UNDER EMBARGO UNTIL MARCH 19, 2013, 12:01 AM ET

# A Cash-Back Rebate Program for Healthy Food Purchases in South Africa
## Results from Scanner Data

Roland Sturm, PhD, Ruopeng An, MPP, MPhil, Darren Segal, BSc, Deepak Patel, MD, MPhil

**Background:** Improving diet quality is a key health promotion strategy. There is much interest in the role of prices and financial incentives to encourage healthy diet, but no data from large population interventions.

**Purpose:** This study examines the effect of a price reduction for healthy food items on household grocery shopping behavior among members of South Africa's largest health plan.

**Methods:** The HealthyFood program provides a cash-back rebate of up to 25% for healthy food purchases in over 400 designated supermarkets across all provinces in South Africa. Monthly household supermarket food purchase scanner data between 2009 and 2012 are linked to 170,000 households (60% eligible for the rebate) with Visa credit cards. Two approaches were used to control for selective participation using these panel data: a household fixed-effect model and a case– control differences-in-differences model.

**Results:** Rebates of 10% and 25% for healthy foods are associated with an increase in the ratio of healthy to total food expenditure by 6.0% (95% CI=5.3, 6.8) and 9.3% (95% CI=8.5, 10.0); an increase in the ratio of fruit and vegetables to total food expenditure by 5.7% (95% CI=4.5, 6.9) and 8.5% (95% CI=7.3, 9.7); and a decrease in the ratio of less desirable to total food expenditure by 5.6% (95% CI=4.7, 6.5) and 7.2% (95% CI=6.3, 8.1).

**Conclusions:** Participation in a rebate program for healthy foods led to increases in purchases of healthy foods and to decreases in purchases of less-desirable foods, with magnitudes similar to estimates from U.S. time-series data.
(Am J Prev Med 2013;xx(x):xxx) © 2013 American Journal of Preventive Medicine

## Introduction

Improving diet quality is a key health promotion strategy worldwide, and the WHO has advocated for price policies to improve diets.[1] Although there is much interest in the role of prices and financial incentives to encourage healthy diet, empirical results from large population interventions are limited. Results come either from historic or regional price and consumption patterns or from small localized trials.[2,3]

In South Africa, a large subsidy program that operates nationwide was started in 2009 by South Africa's largest health insurer. The HealthyFood program provides a rebate of up to 25% on healthy food purchases in more than 400 designated supermarkets across South Africa. The analysis links scanner data from all participating supermarkets to 170,000 households with a credit card. Roughly 60% of those households are eligible for rebates on healthy foods.

## Methods

### HealthyFood Program

The HealthyFood program was launched in 2009 by the South African health insurer Discovery, as part of its health promotion program, Vitality. All Vitality members are eligible for the healthy food benefit at no additional cost to them, but they have to activate the benefit online or by phone. Following activation, members immediately receive a 10% rebate for healthy foods, which increases to a 25% rebate on completion of an online health risk assessment questionnaire. The rebate is capped at a maximum monthly purchase amount of 4000 rands (approximately $480) per household and a limit related to participation in health promotion activities. In 2012, 330,000 individuals were enrolled in the HealthyFood program and could claim rebates at 432 Pick n Pay supermarkets. Pick n Pay is one of the

From the RAND Corporation (Sturm, An), Santa Monica, California; and Discovery Vitality (Segal, Patel), Johannesburg, South Africa

Address correspondence to: Roland Sturm, PhD, RAND Corporation, 1776 Main Street, P.O. BOX 2138, Santa Monica, CA 90407-2138. E-mail: sturm@rand.org.

0749-3797/$36.00
http://dx.doi.org/10.1016/j.amepre.2013.02.011

three largest supermarket chains in South Africa, with a combined 78% market share.[4]

Food items eligible for the rebate program were selected by a panel consisting of nutritionists, physicians, and behavioral scientists based on international guidelines[5,6] on healthy nutrition, including those from South Africa and the U.S.[5,6] A complete list of eligible items (more than 6000) can be found on Discovery's website (www.discovery.co.za) and is also distributed as brochures to program participants. In some instances, items are eligible only in certain forms. For example, raw or minimally processed fruits and vegetables (e.g., canned or frozen) are eligible, but not those prepared with added sugars or salt; only nonfat dairy products are eligible. Participating supermarkets have in-store signs identifying eligible foods; they are also marked on the store receipt. The labeling was implemented prior to the study period and was not changed during the study period.

