Effects of Incentive Items on Participation in a Randomized Chemoprevention Trial

DEBORAH BOWEN, MARK THORNQUIST,
& GARY GOODMAN
*Fred Hutchinson Cancer Research Center, Seattle, WA, and
University of Washington, USA*

GILBERT S. OMENN
University of Michigan, Ann Arbor, MI, USA

KAREN ANDERSON & MATT BARNETT
Fred Hutchinson Cancer Research Center, Seattle, WA, USA

BARBARA VALANIS
Kaiser Permanente, Portland, OR, USA

Abstract

Behavioral research has an important role in increasing and maintaining participation in disease prevention trials, both in interventions and in follow-up visits. We conducted a randomized experiment among participants in the lung cancer chemoprevention trial, CARET (Carotene and Retinol Efficacy Trial) to test the effects of providing two incentives on retention. The items used for this study were a Certificate of Appreciation and one of two lapel pins, provided in a 2 × 2 design. Providing incentives, whether alone or in combination, had no statistically significant effect on retention by the two-year follow-up point. The successful implementation of this randomized incentive study has two implications for future research: (1) study of behavioral interventions and issues is feasible in the context of large controlled trials of disease end-points; and (2) such study is necessary to determine whether selected incentives can increase retention.

Keywords

clinical trials, incentives, retention

109

LARGE-SCALE, long-term primary prevention trials pose challenges for maintaining participation of the participants (Meinert, 1986). The long-term nature of many prevention trials (often five to 10 years of intervention exposure) means that participants must take pills or maintain other intervention changes for years and return for clinical follow-up visits. Other priorities, such as illness, job changes, and other life demands and changes, can interfere with participation over a long follow-up time. Prevention trials often offer no noticeable health benefit for participation, because participants typically are healthy when entering the trial. Because poor adherence can have adverse effects on the statistical power of trials, holding other factors constant (Zelen, 1988) and maintaining participation must be a high priority.

Behavioral researchers have studied methods to increase and sustain participation in prevention trials, for both intervention activities and follow-up visits (Schron, Ockene, McBee, & Shumaker, 1998). In general, incentives are a commonly used component of behavior change interventions and, indeed, have been used to create entire systems of reinforcement (Kazdin, 1984; Warner & Murt, 1984; Winett, King, & Altman, 1989). Winett & colleagues emphasized that, in addition to an external contingency, or outcome such as receiving a small gift for participation in a trial, cognitive mediators of the contingencies are important in understanding the effects of incentives. Specific beliefs about the incentive, values regarding the meaning of the incentive, and internal standards of behavior all play a role in the likelihood that an incentive will alter behavior patterns. These properties of incentives, based on social learning theory (Bandura, 1986), indicated a need to consider the psychological context in which incentives are delivered in research projects. In the context of prevention trials, incentives can serve as reinforcers for participants if the incentives are coherent with participants' beliefs, if they support participants' values, and if they fit current standards of behavior, all as a reminder of the participant's role and identity within the study. We used these principles in designing an incentive program for the CARET (Carotene and Retinol Efficacy Trial).

Unfortunately, these activities receive very little rigorous evaluation in the context of research activities.

The present study, called PRIDE (Participant Retention Item Distribution Evaluation), was conducted among participants in the CARET chemoprevention trial (Omenn et al., 1994). PRIDE was designed to answer two questions: (1) can the efficacy of incentive use be evaluated rigorously in the context of an ongoing large-scale primary prevention trial?; and (2) what are the effects of two specific incentives on adherence and maintenance of active status in CARET?

Methods

CARET is a randomized, multicenter, placebo-controlled trial to determine the safety and efficacy of beta-carotene and retinyl palmitate in preventing lung cancer in two high-risk populations: men and women aged 50 to 69 years at entry who are current or recently quit former smokers with a cigarette smoking history of at least 20 pack-years, and men aged 45 to 69 years at entry who have recent cigarette smoking history and substantial occupational exposure to asbestos (Omenn et al., 1994; Thornquist et al., 1993). CARET was initiated in 1985 in Seattle with randomized pilot studies ($N = 1845$), and expanded in 1988–9 and again in 1990–1 in Seattle and five other study centers around the country, with final recruitment completed in 1994 ($N = 18,314$ total participants) (Goodman et al., 1993; Omenn et al., 1994). Participants were randomized to CARET at their second visit and have follow-up visits at six months (third visit) and then annually, plus two routine telephone contacts per year. Participants will take daily capsules for a mean of six years (range four to 13 years) and receive regular follow-up contacts. The CARET design (Thornquist et al., 1993) assumes that 10 percent of participants would become inactive (stop taking study vitamins or placebos for various reasons) each year for the first three years and then only 1 percent each year would become inactive. Retention rates have, in fact, been much better, with the four-year inactive proportion being 21 percent in Portland and 16 percent in Seattle.

