# Enhancing Patient Engagement in Preventative Care: A Field Experiment on the Effectiveness of Personalized and Generic Digital Outreach Strategies

Authors: Lynne Wang, Mridul Jain, Jesse Cox, and Tejasvi Kalakota
https://github.com/lwang9/patient-outreach-study
April, 2025

## Introduction

Can digital outreach strategies improve patient engagement with preventative healthcare? Despite the growing availability of virtual care services, many patients continue to be disengaged from routine check-ups and screenings. Observational data offers insights into healthcare behaviors but fails to isolate causal effects. Therefore, a randomized field experiment is helpful in understanding the effectiveness of various messaging strategies, particularly personalized versus generic reminders, in encouraging patients to schedule preventive care appointments.

## Justification and Theory

This study builds on the behavioral science literature, notably the Health Belief Model[1] and the Theory of Planned Behavior[2], which suggest that people are more likely to take preventive action when they believe it is beneficial and feel empowered to act. Prior research (e.g., Schwebel & Larimer, 2018)[3] confirms that digital reminders can enhance compliance with healthcare treatments. However, it remains unclear whether personalization adds meaningful value. Our experiment aims to investigate whether customizing outreach (e.g., referencing a patient's name and last visit) enhances patient engagement beyond generic outreach.

## Hypotheses

- *H1:* Generic email reminders will increase patient engagement compared to no outreach.
- *H2:* Personalized email reminders will increase patient engagement more than generic email reminders.

These hypotheses are based on the assumption that both awareness and perceived relevance increase engagement, with personalization being expected to have a more significant impact.

## Research Design

This study is structured as a randomized field experiment designed to evaluate the impact of different digital outreach strategies on patient engagement with preventative care. Additional details about the research design are provided below.

---

[1] Health Belief Model: Champion, V. L., & Skinner, C. S. (2008). The health belief model. In K. Glanz, B. K. Rimer, & K. Viswanath (Eds.), Health behavior and health education: Theory, research, and practice (4th ed., pp. 45–65). Jossey-Bass.
[2] Theory of Planned Behavior: Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.
[3] Theory of Planned Behavior: Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

**Measurement Unit**

The units of measurement in this study are individual patients of the fia.care healthcare provider network. Approximately 469 patients were selected from an existing fia.care patient database. These patients are considered "connected" patients, meaning they have downloaded the fia.care mobile app and logged into it at least once. The fia.care mobile app provides patients with convenient access to a range of digital health services, including booking virtual and in-person appointments, accessing lab results, managing prescriptions, and communicating directly with care providers, making it well-suited for digital outreach interventions like those used in this study.

Each patient record in the study database includes the following fields:

- Id: A unique identifier for the patient used for tracking within the study.
- Email address.
- first_name and last_name.
- variant_assignment: Indicates the group the patient was assigned to in the experiment (e.g., control, placebo, treatment_generic, or treatment_personalized).
- last_seen: The most recent date the patient had contact with fia.care, used to calculate recency.
- Dob (date of birth): Used to derive age and perform a covariate balance test.
- Gender.
- PCP: Primary care provider associated with the patient.
- insurance_carrier.
- HCC(Hierarchical Condition Category).

**Randomization Process**

Patients (N = 469) were randomly assigned to one of four groups. To prevent spillovers, family members were assigned into the same group. To ensure the validity of our randomization process, we conducted a covariate balance check based on age and gender. The results confirmed that age and gender were both well balanced across all experimental groups.

**Treatment:**

Participants were randomly assigned to one of four experimental groups.

- The Control Group received no message at all. The Placebo Group (X0) received a general informational message about the healthcare provider without any prompt for action.
- The Generic Message Group (X1) received outreach promoting general health benefits, screenings, and appointment options. In particular, the email included a direct link to the fia.care mobile app; when clicked, this link opens the app directly on the user's device, facilitating immediate access to appointment booking and other services.
- The Personalized Message Group (X2) received a message tailored with the patient's name and the date of their last visit, along with a direct call to action encouraging them to schedule an appointment. Similar to X1 group, the email also includes a direct link to the fia.care mobile app.

