Preventive Medicine xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Preventive Medicine

journal homepage: www.elsevier.com/locate/ypmed

Comparing five front-of-pack nutrition labels' influence on consumers' perceptions and purchase intentions

Mary T. Gorski Findling[a,*], Paul M. Werth[b], Aviva A. Musicus[c], Marie A. Bragg[d], Dan J. Graham[e], Brian Elbel[f], Christina A. Roberto[g]

[a] Interfaculty Initiative in Health Policy, Harvard University, Cambridge, MA, United States
[b] Department of Psychology, Saint Louis University, St Louis, MO, United States
[c] Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, United States
[d] NYU School of Medicine and College of Global Public Health, New York City, NY, United States
[e] Colorado State University, Fort Collins, CO, United States
[f] NYU School of Medicine and Wagner Graduate School of Public Service, New York City, NY, United States
[g] Medical Ethics & Health Policy, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

ARTICLE INFO

Keywords:
Front-of-package food labels
Nutrition labeling
Food choices
Obesity

ABSTRACT

In 2011, a National Academy of Medicine report recommended that packaged food in the U.S. display a uniform front-of-package nutrition label, using a system such as a 0–3 star ranking. Few studies have directly compared this to other labels to determine which best informs consumers and encourages healthier purchases. In 2013, we randomized adult participants (N = 1247) in an Internet-based survey to one of six conditions: no label control; single traffic light; multiple traffic light; Facts Up Front; NuVal; or 0–3 star ranking. We compared groups on purchase intentions and accuracy of participants' interpretation of food labels. There were no differences in the nutritional quality of hypothetical shopping baskets across conditions ($p = 0.845$). All labels improved consumers' abilities to judge the nutritional quality of foods relative to no label, but the best designs varied by outcomes. NuVal and multiple traffic light labels led to the greatest accuracy identifying the healthier of two products ($p < 0.001$), while the multiple traffic light also led to the most accurate estimates of saturated fat, sugar, and sodium ($p < 0.001$). The single traffic light outperformed other labels when participants compared nutrient levels between similar products ($p < 0.03$). Single/multiple traffic light and Facts Up Front labels led to the most accurate calories per serving estimations ($p < 0.001$). Although front-of-package labels helped participants more accurately assess products' nutrition information relative to no label, no conditions shifted adults' purchase intentions. Results did not point to a clearly superior label design, but they suggest that a 3-star label might not be best for educating consumers.

1. Introduction

Policy makers worldwide are interested in cost-effective approaches to address obesity and related chronic diseases. One popular strategy to encourage healthier eating habits is front-of-pack nutrition labeling on packaged food products or labels on supermarket shelves that provide consumers with clear, easy-to-understand nutrition information (Food and Drug Administration, 2009; Institute of Medicine, 2009; Institute of Medicine, 2011). Several countries have implemented different mandatory or voluntary front-of-package labeling systems, including traffic light labels in the U.K. and Ecuador; (Ecuador Ministry of Public Health, 2013) the Choices checkmark in the Netherlands, Czech Republic, Belgium, and Poland; (The Choices Programme, 2017) the Nordic Keyhole symbol; (Swedish National Food Administration, Danish Veterinary and Food Administration, Norwegian Directorate of Health, Norwegian Food Safety Authority, 2012) Health Stars in Australia; (Commonwealth of Australia, 2014) warning labels in Chile; (U.S. Department of Agriculture, Foreign Agriculture Service, 2015) and a "Healthier Choice" label in Singapore (Singapore Government Health Promotion Board, 2017).

In the United States in 2009, the Food and Drug Administration (FDA) launched a front-of-package labeling initiative to promote a science-based, uniform system (Food and Drug Administration, 2009). For this initiative, the Institute of Medicine (now the National Academy of Medicine; NAM) convened a committee recommending the FDA and U.S. Department of Agriculture make healthier options clearer by

* Corresponding author at: Interfaculty Initiative in Health Policy, Harvard University, 14 Story Street, 4th Floor, Cambridge, MA 02138, United States.
E-mail address: mgorski@mail.harvard.edu (M.T. Gorski Findling).

https://doi.org/10.1016/j.ypmed.2017.10.022
Received 5 August 2017; Received in revised form 7 October 2017; Accepted 18 October 2017

developing a single, standardized front-of-package symbol to translate information from the Nutrition Facts panel. In 2011, the NAM issued a report recommending the symbol appear on all grocery products to maximize its effectiveness by allowing consumers to compare choices within and across categories (Institute of Medicine, 2011). The report provided examples of labeling systems meeting these criteria, including a three-star symbol rating products from less healthy (0 stars) to more healthy (3 stars), modeled after the Energy Star® Program. The U.S. has not yet imposed a mandatory front-of-pack nutrition labeling system, and instead some products display a voluntary system called Facts Up Front, designed and implemented by the food industry (Facts Up Front, 2011).

Prior research has found that certain nutrition label designs are better than others at improving the accuracy of consumers' nutritional judgments of foods (Roberto et al., 2012a; Roberto et al., 2012b). Additional studies demonstrate that point-of-purchase nutrition labels can encourage healthier purchases in cafeterias and supermarkets (Thorndike et al., 2014; Sonnenberg et al., 2013; Cawley et al., 2015; Levy et al., 1985). Although there is some evidence that nutrition labels can influence behaviors, few studies directly compare different label designs to determine which is most effective (Hersey et al., 2013; Watson et al., 2014; Hodgkins et al., 2015). Further, there is little data on the potential influence of the 3-star design proposed in the NAM report, and to our knowledge, no U.S.-based studies that compare it to several common labels.

