Woetzel J, Child P, et al. Overcoming obesity: an initial economic analysis. McKinsey Global Institute; 1 November 2014. http://www.mckinsey.com/insights/economic studies/how the world could better fight obesity (accessed prior to 8 May 2021).

**Duffy 2023**

Duffy F, Peebles I, Maloney E, Robertson MD, Sharpe H. Individuals with restrictive eating disorders' experience of the introduction of calories on menus in England: an interpretative phenomenological analysis study. *European Eating Disorders Review* 2023;**31**(6):802-12. [DOI: 10.1002/erv.3007]

**Dumanovsky 2011**

Dumanovsky T, Huang CY, Nonas CA, Matte TD, Bassett MT, Silver LD, et al. Changes in energy content of lunchtime purchases from fast food restaurants after introduction of calorie labelling: cross sectional customer surveys. *BMJ (Clinical Research Ed.)* 2011;**343**:d4464.

**Eldridge 2020**

Eldridge S, Campbell MK, Campbell MJ, Drahota AK, Giraudeau B, Reeves BC, et al. Revised Cochrane risk of bias tool for randomized trials (RoB 2). Additional considerations for cluster-randomized trials (RoB 2 CRT); 10 November 2020. http://training.cochrane.org/sites/training.cochrane.org/files/public/uploads/Sandra%20Eldridge_Risk%20of%20Bias%20Tool%202%20for%20Crossover%20Trials.pdf (accessed prior to 5 December 2024).

**EPOC 2017**

Cochrane Effective Practice and Organisation of Care (EPOC). Analysis in EPOC reviews. EPOC resources for review authors; 13 March 2019. http://www.epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/Resources-for-authors2017/analysis_in_epoc_reviews.pdf (accessed prior to 5 December 2024).

**European Food Information Council 2018**

European Food Information Council (EUFIC). Global Update on Nutrition Labelling; the 2018 edition. http://www.eufic.org/en/healthy-living/article/global-update-on-nutrition-labelling.

**European Union 2011**

European Union. Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011. Official Journal of the European Union L304 of November 2011:18-63.

**Fanzo 2021**

Fanzo J, Bellows AL, Spiker ML, Thorne-Lyman AL, Bloem MW. The importance of food systems and the environment for nutrition. *American Journal of Clinical Nutrition* 2021;**113**:7-16. [DOI: 10.1093/ajcn/nqaa313]

**Feteira-Santos 2020**

Feteira-Santos R, Fernandes J, Virgolino A, Alarcão V, Sena C, Vieira CP, et al. Effectiveness of interpretive front-of-pack nutritional labelling schemes on the promotion of healthier food choices: a systematic review. *International Journal of Evidence-Based Healthcare* 2020;**18**(1):24-37. [DOI: 10.1097/XEB.0000000000000214]

**Food and Drug Administration 2014**

Food and Drug Administration. Food labeling; nutrition labeling of standard menu items in restaurants and similar retail food establishments. https://www.federalregister.gov/documents/2014/12/01/2014-27833/food-labeling-nutrition-labeling-of-standard-menu-items-in-restaurants-and-similar-retail-food (accessed prior to 5 December 2024).

**Food and Drug Administration 2016**

Food and Drug Administration (FDA). Guide to creating a front of pack (FoP) nutrition label for pre-packaged products sold through retail outlets; updated November 2016. http://www.food.gov.uk/sites/default/files/media/document/fop-guidance_0.pdf (accessed prior to 5 December 2024).

**Food Standards Agency 2021**

Food Standards Agency. Guidance on voluntary energy labelling for out of home businesses in Northern Ireland. https://www.food.gov.uk/sites/default/files/media/document/calorie-wise-technical-guidance_july2021.pdf (accessed prior to 5 December 2024).

**Gardner 2022**

Gardner B, Lally P. Habit and habitual behaviour. *Health Psychology Review* 2022;**17**(3):490-6. [DOI: 10.1080/17437199.2022.2105249]

**Glanz 2005**

Glanz K, Sallis JF, Saelens BE, Frank LD. Healthy nutrition environments: concepts and measures. *American Journal of Health Promotion* 2005;**19**(5):330-3. [DOI: 10.4278/0890-1171-19.5.330]

**Government of Ontario 2023**

Government of Ontario. Calorie display requirements. https://www.ontario.ca/document/guide-menu-labelling-requirements/calorie-display-requirements (accessed prior to 5 December 2024).

**Grummon 2020**

Grummon AH, Hall MG. Sugary drink warnings: a meta-analysis of experimental studies. *PLOS Medicine* 2020;**17**(5):e1003120.

**Grummon 2021**

Grummon A, Petimar J, Zhang F, Rao A, Gortmaker SL, Rimm EB, et al. Calorie labeling and product reformulation: a longitudinal analysis of supermarket prepared foods. *American Journal of Preventive Medicine* 2021;**61**(3):377-85. [DOI: 10.1016/j.amepre.2021.03.013]

**Guyatt 2011**

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction – GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2011;**64**(4):383-94.

**Health Survey for England 2018**

Health Survey for England (HSE). Health Survey for England 2018. Overweight and obesity in adults and children; 3 December 2019. http://files.digital.nhs.uk/52/FD7E18/HSE18-Adult-Child-Obesity-rep.pdf (accessed prior to 5 December 2024).

**Higgins 2021b**

Higgins JP, Eldridge S, Li T, Sterne J (editors), on behalf of the RoB 2 working group on crossover trials. Revised Cochrane risk of bias tool for randomized trials (RoB 2). Additional considerations for crossover trials; 18 March 2021. http://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/rob-2-for-crossover-trials (accessed prior to 5 December 2024).

**Higgins 2023**

Higgins JP, Eldridge S, Li T. Chapter 23: Including variants on randomized trials. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (updated August 2023). Cochrane, 2023. Available from training.cochrane.org/handbook.

**Hollands 2012**

Hollands GJ, Whitwell SC, Parker RA, Prescott NJ, Forbes A, Sanderson J, et al. Effect of communicating DNA based risk assessments for Crohn's disease on smoking cessation: randomised controlled trial. *BMJ (Clinical Research Ed.)* 2012;**345**:e4708. [DOI: 10.1136/bmj.e4708]

**Hollands 2013**

Hollands GJ, Shemilt I, Marteau TM, Jebb SA, Kelly MP, Nakamura R, et al. Altering choice architecture to change population health behaviour: a large-scale conceptual and empirical scoping review of interventions within micro-environments. https://www.repository.cam.ac.uk/handle/1810/245108 (accessed prior to 5 December 2024).

**Hollands 2016**

Hollands GJ, Marteau TM, Fletcher PC. Nonconscious processes in changing health-related behaviour: a conceptual analysis and framework. *Health Psychology Review* 2016;**10**(4):381-94. [DOI: 10.1080/17437199.2015.1138093]

**Hollands 2017**

Hollands GJ, Bignardi G, Johnston M, Kelly MP, Ogilvie D, Petticrew M, et al. The TIPPME intervention typology for changing environments to change behaviour. *Nature Human Behaviour* 2017;**1**:0140.

**Hollands 2019**

Hollands GJ, Carter P, Anwer S, King SE, Jebb SA, Ogilvie D, et al. Altering the availability or proximity of food, alcohol, and tobacco products to change their selection and consumption. *Cochrane Database of Systematic Reviews* 2019, Issue 8. Art. No: CD012573. [DOI: 10.1002/14651858.CD012573.pub2]

**Hollands 2024**

Hollands GJ, South E, Shemilt I, Oliver S, Thomas J, Sowden AJ. Methods used to conceptualise dimensions of health equity impacts of public health interventions in systematic reviews. Journal of Clinical Epidemiology 2024;**169**:111312. [DOI: 10.1016/j.jclinepi.2024.111312]

**Kerins 2020**

Kerins C, McHugh S, McSharry J, Reardon CM, Hayes C, Perry IJ, et al. Barriers and facilitators to implementation of menu labelling interventions from a food service industry perspective: a mixed methods systematic review. *International Journal of Behavioral Nutrition and Physical Activity* 2020;**17**:49. [DOI: 10.1186/s12966-020-00948-1]

**Lacy-Nichols 2022**

Lacy-Nichols J, Marten R, Crosbie E, Moodie R. The public health playbook: ideas for challenging the corporate playbook. *Lancet Global Health* 2022;**10**(7):E1067-72. [DOI: 10.1016/S2214-109X(22)00185-1]

**Lawson 2018**

Lawson GM, Hook CJ, Farah MJ. A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Developmental Science* 2018;**21**:e12529. [DOI: 10.1111/desc.12529]

**Lefebvre 2023**

Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Metzendorf M-I, et al. Technical Supplement to Chapter 4: Searching for and selecting studies. Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (updated August 2023). Cochrane, 2023. Available from training.cochrane.org/handbook.

**Lorenc 2013**

Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. *Journal of Epidemiology and Community Health* 2013;**67**(2):190-3. [DOI: 10.1136/jech-2012-201257]

**Marteau 2012**

Marteau TM, Hollands GJ, Fletcher PC. Changing human behaviour to prevent disease: the importance of targeting automatic processes. *Science* 2012;**337**(6101):1492-5.

**Marteau 2022**

Marteau TM, Hollands GJ, Pechey R, Reynolds JP, Jebb SA. Changing the assortment of available food and drink for leaner, greener diets. *BMJ (Clinical Research Ed.)* 2022;**377**:e069848. [DOI: 10.1136/bmj-2021-069848]

**Marteau 2023a**

Marteau TM, Hollands GJ, Nambiar D, Munafò MR. Changing behaviour at scale to prevent NCDs. In: Noncommunicable Diseases: a Compendium. Abingdon (UK): Routledge, 2023:347-54. [DOI: 10.4324/9781003306689-52]

**Marteau 2023b**

Marteau TM. Evidence-neglect: addressing a barrier to UK health and climate policy ambitions. *Science & Public Policy* 2023;**50**(6):1103-9. [DOI: 10.1093/scipol/scad021]

**McGuinness 2021**

McGuinness LA, Higgins JP. Risk-of-bias VISualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods* 2021;**12**(1):55-61. [DOI: 10.1002/jrsm.1411]

**McKenzie 2023**

McKenzie JE, Brennan SE. Chapter 12: Synthesizing and presenting findings using other methods. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (updated August 2023). Cochrane, 2023. Available from training.cochrane.org/handbook.

**Muller 2016**

Muller L, Prevost M. What cognitive sciences have to say about the impacts of nutritional labelling formats. *Journal of Economic Psychology* 2016;**55**:17-29. [DOI: 10.1016/j.joep.2016.01.005]

**Mytton 2020**

Mytton OT, Boyland E, Adams J, Collins B, O'Connell M, Russell SJ, et al. The potential health impact of restricting less-healthy food and beverage advertising on UK television between 05.30 and 21.00 hours: a modelling study. *PLOS Medicine* 2020;**17**(10):e1003212. [DOI: 10.1371/journal.pmed.1003212]

**NCD Countdown 2020**

NCD Countdown 2030 collaborators. NCD Countdown 2030: pathways to achieving Sustainable Development Goal target 3.4. *Lancet* 2020;**396**(10255):918-34. [DOI: 10.1016/S0140-6736(20)31761-X]

**NHS Digital 2019**

NHS Digital. National Child Measurement Programme, England 2018/19 School Year [NS]; October 2019. http://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme/2018-19-school-year (accessed prior to 8 May 2021).

**Obesity Evidence Hub 2021**

Obesity Evidence Hub. Kilojoule labelling in fast food outlets. https://www.obesityevidencehub.org.au/collections/prevention/kilojoule-labelling-in-fast-food-outlets (accessed 6 January 2025).

**Page 2021**

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)* 2021;**372**:n71. [DOI: 10.1136/bmj.n71]

**Polden 2023**

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**:1088. [DOI: 10.1186/s12889-023-16033-8]

**Public Health England 2018**

Public Health England. Calorie reduction: the scope and ambition for action; March 2018. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/800675/Calories_Evidence_Document.pdf (accessed prior to 5 December 2024).

**Public Health England 2020**

Public Health England. NDNS: results from years 9 to 11 (2016 to 2017 and 2018 to 2019). https://www.gov.uk/government/statistics/ndns-results-from-years-9-to-11-2016-to-2017-and-2018-to-2019 (accessed prior to 5 December 2024).

**Putra 2023**

Putra I, Polden M, Wareing L, Robinson E. Acceptability and perceived harm of calorie labelling and other obesity policies: a cross-sectional survey study of UK adults with eating disorders and other mental health conditions. PsyArXiv 2023. [DOI: 10.31234/osf.io/cxaw7]

**Qureshi 2022**

Qureshi R, Mayo-Wilson E, Li T. Harms in Systematic Reviews Paper 1: an introduction to research on harms. *Journal* *of Clinical Epidemiology* 2022;**143**:186-96. [DOI: 10.1016/j.jclinepi.2021.10.023]

**Reeves 2020**

Reeves B, Deeks J, Higgins JP, Shea B, Tugwell P, Wells GA, on behalf of the Cochrane Non-Randomized Studies of Interventions Methods Group. Chapter 24: Including non-randomized studies on intervention effects. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from training.cochrane.org/handbook/archive/v6.1.

**Rehm 2018**

Rehm J, Guiraud J, Poulnais R, Shield KD. Alcohol dependence and very high risk level of alcohol consumption: a life-threatening and debilitating disease. *Addiction Biology* 2018;**23**(4):961-8. [DOI: 10.1111/adb.12646]

**RevMan 2024 [Computer program]**

Review Manager (RevMan). Version 7.12.0. Copenhagen: The Cochrane Collaboration, 2024. Available at https://revman.cochrane.org.

**Reynolds 2021**

Reynolds JP, Ventsel M, Kosīte D, Dames BR, Brocklebank L, Masterson S, et al. Impact of decreasing the proportion of higher energy foods and reducing portion sizes on food purchased in worksite cafeterias: a stepped-wedge randomised controlled trial. *PLOS Medicine* 2021;**18**(9):e1003743. [DOI: 10.1371/journal.pmed.1003743]

**Robinson 2019**

Robinson E, Burton S, Gough T, Jones A, Haynes A. Point of choice kilocalorie labelling in the UK eating out of home sector: a descriptive study of major chains. *BMC Public Health* 2019;**19**(1):649. [DOI: 10.1186/s12889-019-7017-5]

**Robinson 2021a**

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 (Clinical Research Ed.)* 2021;**372**:n40. [DOI: 10.1136/bmj.n40]

**Robinson 2021b**

Robinson E, Humphreys G, Jones A. Alcohol, calories, and obesity: a rapid systematic review and meta-analysis of consumer knowledge, support, and behavioral effects of energy labeling on alcoholic drinks. *Obesity Reviews* 2021;**22**:e13198. [DOI: 10.1111/obr.13198]

**Robinson 2023a**

Robinson E, Boyland E, Evans R, Finlay A, Halsall L, Humphreys G, et al. Energy labelling of alcoholic drinks: an important or inconsequential obesity policy? *Obesity Science & Practice* 2023;**9**(2):75-86. [DOI: 10.1002/osp4.638]

**Robinson 2023b**

Robinson E, Polden M, Langfield T, Clarke K, Calvert L, Colombet Z, et al. Socioeconomic position and the effect of energy labelling on consumer behaviour: a systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity* 2023;**20**:10. [DOI: 10.1186/s12966-023-01418-0]

**Royal Society for Public Health 2014**

Royal Society for Public Health. Increasing awareness of 'invisible' calories from alcohol; November 2014. http://www.rsph.org.uk/static/uploaded/979245d2-7b5d-4693-a9b3fb1b98b68d76.pdf (accessed prior to 5 December 2024).

**Royal Society for Public Health 2018**

Royal Society for Public Health. Labelling the point: towards better alcohol health information; January 2018. https://www.rsph.org.uk/static/uploaded/4ae31b49-c4d7-4355-ad94a660aba36108.pdf (accessed prior to 5 December 2024).

**Sarink 2016**

Sarink D, Peeters A, Freak-Poli R, Beauchamp A, Woods J, Ball K, et al. The impact of menu energy labelling across socioeconomic groups: a systematic review. *Appetite* 2016;**99**:59-75. [DOI: 10.1016/j.appet.2015.12.022]

**Schünemann 2021**

Schünemann HJ, Higgins JP, Vist GE, Glasziou P, Akl EA, Skoetz N, Guyatt GH. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2. Cochrane, 2021. Available from training.cochrane.org/handbook.

**Shangguan 2019**

Shangguan S, Afshin A, Shulkin M, Ma W, Marsden D, Smith J, et al. A meta-analysis of food labeling effects on consumer diet behaviors and industry practices. *American Journal of Preventive Medicine* 2019;**56**(2):300-14. [DOI: 10.1016/j.amepre.2018.09.024]

**Shaw 2023**

Shaw S, Barrett M, Shand C, Cooper C, Crozier S, Smith D, et al. Influences of the community and consumer nutrition environment on the food purchases and dietary behaviors of adolescents: a systematic review. *Obesity Reviews* 2023;**24**(7):e13569. [DOI: 10.1111/obr.13569]

**Sheeran 2016**

Sheeran P, Webb TL. The intention-behavior gap. *Social and Personality Psychology Compass* 2016;**10**(9):503-18. [DOI: 10.1111/spc3.12265]

**Shelton 2014**

Shelton NJ, Knott CS. Association between alcohol calorie intake and overweight and obesity in English adults. *American Journal of Public Health* 2014;**104**(4):629-31. [DOI: 10.2105/AJPH.2013.301643]

**Shemilt 2013**

Shemilt I, Hollands GJ, Marteau TM, Nakamura R, Jebb SA, Kelly MP, et al. Economic instruments for population diet and physical activity behaviour change: a systematic scoping review. *PLOS One* 2013;**8**(9):e75070. [DOI: 10.1371/journal.pone.0075070]

**Sheron 2016**

Sheron N, Gilmore I. Effect of policy, economics, and the changing alcohol marketplace on alcohol related deaths in England and Wales. *BMJ (Clinical Research Ed.)* 2016;**353**:i1860. [DOI: 10.1136/bmj.i1860]

**Song 2021**

Song J, Brown MK, Tan M, MacGregor GA, Webster J, Campbell NR, et al. Impact of color-coded and warning nutrition labelling schemes: a systematic review and network meta-analysis. *PLOS Medicine* 2021;**18**(10):e1003765. [DOI: 10.1371/journal.pmed.1003765]

**Statistica 2019**

Statistica. UK: population, by age 2019. http://www.statista.com/statistics/281174/uk-population-by-age/ (accessed prior to 8 May 2021).

**Sterne 2016a**

Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. *BMJ (Clinical Research Ed.)* 2016;**355**:i4919.

**Sterne 2016b**

Sterne JA, Higgins JP, Elbers RG, Reeves BC, the development group for ROBINS-I. Risk of Bias In Non-randomized Studies of Interventions (ROBINS-I): detailed guidance; 2016. http://www.riskofbias.info (accessed April 2021).

**Sterne 2019**

Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ (Clinical Research Ed.)* 2019;**366**:i4898. [DOI: 10.1136/bmj.l4898]

**Sterne 2020**

Sterne JA, Hernán MA, McAleenan A, Reeves BC, Higgins JP. Chapter 25: Assessing risk of bias in a non-randomized study. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from training.cochrane.org/handbook/archive/v6.1.

**Swinburn 1999**

Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive Medicine* 1999;**29**(6):563-70. [DOI: 10.1006/pmed.1999.0585]

**Thomas 2020 [Computer program]**

EPPI-Reviewer: advanced software for systematic reviews, maps and other evidence synthesis. Thomas J, Graziosi S, Brunton J, Ghouze Z, O'Driscoll P, Bond M. London: UCL Social Research Institute: EPPI-Centre, 2020.

**Tremmel 2017**

Tremmel M, Gerdtham U-G, Nilsson PM, Saha S. Economic burden of obesity: a systematic literature review. *International Journal of Environmental Research and Public Health* 2017;**14**(4):435. [DOI: 10.3390/ijerph14040435]

**Urban 2011**

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. [DOI: 10.1001/jama.2011.993]

**Ventsel 2022**

Ventsel M, Pechey E, De-Loyde K, Pilling MA, Morris RW, Maistrello G, et al. Effect of health warning labels on motivation towards energy-dense snack foods: two experimental studies. *Appetite* 2022;**175**:106084. [DOI: 10.1016/j.appet.2022.106084]

**Waterlander 2018**

Waterlander WE, Ni Mhurchu C, Eyles H, Vandevijvere S, Cleghorn C, Scarborough P, et al. Food futures: developing effective food systems interventions to improve public health nutrition. *Agricultural Systems* 2018;**160**:124-31. [DOI: 10.1016/j.agsy.2017.01.006]

**Welch 2012**

Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, et al. PRISMA-Equity 2012 Extension: reporting guidelines for systematic reviews with a focus on health equity. *PLOS Medicine* 2012;**9**(10):e1001333. [DOI: 10.1371/journal.pmed.1001333]

**Welch 2016**

Welch V, Petticrew M, Petkovic J, Moher D, Waters E, White H, Tugwell P, PRISMA-Equity Bellagio group. Extending the PRISMA statement to equity-focused systematic reviews (PRISMA-E 2012): explanation and elaboration. *Journal of Clinical Epidemiology* 2016;**70**:68-89. [DOI: 10.1016/j.jclinepi.2015.09.001]

**Winkleby 1992**

Winkleby MA, Jatulis DE, Frank E, Fortmann SP. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *American Journal of Public Health* 1992;**82**(6):816-20.

**World Cancer Research Fund 2021**

World Cancer Research Fund. NOURISHING framework. http://www.wcrf.org/int/policy/policy-databases/nourishing-framework (accessed prior to 8 May 2021).

**World Health Organization 2004**

World Health Organization. Global strategy, on diet, physical activity and health. http://www.who.int/publications/i/item/9241592222 (accessed prior to 8 May 2021).

**World Health Organization 2017**

World Health Organization. Alcohol labelling: a discussion document on policy options; 2017. http://iris.who.int/handle/10665/350744 (accessed prior to 5 December 2024).

**World Health Organization 2018**

World Health Organization. Global status report on alcohol and health. http://www.who.int/publications/i/item/9789241565639 (accessed prior to 8 May 2021).

**World Health Organization 2020**

World Health Organization. Obesity. www.who.int/news-room/facts-in-pictures/detail/6-facts-on-obesity (accessed prior to 8 May 2021).