Example messages for each group are provided in Appendix I.

### ROXO Design

The research is based on the following ROXO structure, where each line represents a group's exposure to repeated observations (O) over time. The R indicates randomization, and the sequences show which treatment was assigned (X0 for placebo, X1 for generic, X2 for personalized), followed by observation points. For example, the line labeled R_X0_O_X0_O_X0_O_X0_..._O indicates that the placebo message group received multiple rounds of messaging, each followed by a measurement of engagement, such as open rate, click-through rate, and/or visit rate. The same structure applies to the generic (X1) and personalized (X2) messaging groups. The final line, R_____O, represents the control group, which received no message but was observed at the end for their visit rate. This repeated-measures design enables the capture of engagement and behavioral responses over time.

```
R_X0_O_X0_O_X0_O_X0. . . _O
R_X1_O_X1_O_X1_O_X1. . . _O
R_X2_O_X2_O_X2_O_X2. . . _O
R___________________. . . _O
```

Fig. 1 presents an overview of the email campaign sent from March 16 to April 16, 2025. Each peak represents a batch of outreach emails sent during the campaign's scheduled delivery days. The graph tracks four key metrics: Accepted, Delivered, Failed, and Opened. Email delivery occurred roughly every 3–4 days, with consistent high-volume spikes indicating scheduled batch sends. Across all dates, nearly all accepted emails were successfully delivered, and a large majority were opened.

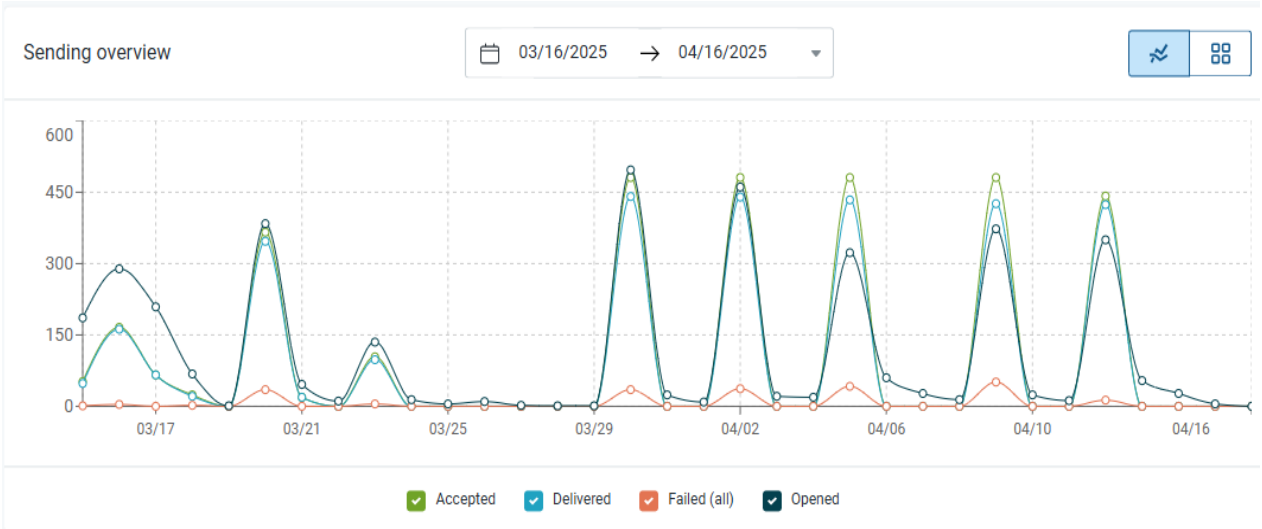



**Fig. 1 Message Sending Overview**

### Binary Outcome Coding

For each patient, engagement outcomes were recorded as binary variables, indicating whether the individual responded to any of the outreach messages during the study period. Specifically, a patient's "open" variable was coded as 1 if they opened at least one email across multiple campaign waves; otherwise, it was coded as 0. Similarly, the "click" variable was coded as 1 if the patient clicked on a link in at least one email, and 0 if no links were clicked in any message received.

