Effect of front-of-package nutrition labeling on food purchases: a systematic review


Received 10 February 2020
Received in revised form 15 June 2020
Accepted 16 June 2020
Available online 28 January 2021

Abstract

Objectives: This study systematically reviewed evidence from interventions on the effect of front-of-package (FOP) nutrition labeling on food purchases. Study design: The study design used in this study is a systematic review. Methods: Keyword search was performed in PubMed, Web of Science, Scopus, and Cochrane Library. Results: Fifteen studies (10 randomized controlled trials, four pre-post studies, and one case-control study) met the eligibility criteria and were included in the review. Five studies were conducted in a controlled setting through the establishment of an online virtual supermarket or physical laboratory food store solely for the intervention. In contrast, the remaining ten studies were conducted in a naturalistic setting where people commonly purchase foods (e.g., supermarket, grocery store, school/hospital cafeteria, or vending machine). FOP labels assessed included traffic lights, health star rating, daily intake guides, health warnings, and high sugar symbol labels. Compared with the control, FOP labels were effective for helping participants make healthier food purchase decisions in five of the 12 studies that assessed traffic lights labels, in one of the two studies that assessed health warning labels, and in one study that assessed high sugar symbol labels. Three assessed health star ratings and one assessed daily intake guide labels, but none revealed an effect on food purchases compared with the control. Conclusions: Findings on the effectiveness of FOP nutrition labels in ‘nudging’ consumers toward healthier food purchases remain mixed and inconclusive. Future studies should examine other types of FOP labels beside the traffic light labels and explore the different effects by consumer affordability, population subgroup, and shopping environment.

Introduction

The nutrition facts label is a crucial source for consumers to obtain nutrition- and health-related information on food products. One objective of the nutrition facts label is to facilitate consumers’ identification and selection of healthier food items that are nutrient-rich and low in energy density. However, nutrition facts label has been criticized as being less visible, difficult to comprehend, and of limited effectiveness, in guiding consumers’ food choices. Front-of-package (FOP) nutrition labels use simple symbols with highly visible logos and numbers to communicate the healthiness of food to consumers. FOP labels seek to help consumers understand and quantify the nutritional value of food items regardless of age, education, and literacy level. A variety of FOP labels have emerged to supply consumers with the information needed to make healthy food choices. The traffic light food labeling system involves labeling foods as red, amber, or green depending on the levels (i.e., amounts) of fat, saturated fat, sugars, and salt. Food with fewer reds, fewer ambers, and more greens can contribute to a healthier diet. By providing a color-coding process...
when selecting foods, consumers can easily detect whether their food choices are healthy or not. Another FOP label is the Australasian health star rating system. The health star rating system presents a one-half (least healthy) to a five-star (most healthy) rating of nutritional quality, with more stars indicating a higher nutritional value of the food. The health star rating system considers energy and contents of the food in terms of both positive nutritional value of the food. The health star rating system of the information presented through the use of highly explicit messages.

### Methods

A systematic review was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses. A systematic review was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Studies that met all of the following criteria were included in the review: (1) Exposure: FOP nutrition labeling; (2) Outcome: food purchases; (3) Study design: intervention; (4) Article type: peer-reviewed publication; (5) Time window of search: from the inception of an electronic bibliographic database to December 1, 2019; and (6) Language: article written in English.

### Study selection criteria

Studies that met any of the following criteria were excluded from the review: (1) An observational study without an intervention component; (2) A mathematical or statistical simulation study; (3) A study examining consumer perception of purchase or intent to purchase rather than an actual purchase that involves monetary transfer; and (4) A study examining other types of food labels rather than FOP nutrition label.

### Search strategy

A keyword search was performed on December 1, 2019, in four electronic bibliographic databases — PubMed, Web of Science, Scopus, and the Cochrane Library. The search algorithm included all possible combinations of keywords from the following two groups: (1) “front-of-pack,” “pack,” “packs,” “package,” “packages,” “signpost,” “traffic light,” “traffic-light,” and “health star”; and (2) “labels,” “labeling,” and “food labeling.” The MeSH term ‘food labeling’ was included in the PubMed search.

