Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Halpern SD, French B, Small DS, et al. Randomized trial of four financial-incentive programs for smoking cessation. N Engl J Med 2015;372:2108-17. DOI: 10.1056/NEJMoa1414293

SUPPLEMENTARY APPENDIX

A Randomized Trial of Four Financial Incentive Programs for Smoking Cessation

Scott D. Halpern, M.D., Ph.D.[1,2,3,4]
Benjamin French, Ph.D.[2,3]
Dylan S. Small, Ph.D.[2,6]
Kathryn Saulsgiver, Ph.D.[2,3]
Michael O. Harhay, M.P.H.[3]
Janet Audrain-McGovern, Ph.D.[2,7]
George Loewenstein, Ph.D.[2,8]
Troyen A. Brennan, M.D., J.D.[9]
David A. Asch, M.D., M.B.A. [1,2,5,10,11]
Kevin G. Volpp, M.D., Ph.D.[1,2,4,5,10,11]

AUTHOR AFFILIATIONS
[1]Department of Medicine, [2]Center for Health Incentives and Behavioral Economics at the Leonard Davis Institute of Health Economics, [3]Department of Biostatistics and Epidemiology, [4]Department of Medical Ethics and Health Policy, and [7]Department of Psychiatry, Perelman School of Medicine at the University of Pennsylvania; [5]Center for Health Care Innovation, University of Pennsylvania Health System; [6]Department of Statistics and [10]Department of Health Care Management, Wharton School, University of Pennsylvania; [8]Center for Behavioral Decision Research, Carnegie Mellon University; [9]CVS/Caremark; [11]Center for Health Equity Research and Promotion, Philadelphia VA Medical Center

SUPPLEMENTARY APPENDIX

Table of Contents

Supplemental Methods:

A. Estimation of the Complier Average Treatment Effect

This section provides a walkthrough of the complier-average treatment effect using instrumental variables (IVs) to estimate efficacy for this study.

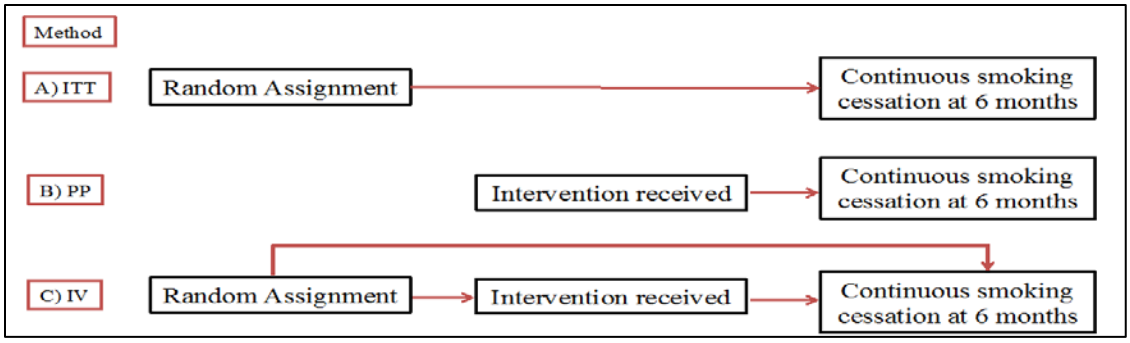
Introduction

In a randomized clinical trial (RCT), there are traditionally two types of analyses (Figure): intention-to-treat (ITT) and per-protocol (PP) analyses. The ITT analysis is used to assess the effect of an intervention among all those who are initially randomized to each study arm (Figure, diagram A). However, randomization does not necessarily indicate acceptance of the assigned intervention or compliance with the treatment. Therefore, PP or as-treated analyses are commonly employed to quantify the treatment effect among those who accept or adhere to their assigned intervention (Figure, diagram B).

The ITT analysis quantifies the effectiveness of an intervention among those randomly assigned to receive it. The PP assesses the efficacy of an intervention among those who chose to accept or adhere to it. This RCT was designed to assess acceptance, efficacy, and effectiveness. As a consequence, participants initially randomized to an intervention arm could choose to accept or not accept their assigned intervention. Traditional PP analyses are often biased because they do not account for post-randomization selection phenomena, such as the possibility that participants who accept or reject their initial arm differ with regard to their underlying propensities to have favorable outcomes.

In this trial we integrate instrumental variable (IV) and principal stratification methods to overcome the biases typically associated with PP analyses (Figure 1, diagram C). The following sections explain how this method works, and describe the calculations of the estimated treatment effects. The IV analysis in this RCT is estimated using the randomization arm (i.e., initial treatment assignment) as the instrument. As a result, the estimated treatment effect of each intervention on smoking cessation is adjusted for the percentage of assigned patients who accepted the treatment arm. This approach differs from a PP analysis, having the key advantage that the IV approach uses the data on all randomized patients, rather than merely those who accept the assigned intervention, and then adjusts for acceptance, thereby attenuating the selection effects.

Methods to quantify treatment effects in an RCT of interventions to improve smoking cessation



Adapted from Sussman & Hayward.[1]

Illustration in current RCT

For illustration, we will use the primary outcome, sustained smoking abstinence through 6 months (coded as 1 if continuous smoking abstinence, and 0 if not). This analysis requires the integration of the acceptance results to help attenuate the potential selection biases associated with successful cessation among those who accept their initial randomization arm compared to those that do not.