### Analysis of Potential Outcomes

This study examines both behavioral and digital engagement outcomes to assess the effectiveness of different messaging strategies. Specifically, we analyze and compare outcomes across four experimental groups, including open rates, unsubscribe rates, click-through rates, and visit rates.

- Open rates serve as a proxy for compliance and are compared across the placebo (X0), generic (X1), and personalized (X2) messaging groups.
- Unsubscribe rates, interpreted as a negative treatment effect, are compared between the generic and personalized message groups to assess whether personalization increases opt-out behavior.
- Click-through rates are treated as an intermediary outcome reflecting active engagement with the message content and are also compared between X1 and X2.
- Finally, visit rates—the primary behavioral outcome—are compared across the control, generic (X1), and personalized message (X2) groups to evaluate whether digital engagement translates into real-world healthcare actions

### CONSORT Flow Diagram

Fig. 2 is a flow diagram illustrating the experimental structure, which begins with a total of 469 patients, all of whom were randomly assigned to one of four experimental groups. The Control Group (N = 108) received no messages. The Placebo Group (N = 132) received general informational emails, while the Generic Message Group (N = 120) and Personalized Message Group (N = 109) received targeted outreach. The treatment groups were exposed to up to seven emails per month, and all treatment groups underwent email engagement monitoring to track opens, clicks, and unsubscribes. Across all groups, including the control group, visit status monitoring was conducted to assess whether patients had visited or were scheduled to visit their healthcare providers at the end of the experiment period.