**Yau 2022**

Yau A, Berger N, Law C, Cornelsen L, Greener R, et al. Changes in household food and drink purchases following restrictions on the advertisement of high fat, salt, and sugar products across the Transport for London network: a controlled interrupted time series analysis. *PLOS Medicine* 2022;**19**(2):e1003915. [DOI: 10.1371/journal.pmed.1003915]

**Zahedi 2023**

Zahedi A, Oroz Artigas S, Swaboda N, Wiers CE, Görgen K, Park SQ. Neural correlates of changing food choices while bypassing values. *NeuroImage* 2023;**274**:120134. [DOI: 10.1016/j.neuroimage.2023.120134]

**Zlatevska 2018**

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. [DOI: 10.1016/j.jretai.2017.09.007]

### References to other published versions of this review

**Clarke 2021b**

Clarke N, Marteau TM, Pilling M, Roberts NW, Jebb SA, Hollands GJ. Energy (calorie) labelling for healthier selection and consumption of food or alcohol. *Cochrane Database of Systematic Reviews* 2021, Issue 6. Art. No: CD014845. [DOI: 10.1002/14651858.CD014845]

* Indicates the major publication for the study

## CHARACTERISTICS OF STUDIES

### Characteristics of included studies *[ordered by study ID]*

**Berry 2019**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-participants quasi-randomised controlled trial<br>Setting: restaurant in South USA<br>Setting type: field setting<br>Recruitment: patrons of the restaurant were asked if they would like to participate in the study immediately after placing their order<br>Allocation to group: calorie labelling on menus or absence of, with menus altered by day of week |
| Participants | Number of enrolled participants: 233<br>Number (%) of enrolled participants completing the study and included in the analysis: 233<br>Study completers: mean age: 47.4 (SD not reported) years<br>Study completers: sex: 67% female<br>Specific social or cultural characteristics: general public, aged > 18 years<br>Socioeconomic status context: low deprivation<br>Ethnicity: not reported<br>BMI: not reported |
| Interventions | Labelling intervention: calorie labelling on menus<br>Control: no calorie labelling on menus<br>Manipulated product type: food<br>Characteristics of manipulated products: all food on menus (in entrées, sides, and drinks) |

**Berry 2019** *(Continued)*

| | |
|---|---|
| | Duration of exposure to intervention: length of 2 meals<br><br>Number of eligible comparisons: 1<br><br>Comparisons included in the analysis: calorie labelling versus no calorie labelling<br><br>Study arms: calorie labelling; no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate |
| Funding source | None reported |
| Notes | We contacted the authors to request numbers per group as data were not reported in the report. There was no response so it was not possible to include the study results in the meta-analysis. |

**Bollinger 2011**

| | |
|---|---|
| ***Study characteristics*** | |
| Methods | Design: controlled before-after study<br><br>Setting: 222 Starbucks coffee shops in New York City (intervention sites), 94 in Boston and Philadelphia (control sites), USA<br><br>Setting type: field setting<br><br>Recruitment: convenience sample of Starbucks customers<br><br>Allocation to group: naturally occurring New York City legislation with Boston and Philadelphia as comparisons |
| Participants | Number of enrolled participants: 222 intervention sites, 94 control sites<br><br>Number (%) of enrolled participants completing the study and included in analysis: 222 intervention sites, 94 control sites (100%)<br><br>Study completers: mean age: not reported<br><br>Study completers: female %: not reported<br><br>Specific social or cultural characteristics: Starbucks customers<br><br>Socioeconomic status context: both high and low<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Labelling intervention: calorie labelling on menus<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: Starbucks menu items<br><br>Duration of exposure to intervention: data from 3 months before and 11 months after calorie labelling<br><br>Study arms: calorie labelling; no calorie labelling |

**Bollinger 2011** *(Continued)*

| | |
|---|---|
| | Number of eligible comparisons: 1<br>Comparisons included in the analysis: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br>Selection outcome analysed: mean energy per sales transaction<br>Timing of selection outcome measurement: immediate |
| Funding source | Not reported |
| Notes | |

**Cawley 2020**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-participants cluster-randomised controlled trial<br>Setting: 2 sit-down restaurants on university campus, USA<br>Setting type: field setting<br>Allocation to group: calorie labelling on menus or absence of, tables randomised on arrival at restaurant by smart phone app and given corresponding menu. Participants on a return visit to the restaurant since the study began were excluded. |
| Participants | Number of enrolled participants (reported as transactions): 8317<br>Number (%) of enrolled participants completing the study (reported as transactions) and included in the analysis: 5551 (66.7%)<br>Study completers: mean age: 34.185 (SD 17.7) years<br>Study completers: sex: 57% female<br>Specific social or cultural characteristics: general public, aged > 18 years. High proportion of students (37.8%)<br>Socioeconomic status context: low deprivation<br>Ethnicity: 65% white, 24% Asian, 1% Hispanic<br>BMI: 7% of the sample had a BMI > 30 kg/m$^2$ and were classified as obese |
| Interventions | Type of labelling intervention: calorie labelling on menus<br>Manipulated product type: food and alcohol<br>Characteristics of manipulated products: all food and alcohol on menus (in appetisers, entrées, desserts, and drinks)<br>Duration of exposure to intervention: length of 1 meal<br>Number of eligible comparisons: 1<br>Comparisons included in the analysis: calorie labelling versus no calorie labelling<br>Study arms: calorie labelling; no calorie labelling |

**Cawley 2020** *(Continued)*

| | |
|---|---|
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate |
| Funding source | The Institute for the Social Sciences, the Institute for Healthy Futures, the Building Faculty Connections Program, and the College of Human Ecology at Cornell University |
| Notes | Contacted authors and received results in format suitable for meta-analyses, from a model that accounts for clustering, and for results reported separately by alcohol and food. |

**Clarke 2023a**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-participants randomised controlled trial<br><br>Setting: online supermarket in the UK<br><br>Setting type: naturalistic laboratory<br><br>Recruitment: participants recruited via a market research agency<br><br>Allocation to group: randomised using an online algorithm within Qualtrics |
| Participants | Number of enrolled participants: 651<br><br>Number (%) of enrolled participants completing the study and included in the analysis: 608<br><br>Study completers: mean age: 35.5 (SD 10.8) years<br><br>Study completers: sex: 55% female<br><br>Specific social or cultural characteristics: general public, aged > 18 years<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: 79% white; 5% Black or African American; 11% Asian; 5% mixed race/ethnicity<br><br>BMI: mean 26.41 $kg/m^2$ |
| Interventions | Labelling intervention: calorie labelling on products<br><br>Control: no calorie information<br><br>Manipulated product type: food (non-alcoholic drinks) and alcohol<br><br>Characteristics of manipulated products: 64 drink options (32 alcoholic, 32 non-alcoholic)<br><br>Duration of exposure to intervention: length of 1 shop<br><br>Study arms: calorie labelling; no calorie labelling; calorie labelling with image-and-text health warning label; calorie labelling with text-only health warning label; image-and-text health warning label only; text-only health warning label only<br><br>Number of eligible comparisons: 1<br><br>Comparisons included in the analysis: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with the intention to purchase |

**Clarke 2023a** *(Continued)*

| | |
|---|---|
| | Selection outcome analysed: energy content (kcal) selected<br><br>Timing of selection outcome measurement: immediate |
| Funding source | The Wellcome Trust |
| Notes | Health warning labels and calorie labels examined in factorial design. Only the cleanest comparison of simple calorie labelling versus no calorie labelling extracted, due to some evidence of a potential interaction between health warning label and calorie information |

**Dubois 2021**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-participants cluster-randomised controlled trial<br><br>Setting: 60 supermarket stores belonging to 3 of the largest retail chains in France (40 intervention stores (10 for each of 4 labelling types) and 20 control stores)<br><br>Setting type: field setting<br><br>Recruitment: intervention was authorised by ministerial decree<br><br>Allocation to group: allocation by supermarket store |
| Participants | Number of enrolled participants: 60 stores (20 control stores, 40 intervention stores (10 for each of 4 labelling types))<br><br>Number (%) of enrolled participants completing the study and included in the analysis: 60 stores<br><br>Study completers: mean age: not reported<br><br>Study completers: sex: not reported<br><br>Specific social or cultural characteristics: general public, customers of these supermarkets<br><br>Socioeconomic status context: a mix of stores from under and lower privileged geographic areas to ensure shoppers from higher and lower socioeconomic status<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Type of labelling intervention: calorie labelling with ≥ 1 other nutrient added as stickers to food products in the supermarket stores<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: all processed or ultraprocessed foods (fresh prepared foods (e.g. pizzas, quiches), pastries (e.g. croissants, brioches), breads (e.g. sliced breads, baguettes), and canned prepared meals (e.g. cooked beans, ravioli))<br><br>Duration of exposure to intervention: 5 weeks<br><br>Study arms: 4 energy labels (SENS, Nutri-Score, NutriRepère, Nutri-Couleurs) versus original packaging including back-of-pack calorie labelling. NutriRepère and Nutri-couleurs labels included energy information<br><br>Number of eligible comparisons: 2 |

**Dubois 2021** *(Continued)*

| | |
|---|---|
| | Comparisons included in the analysis: NutriRepère energy label versus current back of pack and Nutri-Couleurs energy label versus current back of pack |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: energy content (kcal) purchased<br><br>Timing of selection outcome measurement: immediate |
| Funding source | Funding provided by CNAMTS (Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés), FFAS (Fonds Français pour l'Alimentation et la Santé), and the Ministry of Social Affairs and Health. |
| Notes | Contacted authors and received results in a format suitable for meta-analysis, from a model that accounts for clustering.<br><br>The 2 label types with energy information were included as separate comparisons (NutriRepere and NutriCouleur) as this was the format of the results that were obtained from study authors. We could not apply our standard approach to reweighting multiple comparisons from a single study of proportionately reducing the sample size for the control group as these data were not available. Therefore, standard errors for the 2 comparisons from this study were adjusted conservatively to ensure appropriate weighting for the contribution of the study as a whole. We ensured that both comparisons combined gave the same total weight in the meta-analysis that would be contributed should only the single comparison providing the least weight have been entered in the standard (unadjusted) manner. When no such adjustment was performed, and so both comparisons were entered as if they were independent, results did not meaningfully differ. In the forest plots that this study was included in, the n values indicated a total of 20 control clusters (shared between the 2 rows representing 2 comparisons, i.e. 10 per row) and 10 intervention stores per row (to indicate 2 different intervention groups each contributing to 1 comparison). |

**Elbel 2009**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: controlled before-after study<br><br>Setting: 19 fast-food restaurants, belonging to fast-food chains (Wendy's, McDonalds, Burger King, and KFC) located in New York City and Newark. New York City selected as it is the first site in the country to introduce calorie labelling. Newark was selected as a control city<br><br>Setting type: field setting<br><br>Recruitment: all possible customers approached either before or after they entered the restaurant and asked for their receipts<br><br>Allocation to group: naturally allocated based on legislation |
| Participants | Number of enrolled participants: 1156<br><br>Number (%) of enrolled participants completing the study and included in analysis: 1156<br><br>Study completers: mean age: 38 years<br><br>Study completers: female (%): 62%<br><br>Specific social or cultural characteristics: customers of the fast-food restaurants<br><br>Socioeconomic status context: high (targeted restaurants within lower income demographic areas)<br><br>Ethnicity: 65.7% Black/African American; 19.9% Hispanic |

**Elbel 2009** *(Continued)*

| | |
|---|---|
| | BMI: not reported |
| Interventions | Labelling intervention: calorie labelling on menus<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: all regularly available fast food and drinks sold at popular fast-food chains<br><br>Duration of exposure to intervention: data from 2 weeks before and 2 weeks after menu labelling was introduced<br><br>Study arms: calorie labelling; no calorie labelling<br><br>Number of eligible comparisons: 1<br><br>Comparisons included in the analysis: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: mean energy per sales transaction<br><br>Timing of selection outcome measurement: immediate |
| Funding source | Research funded by the Robert Wood Johnson Foundation Healthy Eating Research Program, the Yale Rudd Center for Food Policy and Obesity, and the New York University Wagner Dean's Fund. |
| Notes | |

**Ellison 2013**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-subject cluster-randomised controlled trial<br><br>Setting: 1 restaurant at Oklahoma State University campus, USA<br><br>Setting type: field setting<br><br>Recruitment: diners recruited as they entered the restaurant<br><br>Allocation to group: randomised by party of diners, who were randomly assigned to a table in 1 of the 3 sections which had different menu types. |
| Participants | Number of enrolled participants: 138<br><br>Number (%) of enrolled participants completing the study and included in analysis: 138 (100%)<br><br>Study completers: age: 69% aged 18–34.9 years; 18.1% aged 35–54.9 years; 12.3% aged ≥ 55 years<br><br>Study completers: sex: 55.8% female<br><br>Specific social or cultural characteristics: mainly students<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Labelling intervention: calorie labelling on menus |

**Ellison 2013** *(Continued)*

| | |
|---|---|
| | Control: no energy information on menus<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: all lunch menu items, apart from drinks and the daily special dessert<br><br>Duration of exposure to intervention: 2 weeks (individual participants for 1 meal, but they could return)<br><br>Study arms: calorie labelling; calorie labelling with traffic lights; no calorie labelling<br><br>Number of eligible comparisons: 1<br><br>Comparisons included in the analysis: calorie labelling versus no calorie labelling (combined simple calorie labelling and calorie labelling with traffic light groups) |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate |
| Funding source | Not reported |
| Notes | Diners could choose from 51 menu options including soups, salads, burgers, sandwiches, pasta, vegetarian dishes, meat dishes, deserts, and drinks. Survey data were collected for 2 weeks during the 2010 autumn semester. Data were combined for intervention 1 and intervention 2 and compared with the control.<br><br>We considered that this study was reasonably categorised as a cluster-randomised design for the purposes of describing study characteristics, as participants were clustered by table. However, uncertainty around its methods including the nature of allocation, meant that, consistent with the previous Crockett 2018 version of this review (and using extracted data available from that review), it was treated equivalently to an individually-randomised controlled trial for the purposes of risk of bias assessment (and with the available outcome data having appropriately accounted for potential clustering). |

**Ellison 2014**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-subjects cluster-randomised controlled trial<br><br>Setting: restaurant at Oklahoma State University campus, USA<br><br>Setting type: field setting<br><br>Recruitment: daily lunch receipts collected from the restaurant<br><br>Allocation to group: randomised by party of diners, who were randomly assigned to a table in 1 of the 3 sections, which had different menu types. |
| Participants | Number of enrolled participants: 1532 usable observations (transactions)<br><br>Number (%) of enrolled participants completing the study and included in the analysis: 1532 (100%)<br><br>Study completers: mean age: not reported<br><br>Study completers: sex: not reported |

**Ellison 2014** *(Continued)*

| | |
|---|---|
| | Specific social or cultural characteristics: mainly students<br>Socioeconomic status context: low deprivation<br>Ethnicity: not reported<br>BMI: not reported |
| Interventions | Labelling intervention: calorie labelling on menus<br>Control: no energy information on menus<br>Manipulated product type: food<br>Characteristics of manipulated products: all lunch menu items, including soups, salads, burgers, pasta, and meat dishes.<br>Duration of exposure to intervention: 12 weeks (individual participants for 1 meal, but they could return)<br>Study arms: calorie labelling; calorie labelling with traffic lights; no calorie labelling, with and without pricing intervention<br>Number of eligible comparisons: 1<br>Comparisons included in the analysis: calorie labelling versus no calorie labelling. Combined 2 energy label groups (calorie labelling alone and traffic light labelling) in a factorial design combining the lower/higher price groups |
| Outcomes | Outcomes reported in study: selection with purchasing<br>Selection outcome analysed: energy content (kcal) ordered<br>Timing of selection outcome measurement: immediate |
| Funding source | Not reported |
| Notes | This study also included a price manipulation experiment; we extracted data reported over the entire experiment.<br>We considered that this study was reasonably categorised as a cluster-randomised design for the purposes of describing study characteristics, as participants were clustered by table. However, uncertainty around its methods including the nature of allocation, meant that, consistent with the previous Crockett 2018 version of this review (and using extracted data available from that review), it was treated equivalently to an individually-randomised controlled trial for the purposes of risk of bias assessment (and with the available outcome data having appropriately accounted for potential clustering). |

**Elshiewy 2018**

| *Study characteristics* | |
|---|---|
| Methods | Study design: interrupted time series<br>Setting: major retailer in the UK with approximately 2000 supermarkets nationwide<br>Setting type: field setting<br>Recruitment: purchase data provided by major retailer. The retailed voluntarily introduced front-of-pack Guideline Daily Amount labels |

**Elshiewy 2018** *(Continued)*

| | |
|---|---|
| | Allocation to group: before and after introduction of labels |
| Participants | Number of enrolled participants: 4,131,570 purchase transactions from 188,062 loyalty card members<br><br>Number (%) of enrolled participants completing the study and included in analysis: 188,062<br><br>Study completers: mean age: not reported<br><br>Study completers: female %: not reported<br><br>Specific social or cultural characteristics: customers of the supermarket chain<br><br>Socioeconomic status context: both<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Labelling intervention: calorie labelling on products<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: cookies, breakfast cereals, and soft drinks (store branded only). All other supermarket products and non-store branded products were not included.<br><br>Duration of exposure to intervention: 1 year<br><br>Study arms: calorie labelling with ≥ 1 other nutrient; typical back-of-pack nutrition labelling<br><br>Number of eligible comparisons: 1<br><br>Comparisons included in the analysis: calorie labelling with ≥ 1 other nutrient versus typical back-of-pack nutrition labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: mean energy per sales transaction<br><br>Timing of selection outcome measurement: immediate |
| Funding source | The European Commission provided by the 7th Framework Program for Research and Technological Development for the Small Collaborative Project FLABEL (contract no. 211905). |
| Notes | |

**Fichera 2020**

| | |
|---|---|
| ***Study characteristics*** | |
| Methods | Study design: controlled before-after study<br><br>Setting: 9 of the biggest retailers in the UK; 4 introduced voluntary labelling front-of-pack schemes<br><br>Setting type: field setting<br><br>Recruitment: rolling panel of household data collected by the market research agency Kantar Worldpanel<br><br>Allocation to group: allocation based on voluntary front-of-pack labelling (in 4 supermarkets), compared to supermarkets who did not introduce any labelling scheme (5 supermarkets). |

**Fichera 2020** *(Continued)*

| | |
|---|---|
| Participants | Number of enrolled participants: 20,707 households (data on all grocery purchases used from 2005 to 2008)<br><br>Number (%) of enrolled participants completing the study and included in analysis: 20,707 (100%) households<br><br>Study completers: mean age: 48.6 (SD 15.5) years<br><br>Study completers: female %: 78.5%<br><br>Specific social or cultural characteristics: customers of the supermarket chains<br><br>Socioeconomic status context: low deprivation (83% were described as higher or intermediate social class)<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Labelling intervention: calorie labelling with ≥ 1 other nutrient on products (2 supermarkets introduced traffic light labelling and 2 introduced a hybrid which included traffic lights and guideline daily amounts).<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: those recommended for labelling by the Food Standards Agency – this included ready meals, burgers, pies, breaded/coated meats, pizza, sandwiches and cereals, dairy foods, and cookies. Some foods were not labelled, including some cakes/desserts/cookies – full list of unlabelled foods not provided.<br><br>Duration of exposure to intervention: analysis from July 2005 to July 2008 (3 years) Waitrose, Co-Op, and Marks & Spencer introduced front-of-pack labelling for all 7 types of food in March 2006 (27 months), September 2006 (22 months), and January 2007 (18 months), respectively; and Asda introduced it on 6 types (all except cereals) in September 2007 (10 months)<br><br>Study arms: calorie labelling with ≥ 1 other nutrient; typical back-of-pack nutrition labelling<br><br>Number of eligible comparisons: 1<br><br>Comparisons included in the analysis: calorie labelling with ≥ 1 other nutrient versus typical back-of-pack nutrition labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: total monthly calories from store-brand labelled food purchases<br><br>Timing of selection outcome measurement: immediate |
| Funding source | Financial support from the UK Medical Research Council (MR/K021583/1 and G1002345), the Rank Prize Funds New Investigator Award (von Hinke), and the European Research Council (ERC) under ERC-2009-AdG grant agreement number 249529. |
| Notes | |

**Hammond 2013**

| | |
|---|---|
| ***Study characteristics*** | |
| Methods | Study design: between-subjects cluster-randomised controlled trial |

**Hammond 2013** *(Continued)*

| | |
|---|---|
| | Setting: University in Ontario, Canada<br><br>Setting type: naturalistic laboratory setting<br><br>Recruitment: recruited via newspaper, bus, and online advertisements<br><br>Allocation to group: randomised by group with the intervention changed daily and 10 participants per day |
| Participants | Number of enrolled participants: 666<br><br>Number (%) of enrolled participants completing the study: 635 (95.3%)<br><br>Study completers: age: 19.2% aged 18–24 years, 16.4% aged 25–34 years, 49.1% aged 35–64 years, 15.3% aged ≥ 65 years<br><br>Study completers: sex: 55.8% female<br><br>Specific social or cultural characteristics: aged > 18 years, no food allergies<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: 71% white, 29% other<br><br>BMI: not reported |
| Interventions | Labelling intervention: calorie labelling on menus<br><br>Control: no energy information on menus<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: fast-food menu (subway)<br><br>Duration of exposure to intervention: 1 hour<br><br>Study arms: calorie labelling, calorie labelling with traffic lights, calorie labelling with traffic lights and ≥ 1 other nutrient (salt/sugar), no calorie labelling<br><br>Number of eligible comparisons: 2<br><br>Comparisons included in the analysis: calorie labelling versus no calorie labelling; calorie labelling with ≥ 1 other nutrient present versus absent |
| Outcomes | Outcomes reported in study: selection with purchasing; consumption<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate<br><br>Consumption outcome analysed: energy content (kcal)<br><br>Timing of consumption outcome measurement: immediate |
| Funding source | Canadian Cancer Society Research Institute |
| Notes | 2 calorie label groups (simple calorie labelling and traffic light labelling) were combined.<br><br>We considered that this study was reasonably categorised as a cluster-randomised design for the purposes of describing study characteristics. However, uncertainty around its methods including the nature of allocation, meant that, consistent with the previous Crockett 2018 version of this review (and using extracted data available from that review), it was treated equivalently to an individually randomised controlled trial for the purposes of risk of bias assessment (and with the available outcome data having appropriately accounted for potential clustering). |

### Harnack 2008

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-subjects randomised controlled trial<br>Setting: community conference rooms at local hotels and the basement of a church, Minnesota, USA<br>Setting type: naturalistic laboratory setting<br>Recruitment: via advertisements placed in community newspapers and flyers posted in community locations<br>Allocation to group: randomised by individual, process of randomisation unclear |
| Participants | Number of enrolled participants: 605<br>Number (%) of enrolled participants completing the study and included in analysis: 594 (98.2%)<br>Study completers: mean age: not reported<br>Study completers: sex: 59.4% female<br>Specific social or cultural characteristics: aged > 16 years, general public, weekly fast-food consumers<br>Socioeconomic status context: both low and high deprivation<br>Ethnicity: 3.4% Hispanic, 96.6% other<br>BMI: 42.6% normal weight, 27.9% overweight, 29.6% obese |
| Interventions | Labelling intervention: calorie labelling on menus<br>Control: no energy information on menus<br>Manipulated product type: food<br>Characteristics of manipulated products: fast-food menu (McDonalds)<br>Duration of exposure to intervention: length of 1 meal<br>Study arms: energy labels (present/absent), price (higher/lower)<br>Number of eligible comparisons: 1<br>Comparisons included in the analysis: calorie labelling versus no calorie labelling (groups combined in factorial design) |
| Outcomes | Outcomes reported in study: selection with purchasing; consumption<br>Selection outcome analysed: energy content (kcal) ordered<br>Timing of selection outcome measurement: immediate<br>Consumption outcome analysed: energy content (kcal)<br>Timing of consumption outcome measurement: immediate |
| Funding source | NIDDK (National Institute of Diabetes and Digestive and Kidney Diseases) grant |
| Notes | Price (low/high) combined in a factorial design |

**James 2015**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-subjects randomised controlled trial<br><br>Setting: 1 dining area located in a metabolic kitchen and 1 in a residence occupied by graduate students, Texas Christian University, US<br><br>Setting type: naturalistic laboratory setting<br><br>Recruitment: word of mouth (shared information by oral communication), via fliers placed around campus, through announcements made in class, and via a Texas Christian University newsletter<br><br>Allocation to group: by individual, process of randomisation unclear |
| Participants | Number of enrolled participants: 300<br><br>Number (%) of enrolled participants completing the study and included in analysis: 300 (100%)<br><br>Study completers: mean age: 21.9 (SD 2.31) years<br><br>Study completers: 55.7% female<br><br>Specific social or cultural characteristics: majority young, white college students with normal weight<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: 88% white; 4% Black and African American; 5% Asian; 3% mixed race/ethnicity<br><br>BMI: mean 24.2 $kg/m^2$ |
| Interventions | Labelling intervention: calorie labelling on menus<br><br>Control: no energy information on menus<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: all menu items included (lunch fast-food and beverage options, included 8 burger choices, 3 salad selections with choice of 5 dressings, 4 chicken sandwich choices, chicken nuggets, chicken strips, 4 side selections including French fries, 3 desserts, condiments, water, and a choice of fizzy drinks).<br><br>Duration of exposure to intervention: length of 1 meal<br><br>Study arms: calorie labels; PACE labels; no label<br><br>Number of eligible comparisons: 1<br><br>Comparisons included in the analysis: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection without purchasing<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate<br><br>Consumption outcome analysed: energy content (kcal) consumed<br><br>Timing of consumption outcome measurement: immediate |
| Funding source | Study was partly supported by a Graduate Student Senate grant from Texas Christian University. |