The objective of this study was to compare the NAM-recommended 3-star labeling system to other common front-of-pack nutrition symbols to determine which is easiest for consumers to understand and use, and which is most likely to influence hypothetical purchasing decisions.

2. Methods

2.1. Study sample

A survey firm (Survey Sampling International (SSI)) used a three-stage process to recruit participants. First, randomly selected participants from SSI's online panels were combined with those recruited through websites and telephone and were invited to take a survey, with no details provided to reduce selection bias. After recruitment, potential participants completed proprietary quality control questions before inclusion in the study. Remaining adults were then randomly assigned to surveys they were likely eligible to complete. This survey was administered in June–July 2013 via Qualtrics, an online survey program. Data were analyzed in March 2017. All participants were at least 18 years old and were recruited such that roughly half the sample would be female and participants would approximate the educational profile of the U.S. based on 2010 Census data. The Harvard T.H. Chan School of Public Health Human Subjects Committee approved this study.

2.2. Label conditions

After participants provided informed consent, they were randomized to one of six front-of-package labels, appearing on food images displayed in Fig. 1:

(1) No front-of-package label (control).
(2) Single Traffic Light - calories per serving label and traffic light symbol (red, yellow, or green) reflecting the product's overall nutritional quality.
(3) Multiple Traffic Light - calories per serving label with traffic light symbol (red, yellow, or green), signifying high/medium/low amounts of saturated fat, sodium, and added sugars, with “High/Med/Low” text within the corresponding traffic light circles.
(4) Facts Up Front – “Facts Up Front” label designed by the food industry (Facts Up Front, 2011) with calories, saturated fat, sodium, and sugars per serving; nutrient amounts displayed in grams/mg and percent daily value.
(5) NuVal - label developed by a nutrition researcher (Katz et al., 2010) displaying a 1 to 100 score; higher scores indicate healthier products.
(6) 3-Star - label based on a NAM-recommended design displaying calorie information plus a zero (least healthy) to three (most healthy) star rating, signifying amounts of saturated/*trans* fats, sodium, and added sugars.

We hypothesized that all labels would increase consumer understanding of the nutritional quality of packaged foods and encourage healthier hypothetical purchases relative to no label. Based on previous findings (Roberto et al., 2012a; Roberto et al., 2012b), we also hypothesized that traffic light labels would lead to the most accurate judgments of the nutritional content of packaged foods and the healthiest hypothetical food choices. Such labels should be easily processed because color coding makes them salient, and they leverage automatic associations between “red: stop” and “green: go.” In addition, other research found using “High/Med/Low” text helped consumers better understand labels (Malam et al., 2009). In contrast, we predicted that labeling systems with only numeric information (NuVal, Facts Up Front) would perform worst because they display numbers and/or percentages that require greater cognitive engagement and lack features that would increase saliency (e.g., colors, images). Finally, we predicted that the 3-star symbol would perform better than numeric labels because it uses a simple symbolic presentation rather than numeric information, but worse than traffic light labels because it may be less intuitive and salient than color coded traffic lights. Further, although 1–5 star ranking systems are common (e.g., Amazon and Yelp), the 0–3 star system may be less familiar to consumers.

2.3. Nutrition criteria for labeling systems

Nutrition information for all products was obtained from the Nutrition Facts panel on food packaging or from food manufacturer websites. To test the effectiveness of the front-of-package label design, rather than the underlying nutrition criteria, we used the same nutrient profile model algorithm to assign all foods a healthfulness score (Rayner et al., 2005). The Nutrient Profile Model is an algorithm that adds points for calories and negative nutrients (saturated fat, sugar, sodium), and subtracts points for positive nutrients (fiber, protein) and the percentage of fruit, vegetables, or nuts (not included in calculations because products in this study had minimal amounts) (Lobstein and Davies, 2009). This model has been validated by nutritionists (Lobstein and Davies, 2009; Scarborough et al., 2007; Arambepola et al., 2007) and is used to inform food policies in the U.K. and Australia (Food Standards Australia and New Zealand, 2016; United Kingdom Department of Health, 2011). Consistent with other research (Bragg et al., 2013; Bragg et al., 2012), the nutrient profile model score was converted to a Nutrient Profile Index (NPI) to improve interpretability using the formula: [NPI score = (− 2) × NPM score + 70], where 1 is the worst possible nutrition score and 100 is the best score (Rayner et al., 2005). One limitation of the NPM is that it scores most sugar-sweetened beverages similarly because sugar is the only nutrient. Therefore, for beverages, we used caloric content as a measure of healthfulness, rather than NPI score, to produce more variability across products. *Supplementary Table 1* presents NPI scores for foods and calories for beverages used in the shopping task. For the multiple traffic light system, nutrients were labeled high, medium, or low, based on the UK Food Standards Agency's cut-offs for the NPI (Rayner et al., 2005).

2.4. Survey procedure and main outcomes

2.4.1. Shopping basket score

After providing consent, participants were randomized to a label

Fig. 1. Front-of-package food labels tested. Internet survey of 1247 adults, conducted June–July 2013. Note: Facts Up Front displays calories, saturated fat, sodium, and sugar per serving, and it includes amounts of nutrients in grams/mg and % daily value information. Single Traffic Light uses a calories-per-serving label and a traffic light symbol (red, yellow, or green) reflecting overall nutritional quality. Multiple Traffic Light uses a calories-per-serving label and a modified traffic light symbol with High/Med/Low text, indicating amounts of saturated fat, sodium, and added sugars. NuVal displays a 1–100 score; higher scores indicate healthier products. The National Academy of Medicine 3-star label includes calories per serving and a star rating: 0 stars (least healthy) to 3 stars (most healthy), which represents the amount of saturated and trans fats, sodium, and sugar per serving. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

condition and asked to pretend they were on a typical shopping trip for groceries. Participants selected one out of five products they wanted to buy this week from each of five categories: soups, desserts, cereals, grains, and beverages (*Supplementary Table 1* lists the 25 products). Although we included six categories originally, there was an image error for "milk," so it was excluded from analysis. Participants only saw the front package for all images and did not see the Nutrition Facts Panel. All health claims (e.g., "low sodium") were removed from product images. The primary outcome of shopping basket health score was created by averaging the NPI score of selected foods (not including beverages). Our other main outcome examined mean beverage calories per serving. As an exploratory secondary outcome, we examined mean NPI score within each food category.

