Development of Online Store for Testing Regulatory Food and Nutrition Policies in Brazil

Hypotheses and Analytic Plan

NCT# NCT05789238

Version Date: 05/18/2023

Hypotheses

The overall purpose of this study is to assess the impact of different nutrient profiles for magnifying glass front-of-package labels on consumer choices in an online grocery store.

Primary Outcomes
The primary outcomes are the total sugar (g), saturated fat (g), and sodium (mg) of products purchases.

We hypothesize that:
1. Compared to the no-label control, the ANVISA magnifying glass and PAHO magnifying glass will lead to reduced total sugar of products purchased.
2. Compared to the no-label control, the ANVISA magnifying glass and PAHO magnifying glass will lead to reduced total saturated fat of products purchased.
3. Compared to the no-label control, the ANVISA magnifying glass and PAHO magnifying glass will lead to reduced total sodium of products purchased.

Secondary Outcomes
The secondary outcomes are the percentage of ultra-processed products purchased (as defined by the NOVA classification), perceived healthfulness of unhealthy products, perceived product appeal of unhealthy products, concern about health effects of unhealthy products, ability to identify high sugar/saturated fat/sodium content of unhealthy products (as per the PAHO nutrient profile model, which is stricter than ANVISA's), and noticing of the product label.

We hypothesize that, compared to the no-label control, the ANVISA magnifying glass condition and PAHO magnifying glass condition will lead to:
1. Lower percentage of ultra-processed products (as defined by the NOVA classification) purchased.
2. Lower perceived healthfulness of unhealthy products.
3. Lower perceived product appeal of unhealthy products.
4. Higher concern about health effects of unhealthy products.
5. Higher ability to identify high sugar/saturated fat/sodium content of unhealthy products.
6. Higher noticing of the front-of-package label.

STATISTICAL CONSIDERATION

Statistical Methods

The study is a randomized controlled trial. The analysis will rely on random assignment to identify the effect of each type of front-of-package label nutrient profile model on consumer choices.

We will conduct a soft launch with 100 participants to check randomization and data quality. If randomization is working properly and there are not apparent issues with the data collection, we will resume data collection until the final sample is achieved. The original participants from the soft launch will be included in the final dataset. If randomization is not working properly (as assessed using the tabs command in Stata) and/or there are apparent issues with the data collection (e.g., erroneous skip patterns), we will address all identified issues then resume data collection until the final sample is achieved.

We will descriptively report unadjusted values for primary and secondary outcomes.
We will use a two-sided critical alpha of 0.05 to conduct all statistical tests. Per CONSORT guidelines, we will not test for balance in covariates. Primary analyses will be intent-to-treat, including all eligible participants with non-missing data for the outcome being analyzed.

For all primary, secondary, and other outcomes, we will assess whether the outcomes vary by study arm. For the primary outcomes, we will use two-sample t-tests for means. If there are imbalances in potentially

confounding factors between arms, we will use linear regressions. For the secondary and other outcomes, we will use linear regression.

Sensitivity Analyses

We will describe dropout (defined as entering the store but not completing the purchasing task) by study arm. Dropout is defined as passing the screener survey but not completing the exit survey (i.e., dropping out during the shopping task or during the exit survey). If differential dropout is identified, we will consider sensitivity analyses to handle missing outcome data, such as inverse probability weighting. We will compare the main effects to the primary results if the number of participants excluded from a moderation analysis because of missing demographic data warrants it.

For all analyses, all pairwise comparisons of means will be examined for all outcomes. We will consider a result statistically significant at $p<0.05$.

In exploratory analyses, we will examine whether any of the collected participant characteristics (see list below) moderate the intervention effects on the primary outcomes:

a. Age
b. Gender
c. Region
d. Education
e. Race
f. Household monthly income
g. Household size
h. Parent/caregiver
i. Marital status
j. Experience with online shopping for clothes, electronics, furniture, or home appliances
k. Experience with online grocery shopping
l. Hunger levels at time of survey
m. BMI
n. Self-reported health status
o. Diabetes
p. Heart disease

To test whether these characteristics moderate the effect of each intervention on nutrient content, we will fit a series of regression models (one for each potential moderator), with trial arm, the moderator (specified as dummy variables), and their interaction as predictors. We will use a Wald chunk test to determine the joint significance of the interaction terms. We will quantitatively evaluate the presence of moderation by calculating the marginal effect of each intervention on the outcome at different levels of the moderating variable. If the pattern of main results is similar between intervention arms (magnifying glass label based on ANVISA nutrient profile model, magnifying glass label based on PAHO nutrient profile model), we will consider combining intervention arms for the moderation analysis.

Sample Size and Power

This is an exploratory study and the sample size was maximized at 3,000 participants. Data from a pilot study of 300 participants conducted in December 2022 and January 2023 showed that, accounting for a loss of 6% of participants after excluding the top 2% of values for total saturated fat, sodium and sugar, 3,000 participants would provide at least 90% power to 0.05-level two-sided t-tests for mean differences between ANVISA and the control in total saturated fat and sugar, between PAHO and the control in total saturated fat, and between each arm in the proportion of ultra-processed products checked out (equal and unequal variances).

Outliers and Exclusions

We will examine detailed summary statistics on each continuous outcome measure and exclude observations with unrealistically high values (e.g., in the top 2% of values). Separately, we will also conduct sensitivity analyses excluding individuals who are in the bottom 2 percentile of the expenditures, those who complete the study unusually quickly (e.g., based on the distribution of time to completion of ascertained during a soft launch of the study), and individuals who are non-compliant with the shopping list (<50% of products selected comply with the shopping list).