**James 2015** *(Continued)*

| | |
|---|---|
| Notes | |

**Oliveira 2018**

| *Study characteristics* | |
|---|---|
| Methods | Study design: between-participants cluster-randomised controlled trial<br>Setting: university campus restaurant, Brazil<br>Setting type: field setting<br>Recruitment: recruited via social media and e-mail messages<br>Allocation to group: stratified according to gender, BMI, dietary restrictions, done using computer-generated list and non-involved investigator |
| Participants | Number of enrolled participants: 16 groups (233 people)<br>Number (%) of enrolled participants completing the study and included in analysis: 16 groups (233 people)<br>Study completers: age: 96% aged 20–30 years<br>Study completers: sex: 47.2% female<br>Specific social or cultural characteristics: students aged > 20 years<br>Socioeconomic status context: low deprivation<br>Ethnicity: not reported<br>BMI: not reported |
| Interventions | Labelling intervention: calorie labelling with ≥ 1 other nutrient on menus<br>Control: no energy information on menus<br>Manipulated product type: food<br>Characteristics of manipulated products: all food on lunch menu (including salads (with vegetable oil, vinegar or lemon juice and salt as dressing), rice, beans, meat dishes, potatoes, or other side dishes)<br>Duration of exposure to intervention: 3 hours maximum (length of 1 meal)<br>Study arms: calorie labelling; nutritional quality symbols/ingredients; no calorie labelling<br>Number of eligible comparisons: 1<br>Comparisons included in the analysis: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br>Selection outcome analysed: number of healthy choices<br>Timing of selection outcome measurement: immediate |
| Funding source | Not reported |

**Oliveira 2018** *(Continued)*

| | |
|---|---|
| Notes | Authors were contacted but we were unable to include this study in meta-analysis as data for energy selected was not available. There were no significant differences reported in number of healthy food choices. |

**Petimar 2019**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: controlled before-after study<br>Setting: 82 fast-food restaurants in 4 New England cities, USA<br>Setting type: field setting<br>Allocation to group: McDonalds selected as intervention group as voluntary calorie labelling was implemented in 2012. Compared to 5 control chains that did not label menus – control group of restaurants similar to McDonalds |
| Participants | Number of enrolled participants: 5948<br>Number (%) of enrolled participants completing the study and included in analysis: 5582<br>Study completers: mean age: 27 (SD 16.45) years<br>Study completers: female %: 47<br>Specific social or cultural characteristics: restaurant customers. Reported they chose area as southern states population were more at risk of obesity.<br>Socioeconomic status context: both<br>Ethnicity: 25% white; 33% Black or African American; 5% Asian; 26% Hispanic; 11% other<br>BMI: 25.75 (SD 6.2) kg/$m^2$ |
| Interventions | Labelling intervention: calories on menus/menu boards<br>Manipulated product type: food<br>Characteristics of manipulated products: all restaurant menu items<br>Duration of exposure to intervention: 2 years. Weekly sales data collected before and after introduction (2 years before, 2 years after)<br>Study arms: calorie labelling; no calorie labelling<br>Number of eligible comparisons: 1<br>Comparisons included in the analysis: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br>Selection outcome analysed: energy content (kcal) per transaction<br>Timing of selection outcome measurement: immediate |
| Funding source | Not specified – funding acquired by study author |
| Notes | |

**Petimar 2021**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: interrupted time series<br><br>Setting: restaurant chain in southern USA<br><br>Setting type: field setting<br><br>Recruitment: all sales data obtained from a fast-food franchise with 3 chains in the top 100 largest restaurant chains<br><br>Allocation to group: before and after both voluntary labelling in 1 fast-food franchise and nation-wide implementation of labelling |
| Participants | Number of enrolled participants: 104 restaurants<br><br>Number (%) of enrolled participants completing the study and included in analysis: 104 restaurants<br><br>Study completers: mean age: not reported<br><br>Study completers: female %: not reported<br><br>Specific social or cultural characteristics: areas in southern US selected as that population more at risk of obesity<br><br>Socioeconomic status context: both<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Type of labelling intervention: calories on menus/menu boards.<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: all restaurant menu items including sugar-sweetened beverages<br><br>Duration of exposure to intervention: 1 year<br><br>Study arms: calorie labelling; no calorie labelling<br><br>Number of eligible comparisons: 1<br><br>Comparisons analysed: before and after calorie labels introduced |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: energy per transaction<br><br>Timing of selection outcome measurement: immediate |
| Funding source | Funded by R01DK115492 from the National Institutes of Health awarded to JPB. JP was supported by T32HL098048 from the National Heart, Lung, and Blood Institute. |
| Notes | |

**Platkin 2014**

| *Study characteristics* | |
|---|---|
| Methods | Study design: between-subjects randomised controlled trial<br><br>Setting: private conference room in the University's Graham Center nearby the student union where the students normally eat, South Florida University Campus, US<br><br>Setting type: naturalistic laboratory setting<br><br>Recruitment: 62 overweight or obese females were recruited on a south Florida college campus<br><br>Allocation to group: allocation by individuals, process of randomisation not specified |
| Participants | Number of enrolled participants: 62<br><br>Number (%) of enrolled participants completing the study and included in analysis: 62 (100%)<br><br>Study completers: mean age: 21.9 (SD 3.03) years<br><br>Study completers: sex: 100% female<br><br>Specific social or cultural characteristics: overweight or obese females (BMI ≥ 25 kg/m$^2$ and < 40 kg/m$^2$)<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: 27.33% Black and African American; 45.33% Hispanic; 27.33% other<br><br>BMI: mean 28.41 kg/m$^2$, all classed as overweight or obese |
| Interventions | Labelling interventions: calorie labelling on menus; calorie labelling with PACE<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: lunch from fast-food restaurant (Burger King). Participants were able to choose entrées, a garden salad, side dishes, condiments, and a drink<br><br>Duration of exposure to intervention: length of 1 meal<br><br>Study arms: menu with calorie labels; menu with calorie labels plus PACE; no label<br><br>Number of eligible comparisons: 2<br><br>Comparisons included in the analysis: calorie labelling versus no calorie labelling; calorie labelling with PACE versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing; consumption<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate<br><br>Measurement of consumption outcome: energy content (kcal) consumed<br><br>Timing of consumption outcome measurement: immediate |
| Funding source | None reported |
| Notes | |

**Reynolds 2022**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-participants step-wedged cluster-randomised controlled trial<br>Setting: 10 worksite cafeterias, UK<br>Setting type: field setting<br>Recruitment: through a major UK catering company<br>Allocation to group: allocation by worksite cafeteria |
| Participants | Number of enrolled participants: 10 sites<br>Number (%) of enrolled participants completing the study and included in analysis: 10 sites<br>Study completers: mean age: 40 years<br>Study completers: sex: 28% female<br>Specific social or cultural characteristics: staff at worksite cafeterias<br>Socioeconomic status context: low deprivation<br>Ethnicity: not reported<br>BMI: not reported |
| Interventions | Type of labelling intervention: calorie labelling with PACE. These labels were added in up to 4 locations at each cafeteria: 1. shelf-edge labels, 2. menus next to food and drink displays, 3. individual tent cards next to food and drink displays, and 4. on stickers that were attached to the product packaging<br>Manipulated product type: food<br>Characteristics of manipulated products: categories of food and drink: hot meals, sandwiches, cold drinks, desserts<br>Duration of exposure to intervention: 4–8 weeks<br>Study arms: calorie labels with PACE; no label<br>Number of eligible comparisons: 1<br>Comparisons analysed: calorie labels with PACE; no label. No-label products usually only featured product name and price, but some standardised front-of-pack nutrition labels on branded products (e.g. Coca Cola) and in-house products (e.g. muffins) where energy content was labelled in small print |
| Outcomes | Outcomes reported in study: selection with purchasing<br>Selection outcome analysed: energy content (kcal) ordered<br>Timing of selection outcome measurement: immediate |
| Funding source | Wellcome Trust |
| Notes | Data imputed (see Dealing with missing data): standardised mean difference −0.04 (standard error 0.04) |

**Roberto 2010**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-subjects randomised controlled trial<br><br>Setting: university classroom on the Yale University campus not affiliated with eating or weight research, US<br><br>Setting type: laboratory<br><br>Recruitment: via flyers, word of mouth (shared information by oral communication), newspaper advertisements, and craigslist.com postings<br><br>Allocation to group: allocation by individuals, process of randomisation not specified |
| Participants | Number of enrolled participants: 303<br><br>Number (%) of enrolled participants completing the study and included in analysis: 295 (97.4%)<br><br>Study completers: mean age: 30.5 (SD 12.4) years<br><br>Study completers: sex: 50% female<br><br>Specific social or cultural characteristics: high proportion of students<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: 55% white; 20% Black or African American; 15% Asian; 3% mixed race/ethnicity; 1% minority; 4% other<br><br>BMI: mean 25.2 $kg/m^2$ |
| Interventions | Type of labelling intervention: calorie labelling on menus<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: dinner (menu items included all salads, dressings, sandwiches, wraps, beverages, desserts, mozzarella sticks, French fries, pizza, hamburgers, cheesecakes)<br><br>Duration of exposure to intervention: length of 1 meal<br><br>Study arms: menu with calorie labels; menu with calorie labels plus daily intake information; no label<br><br>Number of eligible comparisons: 1<br><br>Comparisons analysed: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection without purchasing; consumption<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate<br><br>Measurement of consumption outcome: energy content (kcal) consumed<br><br>Timing of consumption outcome measurement: immediate |
| Funding source | Funded by grants from the Rudd Foundation and the Robert Wood Johnson Foundation. |
| Notes | Calorie labelling groups with or without daily intake information combined. |

**Roberto 2012**

| ***Study characteristics*** | |
|---|---|
| Methods | Design: randomised controlled trial<br>Setting: university laboratory not affiliated with eating or weight research, US<br>Recruitment: via flyers, word of mouth (shared information by oral communication), and craigslist.com postings<br>Allocation to groups: randomly generated allocation sequence, stratified by gender |
| Participants | Number of enrolled participants: 243<br>Number (%) of enrolled participants completing the study and included in analysis: 216 (88.9%)<br>Study completers: mean age: 26 (SD 10) years<br>Study completers: sex: 63% female<br>Specific social or cultural characteristics: high proportion of students<br>Socioeconomic status context: low deprivation<br>Ethnicity: 59% white; 11% Black or African American; 22% Asian; 4% Hispanic; 4% other<br>BMI: 22% of participants are classified as overweight or obese |
| Interventions | Type of labelling intervention: calorie labelling on the front of a cereal box<br>Manipulated product type: food<br>Characteristics of manipulated products: cereal<br>Duration of exposure to intervention: length of 1 meal<br>Study arms: calorie labels (present/absent), serving size suggestions (larger, smaller)<br>Number of eligible comparisons: 1<br>Comparisons included in the analysis: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection without purchasing; consumption<br>Selection outcome analysed: total grams of cereal poured, converted to energy content (kcal)<br>Timing of selection outcome measurement: immediate<br>Consumption outcome analysed: total grams of cereal + milk consumed, converted to energy content (kcal)<br>Timing of consumption outcome measurement: immediate |
| Funding source | Supported in part by funding from the Rudd Foundation |
| Notes | |

**Robertson 2020**

| ***Study characteristics*** |
|---|

**Robertson 2020** *(Continued)*

| | |
|---|---|
| Methods | Study design: between-subjects randomised controlled trial<br><br>Setting: Economic and Social Research Institute, Dublin, Ireland<br><br>Setting type: laboratory<br><br>Recruitment: via a market research institute<br><br>Allocation to group: allocation by individuals, process of randomisation not specified |
| Participants | Number of enrolled participants: 145<br><br>Number (%) of enrolled participants completing the study and included in analysis: 142<br><br>Study completers: mean age: 39.65 (SD 14.41) years<br><br>Study completers: sex: 49% female<br><br>Specific social or cultural characteristics: general population, balanced by age, gender, and working status<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Type of labelling intervention: calorie labelling on menus<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: lunch – sandwiches, salads, wraps, pizzas, burgers, fries, hot nachos, chips, chocolate, fruit, soft drinks, fruit juices, and water.<br><br>Duration of exposure to intervention: immediate<br><br>Study arms: calorie labels (present/absent), recommended daily allowance (present/absent), spatial location on menu (left/right)<br><br>Number of eligible comparisons: 1<br><br>Comparisons analysed: calorie label; no label |
| Outcomes | Outcomes reported in study: selection without purchasing; consumption<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate<br><br>Consumption outcome analysed: energy content (kcal) consumed<br><br>Timing of consumption outcome measurement: immediate |
| Funding source | Not specified |
| Notes | All energy labels combined for analysis |

**Temple 2010**

***Study characteristics***

**Temple 2010** *(Continued)*

| | |
|---|---|
| Methods | Study design: between-subjects randomised controlled trial<br><br>Setting: laboratory at University of Buffalo, US<br><br>Setting type: laboratory<br><br>Recruitment: flyers posted around the University at Buffalo North and South campuses<br><br>Allocation to group: allocation by individuals (participants randomly assigned to 1 of 2 groups, process of randomisation not specified) |
| Participants | Number of enrolled participants: 47<br><br>Number (%) of enrolled participants completing the study and included in analysis: 47 (100%)<br><br>Study completers: mean age: 29.9 (SD 1.5) years<br><br>Study completers: sex: 51% female<br><br>Specific social or cultural characteristics: well-educated with more than half (59.5% part- or full-time students)<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: 43% minority<br><br>BMI: mean 25.9 $kg/m^2$ |
| Interventions | Type of labelling intervention: calorie labelling in the form of nutrition labels made using the standard US Department of Agriculture format (which includes calories)<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: foods and non-alcoholic drinks – all items available on the buffet lunch (salad bar, a sandwich bar with different types of bread, rolls, and a variety of meats and cheeses; and several side items and desserts, including potato chips, candy bars, cookies, yogurt, fruit, and pudding. Beverages offered consisted of water, soft drinks, white and chocolate milk, and juice)<br><br>Duration of exposure to intervention: length of 1 meal<br><br>Study arms: menu with calorie labels; menu with calorie labels plus PACE; no label<br><br>Number of eligible comparisons: 1<br><br>Comparisons included in the analysis: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: consumption<br><br>Measurement of consumption outcome: energy content (kcal) consumed<br><br>Timing of consumption outcome measurement: immediate |
| Funding source | Funds from the corresponding author's assistant professor start-up funds through the University at Buffalo |
| Notes | |

**VanEpps 2016**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: between-subjects randomised controlled trial<br><br>Setting: 1 large healthcare company (Humana), USA<br><br>Setting type: field<br><br>Recruitment: sent an initial recruitment e-mail to 1440 randomly selected employees of Humana<br><br>Allocation to group: randomised by individual, process of randomisation unclear |
| Participants | Number of enrolled participants: 453<br><br>Number (%) of enrolled participants completing the study and included in analysis: 249 (54.9%). 204 participants never made a lunch order so were not included in analysis<br><br>Study completers: mean age 40.57 (SD 11.09) years<br><br>Study completers: sex: 60% female<br><br>Specific social or cultural characteristics: employees of a healthcare company<br><br>Socioeconomic status context: low deprivation<br><br>Ethnicity: 81% white; 6% Black or African American; 6% Asian; 1% Hispanic; 2% other<br><br>BMI: 27.51 (SD 6.19) kg/$m^2$ |
| Interventions | Type of labelling intervention: calorie labelling on menus<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: lunch menu for catering company for employees on site. 13 different meal options including sandwiches with side dishes, wraps with side dishes, and entrée-sized salads, and drinks<br><br>Duration of exposure to intervention: 4 weeks (participants could order up to 3 lunches per week on study website)<br><br>Study arms: menu with calorie labels, menu with calorie labels plus traffic lights, traffic lights only, no calorie label<br><br>Number of eligible comparisons: 1<br><br>Comparisons analysed: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate |
| Funding source | Not reported |
| Notes | Combined energy only and energy traffic light group in factorial design (both display only energy information, 1 had an additional traffic light colour). |

**Vasiljevic 2018**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: cluster-randomised stepped-wedge controlled trial<br><br>Setting: worksite cafeterias of companies that are members of the Institute for Grocery Distribution<br><br>Setting type: field<br><br>Recruitment: 6 English worksite cafeterias recruited from the 1027 companies that are members of the Institute of Grocery Distribution<br><br>Allocation to group: worksite cafeterias were randomly allocated to a phase of the stepped wedge design. Randomisation performed by a statistician, with the assistance of computer software. |
| Participants | Number of enrolled participants: 6 sites (5253 people)<br><br>Number (%) of enrolled participants completing the study and included in analysis: 6 sites (100%)<br><br>Study completers: mean age: not reported<br><br>Study completers: sex: 45% female<br><br>Specific social or cultural characteristics: employees of worksite<br><br>Socioeconomic status context: both low and high<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Type of labelling intervention: calorie labelling on menus, along product display shelves and on products<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: all cafeteria products for which energy information was available with their energy content. Salad bars, hot drinks, and vending machine items were excluded from the intervention<br><br>Duration of exposure to intervention: 3–13 weeks depending on the site<br><br>Study arms: menu with calorie labels; no label<br><br>Number of eligible comparisons: 1<br><br>Comparisons analysed: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate |
| Funding source | Department of Health Policy Research Programme (Policy Research Unit in Behaviour and Health [PR-UN-040910109] and the Institute of Grocery Distribution (RG83425)). RP is supported by a Wellcome Trust Research Fellowship in Society and Ethics (106679/Z/ 14/Z). |
| Notes | Data imputed (see Dealing with missing data): standardised mean difference −0.01 (standard error 0.06) |

**Vasiljevic 2019**

| ***Study characteristics*** | |
|---|---|
| Methods | Study design: cluster-randomised stepped-wedge controlled trial<br><br>Setting: worksite cafeterias of companies that were members of the Institute for Grocery Distribution<br><br>Setting type: field<br><br>Recruitment: via a collaboration with the Institute of Grocery Distribution<br><br>Allocation to group: worksite cafeterias were randomly allocated to a phase of the stepped wedge design. Randomisation performed by a statistician using computer software. |
| Participants | Number of enrolled participants: 3 sites (2947 people)<br><br>Number (%) of enrolled participants completing the study and included in analysis: 3 sites (100%)<br><br>Study completers: mean age: not reported<br><br>Study completers: sex: 54.4% female<br><br>Specific social or cultural characteristics: employees of worksite<br><br>Socioeconomic status context: both low and high<br><br>Ethnicity: not reported<br><br>BMI: not reported |
| Interventions | Type of labelling intervention: calorie labelling on menus, along product display shelves and on products<br><br>Manipulated product type: food<br><br>Characteristics of manipulated products: all cafeteria products for which energy information was available with their energy content. Salad bars, deli bars, hot drinks, and vending machine items were excluded from the intervention because of challenges in reliably implementing calorie labelling for these items.<br><br>Duration of exposure to intervention: 8–12 weeks depending on site<br><br>Study arms: menu with calorie labels; no label<br><br>Number of eligible comparisons: 1<br><br>Comparisons analysed: calorie labelling versus no calorie labelling |
| Outcomes | Outcomes reported in study: selection with purchasing<br><br>Selection outcome analysed: energy content (kcal) ordered<br><br>Timing of selection outcome measurement: immediate |
| Funding source | National Institute for Health Research Policy Research Programme (Policy Research Unit in Behaviour and Health [PR-UN-0409-10109]) and the Institute of Grocery Distribution (RG83425), Wellcome Trust Research Fellowship in Society and Ethics (106679/Z/14/Z), Oxford NIHR Biomedical Research Centre and the Oxford NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC). |
| Notes | Data imputed (see Dealing with missing data): standardised mean difference −0.02 (standard error 0.03) |

BMI: body mass index; PACE: physical activity calorie equivalent.

**Characteristics of excluded studies** *[ordered by study ID]*

| Study | Reason for exclusion |
|---|---|
| Acton 2019 | No eligible interventions (within-study comparisons) |
| Acton 2020a | No eligible interventions (within-study comparisons) |
| Acton 2020b | No eligible interventions (within-study comparisons) |
| Acton 2021 | No eligible interventions (within-study comparisons) |
| Allan 2015 | No eligible interventions (within-study comparisons) |
| Allan 2020 | No eligible interventions (within-study comparisons) |
| Al-Otaibi 2021 | Ineligible study design |
| Bailey 2022 | No measurement of eligible selection or consumption outcomes |
| Balasubramanian 2002 | Ineligible study design |
| Baum 2017 | No eligible interventions (within-study comparisons) |
| Bergen 2006 | No eligible interventions (within-study comparisons) |
| Bergman 2021 | No measurement of eligible selection or consumption outcomes |
| Byrd 2017 | No measurement of eligible selection or consumption outcomes |
| Byrd 2021 | No measurement of eligible selection or consumption outcomes |
| Carter 2018 | No measurement of eligible selection or consumption outcomes |
| Cavanagh 2014 | No measurement of eligible selection or consumption outcomes |
| Chu 2009a | Ineligible study design |
| Chu 2009b | Ineligible study design |
| Cioffi 2015 | Ineligible study design |
| Cornil 2017 | No measurement of eligible selection or consumption outcomes |
| Crockett 2014 | No eligible interventions (within-study comparisons) |
| Crosetto 2017 | No eligible interventions (within-study comparisons) |
| Crosetto 2018 | No measurement of eligible selection or consumption outcomes |
| Crosetto 2020 | No measurement of eligible selection or consumption outcomes |
| Dallas 2017 | No measurement of eligible selection or consumption outcomes |
| Dallas 2019 | No measurement of eligible selection or consumption outcomes |

| Study | Reason for exclusion |
|---|---|
| Defago 2017 | No eligible interventions (within-study comparisons) |
| Dubbert 1984 | No eligible interventions (within-study comparisons) |
| Ebneter 2013 | Ineligible study design |
| Egnell 2019 | No measurement of eligible selection or consumption outcomes |
| Erdem 2021 | Ineligible study design |
| Fang 2019 | No eligible interventions (within-study comparisons) |
| Finkelstein 2019a | No measurement of eligible selection or consumption outcomes |
| Finkelstein 2019b | No eligible interventions (within-study comparisons) |
| Finkelstein 2020 | No eligible interventions (within-study comparisons) |
| Finkelstein 2021 | No eligible interventions (within-study comparisons) |
| Girz 2012 | No eligible interventions (within-study comparisons) |
| Goodman 2018 | No measurement of eligible selection or consumption outcomes |
| Guha 2018 | No eligible interventions (within-study comparisons) |
| Gustafson 2018 | Ineligible study design |
| Hartley 2019a | No eligible interventions (within-study comparisons) |
| Hartley 2019b | No eligible interventions (within-study comparisons) |
| Holmes 2013 | Ineligible study design |
| Huseynov 2021 | No measurement of eligible selection or consumption outcomes |
| IRCT20181002041201N1 | No measurement of eligible selection or consumption outcomes |
| ISRCTN90365793 | No eligible interventions (within-study comparisons) |
| Jin 2020 | No eligible interventions (within-study comparisons) |
| Jue 2012 | No eligible interventions (within-study comparisons) |
| Julia 2021 | No eligible interventions (within-study comparisons) |
| Karnik 2018 | No eligible interventions (within-study comparisons) |
| Kral 2002 | No measurement of eligible selection or consumption outcomes |
| Lee 2019 | Ineligible study design |
| Machin 2018 | No measurement of eligible selection or consumption outcomes |
| Machin 2019 | No eligible interventions (within-study comparisons) |

| Study | Reason for exclusion |
|---|---|
| Mantzari 2020 | No measurement of eligible selection or consumption outcomes |
| Maynard 2018 | No measurement of eligible selection or consumption outcomes |
| Mazza 2018 | No eligible interventions (within-study comparisons) |
| McCrickerd 2020 | No eligible interventions (within-study comparisons) |
| McElroy 2016 | Ineligible study design |
| McInerney 2017 | No measurement of eligible selection or consumption outcomes |
| Montford 2017 | No eligible interventions (within-study comparisons) |
| Muller 2020a | No measurement of eligible selection or consumption outcomes |
| Muller 2020b | No measurement of eligible selection or consumption outcome |
| Nassab 2017 | Ineligible study design |
| NCT03553043 | No measurement of eligible selection or consumption outcomes |
| NCT03761342 | No eligible interventions (within-study comparisons) |
| NCT04172337 | No eligible interventions (within-study comparisons) |
| NCT04252898 | No eligible interventions (within-study comparisons) |
| Neal 2017 | No eligible interventions (within-study comparisons) |
| Neuhofer 2020 | No measurement of eligible selection or consumption outcomes |
| Ni Mhurchu 2017 | No eligible interventions (within-study comparisons) |
| Nordström 2020 | No eligible interventions (within-study comparisons) |
| Oh 2020 | No measurement of eligible selection or consumption outcomes |
| Otto 2020 | No eligible interventions (within-study comparisons) |
| Rayner 2017 | No eligible interventions (within-study comparisons) |
| Reinhardt 2019 | No eligible interventions (within-study comparisons) |
| Rising 2017 | No measurement of eligible selection or consumption outcomes |
| Sandoval 2017 | No eligible interventions (within-study comparisons) |
| Scourboutakos 2016 | No eligible interventions (within-study comparisons) |
| Seyedhamzeh 2020 | No measurement of eligible selection or consumption outcomes |
| Shin 2020 | No eligible interventions (within-study comparisons) |
| Shoychet 2023 | No measurement of eligible selection or consumption outcomes |

| Study | Reason for exclusion |
|---|---|
| Silva 2022 | No eligible interventions (within-study comparisons) |
| Tangari 2019 | No measurement of eligible selection or consumption outcomes |
| Tapper 2019 | No measurement of eligible selection or consumption outcomes |
| Tapper 2021a | No measurement of eligible selection or consumption outcomes |
| Thunstrom 2018 | No eligible interventions (within-study comparisons) |
| Thunstrom 2019 | No eligible interventions (within-study comparisons) |
| Turnwald 2019a | No eligible interventions (within-study comparisons) |
| Turnwald 2019b | No eligible interventions (within-study comparisons) |
| Urminsky 2019 | Ineligible study design |
| VanEpps 2021 | No measurement of eligible selection or consumption outcomes |
| Vermeer 2011 | No eligible interventions (within-study comparisons) |
| Viera 2017 | Ineligible study design |
| Viera 2019 | Ineligible study design |
| Wansink 2006 | No eligible interventions (within-study comparisons) |
| Whitt 2017 | Ineligible study design |
| Zhang 2020 | No measurement of eligible selection or consumption outcomes |

