ORIGINAL REPORT

A randomized trial of lottery-based incentives and reminders to improve warfarin adherence: the Warfarin Incentives (WIN2) Trial

Stephen E. Kimmel[1,2,3,4]*, Andrea B. Troxel[2,4], Benjamin French[2,4], George Loewenstein[4,5], Jalpa A. Doshi[3,4], Todd E. H. Hecht[3], Mitchell Laskin[6], Colleen M. Brensinger[2], Chris Meussner[2] and Kevin Volpp[3,4,7,8,9]

[1]Center for Therapeutic Effectiveness Research, University of Pennsylvania, Philadelphia, PA, USA
[2]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[4]Center for Health Incentives and Behavioral Economics, Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA
[5]Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA
[6]Department of Pharmacy Service, Hospital of the University of Pennsylvania, Philadelphia, PA, USA
[7]Center for Health Equity Research and Promotion, Philadelphia Veterans Affairs Medical Center, Philadelphia, PA, USA
[8]Department of Health Care Management, University of Pennsylvania, Philadelphia, PA, USA
[9]Department of Medical Ethics and Health Policy, University of Pennsylvania, Philadelphia, PA, USA

ABSTRACT

Background Previous research has suggested that daily lottery incentives could improve medication adherence. Such daily incentives include implicit reminders. However, the comparative effectiveness of reminders alone versus daily incentives has not been tested.
Methods A total of 270 patients on warfarin were enrolled in a four-arm, multi-center, randomized controlled trial comparing a daily lottery-based incentive, a daily reminder, and a combination of the two against a control group (usual care).
Results Participants in the reminder group had the lowest percentage of time out of target international normalized ratio (INR) range, the primary outcome, with an adjusted odds of an out-of-range INR 36% lower than among those in the control group, 95%CI [7%, 55%]. No other group had a statistically significant improvement in anticoagulation control relative to the control group or to each other. The only group that had significant improvement in incorrect adherence was the lottery group (incorrect adherence: 12.1% compared with 23.7% in the control group, difference of −7.4% 95%CI [−14%, −0.3%]). However, there was no relationship between changes in adherence and anticoagulation control in the lottery group.
Conclusions Automated reminders led to the largest improvements in anticoagulation control, although without impacting measured adherence. Lottery-based reminders improved measured adherence but did not lead to improved anticoagulation control. Copyright © 2016 John Wiley & Sons, Ltd.

INTRODUCTION

The potential benefits of many advances in health care are limited by high rates of non-adherence to medications.[1] Medication non-adherence is of even greater concern for medications with a narrow therapeutic range, because missed doses can rapidly reduce these drugs' effectiveness, while extra doses increase the risk of side effects. Warfarin is an ideal drug for studying the effectiveness of new methods of enhancing adherence. Partly because of its narrow therapeutic range, and partly because of the lack of symptoms associated with the conditions it treats, adherence to warfarin therapy, and anticoagulation control, is generally poor.[2,3] Low rates of adherence not only have direct effects in terms of reduced effectiveness

*Correspondence to: S. E. Kimmel, Professor of Medicine and Epidemiology, University of Pennsylvania School of Medicine, 923 Blockley Hall, 423 Guardian Drive, Philadelphia, PA, USA. E-mail: stevek@mail.med.upenn.edu

and risks associated with poor anticoagulation control but also dissuade many physicians from prescribing warfarin or other anticoagulants to patients who could potentially benefit from them.[2,4]

Previous work has demonstrated the effectiveness of daily lottery-based incentives in increasing weight loss[5] and has suggested that lottery-based incentives might improve medication adherence and INR control.[6,7] Such lottery-type incentives include an implicit reminder, but the comparative effectiveness of daily lottery-based incentives, daily reminders, and combined daily lottery incentives and reminders has never been compared. The Warfarin Incentives (WIN2) Trial was designed to compare the effectiveness of lottery-based approaches, reminders alone, and a combination of incentives and reminders on anticoagulation control.

METHODS

Study design

This study was a multi-center, randomized controlled trial conducted at the Hospital of the University of Pennsylvania (HUP) and the Philadelphia Veterans Affairs Medical Center (PVAMC) from November 2009 to May 2012. Potential participants were recruited from HUP and PVAMC anticoagulation clinics. The Institutional Review Boards of both sites approved the study, and all participants provided written informed consent. The study was registered at clinicaltrials.gov as Randomized Trial of Interventions to Improve Warfarin Adherence, ID # NCT00904982. An independent Data Safety Monitoring Board monitored the trial. The sponsor of the study, the National Heart Lung Blood Institute, had no role in the design of the study, execution of the study, or analysis, interpretation, and writing of the manuscript.

