Public Health Briefs

Evaluating Community-based Nutrition Programs: Assessing the Reliability of a Survey of Grocery Store Product Displays

ALLEN CHEADLE, PHD, BRUCE PSATY, MD, PHD, EDWARD WAGNER, MD, MPH, PAULA DIEHR, PHD, THOMAS KOEPSELL, MD, MPH, SUSAN CURRY, PHD, AND MICHAEL VON KORFF, PHD

Abstract: A pilot test of a survey of grocery store product displays was conducted to measure the amount of health-education information provided and the proportion of the display devoted to "healthier" products. Inter-rater reliability ranged between 0.73 and 0.78 for the healthiness indices and between 0.30 and 0.67 for the health education measures. Test-retest reliability ranged from 0.44 to 1.0. (*Am J Public Health* 1990; 80:709–711.)

Introduction

There have been a number of recent efforts directed at entire communities to modify dietary behavior in order to reduce the risk of cardiovascular disease and cancer.[1–4] Evaluation of these interventions requires "community-level indicators" of program effectiveness—measures of the community environment that are not merely aggregates of individual behavior.[5] We measured the reproducibility of a survey of grocery store product displays to be used as part of the evaluation of community-based nutrition programs. Such surveys are particularly useful when grocery stores and supermarkets play a role in the intervention, as in the Giant Foods project[1] and in the Henry J. Kaiser Family Foundation Community Health Promotion Grants Program (CHPGP).*

Methods

We surveyed four general product areas within each grocery store: fresh produce, meat, milk, and bread. For each major product area, two types of information were recorded: presence or absence of health-promotion items, and the physical dimensions of product displays. The protocol took 30–45 minutes for an experienced rater to complete.

- *Health-promotion items* were defined as health education activities (usually printed material) on or near the displays, which provided nutritional information or increased awareness of healthy food choices, and which were independent of product packaging. Examples of health-promotion items included: shelf labeling, posters, recipes, and store health-information centers.
- *The measurements of the displays* were designed to create an index of the relative amount of space occupied by "healthy" products in each product area, calculated as the proportion of shelf space devoted to such products (e.g., those low in fat) relative to the overall size of the display of similar foods.

The sampling frame consisted of full-service supermarkets, defined as grocery stores with two or more check-out stands, carrying fresh produce and fresh meat, that were listed in the Yellow Pages telephone directory. The sample for the Seattle pilot test of inter-rater reliability included 13 independent stores and 24 chain stores. For this part of the field test, each store was surveyed by two raters who visited the four product areas in a different sequence. Test-retest reliability was evaluated on an additional random sample of 61 stores in five communities that are part of the overall Kaiser Foundation evaluation. These test-retests were carried out by the same rater on some occasions and by different raters on others, with a variety of time intervals separating the visits.

Three indices of display "healthfulness" are presented here: the proportion of poultry and fish in the meat display, the proportion of reduced-fat milk in the milk display, and the proportion of 100% whole wheat bread in the bread display. Health promotion items were coded simply as present or absent for each of the major product areas. To assess reliability, we used the intra-class correlation coefficient[6] for the continuous variables and weighted kappas[7] for the discrete variables.

Results

As Table 1 shows, chain stores (particularly national or regional chains) more often offered health promotion items than did independent stores. By contrast, there was little apparent difference in the three display variables by store type. Table 2, column (1) gives inter-rater reliability coefficients computed for the first visits to the 37 stores in the pilot sample. Reliability for the display variables was fairly high, in the range of .7 to .8, but was lower for the health promotion items. The low figures for the milk and meat health promotion items were partly due to an open-ended format for recording these items in the pilot survey; a closed format used in the later surveys produced better results. The figures in column

From the University of Washington, Seattle, Departments of Health Services (Cheadle, Wagner, Diehr, Koepsell, Curry, Von Korff); Biostatistics (Diehr); Medicine (Psaty); Epidemiology (Psaty, Koepsell); and the Center for Health Studies, Group Health Cooperative of Puget Sound (Wagner, Curry, Von Korff). Address reprint requests to Allen D. Cheadle, PhD, Department of Health Services, JD-43, School of Public Health and Community Medicine, University of Washington, Seattle, WA 98195. This paper, submitted to the Journal March 2, 1989, was revised and accepted for publication November 27, 1989.

*Wagner EH, Kehrer B: The Henry J. Kaiser Family Foundation Community Health Promotion Program: Evaluation design considerations. Presented at the annual meeting of the International Epidemiological Association, Helsinki, Finland, August 6–13, 1987.