Participants

Participants for PRIDE were recruited from the Portland and Seattle Study Centers of CARET.
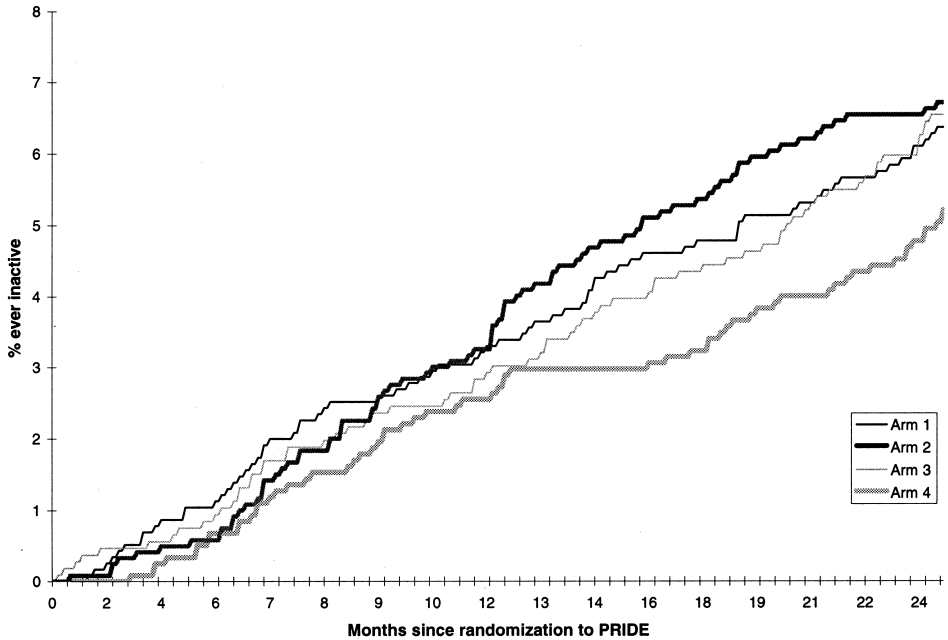
Figure 1. Model of time to inactivity for participants in four arms of PRIDE.

The study population was defined as all participants having a randomization or follow-up visit between 5 August 1991 and 5 February 1992. This time period was chosen to fit the ongoing activities of such a large trial with many demands on staff time. Eligible participants were randomized to one of four arms in PRIDE, receiving a lapel pin, a certificate, a pin and a certificate, or no incentive item. The Coordinating Center and the study centers jointly identified procedures. To minimize interaction between participants in different arms of PRIDE, days, rather than individual participants, were randomized, i.e. all participants with a visit on a given day received the same item(s) (or none). The appropriate incentives were presented during a routine study center visit. Each presentation was accompanied by a statement by study center staff to reinforce the incentive value of the item(s): thanking the participant for past efforts, reinforcing the CARET identity and each participant's attendance to date, and congratulating the participant. We chose these messages because CARET participant data had indicated that the most frequently stated reason for participating in CARET was 'being part of an important research project'.

Study center staff received training and used scripts to deliver this message clearly and convincingly. Participants were followed for two years for the primary outcome, active/inactive status (i.e. whether the participant continued actively to take the CARET study capsules). Both study centers received special quality assurance visits from the coordinating center staff to monitor implementation of PRIDE.