UC=Usual care, IR=Individual rewards, CR= Collaborative Reward, ID=Individual Deposit, CD=Competitive deposit

We begin with the initial results.*

					ITT		***PP***
Arms	**Initially randomized (n)**	**Accepted initial arm (n)**	**% acceptance**	**n quit (ITT)**	**% quit**	**n quit (PP)**	**% quit**
UC	468	468	-	28	6.0%	28	6.0%
IR	498	472	94.8%	76	15.3%	76	16.1%
CR	519	442	85.2%	83	16.0%	80	18.1%
ID	582	75	12.9%	56	9.6%	39	52.0%
CD	471	71	15.1%	52	11.0%	39	54.9%

Due to the similarities in results between the individual and collaborative reward arms and in results between the individual and competitive deposit arms, results were collapsed within the reward and deposit arms separately. Thus, the following data are used.*

					ITT		***PP***
Arms	**Initially randomized (n)**	**Accepted initial arm (n)**	**% acceptance**	**n quit (ITT)**	**% quit**	**n quit (PP)**	**% quit**
UC	468	468	-	28	6.0%	28	6.0%
Rewards	1017	914	89.9%	159	15.6%	156	17.1%
Deposits	1053	146	13.9%	108	10.3%	78	53.4%

UC=Usual care; *note that the percentages reported in this Table are the raw percentages, whereas the percentages reported in the main manuscript are adjusted for the two variables used to stratify the randomization, as is recommended in manuscript reference 30.

This analysis assumes that all participants who would accept deposits if assigned to receive deposits would also accept rewards if assigned to receive rewards, thereby creating three compliance classes or principal strata of subjects.[2,3]

$C_i = 000$, subject *i* would not accept either rewards or deposit

$C_i = 0R0$, subject *i* would accept rewards but not deposit

$C_i = 0RD$, subject *I* would accept both rewards and deposit

Each subject has three potential outcomes that are listed below. Only one of the potential outcomes can be observed, the outcome corresponding to the actual treatment the subject received.

$$Y_i^0 = \text{whether subject i would stop smoking if taking control}$$
$$Y_i^R = \text{whether subject i would stop smoking if taking rewards}$$
$$Y_i^D = \text{whether subject i would stop smoking if taking deposit}$$

Our analysis assumes the exclusion restriction that a subject's potential outcome depends only on the treatment received. In other words, the treatment assignment only influences the potential outcome through the treatment received.[4]

To understand the composition of the groups being compared in estimating these values, we provide the below table. Following the table, we provide the calculations used to estimate the rates for each compliance class. Because this study was randomized we can assume that the rates will be constant across arms, therefore allowing us to infer values in compliance classes that are not directly observed. The study design initially offers each participant entrance to an intervention arm or usual care. The participants can accept or reject the intervention arm, thereby creating a potential bias in comparisons between those who accept versus those who do not. Thus, the composition of each arm can be separated into the three classes mentioned above and displayed in tabular form below. We provide the calculations for the proportions below the table based on the summary tables above.

Initial Arm	Strata / compliance class	Proportion of sample
UC	000	.101
UC	0R0	.760
UC	0RD	.139
Reward	000	.101
Reward	0R0	.760
Reward	0RD	.139
Deposit	000	.101
Deposit	0R0	.760
Deposit	0RD	.139

Note: .101 + .139 + .760 = 100% of all patients.

We use the following notation:

Z = Initial intervention arm A = Treatment taken

1. To estimate the proportion of non-accepters (who receive usual care in this trial) in the reward arms we calculate:

1-((472+442)/(498+519))= .101, 10.1%

In notation: P(000) = P(A=UC | Z=R)

2. To estimate the proportion that would accept either a reward or deposit intervention we calculate:

(75+71)/(582+471) = .139, 13.9

In notation: P(0RD) = P(A=D| Z=D)

3. To estimate the proportion that would accept rewards but not deposits we calculate:

1 - 0.101- 0.139 = .76, 76%

In notation: P(0R0) = 1 – [P(000)+P(0RD)]

With these proportions we can now perform three treatment effect estimates using the weighted averages of these proportions and the observed outcomes using IV techniques.

The analysis was performed using R 3.0.2. The code for this analysis is available upon request.

Estimations

Using this framework for our study, the following estimates were calculated:

Estimate 1:

Effect of rewards vs. usual care for subjects who would accept rewards if offered them ($C_i = 0R0$ or $C_i = 0RD$): 0.107 (95% CI: 0.068, 0.147). This value, presented as a risk difference, is calculated using a weighted average of the usual care outcomes in 0R0 & 0RD compared to the rewards strata with the usual care arm as the reference category.

$$\frac{\hat{P}(Y=1|Z=R)-\hat{P}(Y=1|Z=UC)}{\hat{P}(A=R|Z=R)-\hat{P}(A=R|Z=UC)}=\frac{\hat{P}(Y=1|Z=R)-\hat{P}(Y=1|Z=UC)}{\hat{P}(A=R|Z=R)}$$

Estimate 2:

Effect of deposit vs. usual care for subjects who would accept deposits if offered them ($C_i = 0RD$): 0.308 (95% CI: 0.110, 0.506). This value, presented as a risk difference, is calculated using the usual care outcomes in 0RD compared with outcomes of ORD in the deposit arms.

$$\frac{\hat{P}(Y=1|Z=D)-\hat{P}(Y=1|Z=UC)}{\hat{P}(A=D|Z=D)-\hat{P}(A=D|Z=UC)}=\frac{\hat{P}(Y=1|Z=D)-\hat{P}(Y=1|Z=UC)}{\hat{P}(A=D|Z=D)}$$

Estimate 3:

The final estimate(s) compare the deposit arm to the reward arm in the 0RD class

This comparison is not identified without making an assumption about the following parameter:

e^{β} = given that a subject is either $C = 0\text{RD}$ or $0\text{R}0$,
odds ratio of being $C = 0\text{RD}$ for $Y^R = 1$ vs. $Y^R = 0$

e^{β} therefore reflects the intrinsic difference in underlying probability of quitting among people who would accept deposits (or rewards) vs. people who would accept only rewards. A related measure of this relative quit potential which is identified as:

e^{γ} = given that a subject is either $C = 0\text{RD}$ or $0\text{R}0$, the odds ratio of being $C = 0\text{RD}$ for $Y^0 = 1$ vs. $Y^0 = 0$.