### Data extraction

A standardized data extraction form was used to collect the following methodological and outcome variables from each included study: authors, publication year, country, study design, overall sample size, arm-specific sample size, age distribution, sex distribution, attrition rate, intervention setting, intervention aims, arm-specific intervention components, intervention duration, control variables, measures of purchase behavior, statistical methods, intervention effectiveness on food purchase, the arm-specific direction of impact on purchases, and intervention effectiveness on other secondary outcomes.

Owing to substantial heterogeneity in the outcomes (i.e., types of foods/beverages purchased) and their corresponding measures (e.g., supermarket scanner data, receipts, and self-reported purchase behavior), no two studies included in the review shared the same outcome and measure, which prevented us from conducting meta-analysis. Therefore, this review was limited to a narrative summary of scientific literature.

### Study quality assessment

A study quality assessment tool rated each study based on the following eight criteria: (1) Was the research question or study objective clearly stated? (2) Were the study subjects a population-based sample? (3) Did the study include a control group? (4) Were study subjects randomly assigned to different arms? (5) Was sample size justification (e.g., power analysis) provided? (6) Was purchase behavior objectively measured? (7) Was the intervention setting a natural shopping environment with minimal interference from researchers? (8) Were the statistical procedures appropriate to
address the research question? For each criterion, a score of one was assigned if ‘yes’ was the response, whereas a score of zero was assigned otherwise. A study-specific global score, ranging from zero to eight, was calculated by summing up scores across all criteria.

Results

Study selection

Fig. 1 shows the study selection flow chart. We identified a total of 6513 articles by the keyword search, including 1637 articles from PubMed, 2312 articles from Web of Science, 1799 articles from Scopus, and 765 articles from the Cochrane Library. After removing duplicates, 5185 unique articles entered title and abstract screening, in which 5150 articles were excluded. The full texts of the remaining 35 articles were reviewed against the study selection criteria. Of these, 21 articles were excluded. The reasons for exclusion included the following: observational study design without an intervention, outcomes about consumer perception of purchase or intent to purchase rather than an actual purchase behavior, and interventions focusing on other types of food labels rather than FOP labels. A forward and backward reference search was conducted based on the remaining 14 articles, and one new article was identified that met the study selection criteria. Therefore, a total of 15 articles consisted of the final pool of studies and were included in the review.

Characteristics of the selected studies

Table 1 summarizes the basic characteristics of the 15 studies. All of them were published between 2009 and 2019. They were conducted in the U.S. (n = 4), Australia (n = 2), Canada (n = 2), France (n = 1), Germany (n = 1), New Zealand (n = 1), Taiwan, China (n = 1), U.K. (n = 1), Netherlands (n = 1) and Belgium (n = 1). Study designs included randomized controlled trial (RCT) (n = 10), pre-post study (n = 4), and case-control study (n = 1). Ten studies involved human subjects with a sample size ranging from 109 to 1578 participants. In contrast, the remaining five studies focused on the quality of food or beverage items sold by a vendor (e.g., cafeteria or supermarket). Eight studies recruited people aged 14 years and older, whereas the other studies did not report age distribution. The proportion of women accounted for over half (54%–100%) of the study sample in the eight studies that reported sex distribution. Among the ten studies that reported attrition rate, seven had an attrition rate less than 20% (0%–12%), whereas the remaining three had an attrition rate of 29%, 31%, and 57%. Five of the 15 studies were conducted in a controlled intervention environment through establishing a virtual supermarket online or a physical lab food store solely for the intervention. The remaining studies were performed in a naturalistic environment where people commonly go grocery shopping, such as a supermarket, a small or midsize grocery store, a college or hospital cafeteria, sports or recreational facility, and a vending machine. A variety of FOP labels were assessed in the selected studies, including traffic lights (n = 12), health star rating (n = 3), daily intake guide (n = 1), health warning (n = 2),...
### Table 1
Basic characteristics of front-of-pack nutrition label interventions.