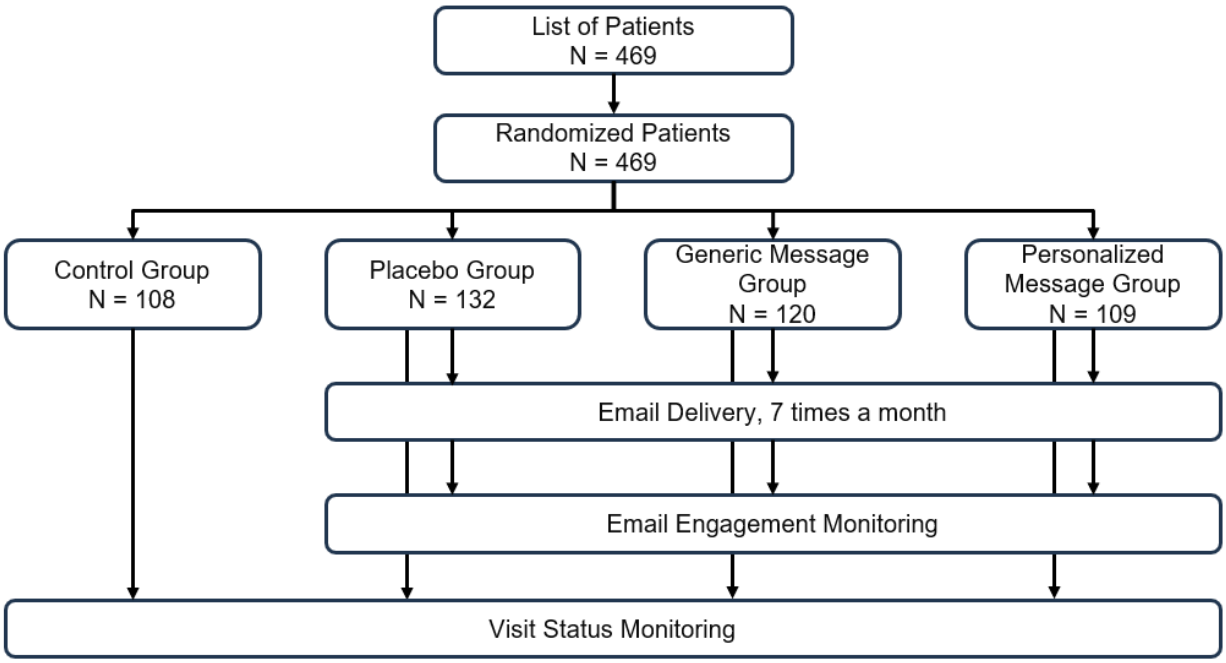



**Fig. 2 Flow Diagram**

## Power Calculation

Fig. 3 illustrates power curves for detecting treatment effects of varying magnitudes—Average Treatment Effects (ATEs) of 0.005, 0.01, and 0.02—across a range of sample sizes. The y-axis represents statistical power, and the x-axis represents sample size per treatment group. These simulations assume a baseline mean of 0.02 for the control group and standard deviations of sd_control = sqrt(0.02 × 0.98) and sd_treat = sqrt(0.02 × 0.98), reflecting the variance expected under a Bernoulli distribution for positive events. The red line shows that even with 5,000 participants per group, power to detect an ATE of 0.005 remains below 0.4. The blue line indicates that detecting an ATE of 0.01 achieves 0.8 power around a sample size of 4,000. The green line shows that an ATE of 0.02 reaches the conventional power threshold (0.8) at approximately 800 participants per group. These curves highlight the limitations of the current study's sample size (~120 per group), which is substantially underpowered to detect small effects. As such, the study's findings are best interpreted as exploratory.

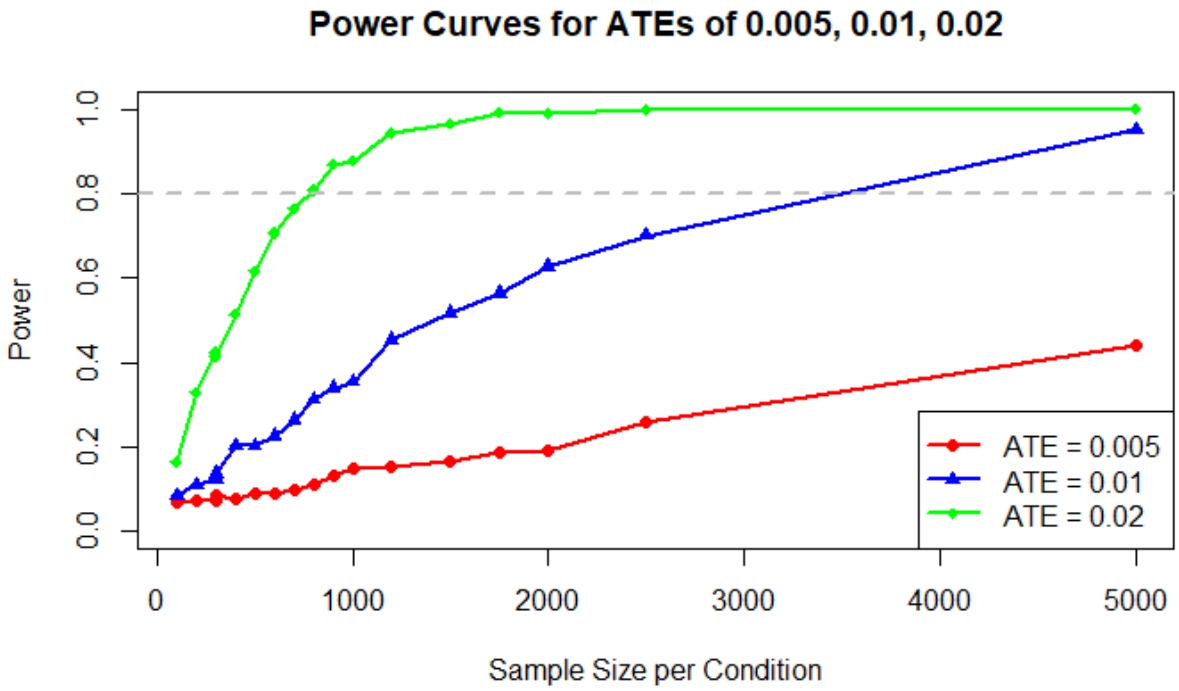



**Fig. 3 Power Curves**

## Data and Measurement

The primary outcome of interest in this study was visit, used as an indicator of real-world healthcare engagement. Secondary outcomes included email open rates, click-through rates, and unsubscribe behaviors, which serve as intermediate measures of digital engagement. Treatments were coded using the variant_assignment variable, which identified the group to which each patient was assigned (control, placebo, generic, or personalized). All outcome variables were coded as binary indicators to facilitate statistical modeling. As a key covariate, the last seen date was used to account for prior patient engagement with the healthcare system, thereby improving precision and supporting subgroup heterogeneous treatment effect (HTE) analysis.