e^{γ} can be estimated as follows. The right hand side of the following three equations can be estimated:

$$P(Y^0=1|C=000)=P(Y=1|Z=R,A=UC)$$
$$P(Y^0=1|C=000 \text{ or } 0R0)=P(Y=1|Z=D,A=UC)$$
$$P(Y^0=1)=P(Y=1|Z=UC,A=UC)$$

The left hand side of the three equations can be written as functions of
$P(Y^0=1|C=000), P(Y^0=1|C=0R0), P(Y^0=1|C=0RD)$ and
$P(C=000), P(C=0R0), P(C=0RD)$. We have presented estimates of
$P(C=000), P(C=0R0), P(C=0RD)$ above. Consequently, the three equations in the three unknowns $P(Y^0=1|C=000), P(Y^0=1|C=0R0), P(Y^0=1|C=0RD)$ can be solved and an estimate of e^{γ}, which is a function of
$P(Y^0=1|C=0R0), P(Y^0=1|C=0RD), P(C=0R0), P(C=0RD)$ can be formed.

e^{β} and e^{γ} both refer to the odds ratio of being C=0RD vs. C=0R0 among those who cease smoking under a certain arm (e^{β} considers smoking cessation under the rewards arm and e^{γ} considers smoking

cessation under the usual care assumption). We make the assumption that $e^{\beta} = e^{\gamma}$ which is equivalent to a claim that the odds ratio effect of rewards vs. usual care on smoking cessation is the same for people who would accept deposits as for those people who would accept rewards only. This assumption of proportional odds is similar to that made in any logistic regression model that does not include an interaction between the treatment and the covariates. We report results for e^{β} at three values: the point estimate of e^{γ}, the upper 95% confidence limit of e^{γ}, and the lower 95% confidence limit of e^{γ}. The outcome remains continuous abstinence through 6 months. Given that a subject is either $C = 0\mathrm{RD}$ or $0\mathrm{R0}$, the odds ratio of being $C = 0\mathrm{RD}$ for $Y^0 = 1$ vs. $Y^0 = 0$ is estimated to be 9.36 with a 95% confidence interval of (2.71, 23.12).

e^{β}	$E(Y_i^D - Y_i^R \mid C_i = 0RD)$ (effect of deposits vs. rewards for rewards+deposit compliers)
2.71	0.258 (0.162, 0.348)
9.36	0.132 (0.031, 0.228)
23.12	0.064 (-0.057, 0.174)

We also find that the largest e^{β} at which the 95% confidence interval for $E(Y_i^D - Y_i^R \mid C_i = 0RD)$ is entirely positive is 12.5. This value is identified by searching over a grid of e^{β} with step size .5. This value also helps to define the sensitivity of our key conclusion to the assumption that $e^{\beta} = e^{\gamma}$. Our base estimate of the superiority of deposits vs. rewards among participants who would accept either is predicated on the most likely scenario, in which $e^{\beta} = e^{\gamma}$. However, the foregoing calculation shows that the effect remains statistically significant across a broad range of estimates of e^{β}, namely as long as e^{β} is less than (12.5/9.36)*e^{γ} = 1.34*e^{γ} (assuming that γ is equal to its point estimate). Thus, even if the odds ratio effect of rewards vs. usual care is different among participants who would accept deposits or rewards vs. participants who would accept only rewards, as long as the odds ratio effect among the former does not exceed 134% of the effect among the latter, the result still holds that the deposit arm is significantly better than the reward arm for participants who would accept either.

B. Details of the power calculation and main protocol modification

We estimated that enrolling 2,185 participants – 437 in each of the 5 arms – would provide 80% power to detect absolute differences of at least 7.5 percentage points in sustained abstinence rates between any one of the novel incentive programs versus the individual reward program previously shown to achieve a 6-month sustained abstinence rate of 14.7%.[1] This calculation was based on two-sided significance testing, adjustment for multiple comparisons using the stepwise Hochberg method where the initial test is at $\alpha = 0.05$ and subsequent tests are at progressively lower α levels, and an allowance

for sample size imbalance across arms of up to 10%. This sample size would also provide > 90% power to detect each incentive's effectiveness compared with usual care (estimated success rate of 5%[1]) and > 80% power to detect differences in acceptance of at least 7.5 percentage points less than acceptance of the individual reward arm, assuming the latter had at least 90% acceptance.

The adaptive randomization protocol enabled early detection of unexpectedly large differences in the interventions' acceptance. With authorization from the Data and Safety Monitoring Board, we implemented caps on the randomization probabilities to the less acceptable arms, while preserving random assignment to all arms. This preserved the ability to enroll the target number of 437 accepting participants in the more acceptable arms (see **Online Protocol**).

