Supplementary Material: Habit Formation Trial

Table of Contents

1 Study design

1.1 Eligibility: self-reported medication adherence

The target population included subjects who had imperfect medication adherence. For the Penn Medicine sample, self-reported imperfect adherence was based on the response to the Morisky Medication Adherence Scale (MMAS-8), an eight-item validated scale including questions about missing medications due to forgetfulness or other reasons, purposefully stopping or reducing medications without consulting a doctor, perceptions about difficulties in taking medication as well as actual adherence in the recent past.

1.2 Schema of Financial Incentives

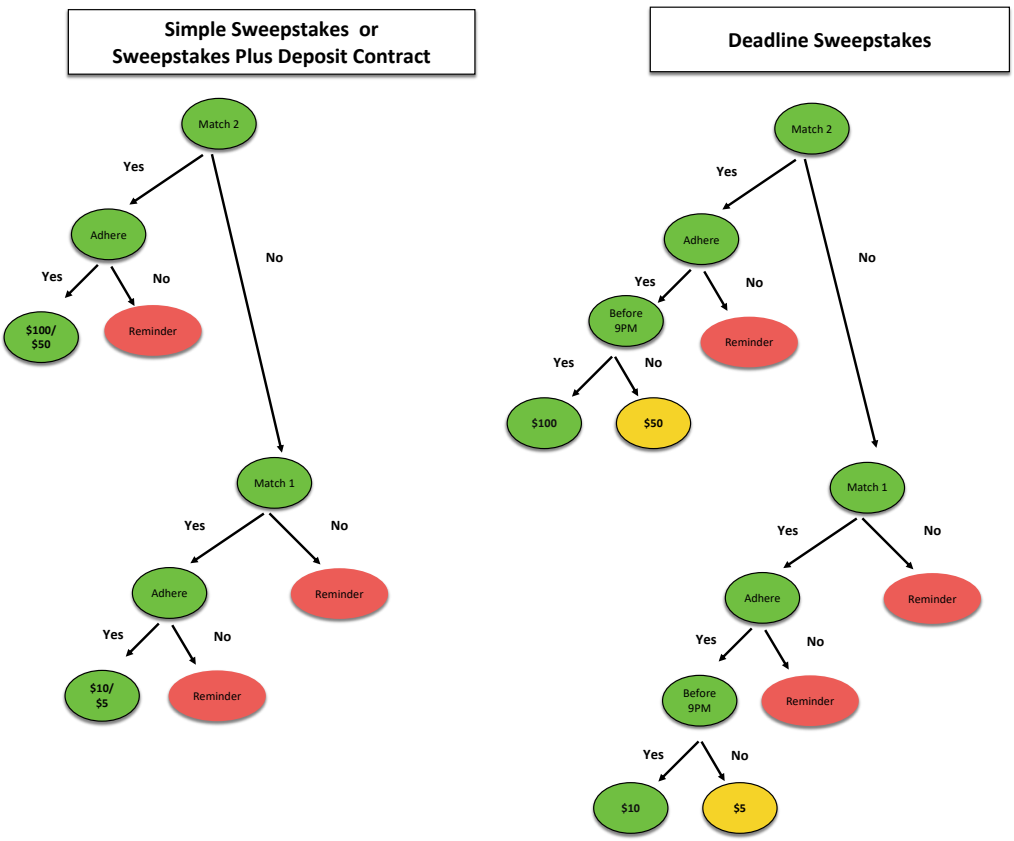

Figure 1: For the Habit Formation Trial, all arms, including control, received daily reminders. The financial incentive arms included Simple Sweepstakes, Sweepstakes Plus Deposit Contract and Deadline Sweepstakes. For Simple Sweepstakes, subjects who were adherent and who matched both digits received a $100 sweepstakes (Green circles). If a subject was adherent and matched one digit they received $10 (Green circles). Sweepstakes Plus Deposit Contract was identical except that the sweepstakes winnings were halved and half of the potential earnings were deposited into a deposit contract. In the Deadline Sweepstakes arm, subjects received the full sweepstakes payout if they were adherent before a pre-determined daily time (e.g. 9 PM) (green circles). They received half of the full sweepstakes payout if they were adherent after the pre-determine daily time (yellow circles). Subjects who had a match of their digits but who were not adherent received a notification of potential winning and a reminder.