After the shopping task, we measured participants' ability to understand and use different front-of-pack labels. Because the next series of outcomes were focused on consumer knowledge, we wanted to provide more information about each labeling system. Therefore, participants viewed a public service advertisement (PSA) for 1 min (fixed across conditions) explaining how to interpret their assigned labeling system (example shown in Supplementary Fig. 1) before making nutritional judgments about products. Control condition participants viewed a similar advertisement encouraging healthy choices without label interpretation instructions (see Supplementary Fig. 2).

2.4.2. Nutrient comparison quiz

After viewing the PSA, participants were asked to identify which of two side-by-side products was higher or lower in certain nutrients (see *Supplementary Table 2* for question wording and product names). Front-of-pack labels were enlarged and displayed next to products to ensure visibility (see Supplementary Fig. 3). Included products were popular foods in the same food category with similar nutritional profiles, except for a particular nutrient. For example, Campbell's® Chicken Noodle Soup was compared to Progresso® Chicken Noodle Soup, and participants were asked to identify which was lower in sodium. To create this outcome, we summed all correct answers on eight questions and created a total percent correct score (out of 100).

2.4.3. Healthier product quiz

Participants were then shown eight pairs of products side-by-side (shown in *Supplementary Table 2*) and asked to select which was healthier. We summed all correct answers on eight questions (based on the healthier NPI score) and created a total percent correct score.

2.4.4. Individual nutrient quizzes

Participants then estimated whether seven food products (shown in *Supplementary Table 2*) had low, medium, or high amounts of saturated fat, sugar, sodium, protein, and fiber. Answers were considered correct if they matched the UK Food Standard Agency's nutritional threshold criteria for "low," "medium," and "high." Total percentage-correct scores were calculated across all seven products.

2.4.5. Calories per serving quiz

Participants then estimated the amount of calories per serving, where exact answers were coded as correct. Total percentage-correct scores for all seven products were calculated.

2.4.6. Product ratings

Participants rated how healthy they thought each of the seven products (in *Supplementary Table 2*) was, how good it would taste, and the likelihood that they would buy it for themselves and their children (among parents) using a 7-point Likert-type scale. Each set of ratings was averaged across all seven products.

2.4.7. Label preferences

At the end of the survey, participants answered questions about their assigned label, including the degree to which they would use it to guide purchasing decisions and whether it was confusing, on a 5-point Likert-type scale used in prior studies (Roberto et al., 2012a; Roberto et al., 2016). A composite score was also created for a variable "too much information and time" by averaging responses to two questions on whether the assigned label had too much information or required too much time to read (see *Supplementary Table 2* for question wording). Each set of label preferences was averaged across all seven products.

2.4.8. Sociodemographic characteristics

Participants were asked about their personal eating and food shopping habits, as well as socio-demographic questions regarding age, gender, race/ethnicity, education, income, parental status, height, weight, and trying to lose weight. Self-reported height and weight were used to calculate body mass index (BMI). How much nutrition labels generally influenced food choices was measured on a 9-point Likert-type scale.

2.5. Statistical analyses

Data analysis was performed using STATA v.14 (Stata, 2015). Chi-squared tests were used to compare group differences for categorical demographic variables. Because of significant group differences in gender ($p < 0.001$), all analyses for main study outcomes controlled for gender. Continuous study outcomes were compared using ANCOVAs, followed by post-hoc Tukey tests with the Tukey-Kramer adjustment to account for unequal sample sizes.

3. Results

3.1. Participants

The study sample includes 1247 adults (out of 1768 recruited; Fig. 2 displays recruitment flow including reasons for exclusion). The mean respondent age was 43 ± 16 years old, with 54% women and about half (49%) with a high school education or less. Median survey completion time was 22 min, with no significant differences between groups. The proportion of individuals excluded was significantly higher among the No Label condition (39.1%) compared to the NuVal (25.9%) and single (24.0%) and multiple traffic light conditions (all p-values < 0.01), but did not significantly differ from the Facts Up Front (32.1%, $p = 0.442$) or 3-star conditions (28.8%, $p = 0.072$). However, the proportion of individuals excluded did not differ across conditions displaying front-of-package labels ($p = 0.120$). We believe the No Label group had a higher exclusion rate because participants did not see any labels food images to help them answer questions. Based on participant feedback at the end of the survey, it appears some control group enrollees were frustrated trying to answer questions about the nutritional profiles of foods without front-of-package labels, so they were less likely to complete the survey. Study groups did not significantly differ by age, race/ethnicity, BMI, education, income, or efforts to lose weight (see Table 1 for sample description). Despite randomization, the NuVal and multiple traffic light groups contained a higher proportion of female respondents (63.4% and 61.3%) compared to Facts up Front (46.2%) and 3-star groups (46.0%), and NuVal had a higher proportion