<table>
<thead>
<tr>
<th>Study ID</th>
<th>First author (year)</th>
<th>Country</th>
<th>Study design</th>
<th>Sample size</th>
<th>Age (years)</th>
<th>Women (%)</th>
<th>Attrition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sacks (2009)[27]</td>
<td>UK</td>
<td>Pre-post</td>
<td>18 types of products</td>
<td>&gt;18</td>
<td>85</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>Sacks (2011)[24]</td>
<td>Australia</td>
<td>Case-control</td>
<td>53 types of foods</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Waterlander (2013)[38]</td>
<td>Netherlands</td>
<td>RCT</td>
<td>109</td>
<td>Study 1: 184</td>
<td>Study 1: 29</td>
<td>Study 1: 79</td>
</tr>
<tr>
<td>5</td>
<td>Olsstad (2015)[26]</td>
<td>Canada</td>
<td>Pre-post</td>
<td>322</td>
<td>&gt;14</td>
<td>49</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>Trudel (2015)</td>
<td>US</td>
<td>RCT</td>
<td>150</td>
<td>40</td>
<td>55</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Julia (2016)[27]</td>
<td>France</td>
<td>RCT</td>
<td>901</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Neal (2017)[29]</td>
<td>Australia</td>
<td>RCT</td>
<td>1578</td>
<td>38</td>
<td>84</td>
<td>17</td>
</tr>
<tr>
<td>10</td>
<td>Ni Mhurchu (2017)[31]</td>
<td>New Zealand</td>
<td>RCT</td>
<td>1357</td>
<td>33</td>
<td>98</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>Chen (2017)[32]</td>
<td>Taiwan, China</td>
<td>Pre-post</td>
<td>481</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Acton (2017)[33]</td>
<td>Canada</td>
<td>RCT</td>
<td>686</td>
<td>&gt;16</td>
<td>54</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>Mazza (2018)[35]</td>
<td>US</td>
<td>Pre-post</td>
<td>Beverages and chips</td>
<td>19,238 beverages</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Study ID Setting Controlled setting Arms Labels Duration (week)

| 1 | Supermarket | No | Arm1: Ready meals | Arm2: Sandwiches | Traffic light label | 8 |
| 2 | Online grocery store | Yes | Arm 1: online intervention store | Arm 2: online control store | Traffic light label | 10 |
| 3 | Web-based supermarket | Yes | Arm 1: Special offer labels | Arm 2: Healthy choice labels | Special offer label, healthy choice label |      |
| 4 | Study 1: Lab store | Yes | Study 1: Arm 1: pasta meals with traffic light colors | Study 2: Arm 1: with traffic light, brand of Corny | Traffic light label |      |
| 5 | Recreation and sports facility | No | 1 arm pre-post | 2 | Traffic light label | 2 |
| 6 | Grocery store | No | Arm 1: red dominant label group | Arm 2: green dominant label group | Traffic light label |      |
| 7 | Lab store Supermarket | Yes | Arm 1: control group | Arm 2: 5-color nutrition label | 5-color FOP nutrition label | 16 |
| 8 | College cafeteria | No | Arm 1: nutrition information panels | Arm 2: traffic light label | Traffic light label | 5 |
| 9 | Retail outlet | No | Arm 1: health star rating | Arm 2: multiple traffic label | Health star rating, multiple traffic label, daily intake guides, health warning | 4 |
| 10 | Supermarket | No | Arm 1: control group | Arm 2: traffic light label | Traffic light label, health star rating | 5 |
| 11 | Canteen and buffet | No | 1 arm pre-post | Traffic light label | High sugar label, health star rating | 44 |
| 12 | Intervention marketplace | Yes | Arm 1: control group | Arm 2: high sugar symbol | Traffic light label, health star rating | 8 |
| 13 | Community supermarket | No | Arm 1: traffic light label, information explanation and financial incentives | Arm 2: control group | Traffic light label | 28 |
| 14 | Hospital cafeteria | No | Arm 1: Soda price | Arm 2: Soda price and traffic light labeling | Traffic light label | <92 |
| 15 | High school vending machine and cafeteria | No | Arm 1: traffic light coding system in vending machines | Arm 2: traffic light coding system in cafeterias | Traffic light coding | 7 |
and high sugar symbol (n = 1). Intervention duration varied from two weeks to 92 weeks in the 12 studies that reported duration. Table 2 reports the measures and outcomes of FOP nutrition label interventions. The majority of the studies (n = 12) used objective purchase measures, one study adopted a subjective purchase measure, and the remaining two used both objective and subjective purchase measures. For objective purchase measures, two studies conducted in a controlled intervention environment used a fictitious checkout to record purchase data. Study participants were asked to approach an online or physical cashier with the food items they selected but without actually making the payment with a credit/debit card or cash. Among the eight studies that recorded actual purchases in a naturalistic setting, four used the total number of sales within a specific time window to measure purchase behavior, and two collected grocery shopping receipts from study participants (through mailing a hard copy of receipts in paid envelopes or submitting a digital copy of receipts). The remaining two studies collected receipts from the cashier. The study that adopted a subjective measure was based on self-reported purchase behavior. The two studies that took both subjective and objective purchase measures used self-reported consumption of foods or beverages in combination with either monthly purchases tracked in a store loyalty or store sales records. Besides food purchases, the other outcomes assessed in the 15 studies consisted of awareness and perceptions of FOP nutrition label, self-control concerning food choices, and health knowledge and attitudes towards healthy food consumption.