The panel identified less-desirable foods and beverages as those that are high in saturated fats, trans-fatty acids, added sugar, salt, or refined starch. The less-desirable food group includes sweets, chocolates, ice cream, sugary foods, chips, sugar-sweetened beverages, and fried items. All foods not specifically classified as either healthy or less desirable were considered neutral and were neither encouraged nor discouraged.

## Analytic Sample

Purchases before and after rebate eligibility were linked to households to assess the effect of a 10% or 25% price change (Figure 1 shows the sample selection). Only purchases made with a Visa credit card issued by Discovery were analyzed, as this is the only identifying information for purchases for which there was no rebate. Approximately one third of all households that are enrolled in Discovery Vitality use this type of credit card.

Scanner data from Pick n Pay for credit card purchases from November 2009 to March 2012 are linked to 169,485 households. About 60% of the Vitality households activated the HealthyFood benefit and initially received a 10% rebate on healthy food purchases; 67% of the households also completed the health risk assessment and became eligible for the 25% rebate. The remaining 40% did not receive a rebate on healthy food purchases. Purchases were collapsed into monthly observations, resulting in a total of 1,909,740 observations (household months). Household food purchases were separated into healthy foods (approximately 20% of total food spending); fruits and vegetables (a subcategory of healthy foods, 10% of total food spending); less-desirable foods (20% of spending); and neutral foods (60%).

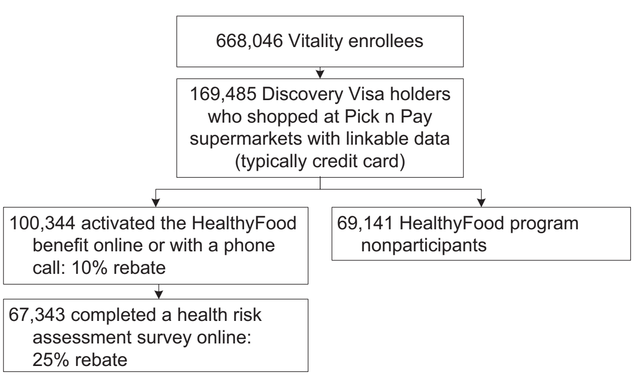


**Figure 1.** Sample selection

## Variable Construction

The three dependent variables are the ratios of healthy food, fruit and vegetables, and less-desirable food to total food expenditure in a household. The dependent variables are ratios, because absolute purchase amounts at Pick n Pay were expected to increase (supermarkets participate to have a competitive advantage). By definition, there are no data for months where individuals had no (linkable) food purchases at Pick n Pay (the ratio is undefined).

The main explanatory variables are two dichotomous variables for receiving a 10% or 25% rebate (mutually exclusive) on healthy food purchases in a month. The reference group is households that were not eligible for any rebate. To analyze whether duration in the program is associated with changes in purchasing patterns (e.g., positive habit formation or a loss of novelty), two interactions between the number of months in a specific rebate policy (either 10% or 25%) and the corresponding dichotomous variable for receiving that rebate rate (either 10% or 25%) were added as regressors.

## Statistical Models

Participation in the program was self-selected, and households are likely to differ in characteristics other than rebate level. Two methods were used to address this problem: household fixed-effects models (the main model) and a case–control difference-in-differences method. The household fixed-effects method uses within-household variations in program enrollment status to identify the rebate effects. This approach deals with any selection biases due to a household-specific component (observable or unobservable) that is constant over time. The approach cannot address differences between eligible and non-eligible households that vary over time.

Time trends and seasonality in grocery shopping pattern were controlled with a set of dichotomous variables for each specific month in a year. For each regression, one main-effect model and one interaction model were estimated. The Eicker-Huber-White sandwich estimator was used to calculate SEs clustered at the household level. All statistical analyses were conducted in STATA 12.

The case–control difference-in-differences method was used to calculate the effect of the rebate effect by subtracting the (change in ratios [of healthy food, fruit/vegetable, or less-desirable food to total food expenditure in a household] among people before and after they became eligible for the rebate) from the (change in ratios over the same time period among nonparticipants). For participants, the before/after period is demarcated by the date they became eligible for the rebate. Each of those households is matched to a household that enrolled in Discovery Vitality on the same month but never became eligible for the healthy foods rebate. This analysis may appear to be similar to a traditional difference-in-differences analysis, but it is a weaker design because the timing of the intervention (eligibility for the rebate) is not exogenously fixed.