**Unsubscribed (Negative Treatment Effect)**

Table 1 shows patients who unsubscribed from email communications, marked with a 1 in the "unsubscribed" column. Across the sample, unsubscribes were observed in all treatment groups, including the placebo, generic, and personalized message groups. Notably, one participant unsubscribed even after clicking through a link in the email.

| id | variant_assignment | last_seen | unsubscribed | opened | click | visit |
|---|---|---|---|---|---|---|
| 21 | placebo | 1/17/2025 | 1 | 1 | 0 | 0 |
| 41 | treatment_generic | a while ago | 1 | 1 | 0 | 0 |
| 44 | placebo | 9/13/2023 | 1 | 1 | 0 | 0 |
| 78 | treatment_generic | a while ago | 1 | 1 | 0 | 0 |
| 105 | treatment_generic | a while ago | 1 | 1 | 0 | 0 |
| 190 | placebo | a while ago | 1 | 1 | 0 | 0 |
| 201 | treatment_personalized | a while ago | 1 | 1 | 0 | 0 |
| 221 | treatment_personalized | a while ago | 1 | 1 | 0 | 0 |
| 232 | treatment_generic | 6/13/2024 | 1 | 1 | 0 | 0 |
| 278 | placebo | 5/8/2024 | 1 | 1 | 0 | 0 |
| 296 | treatment_generic | 4/19/2024 | 1 | 1 | 0 | 0 |
| 363 | placebo | a while ago | 1 | 1 | 0 | 0 |
| 381 | treatment_generic | a while ago | 1 | 1 | 0 | 0 |
| 404 | treatment_personalized | a while ago | 1 | 1 | 0 | 0 |
| 410 | treatment_personalized | a while ago | 1 | 1 | 1 | 0 |
| 415 | treatment_generic | a while ago | 1 | 1 | 0 | 0 |
| 418 | treatment_personalized | a while ago | 1 | 1 | 0 | 0 |
| 425 | placebo | a while ago | 1 | 1 | 0 | 0 |
| 433 | placebo | 2/20/2021 | 1 | 1 | 0 | 0 |
| 438 | treatment_personalized | 1/12/2021 | 1 | 1 | 0 | 0 |
| 448 | placebo | a while ago | 1 | 1 | 0 | 0 |

**Table 1**

**Clicked (Intermediary Treatment Effect)**

Table 2 shows patients who engaged by clicking a link in the email. All individuals in this section have a "click" value of 1. This subset indicates a higher representation from the personalized group. These observations support earlier findings that personalization may increase the likelihood of click-through behavior, an intermediate engagement metric.

| id | variant_assignment | last_seen | unsubscribed | opened | click | visit |
|---|---|---|---|---|---|---|
| 17 | treatment_personalized | a while ago | 0 | 1 | 1 | 0 |
| 95 | treatment_personalized | 12/16/2024 | 0 | 1 | 1 | 0 |
| 107 | placebo | a while ago | 0 | 1 | 1 | 0 |
| 183 | treatment_personalized | a while ago | 0 | 1 | 1 | 0 |
| 342 | treatment_personalized | 12/13/2023 | 0 | 1 | 1 | 0 |
| 383 | treatment_generic | a while ago | 0 | 1 | 1 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 410 | treatment_personalized | a while ago | 1 | 1 | 1 | 0 |
| 422 | treatment_personalized | a while ago | 0 | 1 | 1 | 0 |
| 455 | placebo | 10/31/2022 | 0 | 1 | 1 | 0 |