Incentives

The items used for PRIDE were a Certificate of Appreciation and one of two lapel pins. The Certificate of Appreciation was preprinted on an $8\frac{1}{2} \times 11$ inch bond off-white paper with gold trim and bold, black lettering. The certificate read '[participant's name] in recognition of your contribution to an important national study for the prevention of lung cancer, CARET, Cancer Prevention Study, sponsored by The National Cancer Institute'. The participant's name was computer printed in an attractive font on the certificate. Each certificate had the signatures of the Coordinating Center's principal investigator (G.O.), study center investigator (G.G. or B.V.), and CARET's project officer from the National

111

Cancer Institute (Dr Marjorie Perloff). The lapel pins were each 1″ in size and designed in cloisonné. One pin had six colors with the words 'CARET' and 'NCI Cancer Prevention Study' and had an orange carrot in the middle of the pin. The second pin was done in five colors with the words 'PARTICIPANT, CARET Cancer Prevention Study' and 'Sponsored by NCI'. Participants assigned to groups given the pins were offered a choice of one of the two pins. Each incentive cost approximately $0.25 to purchase, and staff members spent 30 seconds or less during a standard clinic visit providing the incentive. Therefore, these incentives were quite inexpensive in per-participant costs.

Statistical methods

The primary outcome measure in PRIDE was the time of the first inactivation during the two-year follow-up period of PRIDE. The statistical method was the product limit (Kaplan–Meier) estimator (Kalbfleisch & Prentice, 1980), which estimates the time course of the cumulative proportion of individuals who have experienced an event, even if not all individuals have complete follow-up. The log-rank statistic (Kalbfleisch & Prentice, 1980) was used to test whether several Kaplan–Meier curves all estimate the same true time course of cumulative proportion of event occurrence (i.e. inactivation). The effect of prognostic variables on the rate of inactivation was assessed using proportional hazards (Cox) regression (Kalbfleisch & Prentice, 1980) and was tested with likelihood ratio statistics. Parameters in the Cox regression model (in the simple case of dichotomous covariates) estimate the ratio of the event rates (called the relative risk) for those with the covariate to those without the covariate. Cox regression differs from logistic regression by including the time that the event occurred as additional information. As in other regression techniques, Cox regression allows the effects of multiple prognostic variables to be assessed simultaneously.

Both the log-rank statistic and the Cox regression were extended to stratified analyses for the present article. Conceptually, the study population is divided into strata, the statistic computed in each stratum, and the results for all strata pooled. Stratified analysis is indicated if the population has identifiable subgroups with markedly different time courses of events. Where

indicated, test statistics reported in the article are stratified on study center (Portland and Seattle), exposure population (asbestos-exposed workers and smokers), and cohort (randomization 1985–8 and 1989–95).

Results

Effects of incentives

A total of 4728 current CARET participants had study center visits at Portland and Seattle during the randomization period of PRIDE. Of these, 10 were excluded from the analysis because they were not ultimately randomized in CARET; 20 because the randomization to PRIDE was more than two weeks before the randomization to CARET; 48 because they had multiple visits during the PRIDE randomization period and the days of those visits were randomized to different groups; 48 because a member of their household was randomized to a different PRIDE group; and nine because they had become inactive on or before randomization to PRIDE. Thus, 4593 participants were eligible for analysis. The baseline characteristics of the study sample were comparable to those of the larger CARET population: 97 percent Caucasian; 37 percent women; mean age = 58.1 years; 63 percent current smokers. There were no differences in any of these variables among PRIDE study groups.

Figure 1 presents the Kaplan–Meier estimates of the cumulative proportion inactive for the four randomization groups in PRIDE. The estimated proportions inactive two years after randomization to PRIDE were 5 percent, 7 percent, 7 percent, and 6 percent for pin alone, certificate alone, pin and certificate, and comparison conditions, respectively. When pooling of the data from the three groups receiving incentive items, the estimates proportion inactive did not differ between the pooled and control groups (both 6 percent at two years, log-rank $\chi^2 = 0.08$, $p = .78$). The differences among the groups are not statistically significant (log-rank $\chi^2 = 2.84$, $p = .42$). Compared to the group that received no tokens, the estimated relative risks (95 percent confidence intervals) for the groups receiving the pin alone, the certificate alone, and the pin and certificate together, were 0.81 (0.58, 1.14), 1.03 (0.74, 1.42), and 1.06 (0.77, 1.45), respectively. There were also no significant differences at one year.