References

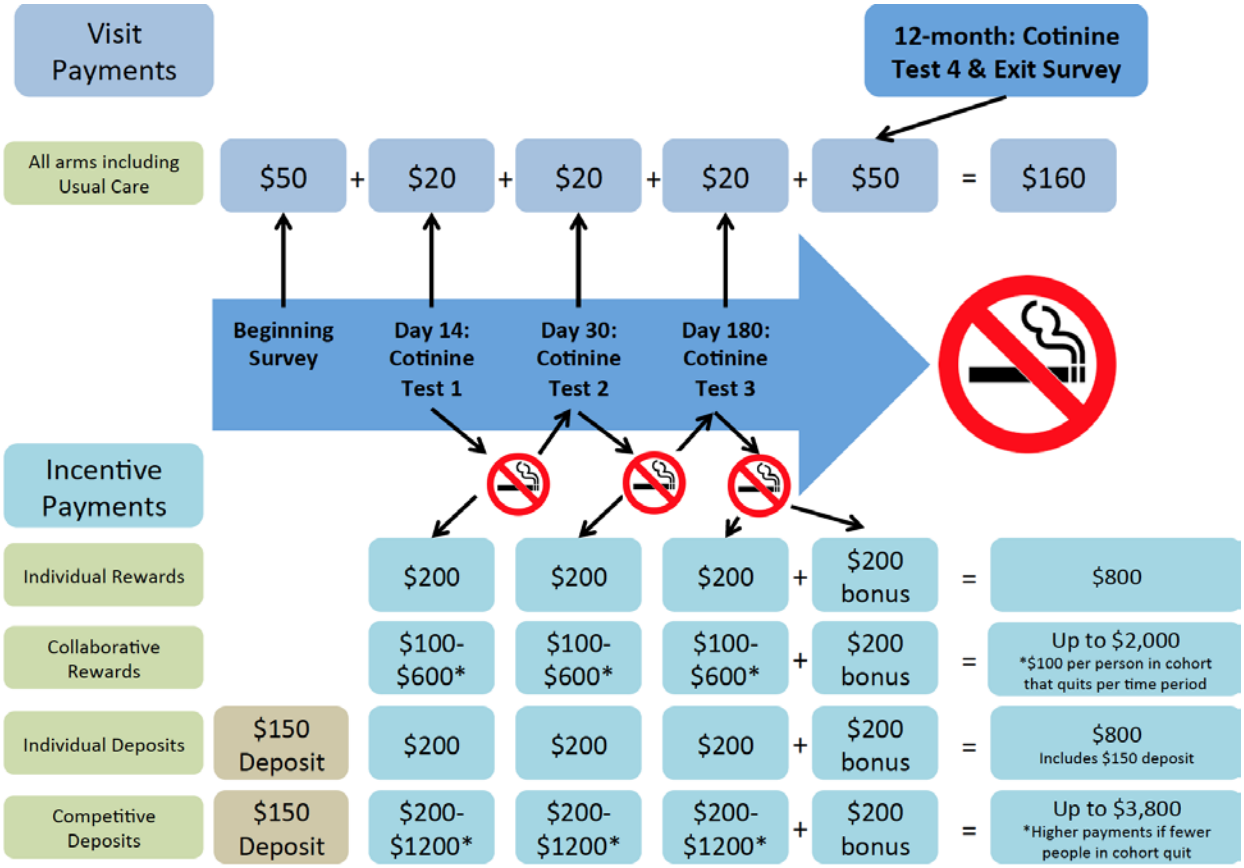
1. Sussman, J.B and Hayward, R.A. (2009). An IV for the RCT: using instrumental variables to adjust for treatment contamination in randomized controlled trials. *British Medical Journal*, 340, 1181-84.
2. Frangakis, C.E. and Rubin, D.B. (2002). Principal stratification in causal inference. *Biometrics*, 58, 21-29.
3. Cheng, J. and Small, D.S. (2006). Bounds on causal effects in three-arm trials with noncompliance. *Journal of the Royal Statistical Society, Series B*, 68, 815-836.
4. Angrist, J., Imbens G, Rubin D. Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association* 1996; 91:444-55.

Figure S1: Timing of recruitment strategies for CVS/Caremark employees and their friends and family members

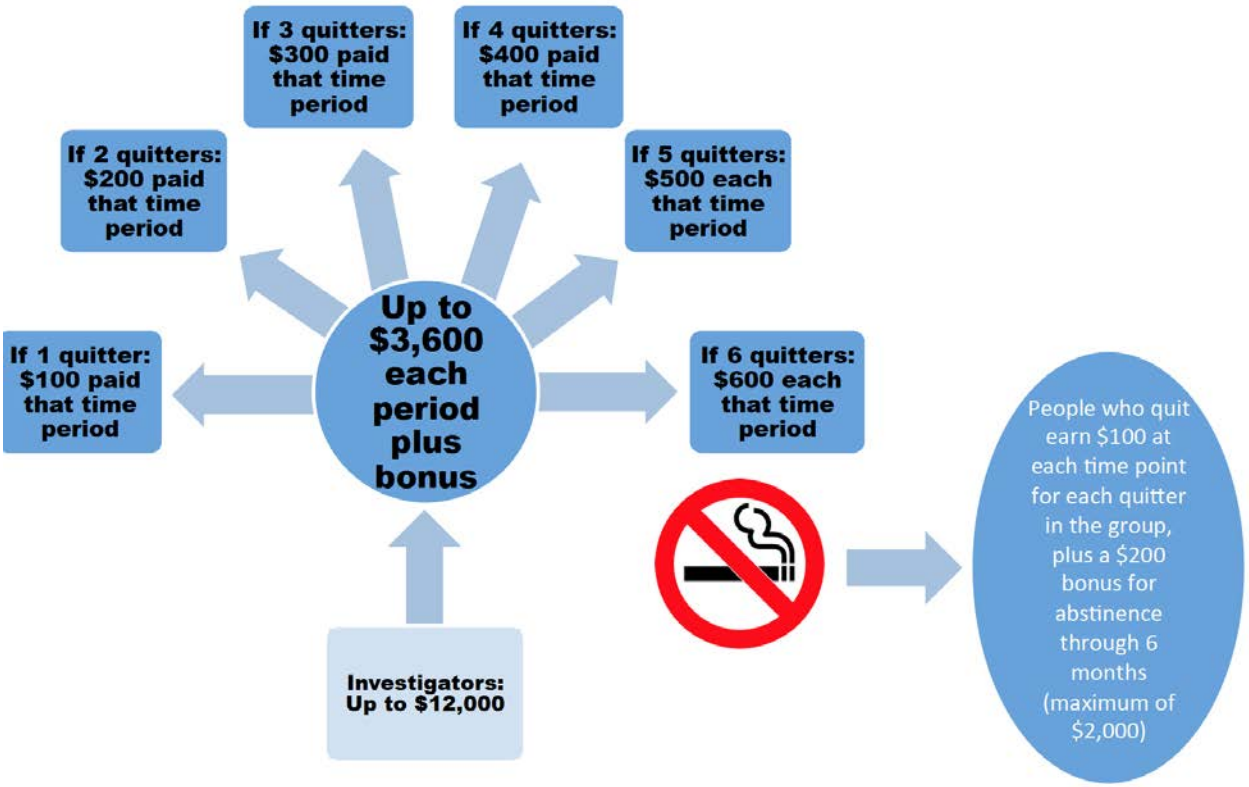
Date	Recruitment Strategies				
Week 1 **2/20/2012**	On-site information tables for smoke-free campus announcement	Email to Human Resources	Human Resource Presentations		
Week 2 **2/27/2012**	Email to cooperate employees	Inclusion in monthly benefits orientation	Email to employees enrolled in wellness program	Posted on employee health care centered website	Posted in web based employee newsletter (full-time employees)
Week 3 **3/5/2012**	Posted in new hire section of web based employee newsletter (full-time employees)	Email to store managers	Paycheck message	Posted on TV Plasma Screens at corporate locations	Posting made through web based employee resources
Week 4 **3/12/2012**	Email to wellness advocates	Posters placed in stores			
Week 5 **3/19/2012**	Postcards sent to all employees	Feature story in employee newsletter (all employees)			
Week 6 **3/26/2012**	Phone message from Chief Medical Officer				
Week 7 **4/2/2012**	Email to cooperate employees	Flyers and information tables at Benefit Fairs (April 2012)			
Weeks 8+	Posted on company wellness partners website (6/1/12)	Second postcard sent to employees (9/7/12)			