2 Study conduct

2.1 Consent for contact and prescreening

The principal investigators negotiated with selected Employers and Insurers to obtain permission for preliminarily-eligible employees/beneficiaries to be contacted about the study by mail. Participants in the Employers and Insurer groups were initially mailed letters of invitation by their pharmacy benefits provider. The letters were prepared by the University of Pennsylvania research team, reviewed by the pharmacy benefits provider, and then mailed to the potential participants by the pharmacy benefits provider staff. When potential subjects received the letter and were interested in participation, they initiated contact with study staff via electronic communication or by phone. No direct contact was made between by the research team and the participant prior to the participant receiving a letter and initiating further contact. This method was agreed to by the Employers and Insurers and approved by the University of Pennsylvania IRB.

For Penn Medicine participants, consent was not required for pre-screening for eligibility and outreach efforts. The issue of permission to be approached is managed through acknowledgement of receipt of a HIPAA (Health Insurance Portability and Accounting Act) form. The form specifically states that as an academic medical center, Penn Medicine supports research and may contact the patient with offers of participation in research activities. Patients can request that they not be contacted for research purposes and Penn Medicine makes reasonable efforts to honor these requests.

2.2 Randomization & blinding

Research coordinators recruited and consented patients while the Way To Health platform performed the randomization. Once an eligible participant consented, the research coordinator mailed the participant an electronic pill bottle. The pill bottles were identical for each participant (outside of systematic device problems). The treatment assignment was revealed to the subject after they plugged in their electronic pill bottle and the Way to Health platform received a signal from the device. The participant accessed their treatment assignment either by electronic text, email, automated phone message or by accessing their Way To Health account. Way To Health generated the randomization sequence using a computer-based random number generator; the research team was blind to this process.

Randomization was stratified by source population (Employer, Insurer, Penn Medicine), income (>50,000 $US per year or not) and gender. Block sizes were 4,8 or 12 subjects.

Except for the research coordinators, the entire research team (principal and co-investigators), project managers, statisticians and students were completely blinded to treatment assignment throughout the study. The research coordinators sometimes had phone contact with participants either to troubleshoot devices or encourage lab visits. While participants sometimes revealed their treatment assignment during these calls, this information was never specifically elicited, nor was it relayed to other members of the research team.

2.3 Daily reminders

Participants received automated electronic reminders from the WTH platform via text, phone, or email; whichever they chose. The automated message stated: 'If you have not done so already, please take your study medication. The timing of the reminder was set to a default of 10 PM but could be individualized based on a query made during the baseline survey asking about the time of day that subject normally took their medication.

2.4 Data safety monitoring

The Data Safety Monitoring board was composed of experts in clinical trials, medical economics, general internal medicine and biostatistics. The DSMB assessed study conduct and outcomes including data quality, participant recruitment, accrual, retention, and adverse events.Each meeting included a recommendation to continue or terminate the study. The recommendation was made by a formal DSMB majority or unanimous vote. If the DSMB had decided to issue a termination recommendation, the full vote of the DSMB would have been required.

2.5 Electronic pill bottles used in the trial

Table 1: Enrollment and switching for three electronic pill bottles used in the trial. N is number of subjects. A total of 51 subjects (6.3%) were switched from their original pill bottle.