Study population

The study population included all patients who were in the maintenance phase of warfarin therapy, defined as stable target INR over two consecutive visits at least 7 days apart. Eligible participants had a working analog telephone line, an expected duration of therapy of at least 6 months, a target INR range within 2.0–3.5, and at least 1 INR out of target range within 90 days prior to enrollment or an INR at enrollment that was below target range. These last two inclusion criteria were used because prior work suggested that these patients were most likely to be non-adherent.[6,7] Exclusion criteria were as follows:

no access to a telephone line (which was required to use the Med-eMonitor, described in the next section); unwillingness to participate or sign a consent form; dementia or any other impairment affecting ability to provide informed consent and/or utilize the Med-eMonitor; enrollment in a different clinical trial of warfarin; illness with anticipated life expectancy of 6 months or less; or INR over the upper limit for the individual's range at the time of enrollment (to avoid possibly exacerbating this over-anticoagulation if a patient's adherence improved during the study).

Randomization and interventions

Eligible participants were randomized using a random number generator and via permuted block randomization with a block size of 4. Randomization was stratified by site (HUP or PVAMC) and by INR status at enrollment (in or below target range). The latter was carried out because a prior study suggested that those below target range might be most likely to benefit from a lottery intervention.[6,7] Neither study staff nor study participants could be blinded because of the nature of the interventions. However, study coordinators were blinded to adherence data, and study investigators and data analysts remained blinded to intervention assignment until all data collection and analyses were completed.

All study participants were given an electronic medication monitoring system (a Med-eMonitor™) to use at home. The electronic medication monitoring system measured the participant's adherence to his/her warfarin regimen throughout the 6-month duration of the study. The electronic medication monitoring system had drawers into which the participant's medication was placed (one drawer was used for those on a single daily dose and two drawers were used for patients taking two different doses on different days in a given week). When a drawer was opened, a message displayed on the monitor and asked the participant if he/she was taking his/her medication for the day. The device registered the answer and sent the information via the participant's telephone line nightly to a secured central server. The device also displayed automated messages of encouragement to participants and provided automated education on the importance of taking their warfarin. The study groups were as follows:

- Reminder group: Participants were given an electronic medication monitoring system with a daily alarm to remind them to take their medication as scheduled.

- Lottery group: Participants were entered into a daily lottery with an expected daily value of about $3. On each study day, participants had a one in five chance of winning $10 and one in 100 chance of winning $100 when taking their warfarin as prescribed. Notification of any lottery winnings, and the amount of those winnings, were sent via the telephone connection to the electronic medication monitoring system overnight so that participants could see their winnings on the electronic medication monitoring system screen the next morning. Participants who did not open a drawer to take their warfarin as directed on a given day were notified if they would have won (if their lottery number was drawn) and how much they would have won had they taken their medication. The system was automated so that no personnel were required to run the lotteries. If participants were told to not take warfarin on a particular day, they would be ineligible to win the lottery if they recorded pill taking for that day. Reminder alarms were disabled for participants in this group. Payments were sent to patients on a monthly basis by money order.
- Lottery + reminder group: Participants were given an electronic medication monitoring system that reminded them to take their medication as scheduled with a daily alarm (as in the alarm group) and were entered into a daily lottery in which they could win money when taking their warfarin as prescribed (as in the lottery group).
- Control group: Participants were given an electronic medication monitoring system, but the alarm feature was not activated and they were not entered into the daily lottery.

Study procedures

Data were collected by structured interviews performed by research coordinators using standardized data collection forms. All participants completed the in-person baseline interview at their anticoagulation clinic. All participants received regular ongoing follow-up from their clinical provider in the anticoagulation clinic as per their regular schedule. At these clinics, routine follow-up is performed monthly for all patients on maintenance dose. The research coordinator entered the date and INR measurement for all of these visits. Study staff contacted participants at 3 months after enrollment to conduct a follow-up interview by telephone, unless the participant specifically requested to have the interview in-person. A final, follow-up in-person interview was conducted at 6 months immediately following a participant's regular clinic visit, or, for participants who were managed by phone, a separate study interview was scheduled at 6 months. The purpose of these study interviews was to collect follow-up data, and there were no other interventions during these visits. The study coordinators who conducted these interviews were unaware of the adherence data collected during the trial.

Outcomes

The primary outcome was a binary indicator for an out-of-range INR (either below or above target range), recorded as a repeated measures at each visit for all participants. The secondary outcomes were the percentage of days that adherence was incorrect (i.e. missed a dose or took an extra dose), calculated each month for all participants; patient-reported bleeding requiring hospitalization or emergency room visit, stroke, transient ischemic attack, or non-CNS thrombosis requiring hospitalization or emergency room visit.