Table 3 summarizes the estimated effects of FOP nutrition labels on food purchases. A total of 12 studies assessed the effectiveness of traffic light labels on at least one subjective or objective food purchase measure. Among them, five found that the traffic light labels were effective in helping participants in the intervention group make healthier food purchases in comparison with those in the control group who were not provided with any FOP nutrition labels.26,32,34–36 Three of the five studies adopted a pre–post design,34,36 and the other two adopted a RCT design.34,36 A variety of statistical models were applied across the five studies, including the Chi-squared test,36,32,34 analysis of covariance (ANCOVA),26 logistic regression,32,36 multivariate regression,35 generalized least squares model with subject-specific random effects,32 generalized estimating equations,26 and Fisher’s exact test.34 Types of populations examined included people aged 14 years and older,26 adults aged 18 years and older with at least one child in the household,34 high school students,36 canteen customers,26 and employees in a pediatric hospital.34 Two studies were conducted in Canada,26,34 and one each in the US,35 Belgium,36 and Taiwan, China.26 One study found that the traffic light labels improved food choices for low self-control consumers, whereas the labels did not affect high self-control consumers.25 Another study found that frequent traffic light label users had significantly healthier food purchases compared with regular nutrition information panel users.33 However, the overall effect of traffic light label use on food purchase was not significantly different from the nutrition information panel group.31 The other four studies did not identify a statistically significant influence on food purchases through the provision of traffic light labels.26,32,34,36 Two of the four studies adopted an RCT design,26,36 one adopted a case-control design,26 and the other adopted a pre–post design.33 Statistical models applied in the four studies included ANOVA,26,34 interrupted time-series analysis,32 Chi-squared test,29 t-test,29 and mixed-effect model.30,33 Types of populations examined included supermarket shoppers,24,37 university students,29 and residents aged 18 years and older in Australia.32 Two studies were conducted in Australia,24,30 and one each in the US29 and the UK.37 A few limitations might lead to null findings, as noted in these four studies. Students might be affected by the intervention and change their eating behavior in cafeterias.25 The study time frame was only a few weeks, and customers might need more time to adjust their grocery shopping habits.24,37 Sample demographics were non-representative of the population.24,30 Different labeling formats on food packs might differentially impact consumers’ comprehension of traffic light labels.30 The food products examined in the studies represented only a small fraction of total sales volumes.24,37

Three studies assessed the effectiveness of health star rating labels on food purchases, and all reported a null finding.26,32,34 Among the two studies that evaluated health warning labels on food purchases,32,33 one identified a significant impact in the expected direction,30 whereas the other reported a null finding.33 All three studies adopted an RCT design.30,31,33 Statistical models applied included t-test,31 Chi-squared test,31 ANOVA,29,33 and mixed-effect model.30,33 Types of populations examined included residents aged 18 years and older,30 and grocery shoppers 16 or 18 years and older.31,33 One each was conducted in Australia,30 New Zealand,31 and Canada.32 A few limitations might lead to null findings, as noted in these three studies. Sample demographics were non-representative of the population.29,31 No training was provided on how to use FOP labels, and the potential impact could be more significant if there were concurrent community education.30 The smartphone app as a medium of intervention delivery may impede consumers’ label reading due to the cumbersome process in scanning, observing, and product checking.31 The sample size could be too small to achieve statistical significance.33 The intervention using daily intake guide labeling (n = 1) revealed no statistically significant impact on food purchases in comparison with the control arm.30 This study was conducted in Australia, adopted an RCT design, recruited residents aged 18 years and older, and applied statistical models, including t-test, ANOVA, and mixed-effect model.