Figure S2: Schematics of the incentive programs

A. All trial arms



B. Collaborative rewards arm in greater detail



C. Competitive deposits arm in greater detail

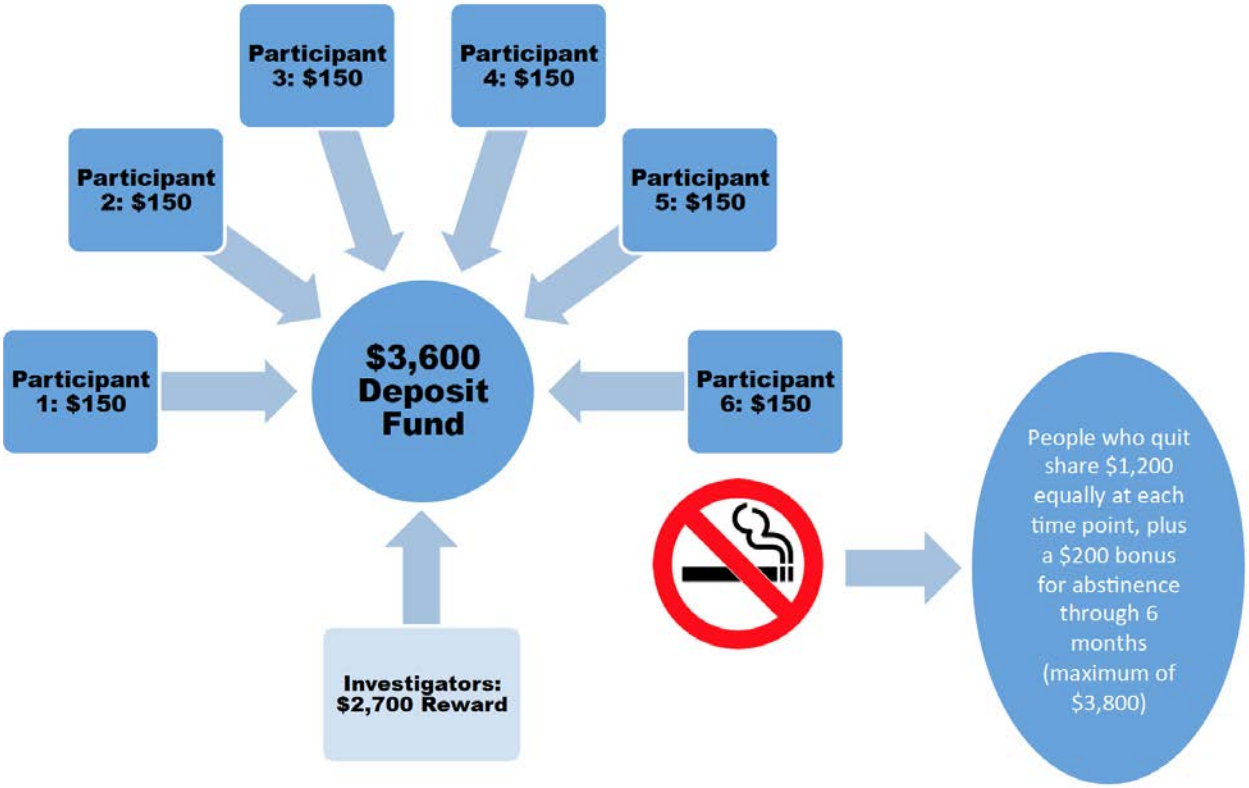
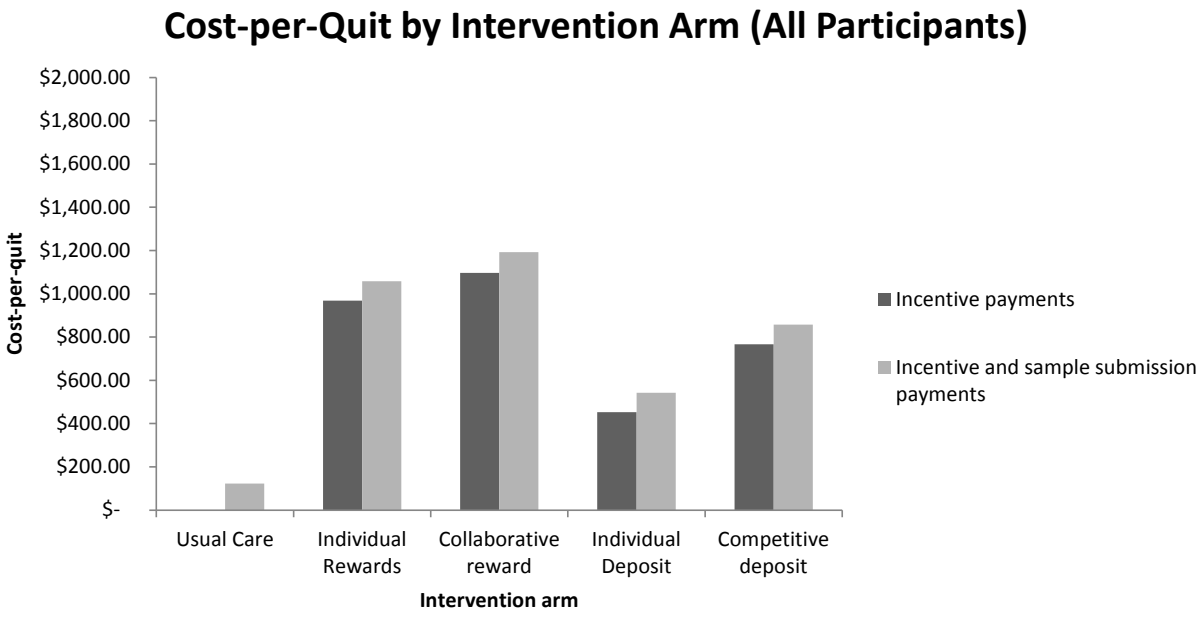