Statistical analyses

Logistic regression models were used to compare the odds of an out-of-range INR across study groups using intention-to-treat analyses. A robust variance estimator was used to account for longitudinal correlation arising from multiple INRs collected on a participant over time.[8] Probability weights equal to the inverse of the total number of INRs for each participant were used to account for heterogeneity in the number of INRs per participant.[9] Linear regression models with a robust variance estimator were used to compare the average percentage of days with incorrect adherence across study groups. The percentage of days with incorrect adherence was positively skewed; a square-root transformation was applied to normalize the distribution. For the primary and secondary outcome, a series of staged models was fit: minimally adjusted, which included variables that stratified randomization (site and INR at enrollment) and day (or month) since randomization; and fully adjusted, which additionally included any covariate (listed in Table 1) found to be imbalanced across the study groups ($p < 0.2$). A priori subgroup analyses were performed among subgroups defined by INR at enrollment by including interaction terms between INR at enrollment (in range or below range) and treatment group in the fully adjusted models. All analyses were completed using SAS 9.3 (SAS Institute, Cary, NC, USA).

Table 1. Characteristics of study participants at randomization

	Control (n = 68)	Lottery (n = 67)	Reminder (n = 67)	Lottery + reminder (n = 68)
Demographic characteristics				
Age, median (IQR), years	64.0 (60.0–71.0)	62.0 (51.0–68.0)	62.0 (55.0–67.0)	61.5 (51.5–68.5)
Female sex, n (%)	23 (34)	21 (31)	17 (25)	25 (37)
African American race, n (%)	46 (68)	54 (81)	44 (67)	52 (78)
Completed high school, n (%)	45 (66)	29 (43)	33 (49)	29 (43)
Employment status, n (%)				
Working	8 (12)	6 (9)	8 (12)	10 (15)
Unemployed	3 (4)	9 (13)	5 (8)	2 (3)
Retired	30 (44)	33 (49)	27 (40)	23 (34)
Disabled	27 (40)	19 (28)	27 (40)	33 (49)
Household income, n (%)				
< 100% of federal poverty level	14 (22)	20 (31)	18 (29)	25 (37)
100–200% of federal poverty level	25 (39)	20 (31)	19 (30)	22 (33)
200–300% of federal poverty level	9 (14)	14 (22)	14 (22)	8 (12)
>300% of federal poverty level	16 (25)	11 (17)	12 (19)	12 (18)
Insurance status*				
Medicaid, n (%)	10 (15)	8 (12)	9 (14)	14 (21)
Medicare, n (%)	40 (59)	33 (51)	27 (42)	27 (40)
Private, n (%)	28 (41)	22 (34)	26 (41)	26 (38)
VA, n (%)	30 (44)	28 (43)	22 (34)	23 (34)
Other, n (%)	0 (0)	3 (5)	1 (2)	0 (0)
None, n (%)	0 (0)	1 (1)	2 (3)	2 (3)
Marital status, n (%)				
Currently married	26 (39)	28 (42)	25 (38)	18 (27)
Separated, divorced, or widowed	27 (40)	27 (40)	30 (46)	34 (51)
Never married	14 (21)	12 (18)	11 (16)	15 (22)
Warfarin therapy				
Indication for warfarin therapy*				
Atrial fibrillation or flutter	37 (56)	31 (47)	33 (50)	30 (45)
Deep vein thrombosis or pulmonary embolism	19 (29)	27 (41)	20 (30)	21 (31)
Mechanical heart valve	9 (14)	0 (0)	7 (11)	5 (8)
Other	20 (30)	21 (32)	21 (32)	24 (36)
Prior warfarin use	13 (19)	16 (24)	14 (21)	16 (24)
Target INR range, n (%)				
2.0–3.0	60 (88)	65 (97)	59 (88)	65 (96)
2.5–3.5	8 (12)	2 (3)	8 (12)	3 (4)
INR at enrollment, n (%)				
Below range	13 (19)	14 (21)	15 (22)	14 (21)
In range	55 (81)	53 (79)	52 (78)	54 (79)
DASS score, median (IQR)	50.0 (42.5–66.0)	50.0 (39.0–65.0)	48.0 (40.0–63.0)	52.0 (44.5–68.4)
Medical history				
Body mass index, median (IQR), kg/m^2	29.4 (25.8–34.8)	29.2 (24.6–32.3)	29.3 (26.7–35.0)	30.6 (26.7–34.7)
Smoking status, n (%)				
Current	7 (10)	15 (22)	15 (22)	15 (22)
Former	36 (53)	28 (42)	27 (40)	32 (47)
Never	25 (37)	24 (36)	25 (37)	21 (31)
History of congestive heart failure, n (%)	26 (38)	23 (34)	28 (42)	19 (28)
History of diabetes mellitus, n (%)	20 (29)	18 (27)	27 (40)	31 (46)
History of hypertension, n (%)	56 (82)	47 (70)	51 (76)	52 (77)
History of myocardial infarction, n (%)	14 (21)	13 (20)	13 (20)	13 (19)
History of stroke, n (%)	18 (27)	14 (21)	13 (19)	9 (13)
General health status, n (%)				
Excellent	1 (1)	3 (4)	1 (2)	2 (3)
Very good	8 (12)	5 (7)	10 (15)	8 (12)
Good	21 (31)	28 (42)	25 (37)	22 (32)
Fair	30 (44)	26 (39)	27 (40)	34 (50)
Poor	8 (12)	5 (7)	4 (6)	2 (3)
Short form (SF-36) health survey				
Physical score, median (IQR)	36.7 (29.0–43.7)	36.6 (28.6–45.9)	39.4 (33.6–47.6)	36.7 (29.1–43.8)
Mental score, median (IQR)	51.6 (41.0–59.0)	53.2 (42.6–59.3)	51.3 (43.4–57.3)	52.3 (46.8–59.1)
CCSE score, median (IQR)	27.0 (23.5–29.0)	26.0 (23.0–29.0)	27.0 (24.0–29.0)	27.0 (24.0–29.0)