TABLE 1—Means of Grocery Store Survey Variables, by Store Ownership Category, Seattle Pilot Test Stores

Variables	Independent (n = 13)	Chains		
		Local (n = 13)	National/Regional (n = 12)	All (n = 37)
Health promotion items[a]				
Produce display	46.1	50.0	100.0	64.9
Meat display	69.2	100.0	100.0	89.2
Milk display	30.8	41.7	50.0	40.5
Display variables[b]				
% Poultry + fish	44.4	50.6	43.8	46.2
	(4.9)	(6.9)	(5.8)	(6.5)
% Reduced-fat milk	67.9	66.6	62.5	65.7
	(8.9)	(6.4)	(9.8)	(8.6)
% Whole-wheat bread	5.9	5.8	4.4	5.4
	(2.7)	(2.1)	(1.2)	(2.1)

NOTES: Standard deviation in parentheses.
[a]Percent of stores with health education materials in the indicated section. Store was counted as having health promotion items if either of the two raters recorded them.
[b]Percent of the display with indicated products. Measurements for each store are the average of the two raters. Definitions for numerator and denominator were as follows:
Meat: Shelf space occupied by poultry and fish/Total
Total = Fresh meat only; excludes frozen, canned, smoked or otherwise processed meat.
Milk: Shelf space occupied by 2% fat and skim milk/Total
Total = 2% fat, skim and whole milk only (half-gallon cartons)
Bread: Shelf space occupied by 100% whole wheat bread/Total
Total = Sliced, packaged loaves; excludes specialty bakery section, desert pastry, french bread, baguettes.

TABLE 2—Reliability Coefficients for Display and Health Promotion Item Measurements*

Variables	Different raters, same occasion[a]	Different raters, different occasions[b]	Same rater, different occasions[c]	Same rater, same occasion[d]
	(n = 37)	(n = 15)	(n = 27)	(n = 19)
Health promotion Items				
Produce display	.67	1.00	.92	—
Meat display	.36	.63	.91	—
Milk display	.30	.46	1.00	—
Display Variables				
% Poultry + fish	.78	.86	.44	.98
% Reduced-fat milk	.75	.68	.58	.91
% Whole-wheat bread	.73	.47	.65	.93

*Statistical methods:
Display Variables: Intra-class correlation coefficients
Health Promotion Items: Weighted kappas
[a]Site: Seattle, WA (n = 37 stores)
[b]Site: Solano County, CA (n = 15)
[c]Sites: San Francisco, CA (n = 15), Orange County, CA (n = 12)
[d]Sites; Santa Cruz, CA (n = 15), Fresno, CA (n = 4)

(2), Table 2, show the correlations between two raters over a one-week time interval. This should approximate contemporaneous inter-rater reliability measures since few changes in health education materials were likely to have been made in this brief period. The revised format raised inter-rater reliability for the meat and milk displays—to 0.63 to 0.46, respectively.

The intra-class correlation coefficients presented in column (4), Table 2, all above 0.9, indicate high reliability in the same rater measuring the same display twice. Thus, most of the variation between occasions by the same rater was due to changes made by stores. The relatively high coefficients for health promotion items in column (3) suggest that these changed infrequently for the time intervals represented here. Display items showed lower intra-class correlation coefficients over time, particularly the meat display.

Discussion

Although sales data would reflect consumption most accurately, these are often unavailable to researchers. In the alternative measures which we used, measurements were made unobtrusively, and all of the stores approached for the pilot test agreed to participate; the response rate for the main study was 88 percent. Inter-rater reliability in the pilot was relatively high, particularly for the display variables. The lower inter-rater reliability figures for the health promotion items may have been due to the one-week interval between rater visits although, as noted earlier, significant changes were unlikely to have been made in that brief period. Test-retest reliability results indicated that both health promotion and display measures were relatively stable over periods ranging from one week to two months. Validity of the approach is currently being assessed, first by comparing the grocery store measures to the socioeconomic characteristics of people in the surrounding area, and second by correlating the survey measures with responses to dietary surveys of randomly sampled individuals living near the stores.

ACKNOWLEDGMENTS

This work was carried out under a grant from the Henry J. Kaiser Family Foundation, Menlo Park, CA (87-4361). Additional support was provided by the National Cancer Institute (90-2118-06). We wish to thank Dennis DuVall, Ginny Scobba, and Jane Steetle for their careful field work.

