MENU Study: Online Experiment

NCT06909019

Statistical Analysis Plan

Version Date: 3/25/25

Research Question and Hypotheses

The primary objective of this study is to test the relative effects of climate-impact menu label designs on the healthfulness of consumers' fast-food meal choices. We predict that, relative to a control label, a traffic light label will lead to the healthiest food choices, followed by a high climate impact "warning" label, then by a summative "grade" label, and finally a low climate impact "positive" label.

We believe the same label performance predictions will apply to our secondary outcomes, including: total greenhouse gas emissions per meal order, total calories per meal order, and other indicators of the nutrition quality/content of meals selected.

Key Variables

Primary outcome: Participants will order two hypothetical meals from two online fast-food menus via a web-based survey. Participants will be randomized to one of five label conditions that will appear on the menu: 1) a control label, 2) a low climate impact label; (3) a high climate impact label; (4) a three-level traffic-light climate label, and 5) a five-level climate impact grade label. Healthfulness of each participant's meal selections based on a modified Nutrient Profile Index (NPI) score: weighted mean NPI score of all items selected per meal (excluding beverages), with each item's NPI score weighted by its proportionate contribution of mass in grams to the total mass of the meal.

Secondary outcomes: Our first secondary outcome is total greenhouse gas emissions of the hypothetical meal ordered measured in kilograms of CO_2 equivalents, excluding drinks. This will be calculated for each menu item using nutritional and recipe information from publicly available information combined with data on associated emissions for related food items from the World Resource Institute (WRI) CoolFood Calculator. Total estimated greenhouse gas emissions per meal will be calculated as the sum of estimated emissions for all menu items ordered in a meal.

We will also assess healthfulness of fast-food main item selections using NPI score, whether any item with red meat was selected (yes/no), total calories ordered, total saturated fat ordered, total sugar ordered, total protein ordered, total fiber ordered, total sodium ordered, whether an SSB was selected (yes/no), and the total price of the meal ordered.

Additionally, we will measure noticeability, perceived message effectiveness of the various climate-impact labels, and knowledge outcomes. Following the meal ordering task, participants will be asked to indicate whether they noticed a label (other than calorie labels) on the menus, and if yes, to recall the type of label noticed. Participants will also respond to four statements about perceived message effectiveness of the climate-impact label for which they were randomized to view on the menu. Statements are derived from the UNC-PME scale and will assess participant levels of discouragement, unpleasantness, and concern regarding high climate impact menu items using a 5-point Likert scale (1= Not at All, 5 = A Great Deal). Responses will be analyzed as the mean score across all items.

We will also assess knowledge of climate impact of foods. Participants will be prompted to numerically rank three meals with varying main proteins by relative climate impact, with "1" equating to smallest impact and "3" equating to largest impact. We will evaluate the % who correctly rank the 3 items by experimental condition. In two questions (one for items from each restaurant), participants will also indicate which of two items has the largest climate impact (or whether items have the same impact). Responses will be analyzed as the % who answer both questions correctly.

Statistical Analyses

This is a between-subjects experiment in which participants will be randomized to one of five arms. All participants will view two fast-food online ordering menus, presented in a random order, with menus for each study arm differing by the labeling scheme applied: (1) a QR code on all items (control); (2) a low climate impact label on low impact items; (3) a high climate impact label on high impact items; (4) traffic light labels (high-, medium-, and low-impact) on all items by level of climate impact; and (5) climate grade labels (A, B, C, D, F) on all items by level of climate impact.

To determine the effects of the four climate-impact labels compared to the control label on our continuous primary outcome of dietary quality, we will regress the outcome onto a categorical indicator representing experimental condition. If the primary outcome is not normally distributed, we will perform appropriate transformations and/or modeling approaches (e.g. gamma distribution and log-link) and use robust standard error estimates to preserve valid standard errors and p- values.

We will use a similar approach for secondary outcomes and regress each secondary outcome onto an indicator of experimental arm. For the binary secondary outcomes, we will use logistic regression and regress the outcome onto a categorical indicator representing experimental condition and will compare the odds of each binary outcome for each experimental condition compared to the control. We will use predicted margins to examine predicted mean values (for continuous outcomes) and predicted proportion of participants (for binary outcomes) of each outcome for each experimental condition. We will analyze results overall (summed across meal selections from both restaurants) and stratified by restaurant.

Statistically significant differences will be considered at $p<0.05$ with Bonferroni-Holm correction for multiple comparisons of the four label conditions compared to control label and within family of outcomes for secondary outcomes, and all tests will be two-tailed.

We will not exclude observations based on values of any of the outcome variables. The survey firm (NORC AmeriSpeak) will remove observations based on individuals who complete the survey in less than one-third of the median response time, individuals who skip or refuse >50% of eligible questions, and who 'straight-line' all question grids. For secondary outcomes we will exclude observations with missing data within family of outcomes.