IQR, interquartile range (25th to 75th percentile); CCSE, Cognitive Capacity Screening Examination; DASS, Duke Anticoagulation Satisfaction Survey.
*Participants could report more than one source of health insurance and more than one indication for warfarin therapy.

Sample size

The sample size was chosen to ensure that clinically meaningful differences in the odds of an out-of-range INR could be detected in any of the study contrasts, including comparison of each group to the control group, as well as comparison of the lottery + reminder group to the lottery and the reminder groups. Unlike in a typical factorial design that relies upon the absence of an interaction between the interventions to have adequate power to test each main effect, we designed this trial to have adequate power both for tests of the main effects and assessment of interactions. A total sample size of 268 participants (67 in each group) was required to detect at least a 30% relative difference in occurrence of an out-of-range INR, assuming a power of 80%, an intraclass correlation coefficient of 0.05, seven INR measurements per person over the 6-month period of the study, and an overall type-1 error rate of 0.01 (to account for the five possible comparisons among study groups). Based on prior studies,[2] an improvement in adherence from around 22% incorrect pills taken to around 14% would correspond to a 30% relative improvement in anticoagulation control.

RESULTS

Study population

Of the 1756 patients screened, 270 were eligible and randomized, 68 to the control group, 67 to the lottery group, 67 to the reminder group, and 68 to the lottery + reminder group (Figure 1). Two participants randomized to the control group withdrew prior to follow-up. Overall, the median age was 62 years, 32% of participants were female, and 30% had household income below the federal poverty level. Approximately 21% of participants had an INR below target range at enrollment. There was some indication that the groups differed with respect to several variables: age, race, education, employment status, insurance, target INR range, mechanical heart valve indication for warfarin, and history of diabetes mellitus differed across the study groups ($p < 0.2$) (Table 1). The median number of follow-up INRs was 7, with no difference among the groups ($p = 0.90$).