Pill Bottle Model	Functionality	Time Period of Enrollment	Started: N (% of Total)	Switched: N (% of Subjects Started)
1	Working as Expected	8/2013-11/2015	541 (67.2%)	34 (6.3%)[a]
	High Failure Rate	12/2015-5/2016		
2	High Failure Rate & Poor Availability	11/2015-7/2016	80 (9.9%)	17 (21.2%)
l3	Working as Expected	4/2016-7/2016	184 (22.9%)	0 (0%)
Total			**805**	**51 (6.3%)**
[a] 31 switched to Model 2 alone; 3 switched to model 2 and then model 3 [b] all 17 switched to Model 3				

3 Outcome measures

3.1 Survey Instruments

In addition to the baseline screening, we administered a series of instruments at 6 and 12 months via the Way to Health portal. These included:

- Stages of Change or Trans-theoretical Model (TTM): A measure of readiness for behavioral change
- SLACK: Subject perception of available income and responsiveness to financial incentives
- PAM: patient activation measure of knowledge, skill and confidence in managing one's health.

The goal of these analyses is exploratory and unrelated to the primary outcome. These instruments will be used to help assess the self-perceived characteristics of participants with respect to assessment of their own health, knowledge, skill and confidence in managing their health, motivation to engage in behavioral change related to their health and to the perceived value of a financial reward. Depending on the results of the study, they may shed light on the underlying mechanism, particularly in the case of meaningful treatment effects.

3.2 Cost effectiveness

The endpoints needed for the cost effectiveness analysis include:

- Costs of the intervention, i.e., electronic pill bottle, staff time and computer platform, financial incentives
- Health outcomes: difference in LDL between control and one or more effective intervention(s), life -years, and quality-adjusted life-years determined from the Coronary Heart Disease Policy Model.

4 Data analysis plans.

4.1 Adherence:

Adherence rate is defined as the number of days with a pill-bottle opening over 6 months or the number of days with a pill bottle opening during the final 30 days of the intervention, divided by either 180 or 30 days.

We will explore measured adherence via pill bottle openings by treatment arm, gender, income, baseline LDL and device model. Some subjects called to report pill-bottle problems. These call-ins were considered valid for the purpose of receiving a financial incentive. The control group did not have a financial incentive tied to pill bottle opening. As a result, they may have been less likely to call in with a problem with their pill bottle. Thus for comparisons of adherence we used only the data from the pill bottle openings.

4.2 Impact of device failure

We did not attempt to determine the faultiness of individual devices. Instead, the staff identified a calendar window of time when the devices underwent high failure rates, specifically 1 December 2015 to 31 May 2016. For the analyses, we have (1) identified this period as 'device failure period' when our experimental protocols were less robust than initially planned, and (2) determined whether each subject was potentially exposed to a faulty device based on their calendar time of enrollment. The choice of dates was based on rapidly increasing numbers of calls from participants reporting device failure. It is possible that Model 1 may have gradually experienced increased failure rates during fall 2015 as the removal of the cell towers commenced—but up until 1 December 2015, failure rates were indistinguishable from baseline rates of failure for this device

4.2.1 Adherence

We will compare rates of adherence by device and treatment arm within and outside of the device failure period

4.2.2 Change in LDL

We intend to explore the possibility that, relative to those with stable devices, subjects with faulty pill bottles altered their behavior by financial incentive arm. For example, subjects who received incentives based on adherence reported from the device to Way to Health may have become frustrated, lost interest in the study or become less adherent to medication if their device was faulty. The control group might have been little impacted by device stability, or alternatively develop heightened interest if their device was switched or they received a call from staff. Any effect of financial incentive might have been altered in the device failure period.

We will explore possible impacts of the device failure period on LDL changes. Our study is not powered for hypothesis tests of these effects. These analyses will be exploratory.