## Characteristics of studies awaiting classification *[ordered by study ID]*

### Clements 2016

| | |
|---|---|
| Methods | Calories ordered in breakfast and lunch meals were measured over 4 different time points (November 2014 (precalorie menu labelling), January 2015 (2 months postcalorie menu labelling), November 2015 (12 months postcalorie menu labelling), and January 2016 (14 months postcalorie menu labelling)) |
| Participants | Staff and visitors in a hospital canteen |
| Interventions | Calorie labelling. Further details not clear from record. |
| Outcomes | Median calories purchased on each day of the week at all time points were compared with precalorie menu labelling |
| Notes | Identified via original searches for this review in 2021. We could not confidently determine eligibility from the available material and we grouped as awaiting classification. We attempted to contact the authors in April 2022 but received no response. |

**Dos Santos 2015**

| | |
|---|---|
| Methods | Longitudinal cohort follow-up study conducted from 21 February to 30 April 2013. After a 1-month baseline period, a display with hot drinks calorie information was posted on coffee machines |
| Participants | Staff at 2 companies (1 blue-collar (manual worker) and 1 white-collar (office worker or professional)) |
| Interventions | Calorie labelling on coffee machines |
| Outcomes | Number of cups purchased evaluated at baseline and end of the study<br><br>Further details unknown |
| Notes | Identified via original searches for this review in 2021. We could not confidently determine eligibility from the available material and we grouped as awaiting classification. We attempted to contact the authors in April 2022 but did not receive a response. |

**Elshiewy 2022**

| | |
|---|---|
| Methods | Interrupted time series design from May 2005 to April 2007. Guideline daily amount labels were introduced in May 2006 |
| Participants | Customers at a large UK retailer |
| Interventions | Calorie labelling (in a guideline daily amount label) on yoghurts and ready meals |
| Outcomes | Healthy food choices |
| Notes | Identified via September 2023 updated searches. We could not confidently determine eligibility from the available material and we have grouped as awaiting classification. We judged that we will need contact the authors and obtain further expert statistical input to determine whether the design and analysis is appropriate and confers eligibility, and whether there is sufficient useable data or whether additional new analysis either by the original authors or the review team would be required. |

**Girz ongoing**

| | |
|---|---|
| Methods | Setting: university laboratory, Canada<br><br>Design: randomised controlled trial<br><br>Recruitment: students enrolled in introductory psychology who consented to participation in the study<br><br>Allocation to groups: a random number generator was used to assign participants to condition |
| Participants | Undergraduate psychology students (49 included in the analysis) |
| Interventions | Intervention: foods (pizza, soup) presented with energy labels (24 participants)<br><br>Control: foods (pizza, soup) presented without energy labels (25 participants) |
| Outcomes | Energy (kcal) consumption |

**Girz ongoing** *(Continued)*

| | |
|---|---|
| Notes | This study was included in the previous version of this Cochrane review as an unpublished study with the authors providing study data (Crockett 2018). This study comprised 6 study groups with combinations of interventions in terms of labelling and ordering behaviour. For this review, we could not confidently determine eligibility from the available material and we grouped as awaiting classification. We attempted to contact the authors in April 2022 but received no response. |

**Lee 2018**

| | |
|---|---|
| Methods | Calorie labelling on menus in fast-food restaurants<br><br>Further details unknown |
| Participants | Customers at fast-food restaurants<br><br>Further details unknown |
| Interventions | Calorie information |
| Outcomes | Menu selections<br><br>Further details unknown |
| Notes | Identified via September 2023 updated searches. We could not confidently determine eligibility from the available material and we have grouped as awaiting classification. We judged that we will need contact the authors and obtain further expert statistical input to determine whether the design and analysis is appropriate and confers eligibility, and whether there is sufficient useable data or whether additional new analysis either by the original authors or the review team would be required. |

**Petimar 2022**

| | |
|---|---|
| Methods | Controlled interrupted time series compared sales 2 years before labelling implementation (April 2015 to April 2017) with sales 7 months after labelling implementation (May 2017 to December 2017). |
| Participants | Data from 173 supermarkets from a supermarket chain with locations in Maine, Massachusetts, New Hampshire, New York, and Vermont |
| Interventions | Calorie labelling on menus |
| Outcomes | Calories per transaction |
| Notes | Study was identified via September 2023 updating searches and is provisionally included but not yet fully integrated into the review. Results and conclusions: 2 of 3 food categories with calorie labelling being applied saw a statistically significant reduction in calories per transaction purchased (there was no change in the third category) and the study concluded that "calorie labeling of prepared foods was associated with small to moderate decreases in calories purchased from prepared bakery and deli items without evidence of substitution to similar packaged foods." |

**van Doorn 2023**

| | |
|---|---|
| Methods | Study 2: between-subjects design: 2 (consumption moment: distant versus immediate) × 2 (traffic light labels present versus not) in a university canteen<br><br>Study 1 is ineligible |
| Participants | 287 students |
| Interventions | Calorie labelling with information about ≥ 1 other nutrient on menus |
| Outcomes | Calories ordered |
| Notes | Identified via September 2023 updated searches. We could not confidently determine eligibility from the available material and we have grouped as awaiting classification. We judged that we will need contact the authors and obtain further expert statistical input to determine whether the design and analysis is appropriate and confers eligibility, and whether there is sufficient useable data or whether additional new analysis either by the original authors or the review team would be required. |

**Zhu 2023**

| | |
|---|---|
| Methods | 132,849 soft drink purchase records assessed from January 2011 to December 2013. This was matched to the introduction of nutritional information panel on various soft drink brands (year of introduction 2004–2013) |
| Participants | 40,000 households |
| Interventions | The introduction of a nutrition information panel (that included calorie information) |
| Outcomes | Healthy soft drink choices |
| Notes | Identified via September 2023 updated searches. We could not confidently determine eligibility from the available material and we have grouped as awaiting classification. We judged that we will need contact the authors and obtain further expert statistical input to determine whether the design and analysis is appropriate and confers eligibility, and whether there is sufficient useable data or whether additional new analysis either by the original authors or the review team would be required. |

## RISK OF BIAS

**Legend:** ✓ Low risk of bias ✗ High risk of bias ~ Some concerns

**Risk of bias for analysis 1.1 Food selection/purchasing (kcal)**

| | Bias | | | | | |
|---|---|---|---|---|---|---|
| **Study** | **Randomisation process** | **Deviations from intended interventions** | **Missing outcome data** | **Measurement of the outcome** | **Selection of the reported results** | **Overall** |
| Cawley 2020 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

| | Bias | | | | | |
|---|---|---|---|---|---|---|
| **Study** | **Randomisation process** | **Deviations from intended interventions** | **Missing outcome data** | **Measurement of the outcome** | **Selection of the reported results** | **Overall** |
| Clarke 2023a | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Dubois 2021 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Dubois 2021 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Ellison 2013 | ~ | ✓ | ~ | ✓ | ~ | ~ |
| Ellison 2014 | ✗ | ✓ | ~ | ✓ | ~ | ✗ |
| Hammond 2013 | ~ | ✓ | ✓ | ✓ | ✓ | ~ |
| Hammond 2013 | ~ | ✓ | ✓ | ✓ | ✓ | ~ |
| Harnack 2008 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| James 2015 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Platkin 2014 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Platkin 2014 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Reynolds 2022 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Roberto 2010 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Roberto 2012 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Robertson 2020 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| VanEpps 2016 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Vasiljevic 2018 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Vasiljevic 2019 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

**Risk of bias for analysis 1.8 Food consumption (kcal)**

| | Bias | | | | | |
|---|---|---|---|---|---|---|
| **Study** | **Randomisation process** | **Deviations from intended interventions** | **Missing outcome data** | **Measurement of the outcome** | **Selection of the reported results** | **Overall** |
| Hammond 2013 | ~ | ✓ | ✓ | ✓ | ✓ | ~ |
| Hammond 2013 | ~ | ✓ | ✓ | ✓ | ✓ | ~ |
| Harnack 2008 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| James 2015 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Platkin 2014 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Platkin 2014 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Roberto 2010 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Roberto 2012 | ~ | ✓ | ✓ | ✓ | ~ | ~ |
| Robertson 2020 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Temple 2010 | ~ | ✓ | ✓ | ✓ | ~ | ~ |

**Risk of bias for analysis 2.1 Alcohol selection/purchasing of energy (kcal)**

| | Bias | | | | | |
|---|---|---|---|---|---|---|
| **Study** | **Randomisation process** | **Deviations from intended interventions** | **Missing outcome data** | **Measurement of the outcome** | **Selection of the reported results** | **Overall** |
| Cawley 2020 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Clarke 2023a | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

## DATA AND ANALYSES

**Comparison 1. Food including non-alcoholic drinks**

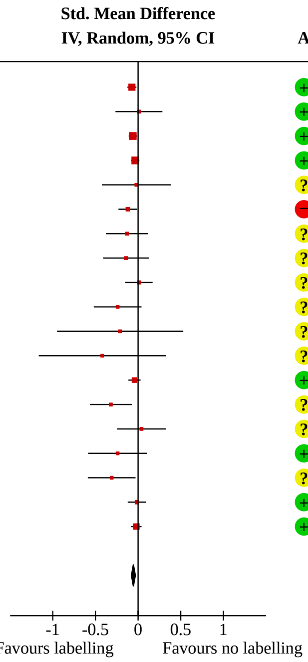
| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1.1 Food selection/purchasing (kcal) | 16 | 9850 | Std. Mean Difference (IV, Random, 95% CI) | -0.06 [-0.08, -0.03] |
| 1.2 Subgroup analysis: food selection/purchasing by setting (kcal) | 16 | 9850 | Std. Mean Difference (IV, Random, 95% CI) | -0.06 [-0.08, -0.03] |
| 1.2.1 Restaurant | 7 | 7466 | Std. Mean Difference (IV, Random, 95% CI) | -0.05 [-0.09, -0.02] |
| 1.2.2 Store | 1 | 40 | Std. Mean Difference (IV, Random, 95% CI) | -0.05 [-0.08, -0.01] |
| 1.2.3 Naturalistic | 6 | 1907 | Std. Mean Difference (IV, Random, 95% CI) | -0.07 [-0.16, 0.03] |
| 1.2.4 Laboratory | 2 | 437 | Std. Mean Difference (IV, Random, 95% CI) | -0.29 [-0.49, -0.09] |
| 1.3 Subgroup analysis: food selection/purchasing by label type (kcal) | 16 | 10057 | Std. Mean Difference (IV, Random, 95% CI) | -0.06 [-0.09, -0.03] |
| 1.3.1 Simple calorie | 14 | 9529 | Std. Mean Difference (IV, Random, 95% CI) | -0.07 [-0.11, -0.03] |
| 1.3.2 Calorie with PACE | 2 | 48 | Std. Mean Difference (IV, Random, 95% CI) | -0.05 [-0.12, 0.03] |
| 1.3.3 Calorie and nutrient | 3 | 480 | Std. Mean Difference (IV, Random, 95% CI) | -0.06 [-0.11, -0.01] |
| 1.4 Subgroup analysis: food selection/purchasing by socioeconomic status (kcal) | 16 | 9850 | Std. Mean Difference (IV, Random, 95% CI) | -0.06 [-0.08, -0.03] |
| 1.4.1 Low | 0 | 0 | Std. Mean Difference (IV, Random, 95% CI) | Not estimable |
| 1.4.2 High | 12 | 9198 | Std. Mean Difference (IV, Random, 95% CI) | -0.08 [-0.12, -0.05] |
| 1.4.3 Both low and high | 4 | 652 | Std. Mean Difference (IV, Random, 95% CI) | -0.04 [-0.07, -0.01] |
| 1.5 Additional exploratory subgroup analysis: food selection/purchasing by placement of labels | 16 | 9850 | Std. Mean Difference (IV, Random, 95% CI) | -0.06 [-0.08, -0.03] |
| 1.5.1 On menus | 12 | 9369 | Std. Mean Difference (IV, Random, 95% CI) | -0.08 [-0.12, -0.03] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1.5.2 On product packaging | 1 | 216 | Std. Mean Difference (IV, Random, 95% CI) | 0.04 [-0.24, 0.32] |
| 1.5.3 Adjacent to product packaging | 1 | 205 | Std. Mean Difference (IV, Random, 95% CI) | 0.01 [-0.26, 0.28] |
| 1.5.4 Both adjacent to and on product packaging | 2 | 60 | Std. Mean Difference (IV, Random, 95% CI) | -0.05 [-0.08, -0.02] |
| 1.6 Sensitivity analysis: food selection/purchasing by low risk of bias (kcal) | 16 | 9850 | Std. Mean Difference (IV, Random, 95% CI) | -0.06 [-0.08, -0.03] |
| 1.6.1 Low risk of bias | 7 | 5976 | Std. Mean Difference (IV, Random, 95% CI) | -0.05 [-0.07, -0.02] |
| 1.6.2 Some concerns or high risk of bias | 9 | 3874 | Std. Mean Difference (IV, Random, 95% CI) | -0.12 [-0.19, -0.06] |
| 1.7 Sensitivity analysis: food selection/purchasing excluding Reynolds 2022, Vasiljevic 2018, Vasiljevic 2019 | 13 | 9812 | Std. Mean Difference (IV, Random, 95% CI) | -0.07 [-0.10, -0.04] |
| 1.8 Food consumption (kcal) | 8 | 2134 | Std. Mean Difference (IV, Random, 95% CI) | -0.19 [-0.33, -0.05] |
| 1.9 Subgroup analysis: food consumption by setting (kcal) | 8 | 2134 | Std. Mean Difference (IV, Random, 95% CI) | -0.19 [-0.33, -0.05] |
| 1.9.1 Restaurant | 0 | 0 | Std. Mean Difference (IV, Random, 95% CI) | Not estimable |
| 1.9.2 Store | 0 | 0 | Std. Mean Difference (IV, Random, 95% CI) | Not estimable |
| 1.9.3 Naturalistic | 5 | 1705 | Std. Mean Difference (IV, Random, 95% CI) | -0.09 [-0.21, 0.02] |
| 1.9.4 Laboratory | 3 | 429 | Std. Mean Difference (IV, Random, 95% CI) | -0.45 [-0.76, -0.14] |
| 1.10 Subgroup analysis: food consumption by label type (kcal) | 8 | 2134 | Std. Mean Difference (IV, Random, 95% CI) | -0.19 [-0.33, -0.05] |
| 1.10.1 Calorie | 8 | 1870 | Std. Mean Difference (IV, Random, 95% CI) | -0.19 [-0.35, -0.02] |
| 1.10.2 Calorie with PACE | 1 | 31 | Std. Mean Difference (IV, Random, 95% CI) | -0.39 [-1.13, 0.36] |
| 1.10.3 Calorie and nutrient | 1 | 233 | Std. Mean Difference (IV, Random, 95% CI) | -0.23 [-0.50, 0.04] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1.11 Subgroup analysis: food consumption by socioeconomic status (kcal) | 8 | 2134 | Std. Mean Difference (IV, Random, 95% CI) | -0.19 [-0.33, -0.05] |
| 1.11.1 Low | 0 | 0 | Std. Mean Difference (IV, Random, 95% CI) | Not estimable |
| 1.11.2 High | 7 | 1540 | Std. Mean Difference (IV, Random, 95% CI) | -0.23 [-0.37, -0.10] |
| 1.11.3 Both low and high | 1 | 594 | Std. Mean Difference (IV, Random, 95% CI) | 0.02 [-0.14, 0.18] |
| 1.12 Sensitivity analysis: food consumption by low risk of bias (kcal) | 8 | 2134 | Std. Mean Difference (IV, Random, 95% CI) | -0.19 [-0.33, -0.05] |
| 1.12.1 Low risk of bias | 1 | 87 | Std. Mean Difference (IV, Random, 95% CI) | -0.68 [-1.13, -0.24] |
| 1.12.2 Some concerns or high risk of bias | 7 | 2047 | Std. Mean Difference (IV, Random, 95% CI) | -0.14 [-0.26, -0.02] |

## Analysis 1.1. Comparison 1: Food including non-alcoholic drinks, Outcome 1: Food selection/purchasing (kcal)

| Study or Subgroup | SMD | SE | Labelling Total | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cawley 2020 | -0.072402 | 0.026854 | 2806 | 2745 | 17.7% | -0.07 [-0.13 , -0.02] | + | + | + | + | + | + |
| Clarke 2023a | 0.01 | 0.140306 | 101 | 104 | 0.7% | 0.01 [-0.26 , 0.28] | + | + | + | + | + | + |
| Dubois 2021[a] | -0.062644 | 0.0245 | 10 | 10 | 21.0% | -0.06 [-0.11 , -0.01] | + | + | + | + | + | + |
| Dubois 2021 | -0.032787 | 0.0245 | 10 | 10 | 21.0% | -0.03 [-0.08 , 0.02] | + | + | + | + | + | + |
| Ellison 2013 | -0.02 | 0.206632 | 108 | 30 | 0.3% | -0.02 [-0.42 , 0.38] | ? | + | ? | + | ? | ? |
| Ellison 2014 | -0.12 | 0.056122 | 1060 | 472 | 4.3% | -0.12 [-0.23 , -0.01] | − | + | ? | + | ? | − |
| Hammond 2013[b] | -0.13 | 0.125 | 321 | 81 | 0.9% | -0.13 [-0.37 , 0.11] | ? | + | + | + | + | ? |
| Hammond 2013 | -0.14 | 0.137755 | 152 | 81 | 0.7% | -0.14 [-0.41 , 0.13] | ? | + | + | + | + | ? |
| Harnack 2008 | 0.01 | 0.081632 | 301 | 293 | 2.0% | 0.01 [-0.15 , 0.17] | ? | + | + | + | ? | ? |
| James 2015 | -0.24 | 0.142857 | 99 | 99 | 0.7% | -0.24 [-0.52 , 0.04] | ? | + | + | + | ? | ? |
| Platkin 2014[c] | -0.21 | 0.377551 | 20 | 11 | 0.1% | -0.21 [-0.95 , 0.53] | ? | + | + | + | ? | ? |
| Platkin 2014 | -0.42 | 0.380102 | 17 | 11 | 0.1% | -0.42 [-1.16 , 0.32] | ? | + | + | + | ? | ? |
| Reynolds 2022 | -0.042178 | 0.0364 | 10 | 10 | 9.9% | -0.04 [-0.11 , 0.03] | + | + | + | + | + | + |
| Roberto 2010 | -0.32 | 0.125 | 200 | 95 | 0.9% | -0.32 [-0.56 , -0.08] | ? | + | + | + | ? | ? |
| Roberto 2012 | 0.04 | 0.145408 | 147 | 69 | 0.6% | 0.04 [-0.24 , 0.32] | ? | + | + | + | ? | ? |
| Robertson 2020 | -0.24 | 0.17602 | 94 | 48 | 0.4% | -0.24 [-0.58 , 0.10] | + | + | + | + | + | + |
| VanEpps 2016 | -0.31 | 0.142857 | 84 | 123 | 0.7% | -0.31 [-0.59 , -0.03] | ? | + | + | + | ? | ? |
| Vasiljevic 2018 | -0.012978 | 0.0555 | 6 | 6 | 4.4% | -0.01 [-0.12 , 0.10] | + | + | + | + | + | + |
| Vasiljevic 2019 | -0.019467 | 0.0306 | 3 | 3 | 13.8% | -0.02 [-0.08 , 0.04] | + | + | + | + | + | + |
| **Total** | | | **5549** | **4301** | **100.0%** | **-0.06 [-0.08 , -0.03]** | | | | | | |




Test for overall effect: Z = 4.72 (P < 0.00001)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² = 0.00; Chi² = 18.32, df = 18 (P = 0.43); I² = 2%

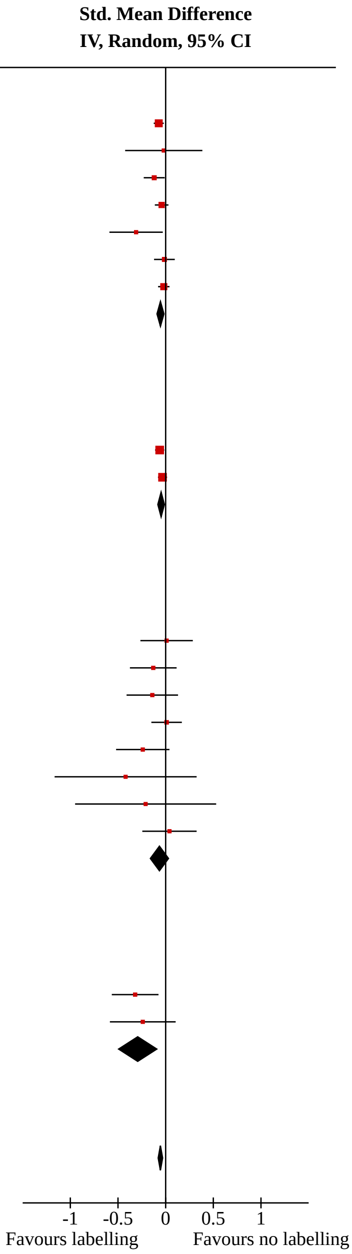
**Footnotes**

[a] Repeated mention of Dubois 2021 in both this and the following row indicates two eligible comparisons from this study
[b] Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[c] Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study