Results of regression models

Stratified Cox regression models were fit to adjust the estimated group relative risks for other prognostic variables. Variables considered in the models included: age at first CARET contact; marital status and general health at the time of randomization to PRIDE; gender; study center; population (asbestos-exposed worker and heavy smoker); cohort (Pilot and Efficacy); visit at which randomization to PRIDE occurred; time since randomization to CARET; reasons the participant gave for participating in CARET; and interactions among these variables. These variables were selected because of their potential relevance to participants' reactions to the PRIDE incentives. Thirty-six participants were excluded from the final model due to missing data, and one additional participant was excluded because the PRIDE randomization occurred at a non-routine CARET contact. The final model selected was stratified on study center, population, and cohort, and included predictor variables for intervention group, study center, the interaction between intervention group and study center, the visit at which randomization to PRIDE

occurred, general health at randomization to PRIDE, and the participant's reason for participating in CARET. Adjusted for these covariates, none of the other variables or interactions considered had statistically significant parameter estimates. Parameter estimates, confidence intervals, and p values are presented in Table 1.

General health at randomization to PRIDE was predictive of the subsequent rate of inactivation; participants who reported poor health were over four times as likely to become inactive as those who reported excellent health. The timing of randomization to PRIDE was also predictive of subsequent inactivation rates; participants who were randomized to PRIDE at the same time as they were randomized to CARET were 2.76 times as likely to become inactive than those who had been participants in CARET for at least a year before randomization to PRIDE. Participants who gave altruistic reasons (labeled 'other' in Table 1) for participating in CARET were numerically less likely to become inactive than those whose reasons weren't altruistic (labeled 'self' in Table 1), but the effect was not statistically significant. Adjusted for these fac-

Table 1. Estimated relative risks (confidence intervals and *p* values) for becoming inactive on study

Prognostic variable	Category	Estimated relative risk	95 % confidence interval	p value
General health	Excellent[a] (*n* = 626)	1		< 0.01
	Very good (*n* = 1883)	1.04	(0.69, 1.55)	
	Good (*n* = 1604)	1.25	(0.84, 1.86)	
	Fair (*n* = 396)	2.54	(1.61, 4.02)	
	Poor (*n* = 47)	4.04	(1.85, 8.83)	
Time of randomization	Second visit (*n* = 1195)	2.76	(1.98, 3.84)	< 0.01
	Third visit (*n* = 908)	2.13	(1.50, 3.04)	
	Annual visit[a] (*n* = 2453)	1		
Reason in CARET	Other (*n* = 4247)	0.70	(0.48, 1.04)	0.09
	Self[a] (*n* = 309)	1		
Tokens received by Portland participants	None[a] (*n* = 449)	1		0.67
	Pin (*n* = 413)	1.26	(0.78, 2.05)	
	Certificate (*n* = 376)	1.31	(0.81, 2.13)	
	Both (*n* = 467)	1.11	(0.69, 1.78)	
Tokens received by Seattle participants	None[a] (*n* = 693)	1		0.26
	Pin (*n* = 756)	0.62	(0.38, 1.01)	
	Certificate (*n* = 680)	0.88	(0.55, 1.39)	
	Both (*n* = 722)	0.88	(0.57, 1.37)	

[a] Comparison category to which other relative risks refer

113

tors, the differences among the four intervention groups were not statistically significant, and there was no evidence of additive effects for the pin and certificate. The groups that received tokens in Seattle had numerically lower inactive rates than the comparison group, while in Portland the same groups had numerically higher rates; the comparison between centers within PRIDE study arms was not statistically significant.

Discussion

Providing incentives, whether alone or in combination, had no statistically significant effect on retention by the two-year follow-up point. Patterns of retention were similar at the one-year point. These findings suggest that manipulating the provision of single or double incentives over and above a carefully planned retention program may not markedly improve retention in a trial like CARET. Each of the study centers had an existing program of retention activities, including standardized procedures, reminders for visits, waiting room activities, national and local newsletters, and previous incentives. This package has resulted in the previously mentioned high retention rates in CARET. PRIDE was superimposed on the existing retention package. Whether the simple PRIDE incentives were not useful in promoting retention or whether the PRIDE incentives got 'lost' in the generally positive atmosphere of the study centers is hard to distinguish. It also could be that the choice of the particular incentives was not the best; we did not measure the reinforcement value of the incentives in a group of participants. Future studies could include process assessment to identify participants' conscious reactions to the incentive. This might allow us to judge the properties of the incentive more thoroughly.