**Table 2**

### Visited (Primary Treatment Effect)

Table 3 shows patients who ultimately scheduled or completed a healthcare appointment, indicated by a 1 in the "visit" column. This is the primary behavioral outcome of interest. Participants who visited included those from all message variants, as well as the control group. Interestingly, some individuals visited despite no email engagement (click = 0), indicating that visit behavior may be through a channel different from clicking the link to open a mobile app of a fia.care. The overall frequency of visits remains low, consistent with previous results indicating no statistically significant impact of the messaging strategies on real-world follow-through.

| id | variant_assignment | last_seen | unsubscribed | opened | click | visit |
|---|---|---|---|---|---|---|
| 25 | control | 1/29/2025 | 0 | 0 | 0 | 1 |
| 28 | placebo | 1/15/2025 | 0 | 1 | 0 | 1 |
| 31 | treatment_generic | a while ago | 0 | 1 | 0 | 1 |
| 97 | treatment_generic | 2/15/2025 | 0 | 1 | 0 | 1 |
| 112 | treatment_personalized | 2/13/2025 | 0 | 1 | 0 | 1 |
| 164 | placebo | 3/8/2025 | 0 | 1 | 0 | 1 |
| 171 | control | 3/3/2025 | 0 | 0 | 0 | 1 |
| 203 | treatment_generic | 7/24/2024 | 0 | 1 | 0 | 1 |
| 248 | treatment_personalized | 1/10/2025 | 0 | 1 | 0 | 1 |
| 383 | treatment_generic | a while ago | 0 | 1 | 1 | 1 |
| 400 | placebo | 1/7/2025 | 0 | 1 | 0 | 1 |

**Table 3**

### Engagement Rates by Variant Assignment

Fig. 4 illustrates the distribution of three key engagement outcomes—unsubscribe, click-through, and visit rates—across the four experimental groups.

- Unsubscribe rates were highest in the placebo group (~6.1%) and slightly lower in the treatment groups (generic ~5.9%, personalized ~5.6%), suggesting all outreach messages were generally well-tolerated.
- Click-through behavior varied more distinctly: the personalized message group had the highest click-through rate (~5.6%), followed by the generic group (~1%), with the placebo group at ~1.5%. This supports the earlier hypothesis that personalization boosts engagement.
- However, visit rates remained low across all groups, with the highest rate in the generic group (~3.3%), followed by the placebo group (~2.4%), the control group (~1.9%), and the personalized group (~1.8%).