**Risk of bias legend**
(A) Bias arising from the randomization process
(B) Bias due to deviations from intended interventions
(C) Bias due to missing outcome data
(D) Bias in measurement of the outcome
(E) Bias in selection of the reported result
(F) Overall bias

## Analysis 1.2. Comparison 1: Food including non-alcoholic drinks, Outcome 2: Subgroup analysis: food selection/purchasing by setting (kcal)

| Study or Subgroup | SMD | SE | Labelling Total | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|
| **1.2.1 Restaurant** | | | | | | |
| Cawley 2020 | -0.072402 | 0.026854 | 2806 | 2745 | 17.7% | -0.07 [-0.13 , -0.02] |
| Ellison 2013 | -0.02 | 0.206632 | 108 | 30 | 0.3% | -0.02 [-0.42 , 0.38] |
| Ellison 2014 | -0.12 | 0.056122 | 1060 | 472 | 4.3% | -0.12 [-0.23 , -0.01] |
| Reynolds 2022 | -0.042178 | 0.0364 | 10 | 10 | 9.9% | -0.04 [-0.11 , 0.03] |
| VanEpps 2016 | -0.31 | 0.142857 | 84 | 123 | 0.7% | -0.31 [-0.59 , -0.03] |
| Vasiljevic 2018 | -0.012978 | 0.0555 | 6 | 6 | 4.4% | -0.01 [-0.12 , 0.10] |
| Vasiljevic 2019 | -0.019467 | 0.0306 | 3 | 3 | 13.8% | -0.02 [-0.08 , 0.04] |
| **Subtotal** | | | **4077** | **3389** | **51.0%** | **-0.05 [-0.09 , -0.02]** |
| Test for overall effect: Z = 2.97 (P = 0.003) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 7.02, df = 6 (P = 0.32); I² = 14% | | | | | | |
| **1.2.2 Store** | | | | | | |
| Dubois 2021[a] | -0.062644 | 0.0245 | 10 | 10 | 21.0% | -0.06 [-0.11 , -0.01] |
| Dubois 2021 | -0.032787 | 0.0245 | 10 | 10 | 21.0% | -0.03 [-0.08 , 0.02] |
| **Subtotal** | | | **20** | **20** | **41.9%** | **-0.05 [-0.08 , -0.01]** |
| Test for overall effect: Z = 2.75 (P = 0.006) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 0.74, df = 1 (P = 0.39); I² = 0% | | | | | | |
| **1.2.3 Naturalistic** | | | | | | |
| Clarke 2023a | 0.01 | 0.140306 | 101 | 104 | 0.7% | 0.01 [-0.26 , 0.28] |
| Hammond 2013[b] | -0.13 | 0.125 | 321 | 81 | 0.9% | -0.13 [-0.37 , 0.11] |
| Hammond 2013 | -0.14 | 0.137755 | 152 | 81 | 0.7% | -0.14 [-0.41 , 0.13] |
| Harnack 2008 | 0.01 | 0.081632 | 301 | 293 | 2.0% | 0.01 [-0.15 , 0.17] |
| James 2015 | -0.24 | 0.142857 | 99 | 99 | 0.7% | -0.24 [-0.52 , 0.04] |
| Platkin 2014[c] | -0.42 | 0.380102 | 17 | 11 | 0.1% | -0.42 [-1.16 , 0.32] |
| Platkin 2014 | -0.21 | 0.377551 | 20 | 11 | 0.1% | -0.21 [-0.95 , 0.53] |
| Roberto 2012 | 0.04 | 0.145408 | 147 | 69 | 0.6% | 0.04 [-0.24 , 0.32] |
| **Subtotal** | | | **1158** | **749** | **5.8%** | **-0.07 [-0.16 , 0.03]** |
| Test for overall effect: Z = 1.35 (P = 0.18) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 4.74, df = 7 (P = 0.69); I² = 0% | | | | | | |
| **1.2.4 Laboratory** | | | | | | |
| Roberto 2010 | -0.32 | 0.125 | 200 | 95 | 0.9% | -0.32 [-0.56 , -0.08] |
| Robertson 2020 | -0.24 | 0.17602 | 94 | 48 | 0.4% | -0.24 [-0.58 , 0.10] |
| **Subtotal** | | | **294** | **143** | **1.3%** | **-0.29 [-0.49 , -0.09]** |
| Test for overall effect: Z = 2.88 (P = 0.004) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 0.14, df = 1 (P = 0.71); I² = 0% | | | | | | |
| **Total** | | | **5549** | **4301** | **100.0%** | **-0.06 [-0.08 , -0.03]** |



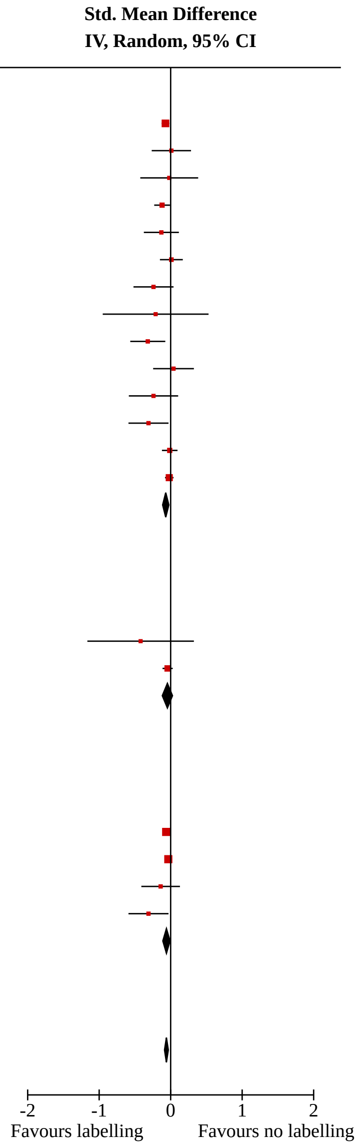

Test for overall effect: Z = 4.72 (P < 0.00001)
Test for subgroup differences: Chi² = 5.68, df = 3 (P = 0.13), I² = 47.2%
Heterogeneity: Tau² = 0.00; Chi² = 18.32, df = 18 (P = 0.43); I² = 2%

**Footnotes**

[a]Repeated mention of Dubois 2021 in both this and the following row indicates two eligible comparisons from this study
[b]Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[c]Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study

## Analysis 1.3. Comparison 1: Food including non-alcoholic drinks, Outcome 3: Subgroup analysis: food selection/purchasing by label type (kcal)

| Study or Subgroup | SMD | SE | Labelling Total | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|
| **1.3.1 Simple calorie** | | | | | | |
| Cawley 2020 | -0.072402 | 0.026854 | 2806 | 2745 | 16.5% | -0.07 [-0.13 , -0.02] |
| Clarke 2023a | 0.01 | 0.140306 | 101 | 104 | 0.9% | 0.01 [-0.26 , 0.28] |
| Ellison 2013 | -0.02 | 0.206632 | 108 | 30 | 0.4% | -0.02 [-0.42 , 0.38] |
| Ellison 2014 | -0.12 | 0.056122 | 1060 | 472 | 5.1% | -0.12 [-0.23 , -0.01] |
| Hammond 2013 | -0.13 | 0.125 | 321 | 81 | 1.1% | -0.13 [-0.37 , 0.11] |
| Harnack 2008 | 0.01 | 0.081632 | 301 | 293 | 2.6% | 0.01 [-0.15 , 0.17] |
| James 2015 | -0.24 | 0.142857 | 99 | 99 | 0.9% | -0.24 [-0.52 , 0.04] |
| Platkin 2014 | -0.21 | 0.377551 | 20 | 11 | 0.1% | -0.21 [-0.95 , 0.53] |
| Roberto 2010 | -0.32 | 0.125 | 200 | 95 | 1.1% | -0.32 [-0.56 , -0.08] |
| Roberto 2012 | 0.04 | 0.145408 | 147 | 69 | 0.8% | 0.04 [-0.24 , 0.32] |
| Robertson 2020 | -0.24 | 0.17602 | 94 | 48 | 0.6% | -0.24 [-0.58 , 0.10] |
| VanEpps 2016 | -0.31 | 0.142857 | 84 | 123 | 0.9% | -0.31 [-0.59 , -0.03] |
| Vasiljevic 2018 | -0.012978 | 0.0555 | 6 | 6 | 5.3% | -0.01 [-0.12 , 0.10] |
| Vasiljevic 2019 | -0.019467 | 0.0306 | 3 | 3 | 13.8% | -0.02 [-0.08 , 0.04] |
| **Subtotal** | | | **5350** | **4179** | **50.2%** | **-0.07 [-0.11 , -0.03]** |
| Test for overall effect: Z = 3.12 (P = 0.002) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 15.80, df = 13 (P = 0.26); I² = 18% | | | | | | |
| **1.3.2 Calorie with PACE** | | | | | | |
| Platkin 2014 | -0.42 | 0.380102 | 17 | 11 | 0.1% | -0.42 [-1.16 , 0.32] |
| Reynolds 2022 | -0.042178 | 0.0364 | 10 | 10 | 10.7% | -0.04 [-0.11 , 0.03] |
| **Subtotal** | | | **27** | **21** | **10.8%** | **-0.05 [-0.12 , 0.03]** |
| Test for overall effect: Z = 1.26 (P = 0.21) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 0.98, df = 1 (P = 0.32); I² = 0% | | | | | | |
| **1.3.3 Calorie and nutrient** | | | | | | |
| Dubois 2021[a] | -0.062644 | 0.0245 | 10 | 10 | 18.6% | -0.06 [-0.11 , -0.01] |
| Dubois 2021 | -0.032787 | 0.0245 | 10 | 10 | 18.6% | -0.03 [-0.08 , 0.02] |
| Hammond 2013 | -0.14 | 0.137755 | 152 | 81 | 0.9% | -0.14 [-0.41 , 0.13] |
| VanEpps 2016 | -0.31 | 0.142857 | 84 | 123 | 0.9% | -0.31 [-0.59 , -0.03] |
| **Subtotal** | | | **256** | **224** | **39.0%** | **-0.06 [-0.11 , -0.01]** |
| Test for overall effect: Z = 2.29 (P = 0.02) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 4.47, df = 3 (P = 0.21); I² = 33% | | | | | | |
| **Total** | | | **5633** | **4424** | **100.0%** | **-0.06 [-0.09 , -0.03]** |



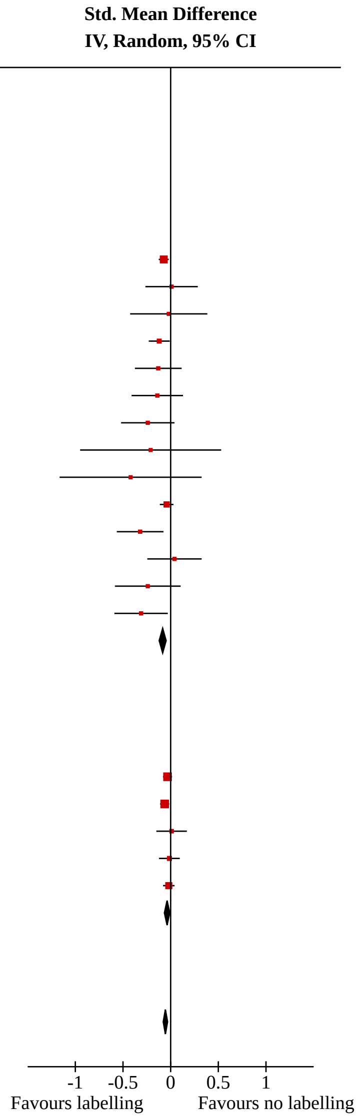

Test for overall effect: Z = 4.40 (P < 0.0001)
Test for subgroup differences: Chi² = 0.31, df = 2 (P = 0.86), I² = 0%
Heterogeneity: Tau² = 0.00; Chi² = 21.50, df = 19 (P = 0.31); I² = 12%

**Footnotes**

[a] Repeated mention of Dubois 2021 in both this and the following row indicates two eligible comparisons from this study

## Analysis 1.4. Comparison 1: Food including non-alcoholic drinks, Outcome 4: Subgroup analysis: food selection/purchasing by socioeconomic status (kcal)

| Study or Subgroup | SMD | SE | Labelling Total | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|
| **1.4.1 Low** | | | | | | |
| **Subtotal** | | | **0** | **0** | | **Not estimable** |
| Test for overall effect: Not applicable | | | | | | |
| Heterogeneity: Not applicable | | | | | | |
| **1.4.2 High** | | | | | | |
| Cawley 2020 | -0.072402 | 0.026854 | 2806 | 2745 | 17.7% | -0.07 [-0.13 , -0.02] |
| Clarke 2023a | 0.01 | 0.140306 | 101 | 104 | 0.7% | 0.01 [-0.26 , 0.28] |
| Ellison 2013 | -0.02 | 0.206632 | 108 | 30 | 0.3% | -0.02 [-0.42 , 0.38] |
| Ellison 2014 | -0.12 | 0.056122 | 1060 | 472 | 4.3% | -0.12 [-0.23 , -0.01] |
| Hammond 2013[a] | -0.13 | 0.125 | 321 | 81 | 0.9% | -0.13 [-0.37 , 0.11] |
| Hammond 2013 | -0.14 | 0.137755 | 152 | 81 | 0.7% | -0.14 [-0.41 , 0.13] |
| James 2015 | -0.24 | 0.142857 | 99 | 99 | 0.7% | -0.24 [-0.52 , 0.04] |
| Platkin 2014[b] | -0.21 | 0.377551 | 20 | 11 | 0.1% | -0.21 [-0.95 , 0.53] |
| Platkin 2014 | -0.42 | 0.380102 | 17 | 11 | 0.1% | -0.42 [-1.16 , 0.32] |
| Reynolds 2022 | -0.042178 | 0.0364 | 10 | 10 | 9.9% | -0.04 [-0.11 , 0.03] |
| Roberto 2010 | -0.32 | 0.125 | 200 | 95 | 0.9% | -0.32 [-0.56 , -0.08] |
| Roberto 2012 | 0.04 | 0.145408 | 147 | 69 | 0.6% | 0.04 [-0.24 , 0.32] |
| Robertson 2020 | -0.24 | 0.17602 | 94 | 48 | 0.4% | -0.24 [-0.58 , 0.10] |
| VanEpps 2016 | -0.31 | 0.142857 | 84 | 123 | 0.7% | -0.31 [-0.59 , -0.03] |
| **Subtotal** | | | **5219** | **3979** | **37.9%** | **-0.08 [-0.12 , -0.05]** |
| Test for overall effect: Z = 4.52 (P < 0.00001) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 12.43, df = 13 (P = 0.49); I² = 0% | | | | | | |
| **1.4.3 Both low and high** | | | | | | |
| Dubois 2021[c] | -0.032787 | 0.0245 | 10 | 10 | 21.0% | -0.03 [-0.08 , 0.02] |
| Dubois 2021 | -0.062644 | 0.0245 | 10 | 10 | 21.0% | -0.06 [-0.11 , -0.01] |
| Harnack 2008 | 0.01 | 0.081632 | 301 | 293 | 2.0% | 0.01 [-0.15 , 0.17] |
| Vasiljevic 2018 | -0.012978 | 0.0555 | 6 | 6 | 4.4% | -0.01 [-0.12 , 0.10] |
| Vasiljevic 2019 | -0.019467 | 0.0306 | 3 | 3 | 13.8% | -0.02 [-0.08 , 0.04] |
| **Subtotal** | | | **330** | **322** | **62.1%** | **-0.04 [-0.07 , -0.01]** |
| Test for overall effect: Z = 2.61 (P = 0.009) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 1.97, df = 4 (P = 0.74); I² = 0% | | | | | | |
| **Total** | | | **5549** | **4301** | **100.0%** | **-0.06 [-0.08 , -0.03]** |



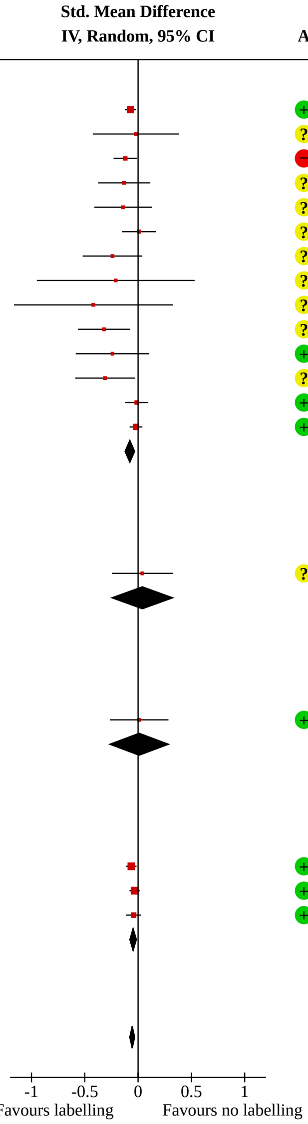

Test for overall effect: Z = 4.72 (P < 0.00001)
Test for subgroup differences: Chi² = 3.92, df = 1 (P = 0.05), I² = 74.5%
Heterogeneity: Tau² = 0.00; Chi² = 18.32, df = 18 (P = 0.43); I² = 2%

**Footnotes**

[a] Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[b] Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study
[c] Repeated mention of Dubois 2021 in both this and the following row indicates two eligible comparisons from this study

## Analysis 1.5. Comparison 1: Food including non-alcoholic drinks, Outcome 5: Additional exploratory subgroup analysis: food selection/purchasing by placement of labels

| Study or Subgroup | SMD | SE | Labelling Total | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1.5.1 On menus** | | | | | | | | | | | | |
| Cawley 2020 | -0.072402 | 0.026854 | 2806 | 2745 | 17.7% | -0.07 [-0.13 , -0.02] | + | + | + | + | + | + |
| Ellison 2013 | -0.02 | 0.206632 | 108 | 30 | 0.3% | -0.02 [-0.42 , 0.38] | ? | + | ? | + | ? | ? |
| Ellison 2014 | -0.12 | 0.056122 | 1060 | 472 | 4.3% | -0.12 [-0.23 , -0.01] | - | + | ? | + | ? | - |
| Hammond 2013[a] | -0.13 | 0.125 | 321 | 81 | 0.9% | -0.13 [-0.37 , 0.11] | ? | + | + | + | + | ? |
| Hammond 2013 | -0.14 | 0.137755 | 152 | 81 | 0.7% | -0.14 [-0.41 , 0.13] | ? | + | + | + | + | ? |
| Harnack 2008 | 0.01 | 0.081632 | 301 | 293 | 2.0% | 0.01 [-0.15 , 0.17] | ? | + | + | + | ? | ? |
| James 2015 | -0.24 | 0.142857 | 99 | 99 | 0.7% | -0.24 [-0.52 , 0.04] | ? | + | + | + | ? | ? |
| Platkin 2014[b] | -0.21 | 0.377551 | 20 | 11 | 0.1% | -0.21 [-0.95 , 0.53] | ? | + | + | + | ? | ? |
| Platkin 2014 | -0.42 | 0.380102 | 17 | 11 | 0.1% | -0.42 [-1.16 , 0.32] | ? | + | + | + | ? | ? |
| Roberto 2010 | -0.32 | 0.125 | 200 | 95 | 0.9% | -0.32 [-0.56 , -0.08] | ? | + | + | + | ? | ? |
| Robertson 2020 | -0.24 | 0.17602 | 94 | 48 | 0.4% | -0.24 [-0.58 , 0.10] | + | + | + | + | + | + |
| VanEpps 2016 | -0.31 | 0.142857 | 84 | 123 | 0.7% | -0.31 [-0.59 , -0.03] | ? | + | + | + | ? | ? |
| Vasiljevic 2018 | -0.012978 | 0.0555 | 6 | 6 | 4.4% | -0.01 [-0.12 , 0.10] | + | + | + | + | + | + |
| Vasiljevic 2019 | -0.019467 | 0.0306 | 3 | 3 | 13.8% | -0.02 [-0.08 , 0.04] | + | + | + | + | + | + |
| **Subtotal** | | | **5271** | **4098** | **46.8%** | **-0.08 [-0.12 , -0.03]** | | | | | | |
| Test for overall effect: Z = 3.38 (P = 0.0007) | | | | | | | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 16.19, df = 13 (P = 0.24); I² = 20% | | | | | | | | | | | | |
| **1.5.2 On product packaging** | | | | | | | | | | | | |
| Roberto 2012 | 0.04 | 0.145408 | 147 | 69 | 0.6% | 0.04 [-0.24 , 0.32] | ? | + | + | + | ? | ? |
| **Subtotal** | | | **147** | **69** | **0.6%** | **0.04 [-0.24 , 0.32]** | | | | | | |
| Test for overall effect: Z = 0.28 (P = 0.78) | | | | | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | | | | | |
| **1.5.3 Adjacent to product packaging** | | | | | | | | | | | | |
| Clarke 2023a | 0.01 | 0.140306 | 101 | 104 | 0.7% | 0.01 [-0.26 , 0.28] | + | + | + | + | + | + |
| **Subtotal** | | | **101** | **104** | **0.7%** | **0.01 [-0.26 , 0.28]** | | | | | | |
| Test for overall effect: Z = 0.07 (P = 0.94) | | | | | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | | | | | |
| **1.5.4 Both adjacent to and on product packaging** | | | | | | | | | | | | |
| Dubois 2021[c] | -0.062644 | 0.0245 | 10 | 10 | 21.0% | -0.06 [-0.11 , -0.01] | + | + | + | + | + | + |
| Dubois 2021 | -0.032787 | 0.0245 | 10 | 10 | 21.0% | -0.03 [-0.08 , 0.02] | + | + | + | + | + | + |
| Reynolds 2022 | -0.042178 | 0.0364 | 10 | 10 | 9.9% | -0.04 [-0.11 , 0.03] | + | + | + | + | + | + |
| **Subtotal** | | | **30** | **30** | **51.8%** | **-0.05 [-0.08 , -0.02]** | | | | | | |
| Test for overall effect: Z = 2.98 (P = 0.003) | | | | | | | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 0.76, df = 2 (P = 0.68); I² = 0% | | | | | | | | | | | | |
| **Total** | | | **5549** | **4301** | **100.0%** | **-0.06 [-0.08 , -0.03]** | | | | | | |




Test for overall effect: Z = 4.72 (P < 0.00001)
Test for subgroup differences: Chi² = 1.87, df = 3 (P = 0.60), I² = 0%
Heterogeneity: Tau² = 0.00; Chi² = 18.32, df = 18 (P = 0.43); I² = 2%

**Footnotes**

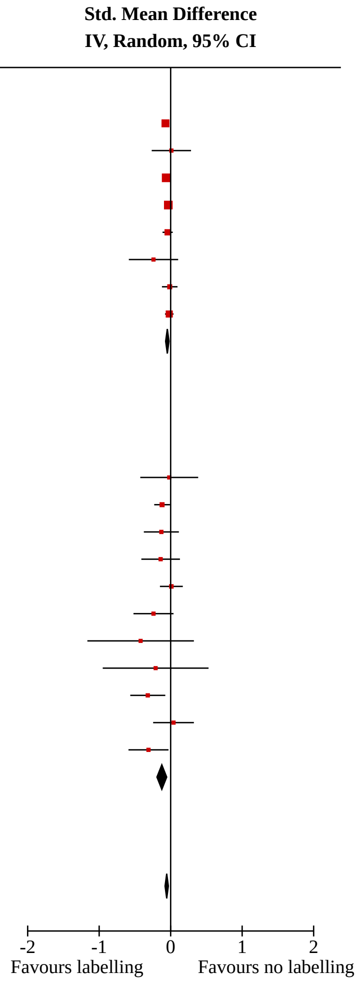
[a] Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[b] Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study
[c] Repeated mention of Dubois 2021 in both this and the following row indicates two eligible comparisons from this study

**Risk of bias legend**

(A) Bias arising from the randomization process
(B) Bias due to deviations from intended interventions
(C) Bias due to missing outcome data
(D) Bias in measurement of the outcome
(E) Bias in selection of the reported result
(F) Overall bias