Risk of out-of-range INR

In the fully adjusted model (Table 2), the odds of an out-of-range INR was 36% lower among participants in the reminder group than among those in the control
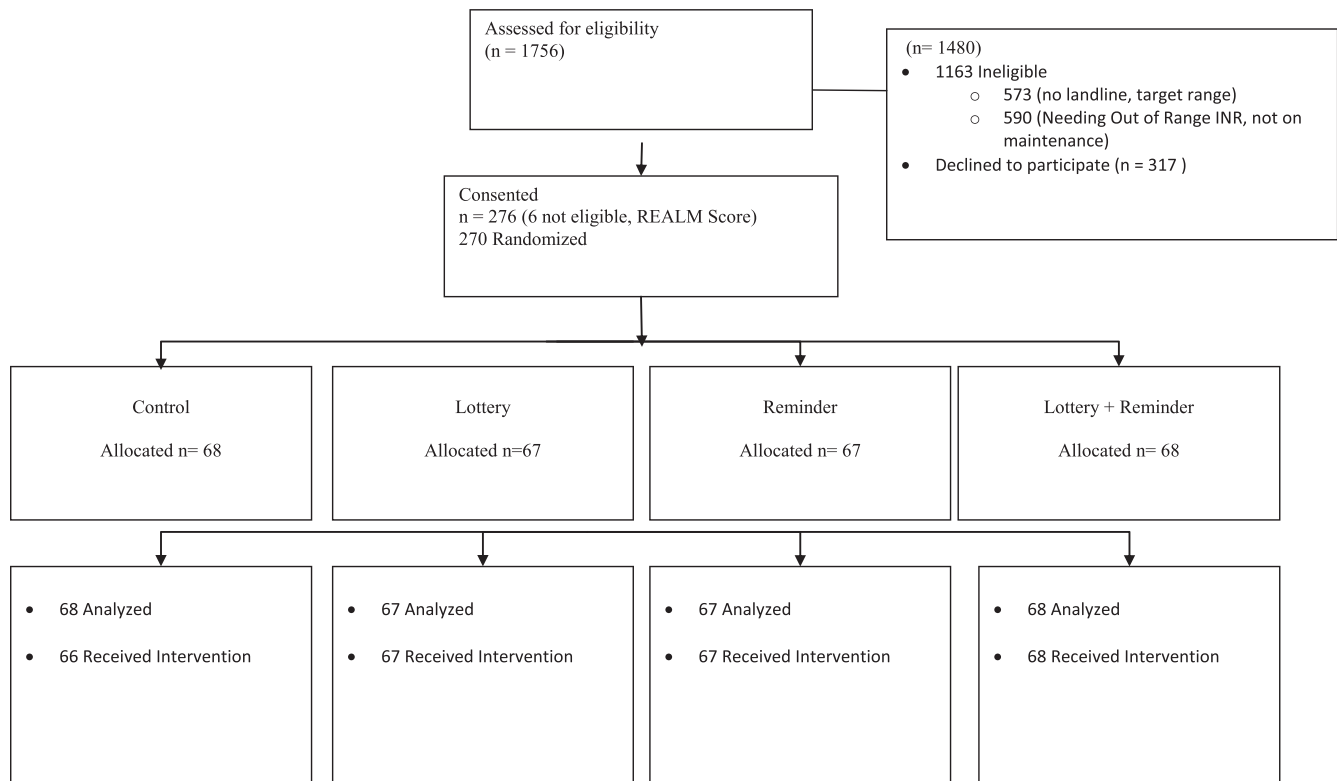

Figure 1. CONSORT diagram

Table 2. Results for the primary outcome of out-of-range INR, among all participants, and stratified by INR at enrollment

	n	% time out of target INR range* Median (IQR)	Minimally adjusted model[†] OR (95%CI)[§]	*p*-value**	Fully adjusted model[‡] OR (95%CI)[§]	*p*-value**
All participants				0.29		0.06
Control	66	31.6 (11.1–50.5)	Referent		Referent	
Lottery	64	30.1 (12.4–46.3)	0.96 (0.69–1.34)		0.98 (0.70–1.38)	
Reminder	64	23.8 (8.8–36.6)	0.71 (0.50–1.03)		0.64 (0.45–0.93)	
Lottery + reminder	67	23.9 (9.9–42.7)	0.86 (0.62–1.20)		0.77 (0.54–1.09)	
Participants with an in-range INR at enrollment[††]				0.28		0.09
Control	53	30.7 (10.2–47.3)	Referent		Referent	
Lottery	50	22.5 (9.0–43.3)	0.89 (0.62–1.30)		0.90 (0.61–1.34)	
Reminder	50	21.1 (3.8–34.1)	0.66 (0.43–1.00)		0.59 (0.39–0.90)	
Lottery + reminder	53	20.5 (9.9–39.5)	0.85 (0.59–1.23)		0.76 (0.51–1.13)	
Participants with a below-range INR at enrollment[††]				0.84		0.48
Control	13	32.6 (21.9–53.7)	Referent		Referent	
Lottery	14	43.6 (30.1–54.3)	1.25 (0.63–2.51)		1.33 (0.71–2.51)	
Reminder	14	39.5 (13.8–49.4)	0.94 (0.46–1.93)		0.86 (0.42–1.78)	
Lottery + reminder	14	38.6 (11.6–58.5)	0.93 (0.46–1.90)		0.82 (0.40–1.69)	