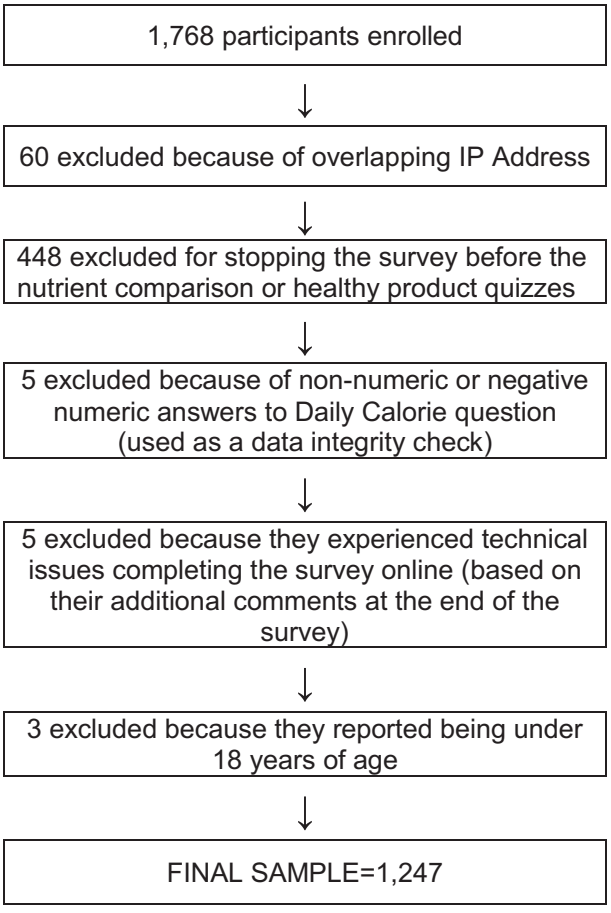

Fig. 2. CONSORT flow diagram of study enrollment and exclusion criteria. Internet survey of 1247 adults, conducted June–July 2013.

Table 1
Sociodemographic information for front-of-package nutrition label Internet study sample[a].

Characteristic	M (± SD) or *n* (col. %)	df	Test statistic	*p*-Value[b]
Age (years, M [SD])	43.4 (15.9)	5;1193	0.26	0.935
BMI (M [SD])	29.1 (8.1)	5;1153	1.23	0.294
Normal weight (< 25)	412 (35.5%)	5	3.87	0.569
Overweight/obese (≥ 25)	747 (64.5%)			
Gender[c]				
Female	673 (54.0%)	5	26.15	< 0.001
Race/ethnicity		15	24.57	0.056
White, non-Hispanic	870 (70.1%)			
Black, non-Hispanic	158 (12.7%)			
Hispanic	106 (8.5%)			
Other	108 (8.7%)			
Education level		20	22.60	0.309
< High school	113 (9.1%)			
High school or GED	497 (40.1%)			
Some college, associate's degree, trade school	287 (23.1%)			
Bachelor's degree	227 (18.3%)			
Graduate/ professional degree	117 (9.4%)			
Income		25	16.25	0.907
< $15,000	172 (13.9%)			
$15,000–$30,000	256 (20.6%)			
$30,001–$45,000	194 (15.6%)			
$45,001–$60,000	200 (16.1%)			
$60,001–$75,000	139 (11.2%)			
$75,000 +	279 (22.5%)			
Married		5	1.23	0.942
Yes	540 (43.5%)			
Currently trying to lose weight		5	2.38	0.794
Yes	639 (51.2%)			
No	608 (48.8%)			
How much nutrition labels generally influence food/ drink choices[d] (M [SD])	6.0 (2.5)	5;1241	0.78	0.566

Internet survey of 1247 adults, conducted June–July 2013. df = degrees of freedom.

[a] Table values are M ± SD for continuous variables and *n* (column %) for categorical variables. Percentages are calculated based on the total number of participants providing data for each variable.

[b] Test statistics and *p*-values are for univariate ANOVAs (continuous variables) or chi-squared tests (categorical variables) comparing the six front-of-package study conditions.

[c] Percent female, by study condition: No Label (47.4%); Single Traffic Light (55.0%); Multiple Traffic Light (61.3%); NuVal (63.4%); Facts Up Front (46.2%); NAM 3-Star Label (46.0%).

[d] Measured on a 9-point Likert scale, 1 = do not influence, 9 = strongly influence.

Table 2
Study outcomes by front-of-package label condition, % (± SE) unless otherwise noted.