## Other Limitations and Sensitivity Analyses

The data exclude purchases from competing supermarkets or other grocery stores and unlinkable cash purchases at Pick n Pay. An implied assumption of the analysis is that the linkable shopping carts at Pick n Pay are representative of total purchases. If shoppers switched from competing stores only for foods that receive a rebate, and not other foods, this would result in an overestimate of the effects of program participation on grocery shopping.

The selective purchasing bias should be greater for individuals who live near alternative supermarkets (where additional time and travel costs for a second shopping trip are small); the bias should be smaller for individuals who live close to a Pick n Pay but far away from a competing store (where additional time and travel costs are large). As a sensitivity analysis, models are re-estimated for people who lived at least 1 kilometer closer to the nearest Pick n Pay than to the nearest competing supermarket. Shrinking coefficients on rebates is an indicator of selective purchasing and that estimates are biased upward (even for the smaller sample of nearby shoppers).

Households with very few linkable data may simply not shop at Pick n Pay very often or may rely more on cash transactions. If linkable data are not representative of purchases and payment mechanisms were related to the type of foods bought (e.g., using a credit card to buy eligible food items and cash to buy soft drinks), results would be biased. The bias would be small if few purchases are unlinkable, even with a strong relationship between transaction type and ratio of healthy foods, but would increase with the share of unlinkable purchases. Another sensitivity analysis re-estimates all models using only loyal Pick n Pay shoppers, defined as having at least 15 months of regular purchases.

## Results

Table 1 provides baseline descriptive statistics on grocery shopping patterns at Pick n Pay supermarkets by (eventual) rebate status, demonstrating the importance of addressing selection biases: Households who became eligible for a rebate during the study period already had a larger proportion of overall food expenditure going toward healthy foods and a smaller proportion toward less-desirable foods at baseline (when nobody was eligible for the rebate) than households that never participated in the program. Participants also lived closer to a Pick n Pay supermarket than to a competing supermarket; the opposite was true for nonparticipants. The difference between nonparticipants and participants is significant at $p<0.001$ for all variables. No meaningful differences were found between the 10%-rebate group and the 25%-rebate group.

Figure 2 displays the case–control difference-in-differences analyses graphically and shows how outcomes change over time from baseline values (Table 1). Among program participants, the ratio of fruit/vegetables (Figure 2, top panel) and healthy foods (Figure 2, middle panel) to total food expenditure increased and the ratio of less-desirable food to total food expenditure (Figure 2, bottom panel) decreased substantially after participants became eligible for a rebate. No similar increase/decrease was found among the matched ineligible households. The change appeared immediately after participants became eligible, and there was neither a convergence back to baseline levels nor an increasing effect over time. For healthy food purchases, there was an increase a few months before the actual rebate eligibility, suggesting some type of anticipation effect.

Table 2 reports the results from household fixed-effect regression models. The ratios of healthy food purchases and of fruit/vegetables to total food purchases increases, whereas the ratio of less-desirable food purchases declines. All estimated coefficients are significant at $p<0.001$. The increase in healthy food purchases is larger than the decline in less-desirable food purchases, which means that some of that shift is due to a decline in the ratio of neutral foods (e.g., a shift from 2% to nonfat milk rather than from ice cream to nonfat milk). That decline is also significant at $p<0.001$.

Demand analyses commonly use relative changes or elasticities (percentage change in demand associated with a 1% change in price). To calculate the relative change, the coefficients in Table 2 were divided by the baseline values in Table 1 (the latter were taken as fixed constants for calculating CIs). Rebates of 10% and 25% for healthy foods were associated with an increase in the ratio of expenditure on healthy foods to total food expenditure by 6.0% (95% CI=5.3, 6.8) and 9.3% (95% CI=8.5, 10.0); an increase in the ratio of expenditure on fruit and vegetables to