High sugar symbol (n = 1) was found to significantly reduce sugar-sweetened beverage consumption and increase the consumption of sugar-free beverages.33 The study was conducted in Canada, adopted an RCT design, recruited grocery shoppers aged 16 years and older, and applied statistical models, including the Chi-squared test and generalized linear mixed-effect model.

Table 4 reports criterion-specific and global ratings from the study quality assessment. The included studies, on average, scored five of eight, with a range from four to seven. All of the 15 studies included in the review clearly stated the research question and objective, employed appropriate statistical procedures to address the research question and employed objective measures on food purchases. Ten studies included a control group, 11 were conducted in a naturalistic shopping environment with minimal interference from researchers, and 10 had study subjects randomly assigned to different arms. In contrast, only three of the 15 studies provided a sample size justification using power analysis, and none of the studies recruited a population-based sample.

Discussion

This study systematically reviewed scientific evidence from interventions on the effect of FOP nutrition labeling on food purchases. A total of 15 studies met the eligibility criteria and were included in the review. A variety of FOP labels were assessed, including traffic lights, health star ratings, daily intake guides, health warnings, and high sugar symbols. Findings on the effectiveness of FOP nutrition labels in nudging consumers toward healthier food and beverage purchases remained mixed and inconclusive. Compared with the control without provision of any FOP label, FOP labels were effective for helping participants make...
healthier food purchase decisions in five of the 12 studies that assessed traffic lights labels, in one of the two studies that assessed health warning labels, and in one study that assessed high sugar symbol labels. Three assessed health star rating labels and one assessed daily intake guide labels, but none revealed an effect on food purchases in comparison with the control.

Previous reviews documented that use of nutrition labels, in general, was associated with a moderately improved health perception and nutrition knowledge. Although those improvements are necessary for forming healthier grocery shopping patterns and better diet quality, the translation is unlikely to be automatic and unobstructed. Indeed, many competing factors are
Despite that, there is evidence indicating a more substantial impact of FOP labels on people with a lower socio-economic status. Time is also a limiting factor in grocery shopping as consumers are likely to impact consumers' food purchase decisions jointly.  

The most salient factor may be affordability as the majority of consumers, especially those of lower socio-economic status, need to balance between food quality and price under a tight budget. Despite that, there is evidence indicating a more substantial impact of FOP labels on people with a lower socio-economic status.  

Table 4
Study quality assessment.

<table>
<thead>
<tr>
<th>Study ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Was the research question or study objective clearly stated?</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2. Were the study subjects a population-based sample?</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3. Did the study include a control group?</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4. Were study subjects randomly assigned to different arms?</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5. Was a sample size justification (e.g., power analysis) provided?</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6. Was purchase behavior objectively measured?</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7. Was the intervention setting a natural shopping environment with minimal interference from researchers?</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8. Were the statistical procedures appropriate to address the research question?</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Global score</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
effectiveness of voluntary FOP labels may indicate that some forms of mandatory FOP labeling such as health warning labels or nutri-score are necessary to result in meaningful changes in people's food purchase behavior.\textsuperscript{40} Finally, the use of nutrition labels varies across population subgroups. Specifically, children, adolescents, and older adults are less likely to use nutrition labels, and they are also more vulnerable to the adverse health effects of nutritional deficiency.\textsuperscript{12–15,41}

Warning labels printed on tobacco packages as required by law have been found effective to discourage consumption.\textsuperscript{32,42} Over the past five years, multiple state legislative attempts across the US have been made to apply what was learned from the tobacco warning labels to developing and implementing FOP warning labels for SSBs.\textsuperscript{43} However, to date no state in the US has passed regulations mandating the adoption of SSB warning labels. In 2016, Chile became the first country to implement an FOP warning label.\textsuperscript{54,40} A similar law was also passed in Mexico in January 2020, but the exact timing for the new labeling requirements to take effect has not been declared.\textsuperscript{47} These policy changes provide a unique opportunity (i.e., a natural experiment) to examine the causal impact of FOP health warning labels on people's food purchase and consumption.