Outcome	No Label (control) (n = 171)[a]		Single Traffic Light (n = 222)[b]		Multiple Traffic Light (n = 235)[c]		NuVal (n = 235)[d]		Facts Up Front (n = 184)[e]		NAM 3-Star Label (n = 200)[f]		F*	p*
	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE		
Shopping basket nutrient profile index scores (1 − 100)[g]														
Overall	58.6	(0.5)	59.0	(0.4)	59.1	(0.4)	58.7	(0.4)	59.5	(0.7)	58.5	(0.7)	0.41	0.845
Soup	67.8	(0.3)	67.2	(0.3)	67.5	(0.3)	67.9	(0.3)	67.5	(0.3)	67.7	(0.3)	1.09	0.365
Dessert	58.8	(0.8)	58.6	(0.7)	59.8	(0.7)	58.1	(0.7)	58.7	(0.9)	59.2	(0.8)	0.65	0.664
Cereal	56.1	(1.1)	57.9	(1.0)	57.3	(1.0)	56.7	(1.0)	56.8	(1.2)	56.0	(1.2)	0.47	0.801
Rice/grains	51.9	(1.1)	52.2	(1.0)	51.9	(0.9)	52.0	(0.9)	52.1	(1.1)	52.3	(1.1)	0.02	1.000
Beverage kcal/serving	56.2	(3.4)	52.5	(3.0)	56.6	(2.9)	59.4	(2.9)	59.3	(3.6)	64.3	(3.5)	1.49	0.190
Quiz scores (0 − 100)														
Healthier product quiz	50.6[c,d,e,f]	(1.4)	47.7[c,d,e,f]	(1.2)	63.2[a,b,e,f]	(1.2)	67.4[a,b,e,f]	(1.2)	56.9[a,b,c,d]	(1.3)	57.6[a,b,c,d]	(1.3)	36.52	< 0.001
Nutrient comparison quiz	41.5[b,d,e,f]	(1.4)	62.5[a,c,d,e,f]	(1.2)	36.8[b,d,e,f]	(1.2)	53.7[a,b,c,f]	(1.2)	56.9[a,b,c,f]	(1.4)	47.3[a,b,c,d,e]	(1.3)	59.73	< 0.001
Saturated fat quiz	45.5[c,e]	(1.4)	46.2[c,e]	(1.2)	89.9[a,b,d,e,f]	(1.2)	44.8[c,e,f]	(1.2)	64.5[a,b,c,d,f]	(1.4)	50.6[c,d,e]	(1.3)	210.39	< 0.001
Sugar quiz	44.6[c]	(1.4)	42.5[c,e]	(1.3)	87.8[a,b,d,e,f]	(1.2)	43.7[c,e]	(1.2)	49.3[b,c,d,f]	(1.4)	43.5[c,e]	(1.3)	206.31	< 0.001
Sodium quiz	44.3[c,e]	(1.5)	42.5[c,e]	(1.3)	87.6[a,b,d,e,f]	(1.2)	40.3[c,e]	(1.2)	52.0[a,b,c,d,f]	(1.4)	40.6[c,e]	(1.3)	214.51	< 0.001
Protein quiz	40.7	(1.7)	41.1	(1.5)	44.3	(1.4)	40.1[e]	(1.4)	46.5[d]	(1.6)	40.6	(1.6)	2.77	0.017
Fiber quiz	43.2	(1.5)	41.5[c,e]	(1.3)	48.5[b,d]	(1.3)	42.7[c,e]	(1.3)	48.6[b,d]	(1.5)	46.7	(1.4)	5.22	< 0.001
Calories per serving quiz	5.8[b,c,e,f]	(2.4)	76.1[a,d,f]	(2.1)	77.1[a,d,f]	(2.1)	7.1[b,c,e,f]	(2.1)	79.6[a,d,f]	(2.3)	60.8[a,b,c,d,e]	(2.2)	255.72	< 0.001
Product ratings (1–7)[h]														
Healthfulness	4.2	(0.1)	4.1	(0.1)	4.0	(0.1)	4.2[e]	(0.1)	3.9[d,f]	(0.1)	4.3[e]	(0.1)	3.35	0.005
Taste	4.9	(0.1)	5.0	(0.1)	4.9	(0.1)	4.9	(0.1)	4.8	(0.1)	5.1	(0.1)	1.06	0.384
Intent to purchase (Self)	3.9	(0.1)	3.8	(0.1)	3.9	(0.1)	3.9	(0.1)	3.8	(0.1)	4.1	(0.1)	1.38	0.229
Intent to purchase (Children)[i]	4.1	(0.1)	3.8	(0.1)	3.7	(0.1)	3.8	(0.1)	3.6	(0.1)	3.8	(0.1)	1.28	0.270
Label preferences (1–5)[j]														
Use label to guide purchasing decisions?	–	–	3.1[c,e]	(0.1)	3.8[b,d,f]	(0.1)	3.3[c,e]	(0.1)	3.8[b,d,f]	(0.1)	3.3[c,e]	(0.1)	29.00	< 0.001
Required too much information & time?	–	–	2.0	(0.1)	2.0	(0.1)	2.0	(0.1)	2.0	(0.1)	2.0	(0.1)	0.44	0.819
Confusing	–	–	2.0	(0.1)	1.8[d]	(0.1)	2.3[c,e]	(0.8)	1.9[d]	(0.1)	2.1	(0.1)	4.71	< 0.001

Internet survey of 1247 adults, conducted June–July 2013. NAM = National Academy of Medicine.

* F statistics and p-values display the effect of front-of-package labeling on outcomes, after controlling for gender. All estimates in this table control for gender.

[a] Significantly different compared to No-Label Control, $p < 0.05$.

[b] Significantly different compared to Single Traffic Light, $p < 0.05$.

[c] Significantly different compared to Multiple Traffic Light, $p < 0.05$.

[d] Significantly different compared to NuVal, $p < 0.05$.

[e] Significantly different compared to Facts Up Front, $p < 0.05$.

[f] Significantly different compared to NAM 3-Star Label, $p < 0.05$.

[g] Nutrient Profile Index (NPI) Score is an algorithm used to assign foods a healthfulness score; NPI score = (− 2) × Nutrient Profile Model score + 70; scale 1–100; 1 indicates least healthful, 100 is most healthful.

[h] Measured on a 7-point Likert scale.

[i] This question was only answered by participants with children, n = 744.

[j] Measured on a 5-point Likert sale. Question not asked of participants in the No Label condition.

of females compared to the No Label control condition (47.4%) (all p-values < 0.03).

3.2. Main outcomes

All results are summarized in Table 2, and Table 3 displays a summary of the study's key findings.

3.2.1. Shopping basket score

The mean NPI score for the hypothetical shopping basket was 58.6 ± 0.5 (within a possible range of 46.5 to 73.0). To put this in context, a score of ≥ 64 is used as the threshold for healthy food products allowed for advertising to children in the U.K. Overall there were no significant differences in mean shopping basket NPI score across conditions (p = 0.845) or for beverage calories per serving (p = 0.190).