IQR, interquartile range (25th to 75th percentile); OR, odds ratio; CI, confidence interval.
*Percentage of time out of target INR range, calculated using linear interpolation between successive INR values.
[†]Adjusted for site (Hospital of the University of Pennsylvania or Philadelphia Veterans Affairs Medical Center), INR at enrollment (in range or below range), and day since randomization. Mechanical heart valve could not be included in the model because there were no participants in the lottery group with this indication.
[‡]Adjusted for site (Hospital of the University of Pennsylvania or Philadelphia Veterans Affairs Medical Center), INR at enrollment (in range or below range), day since randomization, age in years, African American race, completed high school, employment status (working, unemployed, retired, disabled), Medicare insurance, target INR range (2.0–3.0 or 2.5–3.5), and history of diabetes mellitus.
[§]Odds ratios correspond to the estimated odds of an out-of-range INR relative to the control group.
**p-values obtained from multivariable Wald test of the null hypothesis that the lottery, reminder, and lottery + reminder odds ratios are all equal to 1.
[††]There was no evidence that odds ratios differed by INR at enrollment (interaction *p* = 0.68).

group, 95%CI [7%, 55%]. Participants in the lottery + reminder group had an adjusted 23% decreased odds of an out-of-range INR compared with participants in the control group (95%CI [−9%, 46%]), but the difference was not statistically significant. There was no difference in the odds of an out-of-range INR between the lottery and control groups. This lack of an effect was not explained by an increase in over-anticoagulation (OR 0.87, 95%CI 0.53–1.42) because of improved adherence in the lottery arm. There was no significant difference between the lottery + reminder and the reminder groups (adjusted OR 1.19, 95%CI [0.81–1.76]; *p* = 0.37). Although there were some differences in the primary outcome by the presence or absence of a below-range INR at baseline, the test for interaction did not demonstrate a statistically significant difference between participants with an in-range and below-range INR at enrollment (interaction *p* = 0.68).

Adherence

In the overall cohort, participants in the lottery group had the lowest percentage of days with incorrect adherence (Table 3). In the fully adjusted model, the absolute difference in percentage of days with

incorrect adherence was 7.4% lower in the lottery than the control group, 95%CI [−14%, −0.3%]. Participants in the lottery + reminder group also had a lower percentage of days with incorrect adherence compared with the control group, but the difference between these two groups was not statistically significant. There was no significant difference in the percentage of days with incorrect adherence between the reminder and control groups (−2.0%, 95%CI [−8.2, 4.2]), and there was no significant difference between the lottery + reminder and lottery groups (2.8%, 95% CI [−2.0, 7.5]). Although there were some differences in adherence across groups by the presence or absence of a below-range INR at baseline, the test for interaction did not demonstrate a statistically significant difference between participants with an in-range and below-range INR at enrollment (interaction *p* = 0.95).

There was no significant difference in bleeding events across the four study groups (6.2% in control group, 6.2% in lottery group, 6.3% in reminder group, and 7.6% in the lottery + reminder group; *p* = 1.00). There was only one stroke reported in the lottery group (*p* = 0.76) and two non-CNS thrombotic events reported in the lottery group (*p* = 0.18).

Table 3. Results for the secondary outcome of percentage of days that adherence was incorrect, among all participants, and stratified by INR at enrollment

		% incorrect adherence*	Minimally adjusted model[†]		Fully adjusted model[‡]	
	n	Median (IQR)	Difference (95%CI)[§]	p-value**	Difference (95%CI)[§]	p-value **
All participants				0.03		0.008
Control	57	23.7 (8.1–40.5)	Referent		Referent	
Lottery	65	12.1 (6.6–25.0)	−7.9 (−14.1, −1.7)		−7.4 (−14.4, −0.3)	
Reminder	63	21.8 (6.9–39.5)	0.5 (−7.5, 8.5)		−2.0 (−8.2, 4.2)	
Lottery + reminder	66	17.6 (7.0–43.6)	−3.1 (−10.2, 4.1)		−4.6 (−11.1, 1.9)	
Participants with an in-range INR at enrollment[††]				0.03		0.01
Control	46	25.0 (8.1–40.3)	Referent		Referent	
Lottery	51	12.1 (5.9–25.0)	−8.1 (−14.6, −1.6)		−7.5 (−14.7, −0.3)	
Reminder	49	23.5 (9.5–39.5)	1.2 (−7.6, 10.0)		−1.4 (−7.9, 5.1)	
Lottery + Reminder	53	17.1 (7.0–32.6)	−3.9 (−11.5, 3.7)		−5.0 (−11.7, 1.8)	
Participants with a below-range INR at enrollment[††]				0.75		0.73
Control	11	16.9 (7.7–42.3)	Referent		Referent	
Lottery	14	14.5 (8.5–26.5)	−7.7 (−25.9, 10.5)		−7.9 (−24.5, 8.8)	
Reminder	14	20.9 (6.1–33.5)	−2.2 (−22.4, 18.1)		−4.6 (−22.5, 13.3)	
Lottery + reminder	13	34.2 (10.5–44.5)	0.5 (−19.9, 21.0)		−3.6 (−21.1, 14.0)	

IQR, interquartile range (25th to 75th percentile); CI, confidence interval.