## Analysis 1.6. Comparison 1: Food including non-alcoholic drinks, Outcome 6: Sensitivity analysis: food selection/purchasing by low risk of bias (kcal)

| Study or Subgroup | SMD | SE | Labelling Total | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|
| **1.6.1 Low risk of bias** | | | | | | |
| Cawley 2020 | -0.072402 | 0.026854 | 2806 | 2745 | 17.7% | -0.07 [-0.13 , -0.02] |
| Clarke 2023a | 0.01 | 0.140306 | 101 | 104 | 0.7% | 0.01 [-0.26 , 0.28] |
| Dubois 2021[a] | -0.062644 | 0.0245 | 10 | 10 | 21.0% | -0.06 [-0.11 , -0.01] |
| Dubois 2021 | -0.032787 | 0.0245 | 10 | 10 | 21.0% | -0.03 [-0.08 , 0.02] |
| Reynolds 2022 | -0.042178 | 0.0364 | 10 | 10 | 9.9% | -0.04 [-0.11 , 0.03] |
| Robertson 2020 | -0.24 | 0.17602 | 94 | 48 | 0.4% | -0.24 [-0.58 , 0.10] |
| Vasiljevic 2018 | -0.012978 | 0.0555 | 6 | 6 | 4.4% | -0.01 [-0.12 , 0.10] |
| Vasiljevic 2019 | -0.019467 | 0.0306 | 3 | 3 | 13.8% | -0.02 [-0.08 , 0.04] |
| **Subtotal** | | | **3040** | **2936** | **88.8%** | **-0.05 [-0.07 , -0.02]** |
| Test for overall effect: Z = 3.88 (P = 0.0001) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 4.21, df = 7 (P = 0.76); I² = 0% | | | | | | |
| **1.6.2 Some concerns or high risk of bias** | | | | | | |
| Ellison 2013 | -0.02 | 0.206632 | 108 | 30 | 0.3% | -0.02 [-0.42 , 0.38] |
| Ellison 2014 | -0.12 | 0.056122 | 1060 | 472 | 4.3% | -0.12 [-0.23 , -0.01] |
| Hammond 2013[b] | -0.13 | 0.125 | 152 | 81 | 0.9% | -0.13 [-0.37 , 0.11] |
| Hammond 2013 | -0.14 | 0.137755 | 321 | 81 | 0.7% | -0.14 [-0.41 , 0.13] |
| Harnack 2008 | 0.01 | 0.081632 | 301 | 293 | 2.0% | 0.01 [-0.15 , 0.17] |
| James 2015 | -0.24 | 0.142857 | 99 | 99 | 0.7% | -0.24 [-0.52 , 0.04] |
| Platkin 2014[c] | -0.42 | 0.380102 | 20 | 11 | 0.1% | -0.42 [-1.16 , 0.32] |
| Platkin 2014 | -0.21 | 0.377551 | 17 | 11 | 0.1% | -0.21 [-0.95 , 0.53] |
| Roberto 2010 | -0.32 | 0.125 | 200 | 95 | 0.9% | -0.32 [-0.56 , -0.08] |
| Roberto 2012 | 0.04 | 0.145408 | 147 | 69 | 0.6% | 0.04 [-0.24 , 0.32] |
| VanEpps 2016 | -0.31 | 0.142857 | 84 | 123 | 0.7% | -0.31 [-0.59 , -0.03] |
| **Subtotal** | | | **2509** | **1365** | **11.2%** | **-0.12 [-0.19 , -0.06]** |
| Test for overall effect: Z = 3.56 (P = 0.0004) | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 9.71, df = 10 (P = 0.47); I² = 0% | | | | | | |
| **Total** | | | **5549** | **4301** | **100.0%** | **-0.06 [-0.08 , -0.03]** |



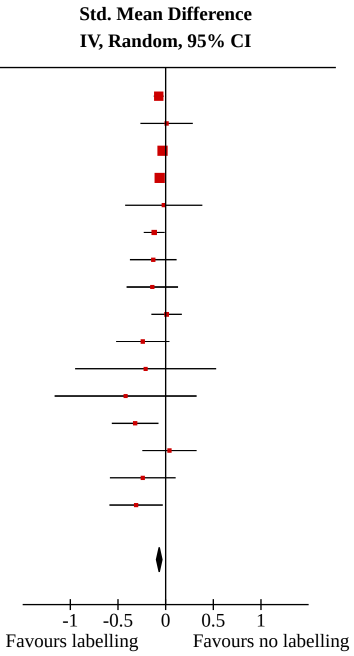

Test for overall effect: Z = 4.72 (P < 0.00001)
Test for subgroup differences: Chi² = 4.40, df = 1 (P = 0.04), I² = 77.3%
Heterogeneity: Tau² = 0.00; Chi² = 18.32, df = 18 (P = 0.43); I² = 2%

**Footnotes**

[a]Repeated mention of Dubois 2021 in both this and the following row indicates two eligible comparisons from this study
[b]Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[c]Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study

## Analysis 1.7. Comparison 1: Food including non-alcoholic drinks, Outcome 7: Sensitivity analysis: food selection/purchasing excluding Reynolds 2022, Vasiljevic 2018, Vasiljevic 2019

| Study or Subgroup | SMD | SE | Labelling Total | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|
| Cawley 2020 | -0.072402 | 0.026854 | 2806 | 2745 | 24.4% | -0.07 [-0.13 , -0.02] |
| Clarke 2023a | 0.01 | 0.140306 | 101 | 104 | 1.1% | 0.01 [-0.26 , 0.28] |
| Dubois 2021[a] | -0.032787 | 0.0245 | 10 | 10 | 28.3% | -0.03 [-0.08 , 0.02] |
| Dubois 2021 | -0.062644 | 0.0245 | 10 | 10 | 28.3% | -0.06 [-0.11 , -0.01] |
| Ellison 2013 | -0.02 | 0.206632 | 108 | 30 | 0.5% | -0.02 [-0.42 , 0.38] |
| Ellison 2014 | -0.12 | 0.056122 | 1060 | 472 | 6.4% | -0.12 [-0.23 , -0.01] |
| Hammond 2013[b] | -0.13 | 0.125 | 321 | 81 | 1.3% | -0.13 [-0.37 , 0.11] |
| Hammond 2013 | -0.14 | 0.137755 | 152 | 81 | 1.1% | -0.14 [-0.41 , 0.13] |
| Harnack 2008 | 0.01 | 0.081632 | 301 | 293 | 3.1% | 0.01 [-0.15 , 0.17] |
| James 2015 | -0.24 | 0.142857 | 99 | 99 | 1.0% | -0.24 [-0.52 , 0.04] |
| Platkin 2014[c] | -0.21 | 0.377551 | 20 | 11 | 0.1% | -0.21 [-0.95 , 0.53] |
| Platkin 2014 | -0.42 | 0.380102 | 17 | 11 | 0.1% | -0.42 [-1.16 , 0.32] |
| Roberto 2010 | -0.32 | 0.125 | 200 | 95 | 1.3% | -0.32 [-0.56 , -0.08] |
| Roberto 2012 | 0.04 | 0.145408 | 147 | 69 | 1.0% | 0.04 [-0.24 , 0.32] |
| Robertson 2020 | -0.24 | 0.17602 | 94 | 48 | 0.7% | -0.24 [-0.58 , 0.10] |
| VanEpps 2016 | -0.31 | 0.142857 | 84 | 123 | 1.0% | -0.31 [-0.59 , -0.03] |
| **Total** | | | **5530** | **4282** | **100.0%** | **-0.07 [-0.10 , -0.04]** |



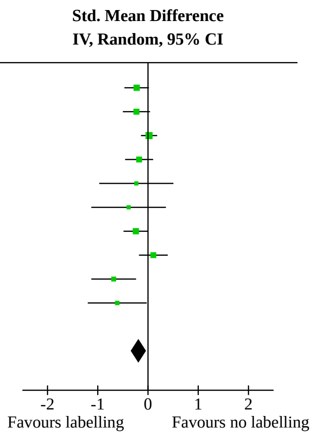

Test for overall effect: Z = 4.63 (P < 0.00001)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² = 0.00; Chi² = 15.65, df = 15 (P = 0.41); I² = 4%

**Footnotes**

[a] Repeated mention of Dubois 2021 in both this and the following row indicates two eligible comparisons from this study
[b] Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[c] Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study

## Analysis 1.8. Comparison 1: Food including non-alcoholic drinks, Outcome 8: Food consumption (kcal)

| Study or Subgroup | Labelling Mean | Labelling SD | Labelling Total | No labelling Mean | No labelling SD | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|---|
| Hammond 2013[a] | 760.043 | 359.6531 | 321 | 839.6 | 318.8 | 81 | 13.8% | -0.23 [-0.47 , 0.02] |
| Hammond 2013 | 764.9 | 326.2 | 152 | 839.6 | 318.8 | 81 | 12.6% | -0.23 [-0.50 , 0.04] |
| Harnack 2008 | 782.9226 | 391.8589 | 301 | 775.2625 | 346.8846 | 293 | 18.2% | 0.02 [-0.14 , 0.18] |
| James 2015 | 722 | 271.63 | 99 | 770 | 269.04 | 99 | 12.2% | -0.18 [-0.46 , 0.10] |
| Platkin 2014[b] | 898.8 | 391.76 | 20 | 995.4 | 429.173042 | 11 | 3.1% | -0.23 [-0.97 , 0.51] |
| Platkin 2014 | 841.3 | 366.715148 | 20 | 995.4 | 429.173042 | 11 | 3.0% | -0.39 [-1.13 , 0.36] |
| Roberto 2010 | 1293.978 | 656.3265 | 200 | 1458.92 | 724.62 | 95 | 13.8% | -0.24 [-0.49 , 0.00] |
| Roberto 2012 | 367.7284 | 225.1295 | 147 | 344.44 | 199.0877 | 69 | 11.9% | 0.11 [-0.18 , 0.39] |
| Robertson 2020 | 481 | 225.07 | 54 | 665 | 324.97 | 33 | 6.9% | -0.68 [-1.13 , -0.24] |
| Temple 2010 | 620.4 | 203.6 | 23 | 822.8 | 408.7 | 24 | 4.5% | -0.61 [-1.20 , -0.03] |
| **Total** | | | **1337** | | | **797** | **100.0%** | **-0.19 [-0.33 , -0.05]** |



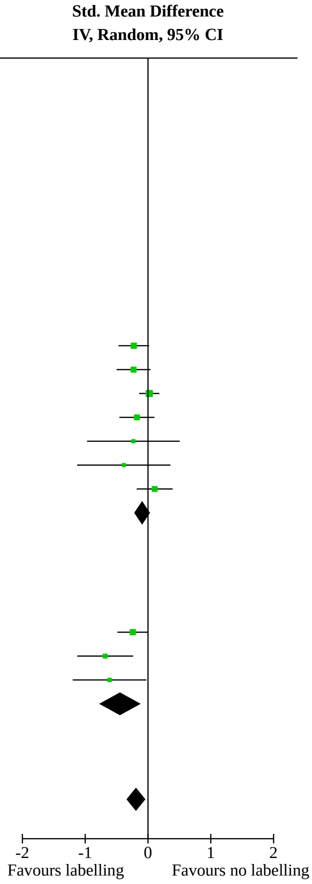

Test for overall effect: Z = 2.72 (P = 0.006)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² = 0.02; Chi² = 17.00, df = 9 (P = 0.05); I² = 47%

**Footnotes**

[a] Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[b] Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study

## Analysis 1.9. Comparison 1: Food including non-alcoholic drinks, Outcome 9: Subgroup analysis: food consumption by setting (kcal)

| Study or Subgroup | Labelling Mean | Labelling SD | Labelling Total | No labelling Mean | No labelling SD | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|---|
| **1.9.1 Restaurant** | | | | | | | | |
| **Subtotal** | | | **0** | | | **0** | | **Not estimable** |
| Test for overall effect: Not applicable | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | |
| **1.9.2 Store** | | | | | | | | |
| **Subtotal** | | | **0** | | | **0** | | **Not estimable** |
| Test for overall effect: Not applicable | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | |
| **1.9.3 Naturalistic** | | | | | | | | |
| Hammond 2013[a] | 760.043 | 359.6531 | 321 | 839.6 | 318.8 | 81 | 13.8% | -0.23 [-0.47 , 0.02] |
| Hammond 2013 | 764.9 | 326.2 | 152 | 839.6 | 318.8 | 81 | 12.6% | -0.23 [-0.50 , 0.04] |
| Harnack 2008 | 782.9226 | 391.8589 | 301 | 775.2625 | 346.8846 | 293 | 18.2% | 0.02 [-0.14 , 0.18] |
| James 2015 | 722 | 271.63 | 99 | 770 | 269.04 | 99 | 12.2% | -0.18 [-0.46 , 0.10] |
| Platkin 2014[b] | 898.8 | 391.76 | 20 | 995.4 | 429.173042 | 11 | 3.1% | -0.23 [-0.97 , 0.51] |
| Platkin 2014 | 841.3 | 366.715148 | 20 | 995.4 | 429.173042 | 11 | 3.0% | -0.39 [-1.13 , 0.36] |
| Roberto 2012 | 367.7284 | 225.1295 | 147 | 344.44 | 199.0877 | 69 | 11.9% | 0.11 [-0.18 , 0.39] |
| **Subtotal** | | | **1060** | | | **645** | **74.8%** | **-0.09 [-0.21 , 0.02]** |
| Test for overall effect: Z = 1.61 (P = 0.11) | | | | | | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 6.96, df = 6 (P = 0.32); I² = 14% | | | | | | | | |
| **1.9.4 Laboratory** | | | | | | | | |
| Roberto 2010 | 1293.978 | 656.3265 | 200 | 1458.92 | 724.62 | 95 | 13.8% | -0.24 [-0.49 , 0.00] |
| Robertson 2020 | 481 | 225.07 | 54 | 665 | 324.97 | 33 | 6.9% | -0.68 [-1.13 , -0.24] |
| Temple 2010 | 620.4 | 203.6 | 23 | 822.8 | 408.7 | 24 | 4.5% | -0.61 [-1.20 , -0.03] |
| **Subtotal** | | | **277** | | | **152** | **25.2%** | **-0.45 [-0.76 , -0.14]** |
| Test for overall effect: Z = 2.81 (P = 0.005) | | | | | | | | |
| Heterogeneity: Tau² = 0.03; Chi² = 3.59, df = 2 (P = 0.17); I² = 44% | | | | | | | | |
| **Total** | | | **1337** | | | **797** | **100.0%** | **-0.19 [-0.33 , -0.05]** |




Test for overall effect: Z = 2.72 (P = 0.006)
Test for subgroup differences: Chi² = 4.34, df = 1 (P = 0.04), I² = 76.9%
Heterogeneity: Tau² = 0.02; Chi² = 17.00, df = 9 (P = 0.05); I² = 47%

**Footnotes**

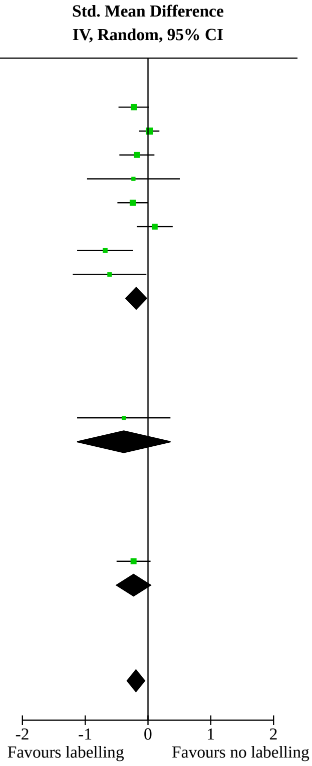
[a]Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study

[b]Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study

## Analysis 1.10. Comparison 1: Food including non-alcoholic drinks, Outcome 10: Subgroup analysis: food consumption by label type (kcal)

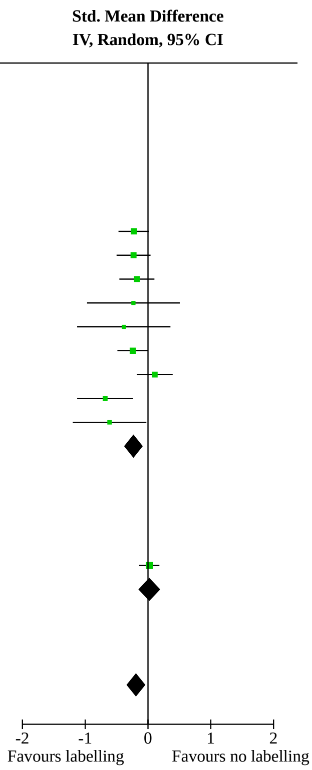
| Study or Subgroup | Labelling Mean | Labelling SD | Labelling Total | No labelling Mean | No labelling SD | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|---|
| **1.10.1 Calorie** | | | | | | | | |
| Hammond 2013 | 760.043 | 359.6531 | 321 | 839.6 | 318.8 | 81 | 13.8% | -0.23 [-0.47 , 0.02] |
| Harnack 2008 | 782.9226 | 391.8589 | 301 | 775.2625 | 346.8846 | 293 | 18.2% | 0.02 [-0.14 , 0.18] |
| James 2015 | 722 | 271.63 | 99 | 770 | 269.04 | 99 | 12.2% | -0.18 [-0.46 , 0.10] |
| Platkin 2014 | 898.8 | 391.76 | 20 | 995.4 | 429.173042 | 11 | 3.1% | -0.23 [-0.97 , 0.51] |
| Roberto 2010 | 1293.978 | 656.3265 | 200 | 1458.92 | 724.62 | 95 | 13.8% | -0.24 [-0.49 , 0.00] |
| Roberto 2012 | 367.7284 | 225.1295 | 147 | 344.44 | 199.0877 | 69 | 11.9% | 0.11 [-0.18 , 0.39] |
| Robertson 2020 | 481 | 225.07 | 54 | 665 | 324.97 | 33 | 6.9% | -0.68 [-1.13 , -0.24] |
| Temple 2010 | 620.4 | 203.6 | 23 | 822.8 | 408.7 | 24 | 4.5% | -0.61 [-1.20 , -0.03] |
| **Subtotal** | | | **1165** | | | **705** | **84.4%** | **-0.19 [-0.35 , -0.02]** |
| Test for overall effect: Z = 2.24 (P = 0.02) | | | | | | | | |
| Heterogeneity: Tau² = 0.03; Chi² = 16.09, df = 7 (P = 0.02); I² = 57% | | | | | | | | |
| **1.10.2 Calorie with PACE** | | | | | | | | |
| Platkin 2014 | 841.3 | 366.715148 | 20 | 995.4 | 429.173042 | 11 | 3.0% | -0.39 [-1.13 , 0.36] |
| **Subtotal** | | | **20** | | | **11** | **3.0%** | **-0.39 [-1.13 , 0.36]** |
| Test for overall effect: Z = 1.02 (P = 0.31) | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | |
| **1.10.3 Calorie and nutrient** | | | | | | | | |
| Hammond 2013 | 764.9 | 326.2 | 152 | 839.6 | 318.8 | 81 | 12.6% | -0.23 [-0.50 , 0.04] |
| **Subtotal** | | | **152** | | | **81** | **12.6%** | **-0.23 [-0.50 , 0.04]** |
| Test for overall effect: Z = 1.67 (P = 0.10) | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | |
| **Total** | | | **1337** | | | **797** | **100.0%** | **-0.19 [-0.33 , -0.05]** |

Test for overall effect: Z = 2.72 (P = 0.006)
Test for subgroup differences: Chi² = 0.30, df = 2 (P = 0.86), I² = 0%
Heterogeneity: Tau² = 0.02; Chi² = 17.00, df = 9 (P = 0.05); I² = 47%



## Analysis 1.11. Comparison 1: Food including non-alcoholic drinks, Outcome 11: Subgroup analysis: food consumption by socioeconomic status (kcal)

| Study or Subgroup | Labelling Mean | Labelling SD | Labelling Total | No labelling Mean | No labelling SD | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|---|
| **1.11.1 Low** | | | | | | | | |
| **Subtotal** | | | **0** | | | **0** | | **Not estimable** |
| Test for overall effect: Not applicable | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | |
| **1.11.2 High** | | | | | | | | |
| Hammond 2013[a] | 760.043 | 359.6531 | 321 | 839.6 | 318.8 | 81 | 13.8% | -0.23 [-0.47 , 0.02] |
| Hammond 2013 | 764.9 | 326.2 | 152 | 839.6 | 318.8 | 81 | 12.6% | -0.23 [-0.50 , 0.04] |
| James 2015 | 722 | 271.63 | 99 | 770 | 269.04 | 99 | 12.2% | -0.18 [-0.46 , 0.10] |
| Platkin 2014[b] | 898.8 | 391.76 | 20 | 995.4 | 429.173042 | 11 | 3.1% | -0.23 [-0.97 , 0.51] |
| Platkin 2014 | 841.3 | 366.715148 | 20 | 995.4 | 429.173042 | 11 | 3.0% | -0.39 [-1.13 , 0.36] |
| Roberto 2010 | 1293.978 | 656.3265 | 200 | 1458.92 | 724.62 | 95 | 13.8% | -0.24 [-0.49 , 0.00] |
| Roberto 2012 | 367.7284 | 225.1295 | 147 | 344.44 | 199.0877 | 69 | 11.9% | 0.11 [-0.18 , 0.39] |
| Robertson 2020 | 481 | 225.07 | 54 | 665 | 324.97 | 33 | 6.9% | -0.68 [-1.13 , -0.24] |
| Temple 2010 | 620.4 | 203.6 | 23 | 822.8 | 408.7 | 24 | 4.5% | -0.61 [-1.20 , -0.03] |
| **Subtotal** | | | **1036** | | | **504** | **81.8%** | **-0.23 [-0.37 , -0.10]** |
| Test for overall effect: Z = 3.32 (P = 0.0009) | | | | | | | | |
| Heterogeneity: Tau² = 0.01; Chi² = 11.20, df = 8 (P = 0.19); I² = 29% | | | | | | | | |
| **1.11.3 Both low and high** | | | | | | | | |
| Harnack 2008 | 782.9226 | 391.8589 | 301 | 775.2625 | 346.8846 | 293 | 18.2% | 0.02 [-0.14 , 0.18] |
| **Subtotal** | | | **301** | | | **293** | **18.2%** | **0.02 [-0.14 , 0.18]** |
| Test for overall effect: Z = 0.25 (P = 0.80) | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | |
| **Total** | | | **1337** | | | **797** | **100.0%** | **-0.19 [-0.33 , -0.05]** |



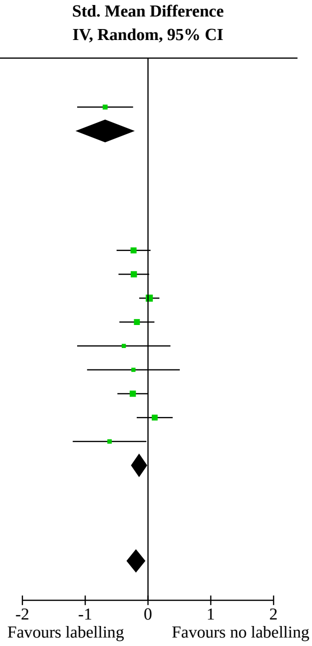

Test for overall effect: Z = 2.72 (P = 0.006)
Test for subgroup differences: Chi² = 5.50, df = 1 (P = 0.02), I² = 81.8%
Heterogeneity: Tau² = 0.02; Chi² = 17.00, df = 9 (P = 0.05); I² = 47%

**Footnotes**

[a] Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[b] Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study