**Table 1.** Baseline household food purchases at designated supermarkets (prior to eligibility for rebate), by eventual rebate status, M (SD)

| Variable | Attribute | No rebate | 10% rebate | 25% rebate |
|---|---|---|---|---|
| Total food expenditure (South African rand) | Continuous | 558 (744) | 1120 (1196) | 1073 (1117) |
| Ratio of healthy to total food expenditure | Continuous | 0.17 (0.13) | 0.21 (0.11) | 0.21 (0.12) |
| Ratio of fruit and vegetable to total food expenditure | Continuous | 0.09 (0.09) | 0.10 (0.08) | 0.10 (0.08) |
| Ratio of less-desirable to total food expenditure | Continuous | 0.22 (0.17) | 0.19 (0.11) | 0.19 (0.12) |
| Distance (km) from home to nearest Shoprite or Woolworth's store | Continuous | 1.72 (2.10) | 2.14 (2.40) | 2.11 (2.39) |
| Distance (km) from home to nearest Pick n Pay store | Continuous | 2.11 (1.99) | 1.96 (1.88) | 1.96 (1.91) |
| Differential distance (km) | Continuous | −0.38 (1.18) | 0.18 (1.77) | 0.15 (1.73) |

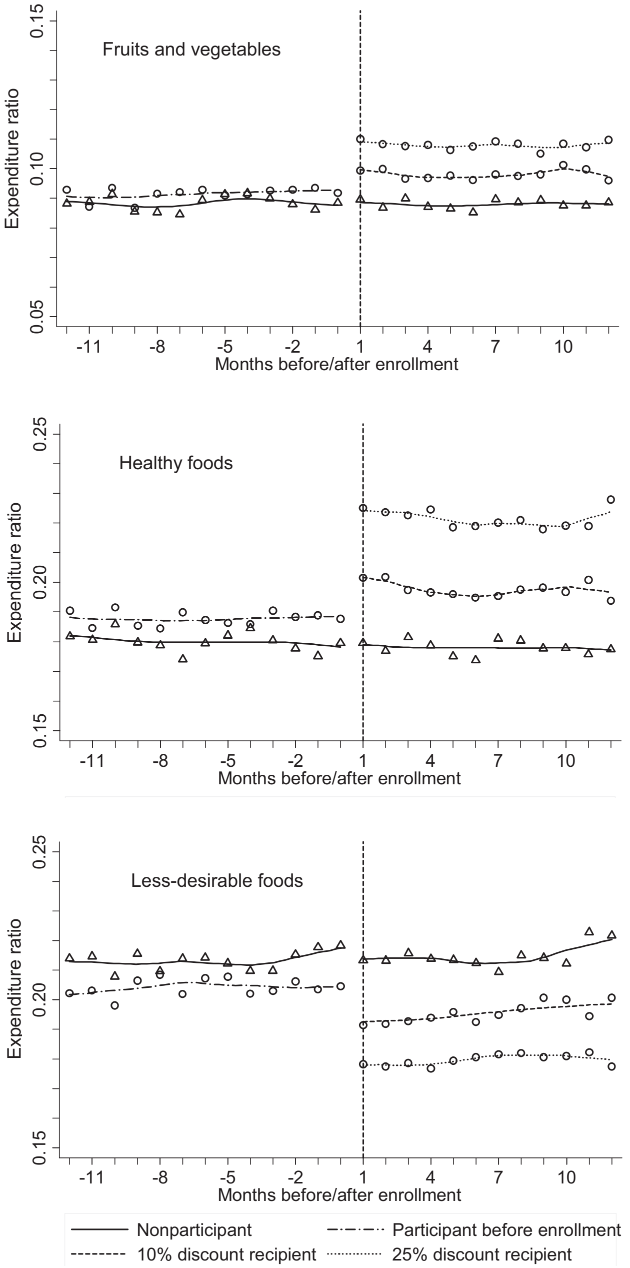


**Figure 2.** Household food purchases among HealthyFood participants and nonparticipants before/after rebate reception

*Note:* Top panel shows ratio of expenditure on fruits and vegetables to total food expenditure; middle panel shows ratio of expenditure on healthy foods to total food expenditure; and bottom panel shows ratio of expenditure on less-desirable foods to total food expenditure. Trends are estimated using nonparametric kernel-weighted local polynomial regressions. Each HealthyFood program participant is matched to a randomly selected nonparticipant who enrolled in the Discovery Vitality program (Johannesburg, South Africa) in the same month as the participant.

total food expenditure by 5.7% (95% CI=4.5, 6.9) and 8.5% (95% CI=7.3, 9.7); and a decrease in the ratio of expenditure on less-desirable to total food expenditure by 5.6% (95% CI=4.7, 6.5) and 7.2% (95% CI=6.3, 8.1).

