Investigating the effect of price discounts on sales of zero-calorie beverages

Introduction

Strong evidence has linked habitual intake of sugar-sweetened beverages (SSBs) with weight gain and a higher risk of type 2 diabetes mellitus, cardiovascular diseases, and even cancer [1]. Pathological mechanisms include additional liquid calories in the diet, an exacerbated increase in blood insulin levels, and possibly the activation of a reward system related to sugar addiction. Multiple policy interventions to inhibit consumption of SSBs have been proposed [2]. These include taxation of SSBs, limitation of access to SSBs in schools and healthcare facilities, and package labelling systems to provide the necessary information for consumers who favour a healthier diet. As new potential interventions arise, there is a growing need for strong evidence of their efficacy.

In this context, the present study aims to evaluate the efficacy of five interventions intended to shift consumer behaviour toward zero-calorie beverages. Three information-based interventions were implemented as posters, flyers, and signs providing information about caloric content in sugared beverages, the amount of time one would need to spend on a treadmill to burn such calories, or both. Two financial interventions were implemented as either a 10% discount on zero-calorie beverages or the same discount plus an informative message about the reason for the discount. Over 30 weeks, investigators alternated interventions in the cafeterias and convenience shops from three different sites (hospitals). Baseline and washout periods were implemented between intervention periods as well. Outcomes include daily sales of bottled sugared beverages and zero-calorie beverages.

Given the above study design, the main statistical question is whether the intervention groups are associated with differences in the average proportion of daily sales of zero-calorie beverages. Other statistical questions include whether there is a significant interaction between intervention and site effects, and whether groups with combined interventions had stronger effects as compared to isolated interventions (e.g., discount+messaging versus discount alone).

Data Description and Summaries

The data come from an interrupted quasi-experimental study on sales of sweetened and zero-calorie beverages which was conducted in three United States hospitals over a 30-week period.

Each observation represents the total daily sales of bottled beverages across multiple storefronts in a single site. Sales were recorded automatically by point of sale terminals in Site A, and manually by cashiers at Sites B and C.

The data includes the number of days since the start of the experiment, the day of the week, the site, and the intervention. The number of days is a non-negative integer, and the other three variables are categorical.

The daily records include the number of bottled zero-calorie and sweetened beverages sold per site. These are each non-negative integer-valued variables. The data also includes the total number of beverages sold per site, which is another non-negative integer value.

Some observations additionally report the daily sales of 100% juice, orange juice, and sports drinks as integer values. However, these counts are only consistently recorded at Site A, and they are hardly recorded for Sites B and C.

There are nine days which are missing counts for the bottled zero-calorie and sugared beverages. Each of these nine observations contains missing data for every sales-related column. Of these nine observations, seven occur in the final week of the study at Site B, and two occur mid-study at Site C.

Exploratory Analysis

Exploratory data analysis can visualize relationships among variables, and informally give some insight on what results can be expected. Based on the data and statistical questions, we recommend the following figures and tables:

1. Side-by-side boxplots of [**ZERO SUGAR RESPONSE**] by site can help identify if there are differences between sites
2. Plots of [**ZERO SUGAR RESPONSE**] over time, stratified by other variables, such as intervention can identify and trends or patterns or differences between interventions.
3. A scatterplot of zero-sugar vs sugary drink sales be used to investigate what kind of relationship they have.

Additionally, summary statistics such as means can be calculated for different variables, such as days of the week to give a numeric measure that can be compared.

Formal Analysis

In this study, an effective treatment will increase the proportion of daily sales that come from zero-calorie beverages among sales that are either zero-calorie beverages or sugary beverages. This proportion increases when an intervention leads to relatively more sales of zero-calorie beverages than of sugary beverages. This avoids the need to perform two separate analyses (one for zero-calorie and another one for sugary beverages) and avoids variation due to sales of other unrelated products.

In light of the above reasoning, the recommended primary analysis is based on a generalized linear mixed effects model (GLMM) with a Poisson distribution. This type of model handles count data and allows estimating the heterogeneity of effects across different sites. The outcome is defined as the counts of zero-calorie beverage sales. Using the (log-transformed) total sales of zero-calorie or sugary beverages as an offset variable, this formulation models the proportion of zero-calorie beverage sales as defined above. The main independent variables of interest are intervention indicators, treating the baseline periods as the reference. Covariates in the model include the site, the days of the week, and the study day. Finally, to assess whether intervention effects differ by site, we recommend testing interaction effects between sites and intervention groups as random slopes, which justify the mixed effects formulation. This assessment can be implemented as a Likelihood Ratio Test.

A major assumption of the above model is the correct specification of the Poisson distribution. To avoid this assumption, a robust version of the above model is recommended. In particular, the same regression specification (offset, main independent variables, and covariates) can be used while employing Poisson estimating equations with a so-called sandwich covariance estimator. This model is known as robust Poisson as it remains valid even if the data do not follow a Poisson distribution [3]. Once exponentiated, coefficients from the GLMM and the robust Poisson model have the same relative change interpretation. For instance, if the exponentiated coefficient for a given intervention is equal to 1.5, then the proportion of zero-calorie beverage sales increases by 50% under the intervention, as compared to the baseline.

Overall, the recommended analysis plan includes a GLMM with a Poisson distribution and a robust Poisson model. While agreement of results between the two analyses is reassuring, the robust Poisson model is preferred in case of qualitative disagreement in model results. Both analyses employ the zero-calorie beverage sales as the primary outcome with the log of the sum of zero-calorie beverage sales and sugary beverage sales as an offset. This specific offset formulation allows interpretable regression coefficients, which are large when interventions lead to more zero-calorie beverage sales relative to sugary beverage sales.

Conclusions

References

Statistical Appendix

[1] Malik VS, Hu FB. The role of sugar-sweetened beverages in the global epidemics of obesity and chronic diseases. Nature Reviews Endocrinology 2022;18. https://doi.org/10.1038/s41574-021-00627-6.

[2] Krieger J, Bleich SN, Scarmo S, Ng SW. Sugar-sweetened beverage reduction policies: Progress and promise. Annual Review of Public Health 2021;42. https://doi.org/10.1146/annurev-publhealth-090419-103005.

[3] Talbot D, Mésidor M, Chiu Y, Simard M, Sirois C. An alternative perspective on the robust poisson method for estimating risk or prevalence ratios. Epidemiology 2022;34:1–7. https://doi.org/10.1097/ede.0000000000001544.