Figure S3: Costs incurred for each participant achieving sustained abstinence from smoking in the 5 study arms



This figure reports the total costs incurred per successful quitter in each arm. These costs reflect the best estimates of the costs that might be incurred by an employer that implemented each of these interventions. Data are reported as the costs of the incentives alone, and as the total costs that would be incurred if an employer required biochemical confirmation of abstinence in the same ways as in this trial. For the latter calculation, the total costs-per-quit were: $122.14 in usual care, $1,058.16 in individual rewards, $1,192.53 in collaborative rewards, $541.85 in individual deposits, and $858.01 in competitive deposits. Because the data are presented at the aggregate (arm) level, the p values were not estimated.

Table S1: Observed incentive payouts to participants in the four incentive arms

	N	Mean	Median	Interquartile Range	Full Range
All participants achieving abstinence for 6 months					
Individual Reward	76	$800	$800	$800-800	$0-800
Collaborative Reward	83	$890.36	$800	$700-1100	$0-1,700
Individual Deposit	56	$557.14	$800	$0-800	$0-800
Competitive Deposit	52	$839.62	$900	$630-1100	$0-1,940
All participants achieving abstinence for 6 months who accepted their incentive program					
Individual Reward	76	$800	$800	$800-800	$800
Collaborative Reward	80	$923.75	$800	$700-1100	$500-1,700
Individual Deposit	39	$800	$800	$800-800	$800
Competitive Deposit	39	$1,119.49	$900	$900-1250	$540-1,940

Table S2: Effectiveness of the four incentive programs versus usual care in producing sustained abstinence at 14 days, 30 days, 6 months and 12 months*

	N Sustained abstinence/ N Randomized	% Sustained abstinence	95% CI	P-value vs. usual care
14 days				
Usual care	66/468	14.2%	(11.0-17.3)	
Individual rewards	116/498	23.5%	(19.8-27.2)	< 0.001
Collaborative rewards	135/519	26.1%	(22.4-29.9)	< 0.001
Individual deposits	86/582	14.5%	(11.6-17.3)	0.89
Competitive deposits	74/471	15.8%	(12.5-19.1)	0.48
30 days				
Usual care	49/468	10.5%	(7.7-13.3)	
Individual rewards	100/498	20.2%	(16.7-23.8)	< 0.001
Collaborative rewards	117/519	22.6%	(19.1-26.2)	< 0.001
Individual deposits	77/582	12.9%	(10.2-15.6)	0.23
Competitive deposits	69/471	14.8%	(11.6-18.0)	0.051
6 months				
Usual care	28/468	6.0%	(3.9-8.2)	
Individual rewards	76/498	15.4%	(12.2-18.5)	< 0.001
Collaborative rewards	83/519	16.0%	(12.9-19.2)	< 0.001
Individual deposits	56/582	9.4%	(7.1-11.8)	0.04
Competitive deposits	52/471	11.1%	(8.3-14.0)	0.006
12 months[$]				
Usual care	16/468	3.4%	(1.8-5.1)	
Individual rewards	37/498	7.5%	(5.2-9.8)	0.007
Collaborative rewards	45/519	8.7%	(6.3-11.2)	0.001
Individual deposits	21/582	3.5%	(2.0-5.0)	0.94
Competitive deposits	29/471	6.2%	(4.0-8.4)	0.052
12 months (based on self-report)				
Usual care	22/468	4.7%	(2.8-6.7)	
Individual rewards	57/498	11.5%	(8.7-14.3)	< 0.001
Collaborative rewards	78/519	15.1%	(12.1-18.2)	< 0.001
Individual deposits	45/582	7.6%	(5.4-9.7)	0.06
Competitive deposits	40/471	8.5%	(6.0-11.0)	0.02

* Abstinence rates are adjusted for two stratifying variables: whether or not participants received their health insurance through CVS/Caremark, and whether their annual household income was $60,000 or greater. [$]Relapse rates at 12 months among those achieving abstinence through 6 months were similar in the usual care group (12/28; 43%) and the four intervention groups combined (135/267; 51%) (p = 0.44).