*Percentage of days that adherence was incorrect, calculated over the entire study period for each participant.

[†]Adjusted for site (Hospital of the University of Pennsylvania or Philadelphia Veterans Affairs Medical Center), INR at enrollment (in range or below range), and month since randomization. Model coefficients transformed back from the square root term so that the reported differences are model based estimates of the mean % incorrect adherence.

[‡]Adjusted for site (Hospital of the University of Pennsylvania or Philadelphia Veterans Affairs Medical Center), INR at enrollment (in range or below range), month since randomization, age in years, African American race, completed high school, employment status (working, unemployed, retired, disabled), Medicare insurance, target INR range (2.0–3.0 or 2.5–3.5), and history of diabetes mellitus. Model coefficients transformed back from the square root term so that the reported differences are model based estimates of the mean % incorrect adherence.

[§]Difference corresponds to the estimated difference in average % incorrect adherence relative to the control group.

**p-values obtained from multivariable Wald test of the null hypothesis that the lottery, reminder, and lottery + reminder differences are all equal to 0.

[††]There was no evidence that odds ratios differed by INR at enrollment (interaction $p = 0.95$).

DISCUSSION

The WIN2 trial demonstrated that a reminder system led to improvement in anticoagulation control but that neither a lottery-based incentive nor a combined lottery-based incentive and reminder system significantly improved anticoagulation control. However, the lottery-based incentive did lead to improvement in measured medication adherence, whereas the reminder group did not. Based on the relative improvement in adherence measured in the lottery group, one might have anticipated a 30% relative improvement in out-of-range INRs,[2] but this was not observed. The trial results highlight the importance of studying a clinically relevant outcome, in this case anticoagulation control, rather than just medication adherence. Further, although a prior study of a lottery-only intervention (WIN trial) suggested that there might be benefit on anticoagulation control and adherence in the subgroup of patients with a low INR at baseline,[2,6] these findings were not confirmed by WIN2.

The lottery-based incentive tested in this trial was based on several important theoretical constructs: frequent (daily) positive reinforcement,[10,12,17,18] past rewards, and the prospect of future rewards[13]; incentives that leverage anticipated regret about not winning an award one could have easily won[11]; and easy scalability. In addition, there is some empirical evidence to support lottery-based incentives on changing health behaviors.[5,14,16]

Anticoagulation control was chosen instead of adherence as the primary outcome because it is a clinically relevant endpoint that clinicians use to judge anticoagulation management. Our study does not allow us to determine why the daily lottery incentive improved measured adherence but not the clinical endpoint, nor why the reminder system improved anticoagulation control but not measured adherence. Based on the study design, this is unlikely to be a case of the study being powered to detect differences in adherence but underpowered to detect significant differences in INR outcomes. One explanation may be that other factors characteristic of warfarin response (dietary adherence, drug–drug interactions) may have diminished the impact of improved adherence. Clinicians may also titrate warfarin dosing to account for a patient's level of adherence. Another explanation is that the measure of adherence relies on participants using the electronic medication monitoring system

every day. It is possible that participants in the reminder group may not have used the electronic medication monitoring system as reliably as those in the lottery arm (in whom the use of the device was directly tied to winning the lottery). This could lead to an underestimate of adherence improvement in the alarm group and an overestimate of adherence in the lottery group. Novel methods of measuring adherence (such as digital pills that emit signals when swallowed) may someday be available to improve our ability to measure adherence. It is less likely, but possible, that the improvement in anticoagulation is a chance finding or that the reminder system, although not changing medication adherence, prompted participants to be more vigilant about their diet and alcohol consumption or avoid interacting medications.

As discussed previously, one limitation of the study was that we could not directly observe pill taking. Although pill compartment openings are considered a valid way to measure adherence,[15] our study emphasizes the importance of measuring clinical outcomes in studies of adherence interventions. Another limitation was that research staff could not be blinded to the intervention. However, INR measurements were performed by clinicians in the anticoagulation clinics who were blinded to the intervention and research staff were blinded to adherence measures. The trial was performed in anticoagulation clinics so the generalizability to other practice settings is unknown.