### Analysis 1.12. Comparison 1: Food including non-alcoholic drinks, Outcome 12: Sensitivity analysis: food consumption by low risk of bias (kcal)

| Study or Subgroup | Labelling Mean | Labelling SD | Labelling Total | No labelling Mean | No labelling SD | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|---|
| **1.12.1 Low risk of bias** | | | | | | | | |
| Robertson 2020 | 481 | 225.07 | 54 | 665 | 324.97 | 33 | 6.9% | -0.68 [-1.13 , -0.24] |
| **Subtotal** | | | **54** | | | **33** | **6.9%** | **-0.68 [-1.13 , -0.24]** |
| Test for overall effect: Z = 3.00 (P = 0.003) | | | | | | | | |
| Heterogeneity: Not applicable | | | | | | | | |
| **1.12.2 Some concerns or high risk of bias** | | | | | | | | |
| Hammond 2013[a] | 764.9 | 326.2 | 152 | 839.6 | 318.8 | 81 | 12.6% | -0.23 [-0.50 , 0.04] |
| Hammond 2013 | 760.043 | 359.6531 | 321 | 839.6 | 318.8 | 81 | 13.8% | -0.23 [-0.47 , 0.02] |
| Harnack 2008 | 782.9226 | 391.8589 | 301 | 775.2625 | 346.8846 | 293 | 18.2% | 0.02 [-0.14 , 0.18] |
| James 2015 | 722 | 271.63 | 99 | 770 | 269.04 | 99 | 12.2% | -0.18 [-0.46 , 0.10] |
| Platkin 2014[b] | 841.3 | 366.715148 | 20 | 995.4 | 429.173042 | 11 | 3.0% | -0.39 [-1.13 , 0.36] |
| Platkin 2014 | 898.8 | 391.76 | 20 | 995.4 | 429.173042 | 11 | 3.1% | -0.23 [-0.97 , 0.51] |
| Roberto 2010 | 1293.978 | 656.3265 | 200 | 1458.92 | 724.62 | 95 | 13.8% | -0.24 [-0.49 , 0.00] |
| Roberto 2012 | 367.7284 | 225.1295 | 147 | 344.44 | 199.0877 | 69 | 11.9% | 0.11 [-0.18 , 0.39] |
| Temple 2010 | 620.4 | 203.6 | 23 | 822.8 | 408.7 | 24 | 4.5% | -0.61 [-1.20 , -0.03] |
| **Subtotal** | | | **1283** | | | **764** | **93.1%** | **-0.14 [-0.26 , -0.02]** |
| Test for overall effect: Z = 2.36 (P = 0.02) | | | | | | | | |
| Heterogeneity: Tau² = 0.01; Chi² = 11.11, df = 8 (P = 0.20); I² = 28% | | | | | | | | |
| **Total** | | | **1337** | | | **797** | **100.0%** | **-0.19 [-0.33 , -0.05]** |




Test for overall effect: Z = 2.72 (P = 0.006)
Test for subgroup differences: Chi² = 5.29, df = 1 (P = 0.02), I² = 81.1%
Heterogeneity: Tau² = 0.02; Chi² = 17.00, df = 9 (P = 0.05); I² = 47%

**Footnotes**

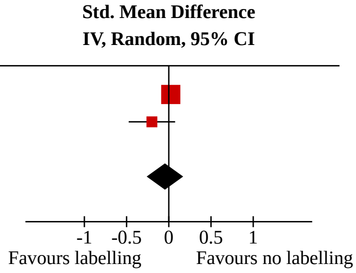
[a] Repeated mention of Hammond 2013 in both this and the following row indicates two eligible comparisons from this study
[b] Repeated mention of Platkin 2014 in both this and the following row indicates two eligible comparisons from this study

## Comparison 2. Alcoholic drinks

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 2.1 Alcohol selection/purchasing of energy (kcal) | 2 | 5756 | Std. Mean Difference (IV, Random, 95% CI) | -0.05 [-0.25, 0.16] |
| 2.2 Alcohol selection/purchasing of alcohol (units) | 1 | | Std. Mean Difference (IV, Random, 95% CI) | Subtotals only |

### Analysis 2.1. Comparison 2: Alcoholic drinks, Outcome 1: Alcohol selection/purchasing of energy (kcal)

| Study or Subgroup | SMD | SE | Labelling Total | No labelling Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|
| Cawley 2020 | 0.023369 | 0.026846 | 2806 | 2745 | 69.0% | 0.02 [-0.03 , 0.08] |
| Clarke 2023a | -0.2 | 0.140306 | 101 | 104 | 31.0% | -0.20 [-0.47 , 0.07] |
| **Total** | | | **2907** | **2849** | **100.0%** | **-0.05 [-0.25 , 0.16]** |



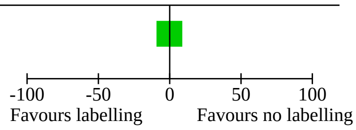

Test for overall effect: Z = 0.44 (P = 0.66)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² = 0.01; Chi² = 2.44, df = 1 (P = 0.12); I² = 59%

### Analysis 2.2. Comparison 2: Alcoholic drinks, Outcome 2: Alcohol selection/purchasing of alcohol (units)

| Study or Subgroup | Labelling Mean | Labelling SD | Labelling Total | No labelling Mean | No labelling SD | No labelling Total | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|
| Clarke 2023a | 19.6 | 20.9 | 101 | 23.7 | 17.5 | 104 | -0.21 [-0.49 , 0.06] |

Test for subgroup differences: Not applicable

Std. Mean Difference
IV, Random, 95% CI



# ADDITIONAL TABLES

**Table 1. Studies not included in the meta-analyses (non-randomised studies and randomised studies that did not report data suitable for meta-analysis)**

| Study ID | Participant characteristics | Study characteristics | Intervention characteristics | Potential key modifiers | Comparison group | Primary outcome | Results | Direction of effect[a] |
|---|---|---|---|---|---|---|---|---|
| **Non-randomised studies** | | | | | | | | |
| Bollinger 2011 | Starbucks customers | Controlled before-and-after study in 222 Starbucks coffee shops in New York City (intervention sites), and 94 Starbucks coffee shops in Boston and Philadelphia (control sites), USA (field setting). Data for all transactions for 3 months before and 11 months after energy posting commenced (over 100 million transactions in the data set) | Calorie labelling on menus | — | Menus without calorie labelling | Selection with purchasing | The study found a 5.8% decrease in average calories per transaction, equivalent to 14.4 calories (P = 0.001). | 1 |
| Elbel 2009 | Fast-food restaurant customers | Controlled before-after study in 19 fast-food restaurants in New York City and Newark, USA (field setting).<br><br>Data for all transactions from 2 weeks before to 2 weeks after. | Calorie labelling on menus | — | Menus without calorie labelling | Selection with purchasing | No significant differences in purchasing behaviour in 5 control sites and 14 intervention sites, over a 2-week period after calorie labelling was implemented, with an increase of 21 calories after labelling in 1 city (New York City) and 3 calories in another that did not implement calorie labelling (Newark). | 0 |
| Elshiewy 2018 | Supermarket customers: 188,062 loyalty card members | Interrupted time series study in 2000 supermarkets in the UK (field setting), 1 year before and 1 year after the introduction of calorie labels | Front of pack calorie labelling on products (store branded cookies, breakfast ce- | — | Typical back of pack nutrition labelling | Selection with purchasing | There was a 9.5% decrease in the number of calories purchased, demonstrated by a significant main effect for calories (β = −0.10, 95% CI −0.12 to −0.08). A simplified front-of-pack nutrition label led to healthier purchases for the labelled products compared with | 1 |

**Table 1. Studies not included in the meta-analyses (non-randomised studies and randomised studies that did not report data suitable for meta-analysis)** *(Continued)*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | reals and soft drinks) | | | | when only back of pack nutrition information was available. | |
| Fichera 2020 | 20,707 households (data on all grocery purchases used from 2005 to 2008) | Controlled before-after study over a 3-year period in 9 supermarket chains in the UK (field setting), before and after the introduction of calorie labels in 4 of these chains. | Front of pack calorie labelling with ≥ 1 other nutrient on products recommended for labelling by the Food Standards Agency (including ready meals, burgers, pies, breaded/coated meats, pizza, sandwiches and cereals, dairy foods, cookies) | Differences by sociodemographic characteristics | Typical back of pack nutrition labelling | Selection with purchasing | Significant reduction in total monthly calories from store-brand labelled food purchases of 588 kcal (standard error 77.22; P < 0.01). There was also a reduction observed in energy purchased when considered across all products (labelled and unlabelled).<br><br>Authors stated that while both low and high social class households reduced the quantity purchased of store-brand labelled foods following the introduction of front-of-pack labels, this reduction was larger for the lower compared to the higher social classes. | 1 |
| Petimar 2019 | 5582 restaurant customers | Controlled before-after study over a 4-year period in McDonalds chains in 4 New England cities, USA (field setting), compared to 5 control chains (restaurants similar to McDonalds). | Calorie labelling on menus | — | Menus without calorie labelling | Selection with purchasing | There was a decrease in energy purchased, but calorie labelling was not associated with a reduction in energy purchased in intervention compared to control chains (adults: −19 calories, 95% CI −112 to 75; children: −13 calories, 95% CI −108 to 135; adolescents: −49 calories, 95% CI −126 to 38). | 1 |
| Petimar 2021 | Fast-food restaurant customers | Interrupted time series study over 4 years in 104 fast-food restaurants in southern USA (field setting) | Calorie labelling on menus | — | Menus without calorie labelling | Selection with purchasing | The model-based predicted mean energy reduction per transaction was 4.7% lower (−73 calories per transaction, 95% CI −81 to −65). | 1 |
| **Randomised studies not included in the meta-analyses** | | | | | | | | |

**Table 1. Studies not included in the meta-analyses (non-randomised studies and randomised studies that did not report data suitable for meta-analysis)** *(Continued)*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Berry 2019[b] | 233 people aged > 18 | Between-participants quasi-randomised control trial in a restaurant in South USA (field setting) | Calorie labelling on menus with menus altered by day of week | — | Menus without calorie labelling | Selection with purchasing | Calorie information<br><br>provision alone had no effect on calories ordered. | — |
| Oliveira 2018[c] | 233 students aged > over 20 years (16 groups) | Between-participants cluster-randomised controlled trial in a university campus restaurant in Brazil (field setting) | Calorie labelling with ≥ 1 other nutrient on menus | — | Menus without calorie labelling | Selection with purchasing | There were no differences reported in number of healthy food choices. The mean number of healthy items<br><br>chosen by participants in the control group was 5.6 items (67.6% of the chosen items), and in the calorie labelling group was 5.4 items (67.5% of the chosen items). | — |

[a] Value for sign test: 1 = favours intervention (benefit); 0 = favours control (harm).

[b] Berry 2019 could not be included in the meta-analysis as numbers per group were not reported and as groups were allocated by day of the week, we could not assume that groups would be adequately randomised. Additionally, it is not clear that the data per group were independent of the other groups (i.e. the same participants may have been randomised to multiple groups). We contacted the study authors but received no response.

[c] Oliveira 2018 could not be included in the meta-analysis as it was not possible to calculate energy purchased (only the number of healthy items).

# APPENDICES

## Appendix 1. Handling of studies identified by the updated 'top-up' search

The updated 'top-up' search conducted in September 2023, covering the period up to 25 August 2023, identified one eligible study: Petimar 2022. While being provisionally accepted into the review, this study is yet to be fully integrated until the next review update. This provisionally accepted study is retained in Studies awaiting classification.

The decision to not fully integrate further results of newly included studies at this stage was a pragmatic one, based on a balance of the likelihood of this study changing results and conclusions, relative to the potential disadvantages of delaying the review further. The former was judged to be highly improbable as it is a non-randomised study that would not be included in the meta-analysis for its outcome (and for which evidence was already judged at high certainty), and its data appear to be consistent with the data for randomised and non-randomised studies that are included in the review: 2/3 food categories with calorie labelling being applied saw a reduction in calories per transaction purchased (there was no change in the third category) and the study concluded that "calorie labeling of prepared foods was associated with small to moderate decreases in calories purchased from prepared bakery and deli items without evidence of substitution to similar packaged foods."

Relative to the lack of any clear benefit from including these data, we considered the potential disadvantages of delaying the review to be greater, given calorie labelling is currently of high policy interest globally, including the active development and implementation of policies in multiple jurisdictions (see Why it is important to do this review).

In addition, the updated 'top-up' searches identified four studies: Elshiewy 2022; Lee 2018; van Doorn 2023; Zhu 2023, for which we could not confidently determine their eligibility from available materials and thus these are also grouped as Studies awaiting classification. Each of these will need correspondence with the study investigators or further expert statistical input to determine whether the design and analysis of the study is appropriate and confers eligibility and whether there is sufficient useable data or whether additional new analysis either by the original authors or the review team would be required. We considered further work to assess and then potentially include these studies would delay the review, including risking that the review would not be completed without a further update to the searches being required.

In summary, the five aforementioned eligible or potentially eligible studies that were identified in the updated 'top-up' search in September 2023, namely Elshiewy 2022; Lee 2018; Petimar 2022; van Doorn 2023; Zhu 2023, are categorised as 'studies awaiting classification' (see Studies awaiting classification table), and they will be further assessed and considered for inclusion at the time of the next review update.

## Appendix 2. Search strategies

**MEDLINE (OvidSP), 1946-present**

**Alcohol labelling search applied to whole period from inception; food labelling search was an updated search and was limited to period from 25th April 2017. Search executed: 9th July 2021**

***Food labelling***

| | |
|---|---|
| 1 | exp Food packaging/ and (label$ or content$ sign$ or symbol$ or ticket$ or sticker$ or diet$ or health$ or calori$ or nutritio$ or guideline daily amount$ or recommended daily amount$ or nutrient reference value$ or nutrient daily value$).ti,ab. |
| 2 | food pack$.ti,ab. |
| 3 | exp Product labelling/ and (food$ or diet$ or health$ or calori$ or nutritio$ or guideline daily amount$ or recommended daily amount$ or nutrient reference value$ or nutrient daily value$ or snack$ or eat$).ti,ab. |
| 4 | exp Food Labeling/ |
| 5 | ((nutritio$ or nutrient$) adj5 (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 6 | (nutrition$ information or nutrient$ information).ti,ab. |
| 7 | (Food$ label$ or food$ content$ label$ or food$ content$ sign$ or food$ content symbol$ or food$ content$ tag$ or food$ content$ ticket $ or food$ content$ sticker$).ti,ab. |

*(Continued)*

| | |
|---|---|
| 8 | traffic light$.ti,ab. |
| 9 | (guideline daily amount$ or nutrient reference value$ or nutrient daily value$).ti,ab. |
| 10 | (recommended dietary allowance$ adj5 (label$ or content$ sign$ or symbol$ or information or ticket$ or sticker$)).ti,ab. |
| 11 | ((Calorific or calorie$ or caloric) and (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 12 | ((Calorific or calorie$ or caloric) adj information).ti,ab. |
| 13 | (menu and (label$ or content$ sign$ or symbol$ or tag$ or ticket$ or sticker$)).ti,ab. |
| 14 | (menu and (nutritional content$ or nutritional information or traffic light or guideline daily amount or GDA or healthy choice or calorie)).ti,ab. |
| 15 | (Label$ adj2 (legislation$ or regulation$ or policies or policy)).ti,ab. |
| 16 | Healthy choice.ti,ab. |
| 17 | exp Product labelling/ and (drink? or beverage? or soda? or flavo?red water? or fruitwater? or cordial? or squash? or juice? or smoothie? or milkshake? or tea or teas or coffee?).ti,ab. |
| 18 | (Drink$ label$ or Drink$ content$ label$ or Drink$ content$ sign$ or Drink$ content symbol$ or Drink$ content$ tag$ or Drink$ content $ ticket$ or Drink$ content$ sticker$).ti,ab. |
| 19 | ((drink? or beverage? or soda? or flavo?red water? or fruit water? or cordial? or squash? or juice? or smoothie? or milkshake? or tea or teas or coffee?) and (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 20 | or/1-19 |
| 21 | exp Food Preferences/ |
| 22 | exp Food Habits/ |
| 23 | exp Feeding Behavior/ |
| 24 | exp Eating/ |
| 25 | exp Diet/ |
| 26 | exp Choice Behavior/ |
| 27 | (intak$ or consume or consumes or consumption or consumed or eat$ or diet$).ti,ab. |
| 28 | (food adj5 (preference$ or habit$ or behavio?r$ or choice$ or decision$ or decid$ or inclin$ or lik$ or choos$ or select$ or pick$)).ti,ab. |
| 29 | ((drink? or beverage?) adj5 (preference$ or habit$ or behavio?r$ or choice$ or decision$ or decid$ or inclin$ or lik$ or choos$ or select $ or pick$)).ti,ab. |
| 30 | or/21-29 |
| 31 | exp Restaurants/ |

*(Continued)*

| | |
|---|---|
| 32 | (purchas$ or buy$ or sale$ or vend$ or sell$).ti,ab. |
| 33 | (shop$ or store$ or supermarket$ or market$ or outlet$ or retailer$ or point of purchase).ti,ab. |
| 34 | (restaurant$ or cafe$ or bar$ or canteen$ or cafeteria$ or dinner hall$ or dining area$ or dining room$ or refector$ or eatery or mess or buffet or bistro$ or eating place$).ti,ab. |
| 35 | or/31-34 |
| 36 | 20 and (30 or 35) |
| 37 | limit 36 to dt=20170425-20210709 |

***Alcohol labelling***

| | |
|---|---|
| 1 | exp Alcoholic Beverages/ |
| 2 | (dr#nk* or beverage* or alcohol* or beer* or lager* or wine* or cider*).ti,ab. |
| 3 | 1 or 2 |
| 4 | exp Product Labeling/ |
| 5 | ((alcohol* or drink*) adj5 (unit? or guideline* or standard drink*)).ti,ab. |
| 6 | ((calorie* or nutrition* or energy or ingredient*) adj5 (label* or inform* or menu* or poster* or glass* or beermat* or bottle* or packag*)).ti,ab. |
| 7 | 4 or 5 or 6 |
| 8 | exp Alcohol Drinking/ |
| 9 | ((purchas* or pour* or select* or consum*) adj5 (alcohol* or drink* or beer* or wine* or lager* or cider*)).ti,ab. |
| 10 | exp Health Knowledge, Attitudes, Practice/ |
| 11 | exp Energy Intake/ |
| 12 | exp Consumer Behavior/ |
| 13 | 8 or 9 or 10 or 11 or 12 |
| 14 | 3 and 7 and 13 |

**Embase (OvidSP), 1974-present**

**Alcohol labelling search applied to whole period from inception; food labelling search was an updated search and was limited to period from 25th April 2017. Search executed: 9th July 2021**

***Food labelling***

| | |
|---|---|
| 1 | food pack$.ti,ab. |
| 2 | Packaging/ and (food$ or diet$ or health$ or calori$ or nutritio$ or guideline daily amount$ or recommended daily amount$ or nutrient reference value$ or nutrient daily value$ or snack$ or eat $).ti,ab. |
| 3 | ((nutritio$ or nutrient$) adj5 (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 4 | (nutrition$ information or nutrient$ information).ti,ab. |
| 5 | (Food$ label$ or food$ content$ label$ or food$ content$ sign$ or food$ content symbol$ or food$ content$ tag$ or food$ content$ ticket $ or food$ content$ sticker$).ti,ab. |
| 6 | traffic light$.ti,ab. |
| 7 | (guideline daily amount$ or nutrient reference value$ or nutrient daily value$).ti,ab. |
| 8 | (recommended dietary allowance$ adj5 (label$ or content$ sign$ or symbol$ or information or ticket$ or sticker$)).ti,ab. |
| 9 | ((Calorific or calorie$ or caloric) and (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 10 | ((Calorific or calorie$ or caloric) adj information).ti,ab. |
| 11 | (menu and (label$ or content$ sign$ or symbol$ or tag$ or ticket$ or sticker$)).ti,ab. |
| 12 | (menu and (nutritional content$ or nutritional information or traffic light or guideline daily amount or GDA or healthy choice or calorie)).ti,ab. |
| 13 | (Label$ adj2 (legislation$ or regulation$ or policies or policy)).ti,ab. |
| 14 | Healthy choice.ti,ab. |
| 15 | exp Product labelling/ and (drink? or beverage? or soda? or flavo?red water? or fruitwater? or cordial? or squash? or juice? or smoothie? or milkshake? or tea or teas or coffee?).ti,ab. |
| 16 | (Drink$ label$ or Drink$ content$ label$ or Drink$ content$ sign$ or Drink$ content symbol$ or Drink$ content$ tag$ or Drink$ content $ ticket$ or Drink$ content$ sticker$).ti,ab. |
| 17 | ((drink? or beverage? or soda? or flavo?red water? or fruit water? or cordial? or squash? or juice? or smoothie? or milkshake? or tea or teas or coffee?) and (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 18 | or/1-17 |
| 19 | exp Feeding Behavior/ |
| 20 | exp Eating/ |
| 21 | exp Diet/ |
| 22 | *decision making/ |
| 23 | (intak$ or consume or consumes or consumption or consumed or eat$ or diet$).ti,ab. |

*(Continued)*

| | |
|---|---|
| 24 | (food adj5 (preference$ or habit$ or behavio?r$ or choice$ or decision$ or decid$ or inclin$ or lik$ or choos$ or select$ or pick$)).ti,ab. |
| 25 | ((drink? or beverage?) adj5 (preference$ or habit$ or behavio?r$ or choice$ or decision$ or decid$ or inclin$ or lik$ or choos$ or select $ or pick$)).ti,ab. |
| 26 | or/19-25 |
| 27 | Restaurant/ |
| 28 | (purchas$ or buy$ or sale$ or vend$ or sell$).ti,ab. |
| 29 | (shop$ or store$ or supermarket$ or market$ or outlet$ or retailer$ or point of purchase).ti,ab. |
| 30 | (restaurant$ or cafe$ or bar$ or canteen$ or cafeteria$ or dinner hall$ or dining area$ or dining room$ or refector$ or eatery or mess or buffet or bistro$ or eating place$).ti,ab. |
| 31 | or/27-30 |
| 32 | 18 and (26 or 31) |
| 33 | limit 32 to dc=20170425-20210709 |

***Alcohol labelling***

| | |
|---|---|
| 1 | exp Alcoholic Beverage/ |
| 2 | (dr#nk* or beverage* or alcohol* or beer* or lager* or wine* or cider*).ti,ab. |
| 3 | 1 or 2 |
| 4 | Packaging/ |
| 5 | ((alcohol* or drink*) adj5 (unit? or guideline* or standard drink*)).ti,ab. |
| 6 | ((calorie* or nutrition* or energy or ingredient*) adj5 (label* or inform* or menu* or poster* or glass* or beermat* or bottle* or packag*)).ti,ab. |
| 7 | 4 or 5 or 6 |
| 8 | exp drinking behavior/ |
| 9 | ((purchas* or pour* or select* or consum*) adj5 (alcohol* or drink* or beer* or wine* or lager* or cider*)).ti,ab. |
| 10 | attitude to health/ or attitude to life/ |
| 11 | Caloric Intake/ |
| 12 | consumer attitude/ |
| 13 | 8 or 9 or 10 or 11 or 12 |

*(Continued)*

| | |
|---|---|
| 14 | 3 and 7 and 13 |

**PsycINFO (OvidSP), 1806-present**

**Alcohol labelling search applied to whole period from inception; food labelling search was an updated search and was limited to period from 25th April 2017. Search executed: 9th July 2021**