Several limitations should be noted about the review and the included studies. Despite a large body of literature about nutrition labeling, the number of studies that focused on the impact of FOP labels on food purchases remains limited. Except for the traffic lights labels that were examined by 12 studies, all of the other FOP labels were assessed by only one or a few studies, which deserve more thorough investigations in future research. No study had explored the potential differential responses to FOP labels by gender, age group, race/ethnicity, or education level. The possible interaction between affordability and the use of FOP labels in jointly influencing purchase behaviors was not examined. No two studies shared the same outcome and measure, which prevented a meta-analysis. The backward and forward search was limited to the 14 studies identified by the title/abstract and full-text review. It would be more comprehensive if we conducted the backward and forward search for all studies harvested from the keyword search, but it was practically infeasible due to the sheer number of studies that would be included. Most studies included in the review were based on developed countries. In contrast, studies focusing on developing countries tend to be scarce. Nutri-score, also known as the 5–Color Nutrition label, is a nutrition label that was selected by the French government in 2017 to be displayed on food products.\textsuperscript{40} However, no study evaluating the effect of nutri-score on food purchases was identified from the keyword and reference search. Some studies adopted subjective rather than objective measures for food purchases, which may be prone to social desirability bias. This study reviewed articles written in English only. Studies evaluating the impact of FOP labels on food purchases that were written in other languages were excluded due to our review team's capacity limit. People's food purchase decisions may be partially determined by the setting or purchase environment (e.g., supermarket, convenience store, cafeteria, or vending machine), and the influence of FOP labels may differ across those settings. Finally, the dynamic interplay between the food industry and government in jointly determining the possible legislation on mandating FOP labels was seldom studied. For instance, what are the preferences of the industry over different types of FOP labels? How likely will the industry support or deter specific legislative attempts through lobbying and campaign? To what extent will the legislation effort and public demand drive food reformulation? How could FOP labels be used in combination with other policies such as a soda tax, healthy food subsidy, and nutrition education, to achieve a higher impact in nudging people toward a healthier diet?

For advancing research in this field, a few focal points should be noted. First, large-scale, population-representative study samples should be recruited to produce generalizable results and allow subpopulation-specific estimates by gender, age group, race/ethnicity, and other socio-economic dimensions. Second, studies should adopt a randomized controlled study design and measure purchases in real-world settings. Third, multi-arm factorial designs are encouraged to compare the relative effectiveness of various FOP label types. Fourth, studies should follow participants for an extended period to access the mid-to-long-term efficacy of FOP labels on purchase behavior modification. Finally, future research may focus on the effectiveness of health warning labels and nutri-score, which may hold some potential but are currently understudied.

In conclusion, this study reviewed scientific evidence from interventions that examined the effects of various FOP label types on food purchases. FOP labels were found to be effective for participants to make healthier food purchase decisions in five of the 12 studies assessing traffic lights labels, in one of the two studies evaluating health warning labels, and in one study evaluating high sugar symbol labels. Three studies evaluating health star ratings and one assessing daily intake guide labels revealed no effect. Large-scale, population-representative RCTs with factorial designs are warranted to examine the effect of FOP labels by population subgroups and the mid-to-long-term efficacy on purchase behavior modification.

Author statements

Ethical approval

Not required.

Funding

The authors have no funding source to report.

Competing interests

The authors have no conflict of interests to declare.


(“front-of-pack” OR “pack” OR “packs” OR “package” OR “packages” OR “signpost” OR “traffic light” OR “traffic-light” OR “health star”) AND (“label” OR “labels” OR “labeling” OR “labelling” OR “food labeling”[MeSH]) AND (“humans”[MeSH Terms] AND English [lang]).

References

6. Department of health, the food standards agency in scotland, Northern Ireland and Wales. Guide to creating a front of pack (Fop) nutrition label for pre-packed products sold through retail outlets. 2016. Available at: https://www.food.gov.