Adherence will remain an important, and perhaps larger, concern for direct oral anticoagulants. These medications do not have an accepted method to monitor their level of anticoagulation control as a surrogate for potential non-adherence. Therefore, the importance of developing and testing novel methods to ensure proper adherence will continue to grow.

CONCLUSION

A lottery-based incentive system improved measured adherence but did not improve anticoagulation control, while an electronic reminder system improved anticoagulation control without impacting measured adherence. This trial highlights the challenge of connecting clinically meaningful outcomes with improved adherence, particularly with complex medications such as warfarin.

The study was approved by the Institutional Review Boards of the Hospital of the University of Pennsylvania and the Philadelphia Veterans Affairs Medical Center.

KEY POINTS

- Adherence to warfarin, a narrow therapeutic index medication, is suboptimal
- Improving adherence and anticoagulation control could result in substantial benefits to patients but is difficult to achieve
- Novel approaches, including economic incentives, could improve anticoagulation control by improving medication adherence
- A simple reminder system improved anticoagulation control but a novel, more costly and complex, lottery-based intervention did not
- There was a disconnect between measured adherence and the clinical outcome of anticoagulation control, highlighting the great importance of studying clinically meaningful outcomes when examining interventions targeted at improving medication adherence

REFERENCES

1. Kripalani S, Yao X, Haynes RB. Interventions to enhance medication adherence in chronic medical conditions: a systematic review. *Arch Intern Med* 2007; **167**(6): 540–550.
2. Kimmel SE, Chen Z, Price M, *et al.* The influence of patient adherence on anticoagulation control with warfarin. *Arch Intern Med* 2007; **167**: 229–235.
3. Parker CS, Chen Z, Price M, *et al.* Adherence to warfarin assessed by electronic pill caps, clinician assessment, and patient reports: results from the IN-RANGE study. *J Gen Intern Med* 2007; **22**(9): 1254–1259.
4. Go AS, Hylek EM, Borowsky LH, *et al.* Warfarin use among ambulatory patients with nonvalvular atrial fibrillation: the Anticoagulation and risk factors in atrial fibrillation (ATRIA) study. *Ann Intern Med* 1999; **131**(12): 927–934.
5. Volpp KG, John LK, Troxel AB, *et al.* Financial incentive-based approaches for weight loss: a randomized trial. *JAMA* 2008; **300**(22): 2631–2637. doi:10.1001/jama.2008.804.
6. Kimmel SE, Troxel AB, Loewenstein G, *et al.* Randomized trial of lottery-based incentives to improve warfarin adherence. *Am Heart J* 2012; **164**(2): 268–274. doi:10.1016/j.ahj.2012.05.005.
7. Volpp KG, Loewenstein G, Troxel AB, *et al.* A test of financial incentives to improve warfarin adherence. *BMC Health Serv Res* 2008; **8**: 272–277. doi:10.1186/1472-6963-8-272.
8. Liang KY, Zeger SL. Longitudinal data analysis using general linear models. *Biometrika* 1986; **73**: 13–22.
9. Williamson JM, Datta S, Satten GA. Marginal analyses of clustered data when cluster size is informative. *Biometrics* 2003; **59**(1): 36–42.
10. Loewenstein G, Prelec D. Anomalies in intertemporal choice: evidence and an interpretation. *Q J Econ* 1992; **107**: 573–597.
11. Connolly T, Butler DU. Regret in economic and psychological theories of choice. *J Behav Decis Mak* 2006; **19**(2): 148–158.

12. Kirby K. Bidding on the future:evidence against normative discounting of delayed rewards. *J Exp Psychol* 2009; **126**: 54–70.
13. Camerer D, Ho T-H. Experience-weighted attraction learning in normal form games. *Econometrica* 1999; **67**: 837–874.
14. Haisley E, Volpp KG, Pellathy T, Loewenstein G. The impact of alternative incentive schemes on completion of health risk assessments. *Am J Health Promot* 2012; **26**(3): 184–188. doi:10.4278/ajhp.100729-ARB-257.
15. Fallab-Stubi CL, Zellweger JP, Sauty A, Uldry C, Iorillo D, Burnier M. Electronic monitoring of adherence to treatment in the preventive chemotherapy of tuberculosis. *Int J Tuberc Lung Dis* 1998; **2**(7): 525–530.
16. Haff N, Patel MS, Lim R, *et al.* The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. *Am J Health Promot* 2015; **29**(5): 314–323. doi:10.4278/ajhp.140714-LIT-333.
17. Thaler RH. Some empirical evidence on time inconsistency. *Rev Econ Stud* 1981; **23**: 165–180.
18. Ainslie G. Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychol Bull* 1975; **82**(4): 463–496.