The third column in Table 2 shows the sensitivity analysis that restricts the sample to people who lived at least 1 kilometer closer to the nearest Pick n Pay supermarket than to the nearest competing store. It is a small group, but should be the group least likely to shop for discounted healthy foods only at Pick n Pay. If such a bias were to exist among shoppers generally, this sensitivity analysis should show smaller effect sizes, but point estimates are larger than for the full sample (although not significantly larger for fruit/vegetables or less-desirable foods). Another sensitivity analysis repeats the analysis for "loyal" shoppers only, but without substantive changes.

An alternative specification of the fixed-effects models allows for an interaction between duration in the program and rebate. Time in the program does not appear to play a major role, and the estimated coefficients are small (not shown). This finding suggests that the price effects on food purchase remain stable over time, that is, people respond immediately and permanently to the price effect.

## Discussion

This study examines the relationship between a price reduction for healthy food items and supermarket shopping in a nationwide program. Rebates predict a higher ratio of expenditure on healthy to total food expenditure and a lower ratio of expenditure on less-desirable food to total food expenditure. A 10% rebate predicts a 6.0% increase in the ratio of expenditure on healthy foods to total food expenditure, a 5.7% increase in the ratio of expenditure on fruits and vegetables (a subcategory of healthy foods) to total food expenditure, and a 5.6% decrease in the ratio of expenditure on less-desirable foods to total food expenditure. A 25% rebate predicts a 9.3% increase in expenditure on healthy foods; an 8.5% increase in expenditure on fruits/vegetables, and a 7.2% decrease in expenditure on less-desirable foods (all increases are as a proportion of baseline ratios). The price effects remain stable over time.

Sensitivity analyses to identify biases that could be a consequence of data collection (only purchases paid by credit card) or strategic consumer behavior (more shopping at Pick n Pay, but only for foods eligible for rebates) suggest that biases due to payment type or strategic shopping are negligible. Rebate eligibility is not randomly assigned, but households have to opt into the program even though the program is free to them. Program participation results from the interaction between program rules, geographic location, and taste and lifestyle preferences.

A direct comparison of purchasing patterns between households eligible for the rebate and those not eligible

**Table 2.** Estimated rebate effects on household food purchases at designated supermarkets

| | All shoppers | | Nearby shoppers | |
|---|---|---|---|---|
| **Dependent variable** | **10% rebate** | **25% rebate** | **10% rebate** | **25% rebate** |
| Ratio of healthy to total food expenditure | **0.0127**** (0.0008) | **0.0195**** (0.0008) | **0.0182**** (0.0032) | **0.0254**** (0.0032) |
| Ratio of fruit/vegetable to total food expenditure | **0.0057**** (0.0006) | **0.0085**** (0.0006) | **0.0063*** (0.0022) | **0.0091**** (0.0022) |
| Ratio of less-desirable to total food expenditure | −**0.0106**** (0.0009) | −**0.0137**** (0.0009) | −**0.0145**** (0.0032) | −**0.0165**** (0.0032) |

*Note:* Boldface indicates significance. Reported parameters (SE in parentheses) are estimated using linear household fixed-effect models, controlling for month/year fixed effects. Nearby shoppers ($n$=56,908 households) are defined as those who lived at least 1 kilometer closer to the nearest Pick n Pay supermarket relative to the nearest Shoprite or Woolworth's supermarket. The Eicker-Huber-White sandwich estimator is used to calculate SEs clustered at the household level.
*$p$<0 .01, **$p$<0.001

would confound price effects with determinants of participation selection. Such confounding cannot be satisfactorily addressed in cross-sectional data where adjustment is based on observed variables. With panel data, however, confounding factors (even if unobservable) can be controlled as long as they are constant (e.g., if a participating household always bought x% more healthy foods than a nonparticipating household at the same set of prices during the study period). Controlling for factors in this way is not a complete solution for any possible selection effects as the model cannot control for unobserved time-varying differences between participating and nonparticipating households.