REFERENCES

1. Ernst ND, Wu M, Frommer P, Katz E, Matthews O, Moskowitz J, Pinsky JL, Pohl S, Schreiber GB, Sondik E, *et al*: Nutrition education at the point of purchase: the foods for health project evaluated. Prev Med 1986; 15(1):60–73.
2. Jacobs DR Jr, Luepker RV, Mittelmark MB, Folson AR, Pirie PL, Mascioli SR, Hannan PJ, Pechacek TF, Bracht NF, Carlaw RW, *et al*: Community-wide prevention strategies: evaluation design of the Minnesota Heart Health Program. J Chronic Dis 1986; 39(10):775–788.
3. Lasater TM, Lefebvre RC, Carleton RA: The Pawtucket Heart Health Program: Community level programming for heart health. RI Med J 1988; 71(1):31–34.
4. Kottke TE, Puska P, Salonen JT, Tuomilehto J, Nissinen A: Changes in perceived heart disease risk and health during a community-based heart disease prevention program: the North Karelia project. Am J Public Health 1984; 74:1404–1405.
5. Wallack LM: Mass media campaigns: the odds against finding behavior change. Health Educ Q 1981; 8:209–260.
6. Bartko JJ: The intraclass correlation coefficient as a measure of reliability. Psychol Rep 1966; 19:3–11.
7. Fleiss JL: Statistical Methods for Rates and Proportions, 2nd Ed. New York: John Wiley, 1981.

Poison Control Center Follow-up of Occupational Disease

EDDY A. BRESNITZ, MD, MS

Abstract: We followed up 73 of 372 calls to a Regional Poison Control Center (RPCC) that involved workplace disease/exposure(s); most other calls were not made by the workers. An average of 12 additional people per workplace were potentially exposed. Six of the 73 contacted a government agency for investigation of the hazard/illness. Twenty-five percent of callers were still exposed an average of seven months after the original call. The results indicate that poison control centers should develop a public health component to calls about possible workplace poisonings. (*Am J Public Health* 1990; 80:711–712.)

Introduction

An accurate estimate of the incidence of occupational disease in the United States is hampered by the lack of an active surveillance system.[1–4] Investigators have proposed using existing data systems including regional poison control centers[5,6] for passive surveillance of occupational disease. A call may represent a group of affected/exposed workers and serve as a sentinel health event that identifies hazardous worksites.[7]

We conducted a follow-up survey of individuals with occupational exposure/disease who contacted the Delaware Valley Regional Poison Control Center (DVRPCC), over a 12-month period.

Methods

Approximately 7,760 individuals over the age of 17 contacted the Delaware Center between July 1, 1986 and June 30, 1987. Information was recorded on the American Association of Poison Control Centers' precoded report form. Referral guidelines listing the telephone numbers of outside agencies to which occupationally related callers may be referred are available to staff. The guidelines indicate that calls or complaints about known Occupational Safety and Health Administration (OSHA) regulated substances should be referred to OSHA and complaints about substances not covered or enforced by OSHA should be referred to the National Institute for Occupational Safety and Health (NIOSH).

A workplace exposure was recorded on the intake form of 372 calls (4.8 percent). Trained personnel, unrelated to the poison control center, abstracted data from these intake forms and called the telephone number obtained during the initial encounter. These follow-up calls were made between July 1 and November 1, 1987, an average of seven months after the initial call.

The interviewer explained the purpose of the study to the affected worker and obtained informed consent. To proceed, the interview was then either administered immediately or within one week at a prearranged time. In three cases, the interviewee was the sentinel case's employer who answered the questions from his/her perspective. Interviewers used a pre-tested, pre-coded form containing predominantly closed-ended questions. The analysis included 95% confidence intervals for differences in proportions.[8]

The study protocol was reviewed and approved by the Committee for the Protection of Human Subjects of the Medical College of Pennsylvania and the Medical Advisory Panel of the DVRPCC.

Results

The interviewers successfully interviewed 73 of the 372 work-related calls. Table 1 compares the characteristics of the 73 workers interviewed and the 299 workers not interviewed from information on the Center's intake form. Of those not interviewed, 209 (69.9 percent) calls originating from medical personnel or hospitals were excluded, because we did not believe that our questions would be answered without the patient's consent. Forty-nine callers (16.4 percent) either left no telephone number or gave a disconnected or false number. Of the remainder, 16 (5.4 percent) had been terminated from their jobs, 15 (5 percent) refused to answer questions and one (.3% percent) was an informational call.

The interviewed workers (responders) had acute exposures to more than 60 different chemicals at 73 different worksites; 19 individuals were exposed to two or more substances. All but four of the affected workers were symptomatic at the time of the initial call.

The responders had worked an average of 5.5 years at the worksite where the exposure occurred at the time of the follow-up call. Forty-nine (67 percent) of the 73 responders indicated that other workers (mean of 12) had been exposed

Address reprint requests to Eddy A. Bresnitz, MD, Division of Occupational and Environmental Health, Department of Community & Preventive Medicine, Medical College of Pennsylvania, 3300 Henry Avenue, Philadelphia, PA 19129. This paper, submitted to the Journal December 27, 1988, was revised and accepted for publication November 8, 1989.