3.2.2. Healthier product quiz

All labels led to greater accuracy in consumers' ability to identify the healthier of two products relative to no label (all p-values < 0.02), except for the single traffic light (p = 0.624). The NuVal and multiple traffic light systems led to the highest scores on this outcome (67.4% and 63.2%) and did not differ from one another (p = 0.125).

3.2.3. Nutrient comparison quiz

All front-of-package labels significantly increased consumers' ability to compare two products and identify which was higher or lower in certain nutrients compared to no label (mean quiz score: 41.5% correct), except for the multiple traffic light label (36.8%, p = 0.116). The single traffic light label led to the highest scores on this quiz (62.5% correct), and was the only label to significantly outperform all other conditions (all p-values < 0.03).

3.2.4. Individual nutrient quizzes

The multiple traffic light group significantly outperformed other

Table 3
Summary of front-of-package label performance for main outcomes[a].

	No Label (control)	Single Traffic Light	Multiple Traffic Light	NuVal	Facts Up Front	NAM 3-Star Label
Shopping basket score						
Overall						
Soup						
Dessert						
Cereal						
Rice						
Beverage						
Nutrient quizzes						
Nutrient comparison		X				
Saturated fat			X			
Sugars			X			
Sodium			X			
Fiber						
Protein						
Calories per serving		X	X		X	
Healthier product quiz			X	X		
Perceptions						
Healthfulness						
Taste						
Intent to purchase						
Label preferences						
Would use label to guide purchasing decisions			X		X	
Required too much information and time						
Confusing						

Internet survey of 1247 adults, conducted June–July 2013. NAM = National Academy of Medicine.

[a] Each X indicates the label(s) that significantly outperformed the others on that outcome. Multiple Xs indicate that those front-of-package label groups outperformed the other groups and did not significantly differ from one another. An absence of Xs indicates that none of the groups differed from the control group and/or one another.

groups by > 25 percentage points (all p-values < 0.001) on quizzes that asked participants to estimate whether products had low, medium, or high amounts of saturated fat, sugar, and sodium. For the protein and fiber quizzes, label groups were not significantly different from the control.

3.2.5. Calories per serving quiz

Every label significantly improved consumers' ability to accurately estimate the calories per serving in products relative to no label (no label mean score was 5.8% correct), except for NuVal (7.1%, $p = 1.000$), which was the only label not displaying calorie information. In addition, the single traffic light (76.1%), multiple traffic light (77.1%), and Facts Up Front labels (79.6%) led to significantly higher quiz scores compared to NuVal (7.1%) and the 3-star (60.8%) labels, but were not significantly different from each other.

3.2.6. Additional perception ratings

There were no significant differences between label groups and the control in perceptions of healthfulness, taste, or intent to purchase products for oneself or one's children. Among label conditions, Facts Up Front led to lower perceptions of product healthfulness (3.9) compared to the NuVal (4.2, $p = 0.021$) and the 3-star symbol (4.3, $p = 0.007$).

3.2.7. Label preferences

Adults in the multiple traffic light and Facts Up Front groups indicated they would use these labels to guide purchasing decisions more than those in other label groups (all p-values < 0.001). Adults viewing NuVal labels rated them as more confusing than those viewing multiple traffic light and Facts Up Front labels (all p-values < 0.01). There were no group differences in whether a label required too much information and time to use ($p = 0.819$).

4. Discussion

Our first key finding is that seeing any front-of-package labels significantly improved adults' ability to understand and compare the nutritional quality of foods and beverages relative to no label. This is consistent with other research showing front-of-package nutrition labels improve consumers' understanding of the nutritional quality of food (Roberto et al., 2012a; Roberto et al., 2012b; Thorndike et al., 2014; Hersey et al., 2013; Watson et al., 2014; Hodgkins et al., 2015), and it suggests that policies requiring front-of-pack labels are likely to improve consumer knowledge. However, identifying an optimal front-of-package design based on our results was difficult because results varied by outcome. Arguably, it is most important for consumers to be able to compare two products and determine which is healthier. For that outcome, the NuVal and multiple traffic light labels performed best, while the single traffic light performed worst. However, when consumers compared products on certain levels of nutrients (i.e. which product is lower in sodium, sugar, saturated fat, calories), the single traffic light performed best, while the multiple traffic light and no-label conditions performed worst. We think the multiple traffic light labels were less helpful because many products had the same color traffic light for certain nutrients. Yet when we asked participants to view an individual product and estimate levels of certain nutrients (e.g., saturated fat, sodium), they did far better when viewing a multiple traffic light compared to all other labels, while other labels did not perform better than no label (except for Facts Up Front on two quizzes). Finally, all labels displaying calories helped consumers estimate calories, although the 3-star label performed worst of all labels showing calorie information.

However, despite increased consumer knowledge, none of the labels influenced the healthfulness of adults' hypothetical purchases. There are very few real-world studies of front-of-package nutrition labels for comparison with these results. One study of multiple traffic light labels found no significant increase in purchases of healthier sandwiches or ready-to-eat meals after labels were introduced (Sacks et al., 2009), while other real-world studies observed associations between labeling systems and increased purchases of healthier items (Levy et al., 1985) or decreased purchases of less healthy items (Cawley et al., 2015). It is possible that existing brand preferences, the hypothetical setting, products used, and/or sampling explain differences across studies. It is also possible that viewing a PSA prior to the shopping task might have increased label effectiveness, given previous research findings that viewing cues like PSAs before a shopping task increases the salience of front-of-package labels (Graham et al., 2015). Further, because participants were forced to choose an item in each category, we were unable to assess whether the labels would have dissuaded them from selecting certain foods altogether.