A limitation to this study is that fewer participants were randomized than the 6500 specified in the initial study design. Only 4728 participants were enrolled in the study, primarily due to an overestimate of the number of participants who would be enrolled in CARET during the six-month PRIDE randomization period. This decrease in sample size compromises the ability to detect small differences in retention rates; and given the high retention rates in CARET, any increase in rates is likely to be small, even if the

incentives have a strong effect on retention. However, despite the smaller sample size, the PRIDE study had 78 percent power with a (two-sided) significance level of 0.05 to detect a 2 percent difference in the retention rates between the groups receiving incentive items and the control group. This gives us some confidence that we have provided a reasonable test of the PRIDE hypothesis.

Receiving incentives was associated with (not statistically significant) lower inactivation in Seattle, but not in Portland. A review of study documentation for the relevant time indicated that Portland Study Center staff reported mixed levels of enthusiasm for the PRIDE activities, compared to Seattle Study Center staff. Portland staff may have delivered different messages about the value of the participants' receipt of the incentive, compared to Seattle staff. One area for future research is that staff perspective on a given retention item could be important to the success of retention activities. This hypothesis fits with the theory that cognitive mediators may influence the efficacy of incentives.

The successful conduct of this randomized incentive study has two implications for future research: (1) the study of behavioral interventions and issues in the context of large controlled trials of disease end-points is feasible; and (2) such study is necessary to investigate both the nature of items that may be effective and the timing for use of such incentives. Based on the findings of PRIDE, behavioral strategies need more evaluation before one assumes that they are efficacious.

References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive approach*. Englewood Cliffs, NJ: Prentice Hall.

Goodman, G., Omenn, G. S., Thornquist, M., Lund, B., Metch, B., & Gylys-Colwell, I. (1993). The Carotene and Retinol Efficacy Trial (CARET) to prevent lung cancer in high-risk populations: pilot study with cigarette smokers. *Cancer Epidemiology Biomarkers and Prevention, 2*, 389–396.

Kalbfleisch, J. D., & Prentice, R. L. (1980). *The statistical analysis of failure time data*. New York: Wiley.

Kazdin, A. E. (1984). *Behavior modification in applied settings* (3rd ed.). Homewood, IL: Porsey.

Meinert, C. L. (1986). *Clinical trials: Design, conduct, and analysis*. New York: Oxford University Press.

Omenn, G. S., Goodman, G. E., Thornquist, M., Grizzle, J., Rosenstock, L., Barnhart, S., Balmes, J., Cherniack, M., Cullen, M., Glass, A., Keogh, J., Meyskens, F. Jr., Valanis, B., & Williams, J. Jr. (1994). The β–Carotene and Retinol Efficacy Trial (CARET) for chemoprevention of lung cancer in high-risk populations: Smokers and asbestos-exposed workers. *Cancer Research, 54*, 2038S–2043S.

Schron, E., Ockene, J., McBee, W. L., & Shumaker, S. A. (Eds.) (1998). *The handbook of health behaviour change. Second Edition*. New York: Springer.

Thornquist, M. D., Omenn, G. S., Goodman, G. E., Grizzle, J. E., Rosenstock, L., Barnhart, S., Anderson, G. L., Hammar, S., Balmes, J., Cherniack, M., Cone, J., Cullen, M. R., Glass, A., Keogh, J. P., Meyskens, F. L. Jr., Valanis, B., & Williams, J. H. Jr. (1993). Statistical design and monitoring of the Carotene and Retinol Efficacy Trial (CARET). *Controlled Clinical Trials*, 14, 308–324.

Warner, K. E., & Murt, H. A. (1984). Economic incentives for health. In L. Breslow, L. B. Laver, & J. E. Fielding (Eds.), *Annual review of public health* (Vol. 5, pp. 107–133). Palo Alto, CA: Annual Reviews.

Winett, R. A., King, A. C., & Altman, D. C. (1989). *Health psychology and public health: An integrative approach*. Elmsford, NY: Pergamon Press.

Zelen, M. (1988). Are primary cancer prevention trials feasibile? *Journal of the National Cancer Institute, 80*, 1442–1444.

115