Table S3: Self-reported versus biochemically verified abstinence at 14 days, 30 days, 6 months and 12 months

	Usual Care N = 468	Individual Reward N = 498	Collaborative Reward N = 519	Individual Deposit N = 582	Competitive Deposit N = 471
14 days					
Self-reported quit	112	162	188	115	103
Tested negative	65	116	134	86	72
% of self reports confirmed	*58%*	*72%*	*71%*	*75%*	*70%*
30 days					
Self-reported quit	59	109	126	85	73
Tested negative	48	100	117	77	68
% of self reports confirmed	*81%*	*92%*	*93%*	*91%*	*93%*
6 months					
Self-reported quit	33	90	100	66	61
Tested negative	28	76	83	56	52
% of self reports confirmed	*85%*	*84%*	*83%*	*85%*	*85%*
12 months					
Self-reported quit	22	57	78	45	40
Tested negative	16	37	45	21	29
% of self reports confirmed	*73%*	*65%*	*58%*	*47%*	*73%*

Table S4: Effectiveness of reward-based versus deposit-based incentives in producing sustained abstinence at 14 days, 30 days, 6 months and 12 months*

	N Sustained abstinence/ N Randomized	% Sustained abstinence	95% CI	P-value reward versus deposit	P-value versus usual care
14 days				< 0.001	
Reward-based incentives	251/1017	24.8%	(22.2-27.5)		< 0.001
Deposit-based incentives	160/1053	15.1%	(12.9-17.2)		0.65
30 days				< 0.001	
Reward-based incentives	217/1017	21.5%	(19.0-24.0)		< 0.001
Deposit-based incentives	146/1053	13.7%	(11.7-15.8)		0.08
6 months				< 0.001	
Reward-based incentives	159/1017	15.7%	(13.5-18.0)		<0.001
Deposit-based incentives	108/1053	10.2%	(8.4-12.0)		0.009
12 months				0.002	
Reward-based incentives	82/1017	8.1%	(6.4-9.8)		0.001
Deposit-based incentives	50/1053	4.7%	(3.4-6.0)		0.26
12 months (based on self-report)				< 0.001	
Reward-based incentives	135/1017	13.4%	(11.3-15.5)		< 0.001
Deposit-based incentives	85/1053	8.0%	(6.4-9.6)		0.023

*Abstinence rates are adjusted for two stratifying variables: whether or not participants received their health insurance through CVS/Caremark, and whether their annual household income was $60,000 or greater.

Table S5: Effectiveness of individual-based versus group-based incentives in producing sustained abstinence at 14 days, 30 days, 6 months and 12 months*

	N Sustained abstinence/ N Randomized	% Sustained abstinence	95% CI	P-value group versus individual	P-value versus usual care
14 days				0.13	
Individual-based incentives	202/1080	18.6%	(16.3-20.9)		0.034
Group-based incentives	209/990	21.2%	(18.7-23.8)		0.001
30 days				0.12	
Individual-based incentives	177/1080	16.3%	(14.1-18.5)		0.003
Group-based incentives	186/990	18.9%	(16.5-21.3)		< 0.001
6 months				0.29	
Individual-based incentives	132/1080	12.1%	(10.2-14.1)		< 0.001
Group-based incentives	135/990	13.7%	(11.6-15.8)		< 0.001
12 months				0.043	
Individual-based incentives	58/1080	5.3%	(4.0-6.7)		0.11
Group-based incentives	74/990	7.5%	(5.9-9.2)		0.003
12 months (based on self-report)				0.055	
Individual-based incentives	102/1080	9.4%	(7.6-11.1)		0.002
Group-based incentives	118/990	12.0%	(10.0-14.0)		< 0.001

*Abstinence rates are adjusted for two stratifying variables: whether or not participants received their health insurance through CVS/Caremark, and whether their annual household income was $60,000 or greater.

Table S6: Proportions of participants achieving sustained abstinence through 6 months by intervention arm and randomization stratum

	Usual Care *n=468*	**Individual Reward** *n=498*	**Collaborative Reward** *n=519*	**Individual Deposit** *n*=582	**Competitive Deposit** *n=471*
Randomization stratum – no. (%)					
High income (≥$60,000) without benefits	5/70 (7.1)	13/69 (18.8)	14/82 (17.1)	11/96 (11.5)	10/51 (19.6)
High income with benefits	2/52 (3.8)	10/54 (18.5)	14/55 (25.5)	12/83 (14.5)	11/62 (17.7)
Low income without benefits	17/210 (8.1)	29/222 (13.1)	31/234 (13.2)	17/244 (7.0)	17/216 (7.9)
Low income with benefits	4/136 (2.9)	24/153 (15.7)	24/148 (16.2)	16/159 (10.1)	14/142 (9.9)
6 month sustained abstinence total*	28/468 (6.0)	76/498 (15.4)	83/519 (16.0)	56/582 (9.4)	52/471 (11.1)

*The overall 6 month sustained abstinence rates are adjusted for two stratifying variables: whether or not participants received their health insurance through CVS/Caremark, and whether their annual household income was $60,000 or greater.