This study has several limitations. Although participants were recruited to reflect the gender and educational profile of the U.S., we did not recruit a randomly sampled, nationally representative sample. Further, due to group differences by gender, some groups had more women than men. Although we controlled for gender in our analyses, this may have influenced our results. In addition, in this hypothetical setting, we only tested a small number of products (although a range of product types were tested), and participants only made decisions at one point in time, in contrast to real-world shopping with repeated exposure to labels. In addition, the control group had higher attrition than some other groups, which may bias our results in either direction, depending on baseline nutrition knowledge. However, we did not see differential

attrition across the five front-of-package label groups, increasing our confidence in being able to compare results across label groups. Finally, we examined perceptions after participants received information about how to interpret the labeling systems. In real-world settings, consumer education efforts may be limited, and consumers can be simultaneously exposed to different labels and health claims that compete for their attention.

Despite these limitations, this study adds to the literature in several ways. It is the first U.S. study, to our knowledge, comparing a 3-star labeling system recommended by NAM with other common systems. It also includes a randomized design, a large sample approximating the educational profile of the U.S., and a range of products and perceptual outcomes.

More than eight years after the FDA announced it would take action on front-of-package labels, there is still no mandatory labeling system in the U.S. Our results suggest that a uniform front-of-package labeling system may effectively improve consumer knowledge, but its impact on consumer behavior without accompanying cues, education, or promotion efforts is less clear. There is still no label design that has been identified as the "best" at educating consumers and improving the nutritional profile of what people buy, so researchers and government agencies should continue to test and compare consumer responses to science-based label designs. Because consumers are more likely to use front-of-package nutrition labels than Nutrition Facts labels (Graham et al., 2015), front-of-package labels present an important opportunity to inform and persuade consumers at the point-of-purchase. Our results strongly suggest that uniform front-of-pack labeling systems can improve consumer judgments compared to no label, but the impact of the label varies based on the outcome measured. Both the NuVal and multiple traffic light labels helped consumers identify the healthier of two products, and the multiple traffic light also improved their ability to estimate certain nutrient levels in products. In contrast, the 3-star symbol did not outperform any of the other labels on any outcomes, suggesting it might not be the best label design. Future research should explore why labels differ across nutrition outcomes and identify which outcomes are most important and relevant to consumers' knowledge and real-world decision making.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ypmed.2017.10.022.

Conflict of interest statement

This work was supported by an RWJF Health and Society Scholars grant. The funding source had no role in the design and conduct of the study, collection, management, analysis and interpretation of the data and preparation, review, or approval of the manuscript. M.T. Gorski Findling was supported by grant number T32HS000055 from the Agency for Healthcare Research and Quality. A.A. Musicus is supported by the National Institute of Environmental Health Sciences of NIH under Training Grant 5T32ES007069-37. C. Roberto is supported by the National Institute on Aging of the National Institutes of Health under Award Number P30AG034546. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Agency for Healthcare Research and Quality.

Financial disclosures

Marie A. Bragg has no financial disclosures.
Brian Elbel has no financial disclosures.
Mary T. Gorski Findling has no financial disclosures.
Dan J. Graham has no financial disclosures.
Aviva Musicus has no financial disclosures.
Christina A. Roberto has no financial disclosures.
Paul M. Werth has no financial disclosures.

Author contributions

M.T. Gorski Findling conducted the statistical analyses and led the writing of this manuscript.

P.M. Werth led data acquisition and provided feedback on drafts of the manuscript.

C.A. Roberto obtained funding, originated the study idea and design, oversaw data acquisition, and provided critical feedback on drafts of the manuscript.

A.A. Musicus created survey images and provided critical feedback on drafts of the manuscript.

M. Bragg, D. Graham, and B. Elbel provided critical feedback on drafts of the manuscript.

Transparency document

The Transparency document associated with this article can be found, in online version.

References

Arambepola, C., Scarborough, P., Rayner, M., 2007. Validating a nutrient profile model. Public Health Nutr. 11, 371–378.

Bragg, M.A., Liu, P.J., Roberto, C.A., Sarda, V., Harris, J.L., Brownell, K.D., 2012. The use of sports references in marketing of food and beverage products in supermarkets. Public Health Nutr. 16 (4), 738–742.

Bragg, M.A., Yanamadala, S., Roberto, C.A., Harris, J.L., Brownell, K.D., 2013. Athlete endorsements in food marketing. Pediatrics 132 (5), 1–6.

Cawley, J., Sweeney, M.J., Sobal, J., Just, D.R., 2015. The impact of a supermarket nutrition rating system on purchases of nutritious and less nutritious foods. Public Health Nutr. 18 (1), 8–14.

Commonwealth of Australia, 2014. Health Star Rating System. Available online: http://healthstarrating.gov.au/internet/healthstarrating/publishing.nsf/content/home, Accessed date: 1 August 2017.

Ecuador Ministry of Public Health, November 2013. No. 4522, El Reglamento de Etiquetado de Alimentos Procesados.

Facts Up Front, 2011. Available online: http://www.factsupfront.org/Newsroom/2, Accessed date: 1 August 2017.

Food and Drug Administration, 2009. Background Information on Point of Purchase Labeling. https://www.fda.gov/Food/IngredientsPackagingLabeling/LabelingNutrition/ucm187320.htm#front, Accessed date: 1 August 2017.

Food Standards Australia and New Zealand, May 2016. Short Guide for Industry to the Nutrient Profiling Scoring Criterion in Standard 1.2.7—Nutrition, Health and Related Claims. http://www.foodstandards.gov.au/industry/labelling/Documents/Short-guide-for-industry-to-the-NPSC.pdf, Accessed date: 1 August 2017.