We will ask whether exposure to the device failure period impacted (1) LDL change from baseline, and (2) the effect of financial incentives on LDL change. The following analyses will be conducted:

1. Participants will be assigned to strata on the basis of whether their time on-study during the intervention included the 'high device failure' period. For each stratum and for each financial incentive arm, we will create estimates and 95% CI of the mean change in LDL. This will provide an estimate of the alteration in effect size (difference in mean LDL reduction between arms) due to exposure to faulty devices.
2. The approach is identical to (1) except that rather than determining strata we will create a quantitative variable, *X*, for each subject. This variable quantifies the proportion of their intervention time in the device failure period. For 754 subjects who were not exposed to the device failure period, this variable will be 0, but for subjects with say 2 months in the device failure period, this would be 60/180=.333. A variable, *Z*, will indicate whether a subject experienced at least one device swap (e.g. Model 2 replaced by Model 3). Let Δ_{LDL} be the change from baseline and $TrtArm$ be a set of indicators for the arms. Here the mean models will be:

a. Proportion of time in high device failure window

$$\Delta_{LDL} = \beta_0 + TrtArm + \beta_1 X + \beta_2 (X \times TrtArm)$$

b. Device swap

$$\Delta_{LDL} = \beta_0 + TrtArm + \beta_1 Z + \beta_2 (Z \times TrtArm)$$

c. Both device swap and time in high device failure window

$$\Delta_{LDL} = \beta_0 + TrtArm + \beta_1 X + \beta_2 X \times TrtArm + \beta_1 Z + \beta_2 Z \times TrtArm$$

The estimates of the β coefficients, and their 95% CI, will be evaluated for possible effect modification by exposure to the period of high device failure.

3. We will compare the probability of completing the study (lab visit at 12 months) by whether the subject spent time in the window of a poorly functioning device.
4. Lastly, we used three different devices in this study. Of interest is how Δ_{LDL}varied as a function of device type, and whether there was effect modification of Δ_{LDL} between device and treatment arm.

4.3 Missing data & multiple imputation:

1. Exploration: An initial step will use graphical methods and summary statistics to explore the distribution of LDL at 12 months as a function of baseline covariates, and to explore the extent and patterns of missingness in both the baseline and outcome (LDL at 12 months) variables.
2. Multiple Imputation Procedure: We will assume that the data are missing at random (MAR). In addition to missingness in the 12-month LDL variable, several baseline covariates include some incomplete data. The missingness pattern is arbitrary. We will thus use a Fully Conditional Specification (FCS) or sequential regression method (PROC MI in SAS) to impute the data. All baseline covariates (baseline LDL, age, sex, race, income, education, population (Employers, Insurers, Penn Medicine), study arm) will be included in the imputation. Convergence of the Gibbs sampler will be assessed using trace plots. The imputed data will be examined for credibility e.g. LDL values above 30 mg/dl. Cross-tabulations and graphical procedures will be used to compare the distribution of the observed and imputed data. A total of 100 imputed datasets will be constructed. The substantive model (12 month change in LDL as a function of treatment and baseline LDL) will be fit for each imputed dataset (PROC MIXED in SAS) and the results combined using standard 'Rubin' formulae to account for both within- and

between imputation variance. (PROC MIANALYZE in SAS). (1) Relative efficiency for the finite-sample imputation procedure will be reported.

4.4 Cost effectiveness

Cost-effectiveness ratios are defined as the difference in costs between two alternatives divided by the difference in health outcomes; these endpoints will permit calculation of the following cost-effectiveness ratios for each intervention compared to control:

- The difference in trial costs per unit reduction in LDL ascribed to the intervention;
- The difference in lifetime medical costs per life-year gained from the intervention,
- The difference in lifetime medical costs per QALY gained from the intervention

The short term analysis will determine the cost of the intervention (electronic pill bottle, staff time and computer platform, financial incentives) per unit reduction in LDL of the intervention compared with control. For the longer term, the lifetimes of participants, we will use the Coronary Heart Disease Policy Model to project the effects of the reduction in LDL on years of life, quality-adjusted years of life (QALYs), and medical costs (including intervention costs). Thus, the health endpoints will be the difference in LDL between intervention and control and the projected differences in life-years and QALYs resulting from the difference in LDL. The cost endpoints will be the difference in intervention costs during the trial and the projected differences in medical costs beyond the trial.

References

1. Little, R., Rubin, D. Statistical Analysis with Missing Data. Second. Wiley & Sons; 1987. 381 p. (Wiley Series in Probability and Statistics).