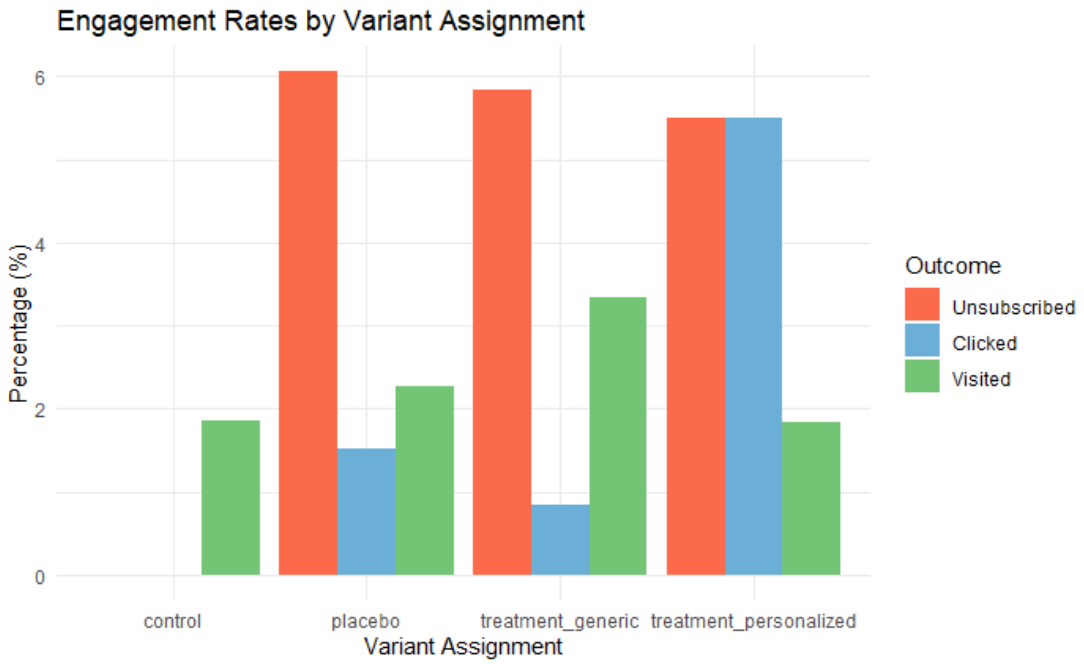



**Fig. 4 Bar chart comparing outcomes between control and each treatment group (X0, X1, and X2)**

## Analysis

Our analysis assesses the impact of various outreach strategies on patient engagement, examining both digital (via email interactions) and behavioral (via healthcare visit rates) aspects. We begin with a placebo test that confirms all message variants—placebo (X0), generic (X1), and personalized (X1)—elicited similarly high open rates (~100%), validating system functionality and treatment compliance. This high engagement translates to a complier rate of 1, with no attrition observed. The most notable effect was observed in click-through behavior, where personalized messages significantly outperformed generic messages ($p = 0.0478$), suggesting that tailored content enhances digital engagement. However, unsubscribe rates were unaffected by personalization, suggesting that both message types were similarly well-tolerated.

Despite these differences in intermediate outcomes, visit rates did not differ significantly between any groups, suggesting that enhanced digital engagement did not translate into increased real-world healthcare utilization. Furthermore, a heterogeneous treatment effect (HTE) analysis revealed no clear evidence that patient recency (i.e., the time since their last visit) modified the impact of outreach. Importantly, the study's statistical power was limited by a relatively small sample size, particularly given the low base rate of visit behavior. As a result, even if the outreach strategies had modest effects on visits, the experiment may not have been able to detect them. Future studies will require larger samples to reliably detect effects on behavioral outcomes, such as appointment scheduling.

Additional details about each of these tests are further described below.

### Placebo Test – Intent-to-Treat (ITT)

To verify the validity of our engagement measures and ensure our email system was functioning as expected, we conducted a placebo test by comparing open rates across the placebo, generic, and personalized messaging groups (shown below). The results showed no statistically significant differences in open rates: the coefficients for both generic

and personalized messages relative to the placebo group were approximately 0.0076 with p-values of 0.3182. These findings suggest that all outreach messages elicited similarly high open rates (~99%), and that message content did not meaningfully influence this behavior.

```
                                    Placebo Test Results
T test of coefficients:
                                          Estimate Std. Error  t value Pr(>|t|)
(Intercept)                               0.9924242  0.0075786 130.9515   <2e-16 ***
variant_assignmenttreatment_generic       0.0075758  0.0075786   0.9996   0.3182
variant_assignmenttreatment_personalized 0.0075758  0.0075786   0.9996   0.3182
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### Compliance and Attrition

Across all treatment groups, the email open rates approached or reached 100%, indicating near-perfect compliance with the digital treatment. Specifically, the open rate for the placebo group was 99.2%, while both the generic and personalized message groups exhibited 100% open rates. This high engagement confirms that all participants in the treatment groups were successfully exposed to the treatment. As such, the complier rate is effectively 1, and no attrition was observed during the course of the study. These results indicate that there is no need for adjustments due to noncompliance or attrition.

| variant_assignment <chr> | open_rate <dbl> | count <int> |
|---|---|---|
| placebo | 0.9924242 | 132 |
| treatment_generic | 1.0000000 | 120 |
| treatment_personalized | 1.0000000 | 109 |

### Unsubscribe Rate: Generic (X1) vs Personalized Messaging (X2)

The comparison of unsubscribe rates between the personalized and generic message groups indicates no statistically significant difference. The coefficient estimate for the personalized group was -0.0033 with a p-value of 0.9148. These results indicate that recipients were not more likely to opt out after receiving personalized messages compared to generic ones. In practical terms, this suggests that personalization did not provoke negative reactions or increase opt-out behavior, and that both message formats were similarly acceptable to patients in terms of continued email engagement.