Graham, D.J., Heidrick, C., Hodgin, K., 2015. Nutrition label viewing during a food selection task: front-of-package labels vs. Nutrition Facts Labels. JAND 115 (10), 1636–1646.

Hersey, J.C., Wohlgenant, K.C., Arsenault, J.E., Kosa, K.M., Muth, M.K., 2013. Effects of front-of-package and shelf nutrition labeling systems on consumers. Nutr. Rev. 71 (1), 1–14.

Hodgkins, C.E., Raats, M.M., Fife-Shaw, C., et al., 2015. Guiding healthier food choice: systematic comparison of four front-of-pack labelling systems and their effect on judgments of product healthiness. Br. J. Nutr. 113, 1652–1663.

Institute of Medicine, October 2011. Front-of-Package Nutrition Rating Systems and Symbols: Promoting Healthier Choices. http://www.nap.edu/catalog/13221/front-of-package-nutrition-rating-systems-and-symbols-promoting-healthier, Accessed date: 1 August 2017.

Institute of Medicine, 2009. Examination of Front-of-Package Nutrition Rating Systems and Symbols. http://iom.nationalacademies.org/Activities/Nutrition/NutritionSymbols.aspx, Accessed date: 1 August 2017.

Katz, D.L., Njike, V.Y., Rhee, L.Q., Reingold, A., Ayoob, K.T., 2010. Performance characteristics of NuVal and the Overall Nutrition Quality Index (ONQI). Am. J. Clin. Nutr. 91 (4), 1102S–1108S.

Levy, A.S., Mathews, O., Stephenson, M., et al., 1985. The impact of a nutrition information program on food purchases. J. Public Policy Mark. 4, 1–13.

Lobstein, T., Davies, S., 2009. Defining and labeling 'healthy' and 'unhealthy' food. Public Health Nutr. 12, 331–340.

Malam, S., Clegg, S., Kirwin, S., et al., 2009. Comprehension and Use of UK Nutrition Signpost Labelling Schemes. http://www.food.gov.uk/multimedia/pdfs/pmpreport.pdf, Accessed date: 1 August 2017.

Rayner, M., Scarborough, P., Boxer, A., Stockley, L., 2005. Nutrient Profiles: Development of Final Model. British Heart Foundation Health Promotion Research Group. http://tna.europarchive.org/20120102091535/http://www.food.gov.uk/multimedia/pdfs/nutprofr.pdf (Accessed August 1, 2017).

Roberto, C.A., Bragg, M.A., Schwartz, M.B., et al., 2012a. Facts up front versus traffic light food labels: a randomized controlled trial. Am. J. Prev. Med. 43 (2), 134–141.

Roberto, C.A., Bragg, M.A., Seamans, M.J., Mechulan, R.L., Novak, N., Brownell, K.D., 2012b. Evaluation of consumer understanding of different front-of-package nutrition

labels, 2010–2011. Prev. Chronic Dis. 9, 120015.
Roberto, C.A., Wong, D., Musicus, A., Hammond, D., 2016. The influence of sugar-sweetened beverage health warning labels on parents' choices. Pediatrics 137 (2), e20153185.
Sacks, G., Rayner, M., Swinburn, B., 2009. Impact of front-of-pack 'traffic-light' nutrition labelling on consumer food purchases in the U.K. Health Promot. Int. 24 (4), 344–352.
Scarborough, P., Boxer, A., Rayner, M., Stockley, L., 2007. Testing nutrient profile models using data from a survey of nutrition professionals. Public Health Nutr. 10, 337–345.
Singapore Government Health Promotion Board, 2017. Healthier Choice Symbol Programme. Available online: https://www.hpb.gov.sg/food-beverage/healthier-choice-symbol.
Sonnenberg, L., Gelsomin, E., Levy, D.E., Riis, J., Barraclough, S., Thorndike, A.N., 2013. A traffic light food labeling intervention increases consumer awareness of health and healthy choices at the point-of-purchase. Prev. Med. 57 (4), 253–257.
StataCorp, 2015. Stata Statistical Software: Release 14. StataCorp LP., College Station, TX.
Swedish National Food Administration, Danish Veterinary and Food Administration, Norwegian Directorate of Health, Norwegian Food Safety Authority, 2012, June 25. Design Manual for the Keyhole Logo. Available online: http://www.livsmedelsverket.se/globalassets/produktion-handel-kontroll/livsmedelsinformation-markning-halsopastaenden/nyckelhalet/design-manual-for-the-keyhole-logo.-2012.-livsmedelsverket-m-fl.pdf.
The Choices Programme, 2017. Where is Choices Programme Active? Available online: http://www.choicesprogramme.org.
Thorndike, A.N., Riis, J., Sonnenberg, L.M., Levy, D.E., 2014. Traffic-light labels and choice architecture: promoting healthy food choices. Am. J. Prev. Med. 46 (2), 143–149 (Erratum in Prev Med. 2014 Feb; 59: 86).
U.S. Department of Agriculture, Foreign Agriculture Service, 2015, April 22. Global Agricultural Information Network Report Number CI1506. New Nutritional Labeling Regulation, Chile Available online: https://gain.fas.usda.gov/Recent%20GAIN%20Publications/New%20Nutritional%20Labeling%20Regulation_Santiago_Chile_4-22-2015.pdf.
United Kingdom Department of Health, January 2011. Nutrient Profiling Technical Guidance. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216094/dh_123492.pdf, Accessed date: 1 August 2017.
Watson, W.L., Kelly, B., Hector, D., Hughes, C., King, L., Crawford, J., Sergeant, J., Chapman, K., 2014. Can front-of-pack labelling schemes guide healthier food choices? Australian shoppers' responses to seven labelling formats. Appetite 72, 90–97.