Table S7: Efficacy of reward-based and deposit-based incentives conditional on acceptance of the assigned intervention (i.e., per-protocol analysis)*

	N Sustained abstinence/ N Randomized	% Sustained abstinence	95% CI	P-value reward versus deposit	P-value versus usual care
14 days				< 0.001	
Reward-based incentives	245/914	26.9%	(24.1-29.8)		< 0.001
Deposit-based incentives	108/146	72.4%	(64.9-79.9)		< 0.001
30 days				< 0.001	
Reward-based incentives	212/914	23.3%	(20.5-26.0)		< 0.001
Deposit-based incentives	102/146	68.2%	(60.4-76.0)		< 0.001
6 months				< 0.001	
Reward-based incentives	156/914	17.1%	(14.7-19.6)		< 0.001
Deposit-based incentives	78/146	52.3%	(44.0-60.6)		< 0.001
12 months				0.002	
Reward-based incentives	80/914	8.8%	(7.0-10.6)		< 0.001
Deposit-based incentives	27/146	17.2%	(11.1-23.3)		< 0.001
12 months (based on self-report)				< 0.001	
Reward-based incentives	132/914	14.5%	(12.2-16.8)		< 0.001
Deposit-based incentives	56/146	36.5%	(28.6-44.4)		< 0.001

*Abstinence rates are adjusted for two stratifying variables: whether or not participants received their health insurance through CVS/Caremark, and whether their annual household income was $60,000 or greater.

Table S8: Results of screening for smoking status at trial enrollment across arms

Arm	Selected for screen		Noncompliant		Negative*		Positive*		Negative and Noncompliant	
	N	%	N	%	N	%	N	%	N	%
Usual Care	47	31%	10	21%	2	4%	35	74%	12	26%
Individual Reward	48	32%	4	8%	4	8%	40	83%	8	17%
Collaborative Reward	47	31%	6	13%	3	6%	38	81%	9	19%
Individual Deposit	4	3%	1	25%	0	0%	3	75%	1	25%
Competitive Deposit	4	3%	0	0%	0	0%	4	100%	0	0%

*Negative and positive refer to the results of submitted cotinine assays, such that negative tests indicated likely non-smokers, and positive tests confirmed active smoking status. Fewer participants were selected for screening in the deposit contract arms because we aimed to screen a 5% random sample of participants who accepted their assigned intervention, and fewer participants accepted deposit contracts.

Table S9: Sensitivity analysis of sustained abstinence at 6 months to account for potential enrollment of non-smokers

Scenario	Observed			Assuming 6% of successful participants were ineligible			Assuming 20% of successful participants were ineligible		
	N Sustained abstinence/ N Randomized	% Sustained abstinence (95% CI)	P-value vs. UC arm	N Sustained abstinence/ N Randomized	% Sustained abstinence	P-value vs. UC arm	N Sustained abstinence/ N Randomized	% Sustained abstinence	P-value vs. UC arm
Arm									
Usual Care	28/468	6.0% (3.9-8.2)	--	26 / 466	5.6%	--	22 / 462	4.8%	--
Individual Rewards	76/498	15.4% (12.2-18.5)	< 0.001	71 / 493	14.4%	< 0.001	61 / 483	12.6%	< 0.001
Collaborative Rewards	83/519	16.0% (12.9-19.2)	< 0.001	78 / 514	15.2%	< 0.001	66 / 502	13.1%	< 0.001
Individual Deposits	56/582	9.4% (7.1-11.8)	0.041	53 / 579	9.2%	0.029	45 / 571	7.9%	0.043
Competitive Deposits	52/471	11.1% (8.3-14.0)	0.006	49 / 468	10.5%	0.006	42 / 461	9.1%	0.009

We screened a random sample of 5% of enrolled participants to ensure that they were truly smokers at the time of enrollment (n = 150). These participants were offered $100 to complete a cotinine sample to confirm smoking status. Of the 150 participants eligible for screening, 129 (86%) returned samples. Of the samples returned, 9 (7% of the 129 who returned samples, and 6% of the 150 who were selected for screening) were negative for cotinine, suggesting that these participants were non-smokers (although these could also represent false-negative assays). If we assume that all 9 participants who returned samples negative for cotinine were truly non-smokers, then the proportion of participants in the trial who enrolled despite being non-smokers (and were hence ineligible) would be 6%. Assuming that all participants in this 6% would have been classified as Assuming that all participants in this 6% would have been classified as having successfully quit (i.e., they did not newly begin smoking during the incentives trial), we deduct 6% from the numerator of each quit rate in the first sensitivity analysis, and deduct the same number of participants from the denominator (because these individuals would have been ineligible). The second sensitivity analysis uses the

same approach, but uses the more conservative assumption that both the 9 (6%) with negative assays, and the 21 (14%) who did not return samples (total = 30/150, 20%) were all non-smokers. All comparisons of the interventions' quit rates to usual care were performed using the intention-to-treat approach.

Table S10: Self-reported abstinence and verified abstinence rates at 14 days, 30 days, 6 months and 12 months across participants as assigned (ITT)

	Usual Care N = 468	Individual Reward N = 498	Collaborative Reward N = 519	Individual Deposit N = 582	Competitive Deposit N = 471	All arms N = 2538
Self-report Quit on 14-day	112	162	188	115	103	680
% Total	24%	33%	36%	20%	22%	27%
Tested Negative on 14-day	65	116	134	86	72	473
% Total	*14%*	*23%*	*26%*	*15%*	*15%*	*19%*
Self-report Quit on 30-day	59	109	126	85	73	452
% Total	13%	22%	24%	15%	15%	18%
Tested Negative on 30-day	48	100	117	77	68	410
% Total	*10%*	*20%*	*23%*	*13%*	*14%*	*16%*
Self-report Quit on 6-month	33	90	100	66	61	350
% Total	7%	18%	19%	11%	13%	14%
Tested Negative on 6-month	28	76	83	56	52	295
% Total	*6%*	*15%*	*16%*	*10%*	*11%*	*12%*
Self-report Quit on 12-month	22	57	78	45	40	242
% Total	5%	11%	15%	8%	8%	10%
Tested Negative on 12-month	16	37	45	21	29	148
% Total	*3%*	*7%*	*9%*	*4%*	*6%*	*6%*