***Food labelling***

| | |
|---|---|
| 1 | food pack$.ti,ab. |
| 2 | ((nutritio$ or nutrient$) adj5 (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 3 | (nutrition$ information or nutrient$ information).ti,ab. |
| 4 | (Food$ label$ or food$ content$ label$ or food$ content$ sign$ or food$ content symbol$ or food$ content$ tag$ or food$ content$ ticket $ or food$ content$ sticker$).ti,ab. |
| 5 | traffic light$.ti,ab. |
| 6 | (guideline daily amount$ or nutrient reference value$ or nutrient daily value$).ti,ab. |
| 7 | (recommended dietary allowance$ adj5 (label$ or content$ sign$ or symbol$ or information or ticket$ or sticker$)).ti,ab. |
| 8 | ((Calorific or calorie$ or caloric) and (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 9 | ((Calorific or calorie$ or caloric) adj information).ti,ab. |
| 10 | (menu and (label$ or content$ sign$ or symbol$ or tag$ or ticket$ or sticker$)).ti,ab. |
| 11 | (menu and (nutritional content$ or nutritional information or traffic light or guideline daily amount or GDA or healthy choice or calorie)).ti,ab. |
| 12 | (Label$ adj2 (legislation$ or regulation$ or policies or policy)).ti,ab. |
| 13 | Healthy choice.ti,ab. |
| 14 | exp Product labelling/ and (drink? or beverage? or soda? or flavo?red water? or fruitwater? or cordial? or squash? or juice? or smoothie? or milkshake? or tea or teas or coffee?).ti,ab. |
| 15 | (Drink$ label$ or Drink$ content$ label$ or Drink$ content$ sign$ or Drink$ content symbol$ or Drink$ content$ tag$ or Drink$ content $ ticket$ or Drink$ content$ sticker$).ti,ab. |
| 16 | ((drink? or beverage? or soda? or flavo?red water? or fruit water? or cordial? or squash? or juice? or smoothie? or milkshake? or tea or teas or coffee?) and (label$ or content$ sign$ or symbol$ or ticket$ or sticker$)).ti,ab. |
| 17 | or/1-16 |
| 18 | Food Preferences/ |
| 19 | Food Intake/ |

*(Continued)*

| | |
|---|---|
| 20 | Eating behavior/ or Dietary Restraint/ |
| 21 | Eating Attitudes/ or "obesity (attitudes toward)"/ |
| 22 | exp Diets/ |
| 23 | exp Choice Behavior/ |
| 24 | (intak$ or consume or consumes or consumption or consumed or eat$ or diet$).ti,ab. |
| 25 | (food adj5 (preference$ or habit$ or behavio?r$ or choice$ or decision$ or decid$ or inclin$ or lik$ or choos$ or select$ or pick$)).ti,ab. |
| 26 | ((drink? or beverage?) adj5 (preference$ or habit$ or behavio?r$ or choice$ or decision$ or decid$ or inclin$ or lik$ or choos$ or select $ or pick$)).ti,ab. |
| 27 | or/18-26 |
| 28 | Hospitality Industry/ |
| 29 | (purchas$ or buy$ or sale$ or vend$ or sell$).ti,ab. |
| 30 | (shop$ or store$ or supermarket$ or market$ or outlet$ or retailer$ or point of purchase).ti,ab. |
| 31 | (restaurant$ or cafe$ or bar$ or canteen$ or cafeteria$ or dinner hall$ or dining area$ or dining room$ or refector$ or eatery or mess or buffet or bistro$ or eating place$).ti,ab. |
| 32 | or/28-31 |
| 33 | 17 and (27 or 32) |
| 34 | limit 33 to up=20170425-20210709 |

***Alcohol labelling***

| | |
|---|---|
| 1 | exp Alcoholic Beverages/ |
| 2 | (dr#nk* or beverage* or alcohol* or beer* or lager* or wine* or cider*).ti,ab. |
| 3 | 1 or 2 |
| 4 | ((alcohol* or drink*) adj5 (unit? or guideline* or standard drink*)).ti,ab. |
| 5 | ((calorie* or nutrition* or energy or ingredient*) adj5 (label* or inform* or menu* or poster* or glass* or beermat* or bottle* or packag*)).ti,ab. |
| 6 | 4 or 5 |
| 7 | exp Alcohol Drinking Patterns/ or Alcohol Drinking Attitudes/ |
| 8 | ((purchas* or pour* or select* or consum*) adj5 (alcohol* or drink* or beer* or wine* or lager* or cider*)).ti,ab. |

*(Continued)*

| | |
|---|---|
| 9 | attitude change/ or exp health attitudes/ or "obesity (attitudes toward)"/ |
| 10 | exp consumer attitudes/ |
| 11 | exp Consumer Behavior/ |
| 12 | 7 or 8 or 9 or 10 or 11 |
| 13 | 3 and 6 and 12 |

**CENTRAL (Cochrane Library), from inception to Issue 7 of 12, July 2021**

**Alcohol labelling search applied to whole period from inception; food labelling search was an updated search and was limited to period from 25th April 2017. Search executed: 9th July 2021**

***Food labelling***

| | |
|---|---|
| #1 | MeSH descriptor: [Food Packaging] explode all trees |
| #2 | MeSH descriptor: [Product Labeling] this term only |
| #3 | MeSH descriptor: [Food Labeling] explode all trees |
| #4 | (((nutritio* or nutrient*) NEAR (label* or "content sign*" or "contents sign*" or symbol* or ticket* or sticker*))):ti,ab,kw OR ("nutrition information" or "nutritional information" or "nutrient information" OR "nutrients information"):ti,ab,kw OR ((food* NEXT (label* or "content label*" or "contents label*" or "content sign*" or "contents sign*" or "content symbol*" or "contents symbol*" or "content tag*" or "contents tag*" or "content ticket*" or "contents ticket*" or "content sticker*" or "contents sticker*"))):ti,ab,kw OR ("traffic light*"):ti,ab,kw |
| #5 | ("guideline daily amount*" or "nutrient reference value*" or "nutrient daily value*"):ti,ab,kw OR (("recommended dietary allowance*" NEAR/5 (label* or "content sign*" or "contents sign" or symbol* or information or ticket* or sticker*))):ti,ab,kw OR (((Calorific or calorie* or caloric) and (label* or "content sign*" or "contents sign" or symbol* or information or ticket* or sticker*))):ti,ab,kw OR (((Calorific or calorie* or caloric) NEXT information)):ti,ab,kw |
| #6 | ((menu and (label* or "content sign*" or "contents sign" or symbol* or information or ticket* or sticker*))):ti,ab,kw OR ((menu and ("nutritional content*" or "nutritional information" or "guideline daily amount" or GDA or "healthy choice" or calorie*))):ti,ab,kw OR ((Label* NEAR/2 (legislation* or regulation* or policies or policy))):ti,ab,kw OR ("healthy choice"):ti,ab,kw |
| #7 | (((drink or beverage* or soda* or "flavored water*" or "flavoured water*" or fruit water* or cordial* or squash* or juice* or smoothie* or milkshake* or tea or teas or coffee*) NEAR/5 (label* or "content sign*" or "contents sign" or tag* or symbol* or ticket* or sticker*))):ti,ab,kw |
| #8 | #1 or #2 or #3 or #4 or #5 or #6 or #7 |
| #9 | MeSH descriptor: [Food Preferences] explode all trees |
| #10 | MeSH descriptor: [Feeding Behavior] explode all trees |
| #11 | MeSH descriptor: [Eating] this term only |
| #12 | MeSH descriptor: [Diet] explode all trees |

*(Continued)*

| | |
|---|---|
| #13 | MeSH descriptor: [Choice Behavior] this term only |
| #14 | ((intak* or consume or consumes or consumption or consumed or eat* or diet*)):ti,ab,kw OR ((food NEAR/5 (preference* or habit* or behavior* or behaviour* or choice* or decision* or decid* or in-clin* or lik* or choos* or select* or pick*))):ti,ab,kw OR (((drink* or beverage*) NEAR/5 (preference* or habit* or behavior* or behaviour* or choice* or decision* or decid* or inclin* or lik* or choos* or select * or pick*))):ti,ab,kw |
| #15 | MeSH descriptor: [Restaurants] explode all trees |
| #16 | (purchas* or buy* or sale* or vend* or sell*):ti,ab,kw OR (shop* or store* or supermarket* or mar-ket* or outlet* or retailer*):ti,ab,kw OR (restaurant* or cafe* or bar* or canteen* or cafeteria* or "dinner hall*" or "dining area*" or "dining room*" or refector* or eatery or mess or buffet or bistro* or "eating place*"):ti,ab,kw |
| #17 | #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 |
| #18 | #8 and #17 with Publication Year from 2017 to 2021, in Trials |

***Alcohol labelling***

| | |
|---|---|
| #1 | MeSH descriptor: [Alcoholic Beverages] explode all trees |
| #2 | (drink* or drunk* or beverage* or alcohol* or beer* or lager* or wine* or cider*):ti,ab,kw |
| #3 | #1 or #2 |
| #4 | MeSH descriptor: [Product Labeling] explode all trees |
| #5 | (((alcohol* or drink*) NEAR/5 (unit or units or guideline* or "standard drink*"))):ti,ab,kw OR (((calo-rie* or nutrition* or energy or ingredient*) NEAR/5 (label* or inform* or menu* or poster* or glass* or beermat* or bottle* or packag*))):ti,ab,kw |
| #6 | #4 or #5 |
| #7 | MeSH descriptor: [Health Knowledge, Attitudes, Practice] explode all trees |
| #8 | MeSH descriptor: [Energy Intake] explode all trees |
| #9 | MeSH descriptor: [Consumer Behavior] explode all trees |
| #10 | MeSH descriptor: [Alcohol Drinking] explode all trees |
| #11 | (((purchas* or pour* or select* or consum*) NEAR/5 (alcohol* or drink* or beer* or wine* or lager* or cider*))):ti,ab,kw |
| #12 | #7 or #8 or #9 or #10 or #11 |
| #13 | #3 and #6 and #12 |

**ASSIA (ProQuest), 1952-present**

**Alcohol labelling search applied to whole period from inception; food labelling search was an updated search and was limited to period from 25th April 2017. Search executed: 9th July 2021**

***Food labelling***

| | |
|---|---|
| 1 | noft("food pack*") OR noft(((nutritio* or nutrient*) NEAR/5 (label* or "content sign*" or "contents sign*" or symbol* or ticket* or sticker*))) OR noft(("nutrition information" or "nutritional information" or "nutrient information" OR "nutrients information")) OR noft((food* NEXT (label* or "content label*" or "contents label*" or "content sign*" or "contents sign*" or "content symbol*" or "contents symbol*" or "content tag*" or "contents tag*" or "content ticket*" or "contents ticket*" or "content sticker*" or "contents sticker*" or "content information" or "contents information"))) OR noft("traffic light*") OR noft("guideline daily amount*" or "nutrient reference value*" or "nutrient daily value*") OR noft(("recommended dietary allowance*" NEAR/5 (label* or "content sign*" or "contents sign" or symbol* or information or ticket* or sticker*))) OR noft(((Calorific or calorie* or caloric) and (label* or "content sign*" or "contents sign" or symbol* or information or ticket* or sticker*))) OR noft(((Calorific or calorie* or caloric) NEAR/1 information)) |
| 2 | noft((menu AND (label* OR "content sign*" OR "contents sign" OR symbol* OR information OR ticket* OR sticker*))) AND noft((menu AND ("nutritional content*" OR "nutritional information" OR "guideline daily amount" OR GDA OR "healthy choice" OR calorie*))) OR noft((Label* NEAR2 (legislation* OR regulation* OR policies OR policy))) OR noft("Healthy choice") OR noft(((drink* OR beverage* OR soda* OR "flavored water*" OR "flavoured water*" OR "fruit water*" OR cordial* OR squash* OR juice* OR smoothie* OR milkshake* OR tea OR teas OR coffee*) NEAR5 (label* OR "content sign*" OR "contents sign" OR tag* OR symbol* OR ticket* OR sticker*))) |
| 3 | 1 OR 2 |
| 4 | 1 OR 2 - LIMITS APPLIED 2017-2021 |

***Alcohol labelling***

| | |
|---|---|
| 1 | (drink* or drunk* or beverage* or alcohol* or beer* or lager* or wine* or cider*) |
| 2 | noft(((alcohol* or drink*) NEAR/5 (unit or units or guideline* or "standard drink*"))) OR noft(((calorie* or nutrition* or energy or ingredient*) NEAR/5 (label* or inform* or menu* or poster* or glass* or beermat* or bottle* or packag*))) |
| 3 | noft(((purchas* or pour* or select* or consum*) NEAR/5 (alcohol* or drink* or beer* or wine* or lager* or cider*))) |
| 4 | 1 AND 2 AND 3 |

**Science Citation Index; Social Science Citation Index; Conference Proceedings Citation Index - Science; Conference Proceedings Citation Index - Social Science & Humanities (Web of Science), 1900-present**

**Alcohol labelling search applied to whole period from inception; food labelling search was an updated search and was limited to period from 25th April 2017. Search executed: 9th July 2021**

***Food labelling***

| | |
|---|---|
| 1 | "food pack*" (Topic) or ((nutritio* or nutrient*) NEAR/5 (label* or "content sign*" or "contents sign*" or symbol* or ticket* or sticker*)) (Topic) or ("nutrition information" or "nutritional information" or "nutrient information" OR "nutrients information") (Topic) or (food* NEXT (label* or "content label*" or "contents label*" or "content sign*" or "contents sign*" or "content symbol*" or "contents symbol*" or "content tag*" or "contents tag*" or "content ticket*" or "contents ticket*" or "content sticker*" or "contents sticker*" or "content information" or "contents information")) (Topic) or "traffic light*" (Topic) or "guideline daily amount*" or "nutrient reference value*" or "nutrient daily value*" (Topic) or ("recommended dietary allowance*" NEAR/5 (label* or "content sign*" or "contents sign" or symbol* or information or ticket* or sticker*)) (Topic) or ((Calorific or calorie* or caloric) and (label* or "content sign*" or "contents sign" or symbol* or information or ticket* or sticker*)) (Topic) or ((Calorific or calorie* or caloric) NEXT information) (Topic) |
| 2 | (menu and (label* or "content sign*" or "contents sign" or symbol* or information or ticket* or sticker*)) (Topic) or (menu and ("nutritional content*" or "nutritional information" or "guideline daily amount" or GDA or "healthy choice" or calorie*)) (Topic) or (Label* NEAR/2 (legislation* or regulation* or policies or policy)) (Topic) or "Healthy choice" (Topic) or ((drink* or beverage* or soda* or "flavored water*" or "flavoured water*" or "fruit water*" or cordial* or squash* or juice* or smoothie* or milkshake* or tea or teas or coffee*) NEAR/5 (label* or "content sign*" or "contents sign" or tag* or symbol* or ticket* or sticker*)) (Topic) |
| 3 | 1 or 2 |
| 4 | (intak* or consume or consumes or consumption or consumed or eat* or diet*) (Topic) or (food NEAR/5 (preference* or habit* or behavior* or behaviour* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)) (Topic) or ((drink* or beverage*) NEAR/5 (preference* or habit* or behavior* or behaviour* or choice* or decision* or decid* or inclin* or lik* or choos* or select * or pick*)) (Topic) or purchas* or buy* or sale* or vend* or sell* (Topic) or shop* or store* or supermarket* or market* or outlet* or retailer* (Topic) or restaurant* or cafe* or bar* or canteen* or cafeteria* or "dinner hall*" or "dining area*" or "dining room*" or refector* or eatery or mess or buffet or bistro* or "eating place*" (Topic) |
| 5 | 3 and 4 |
| 6 | 3 and 4 2017 or 2018 or 2019 or 2020 or 2021 (Publication Years) |

***Alcohol labelling***

| | |
|---|---|
| | Alcohol labelling |
| 1 | TS=((drink* or drunk* or beverage* or alcohol* or beer* or lager* or wine* or cider*)) |
| 2 | TS=((alcohol* or drink*) NEAR/5 (unit or units or guideline* or "standard drink*")) OR TS=((calorie* or nutrition* or energy or ingredient*) NEAR/5 (label* or inform* or menu* or poster* or glass* or beermat* or bottle* or packag*)) |
| 3 | TS=((purchas* or pour* or select* or consum*) NEAR/5 (alcohol* or drink* or beer* or wine* or lager* or cider*)) |
| 4 | 1 and 2 and 3 |

**MAG - Custom Search (EPPI Reviewer), 1800 to 2 August 2021**

**Food labelling search only; this was an updating search limited to the period from 1st January 2017.**

**Search executed: 1st September 2021**

| | |
|---|---|
| AND(OR(AND(Composite(F.FId=2778297792),D>'2017-01-01'),AND(Composite(F.FId=3017654825),D>'2017-01-01'),AND(Composite(F.FId=3017781192),D>'2017-01-01'),AND(Composite(F.FId=63412515),D>'2017-01-01'),AND(Composite(F.FId=96105989),D>'2017-01-01'),AND(Composite(F.FId=3017461897),D>'2017-01-01'),AND(Composite(F.FId=3018629635),D>'2017-01-01'),AND(Composite(F.FId=3018039482),D>'2017-01-01'),AND(Composite(F.FId=3020115395),D>'2017-01-01'),AND(Composite(F.FId=3020577620),D>'2017-01-01'),AND(Composite(F.FId=2909101193),D>'2017-01-01'),AND (Composite(F.FId=2993824933),D>'2017-01-01'),AND(Composite(F.FId=2910706634),D>'2017-01-01')),D=['2017-01-01, 2021-07-09']) | |
| Composite(F.FId=2910706634) | Topic: Food labelling |
| Composite(F.FId=2993824933) | Topic: Food labeling |
| Composite(F.FId=2909101193) | Topic: Food label |
| Composite(F.FId=3020577620) | Topic: Menu labelling |
| Composite(F.FId=3020115395) | Topic: Menu labeling |
| Composite(F.FId=3018039482) | Topic: Nutritional labelling |
| Composite(F.FId=3018629635) | Topic: Front of package |
| Composite(F.FId=3017461897) | Topic: Front of pack |
| Composite(F.FId=96105989) | Topic: Guideline Daily Amount |
| Composite(F.FId=63412515) | Topic: Nutrition facts label |
| Composite(F.FId=3017781192) | Topic: Nutritional information |
| Composite(F.FId=3017654825) | Topic: Nutrition information |
| Composite(F.FId=2778297792) | Topic: Nutrition Labeling and Education Act of 1990 |

**OpenAlex - Custom Search (EPPI Reviewer), 1800 to 25th August 2023**

**Food labelling search only; this was an updating search limited to the period from 1st January 2017.**

**Search executed: 25th August 2023**

| | |
|---|---|
| 1 | concepts.id:https://openalex.org/C2910706634,from_publication_date:2017-01-01 [Topic: Food labelling] |
| 2 | concepts.id:https://openalex.org/C2993824933,from_publication_date:2017-01-01 [Topic: Food labelling] |
| 3 | concepts.id:https://openalex.org/C2910069897,from_publication_date:2017-01-01 [Topic: Nutrition Labeling] |
| 4 | concepts.id:https://openalex.org/C63412515,from_publication_date:2017-01-01 [Topic: Nutrition facts label] |

*(Continued)*

| | |
|---|---|
| 5 | concepts.id:https://openalex.org/C3017654825,from_publication_date:2017-01-01 [Topic: Nutrition information] |
| 6 | concepts.id:https://openalex.org/C3017781192,from_publication_date:2017-01-01 [Topic: Nutritional information] |
| 7 | concepts.id:https://openalex.org/C112625547,from_publication_date:2017-01-01 [Topic: Food composition data] |
| 8 | concepts.id:https://openalex.org/C152572085,from_publication_date:2017-01-01 [Topic: Reference Daily Intake] |
| 9 | 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 |

## HISTORY

Protocol first published: Issue 6, 2021

## CONTRIBUTIONS OF AUTHORS

Drafted the protocol: NC, EP, TMM, MP, NWR, SAJ, GJH

Developed the search strategy: NWR, NC, IS, GJH

Searched for studies: NC, EP, NWR, IS, GJH

Obtained full-text reports of studies: NC, EP, NWR, IS, GJH

Selected studies for inclusion: NC, EP, IS, GJH

Extracted data from studies: NC, EP, GJH

Assessed risk of bias in included studies: NC, EP, GJH

Entered data into Review Manager: NC, GJH

Carried out the analysis: MP, NC, GJH

Interpreted the analysis and drafted/revised the final review: all authors

## DECLARATIONS OF INTEREST

NC: no relevant interests; involved in Clarke 2023a (funding source: Wellcome Trust) – lead author, involved in design, data collection, and write-up of the study, University of Cambridge.

EP: no relevant interests; involved in Clarke 2023a; Reynolds 2022 – studies undertaken as part of a programme of work as a research assistant at the University of Cambridge (one study looked at the impact of graphic health warning labels and the other on PACE labels).

IS: none.

MP: no relevant interests; involved in Clarke 2023a; Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019.

NWR: none.

TMM: no relevant interests; involved in Clarke 2023a (funder: Wellcome Trust, Cancer Research UK); Reynolds 2022 (funder: Department of Health Policy Research Programme and the Institute of Grocery Distribution); Vasiljevic 2018 (funder: National Institute for Health Research Policy Research Programme and the Institute of Grocery Distribution); Vasiljevic 2019.

SAJ: Food Standards Agency (employment); published general expression of views about complex evidence base relating to labelling – no fixed position; involved in Reynolds 2022; Vasiljevic 2018; Vasiljevic 2019 – primary research studies on effectiveness of labelling to change food purchasing behaviour.

GJH: no relevant interests; involved in Clarke 2023a (funded in whole or in part by the Wellcome Trust (ref: 206853/Z/17/Z)); Reynolds 2022 (funded by a collaborative Award in Science from Wellcome Trust (Behaviour Change by Design: 206853/Z/17/Z) awarded to Theresa Marteau, Paul Fletcher, Gareth Hollands, and Marcus Munafò); Vasiljevic 2018 (funded by the Department of Health Policy Research Programme – Policy Research Unit in Behaviour and Health [PR-UN-0409-10109] and the Institute of Grocery Distribution [RG83425]); Vasiljevic 2019 (funded by the National Institute for Health Research Policy Research Programme (Policy Research Unit in Behaviour and Health [PR-UN-0409-10109]) and the Institute of Grocery Distribution [RG83425]) – part of the research team, based at the University of Cambridge for each of these four studies.

Review authors who were involved in the conduct of a study that was eligible for inclusion in this review (NC, EP, MP, TM, SJ, GH) did not make any study eligibility decisions about, extract data from, carry out the risk of bias assessment for, or perform GRADE assessments for the studies they were involved in.

## SOURCES OF SUPPORT

### Internal sources

- University of Cambridge, UK

  Computer provision, software access, database access (NC, EP, TMM, MP, GJH)
- University College London, UK

  Computer provision, software access, database access (IS, GJH)
- Bath Spa University, UK

  Computer provision, software access, database access (NC)
- University of Oxford, UK

  Computer provision, software access, database access (NWR, SAJ)

### External sources

- Wellcome Trust, UK

  Production of this review was funded in part by a Collaborative Award in Science from the Wellcome Trust (ref: 206853/Z/17/Z). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We changed the title from "Energy (calorie) labelling for healthier selection and consumption of food or alcohol" to "Calorie (energy) labelling for changing selection and consumption of food or alcohol", due to first, considering 'calorie labelling' to be a more widely used and understood term than energy labelling. Second, the use of the wording 'healthier selection and consumption', being inherited from the previous version of this review, was considered misleading because reduced calorie selection and consumption is not necessarily healthier.

We clarified the definitions of eligible outcomes included in the published protocol (Clarke 2021b), prior to screening studies for inclusion in the current review. Although we had previously referred to 'hypothetical' selection as an ineligible outcome, we have now operationalised this in more detail.

We planned to repeat the meta-analyses by instead entering a single effect estimate for each multi-arm study using the mean standardised mean difference and mean variance across multiple comparisons from that study. However, this was not possible for all studies where this applied because the necessary data were not available.

Due to aiming to simplify the logic model and prioritise the data considered most important to extract (and informed by the scarcity of relevant data as eligible studies were encountered), we removed the following constructs from the logic model as presented at the protocol stage and did not extract any corresponding data: behavioural characteristics (e.g. dietary restraint; dietary disinhibition; level of intake or dependence for targeted product), and biological state (e.g. hunger).

We included an additional sensitivity analysis of only selection outcome data that included purchasing (to compare to selection outcome data irrespective of purchasing), and an additional subgroup analysis concerning where labels were placed.

## INDEX TERMS

### Medical Subject Headings (MeSH)

*Alcohol Drinking; *Alcoholic Beverages [statistics & numerical data]; Bias; *Energy Intake; *Food Labeling [methods]; Randomized Controlled Trials as Topic

### MeSH check words

Humans