Studying a large real-life program entails compromises resulting from the day-to-day operations of supermarkets and health plans. At times, it allows unexpected ways to strengthen an analysis (in this case, creating panel data of credit card holders). At other times, even seemingly simple analyses become infeasible. For example, software transitions at the supermarket made it impossible to obtain data prior to 2009 (i.e., before and after stores added signs for healthier food choices). There were no changes in signage during the study period.

Economic studies typically express price responses in terms of elasticities (i.e., the percentage change in sales when prices change by 1%). A recent systematic review compiled findings from 160 studies, primarily time-series analyses of aggregated data (99 studies) or surveys (34 studies) in the U.S.[7] Scanner data have not yet been used much, as this is a relatively recent technology and no research has distinguished healthier versus less healthy food purchase patterns beyond a single food group (e.g., substitution of whole milk with low-fat milk).

The most studied food groups were meats and milk, although there were 20 U.S.-based studies for fruits and vegetables.[7] The average price elasticity for fruit was 0.70 and for vegetables 0.58. Assuming that total purchases remain constant and only the composition changes, the 5.7% increase in fruit/vegetable consumption associated with the 10% rebate implies an elasticity of 0.57. Another systematic review, not based on interventions, suggested that a 20% price discount could increase fruit and vegetable consumption by 10%.[3] The current analysis found that a 25% rebate is associated with an 8.5% increase in fruit and vegetable purchases.

The comparability is limited in that this study looks at ratios, not total amounts. Rebates and price subsidies for healthy food can also have an income effect. Even if the share of expenditure on healthy foods increases, total purchases or consumption, including that of less-desirable items, may also increase, although there seems to be no evidence for such an effect.[3,8] Subsidizing healthier foods may improve diet quality, but that does not necessarily indicate a change in total calorie consumption (or obesity rates).

Although the HealthyFood program is timely in addressing a topical worldwide policy question, its generalizability to other populations remains uncertain. The program is unique due to its size and geographic scope and because it is the only large price intervention led by the private sector on an ongoing basis. That factor makes the current study novel and interesting but also contributes to its limitations by weakening the study design. Nevertheless, this study serves as an important data point in the ongoing effort to quantify the influence of prices on dietary behaviors.

The results from this rebate program suggest that reducing the costs of healthy food purchases is likely to change purchasing patterns in a meaningful way. However, it is not a cheap way to achieve major changes in population diets. Changes in purchases are commensurate with price changes, but even a large price change for healthy foods (e.g., 25%) can at best address a small part of the discrepancy between population dietary patterns and dietary guidelines.

This study is funded by the National Cancer Institute (grant R21CA161287); the National Institute of Child Health & Human Development (grant R21HD071568); and the Anne and James Rothenberg Dissertation Award, 2011–2012.

No financial disclosures were reported by the authors of this paper.

## References

1. WHO. Global strategy on diet, physical activity and health. 2004. apps.who.int/gb/ebwha/pdf_files/WHA57/A57_R17-en.pdf.
2. An R. Effectiveness of subsidies in promoting healthy food purchases and consumption: a review of field experiments. Public Health Nutr 2012;5:1–14. dx.doi.org/10.1017/S1368980012004715.
3. Powell LM, Chriqui JF, Khan T, Wada R, Chaloupka FJ. Assessing the potential effectiveness of food and beverage taxes and subsidies for improving public health: a systematic review of prices, demand and body weight outcomes. Obes Rev 2013;14(2):110–28. 10.1111/obr.12002.
4. U.S. Department of Agriculture Foreign Agricultural Service. 2010 annual retail food sector report of South Africa. 2011. gain.fas.usda.gov/Recent%20GAIN%20Publications/Retail%20Foods_Pretoria_South%20Africa%20-%20Republic%20of_1-5-2011.pdf.
5. DHHS and U.S. Department of Agriculture. Dietary guidelines for Americans, 2005, 6th ed. Washington DC: U.S. Government Printing Office, 2005.
6. Vorster HH, Love P, Browne C. Development of food-based dietary guidelines for South Africa—the process. South African J Clin Nutr 2001;14(3S):S3–S6.
7. Andreyeva T, Long MW, Brownell KD. The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. Am J Public Health 2010;100(2):216–22.
8. Sturm R, Datar A. Regional price differences and food consumption frequency among elementary school children. Public Health 2011;125(3):136–41.