```
                              Unsubscribe Rate Test Results
t test of coefficients:

                                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                               0.0583333  0.0214892  2.7145 0.007147 **
variant_assignmenttreatment_personalized -0.0032875  0.0307116 -0.1070 0.914849
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### Click-Through Rate: Generic (X1) vs. Personalized Messaging (X2)

The regression analysis comparing click-through rates between the generic and personalized message groups shows that personalization had a statistically significant positive effect. Specifically, the estimated effect size for the personalized message group was 0.0467 with a p-value of 0.0478. This result crosses the conventional threshold for statistical significance at the 5% level, indicating that patients who received personalized messages were more likely

to click on embedded links compared to those receiving generic reminders. This finding suggests that personalization can effectively increase digital engagement by making messages feel more relevant and actionable to recipients.

**Click-Through Rate Test Results**

```
t test of coefficients:

                                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                              0.0083333  0.0083350  0.9998  0.31847
variant_assignmenttreatment_personalized 0.0467125  0.0234710  1.9902  0.04777 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### Visit Rate: Control vs. Generic (X1) vs. Personalized Messaging (X2)

The analysis of visit behavior showed no statistically significant effect from either generic or personalized outreach. The estimated coefficients were 0.0148 (p = 0.4809) for the generic message group and -0.0002 (p = 0.9926) for the personalized group, relative to the control. These small and statistically insignificant differences suggest that neither form of messaging meaningfully increased real-world healthcare engagement in terms of scheduling visits. Overall visit rates were low across all groups, which underscores the limited statistical power to detect meaningful differences in this outcome.

**Visit Rate Test Results**

```
t test of coefficients:

                                           Estimate  Std. Error t value Pr(>|t|)
(Intercept)                              0.01851852  0.01303089  1.4211   0.1562
variant_assignmenttreatment_generic      0.01481481  0.02099368  0.7057   0.4809
variant_assignmenttreatment_personalized -0.00016989  0.01834491 -0.0093   0.9926
```

### Heterogeneous Treatment Effects (HTE) by Patient Recency

To explore whether the effect of outreach varied by prior patient engagement, we conducted an HTE analysis interacting treatment assignment with a binary indicator for whether a patient had not been seen recently (i.e., not visited within the past year). The results showed no statistically significant interactions between treatment assignment and patient recency. Specifically, the interaction terms for generic × not seen recently (estimate = 0.0413, p = 0.5772) and personalized × not seen recently (estimate = 0.0119, p = 0.8743) were both not statistically significant. However, increasing the sample size may enable the detection of the effect with statistical significance.

**HTE Visit Rate Test Results**

```
t test of coefficients:

                                                        Estimate Std. Error t value Pr(>|t|)
(Intercept)                                             0.083333   0.056926  1.4639   0.1442
variant_assignmenttreatment_generic                    -0.018817   0.072269 -0.2604   0.7947
variant_assignmenttreatment_personalized               -0.011905   0.075182 -0.1583   0.8743
not_seen_recently                                      -0.083333   0.056926 -1.4639   0.1442
variant_assignmenttreatment_generic:not_seen_recently   0.041289   0.073987  0.5581   0.5772
variant_assignmenttreatment_personalized:not_seen_recently 0.011905 0.075182 0.1583   0.8743
```

## Limitations

This study has several limitations. First, the small sample size limited the statistical power to detect modest treatment effects, particularly for low-probability outcomes, such as healthcare visits. Second, the short duration of the treatment—just one month—may not have been sufficient to influence scheduling behavior for preventative care, which often occurs on longer timelines. Third, the distribution of outcomes was highly skewed, with the vast majority of patients not booking appointments, thereby reducing the sensitivity of visit-based measures. Finally, there may be unmeasured confounding in participants' decisions to unsubscribe, such as prior experiences with the provider or differing baseline attitudes toward digital communication, which were not captured in the data.

## Conclusion

While personalized messaging increased digital engagement, as evidenced by click-through rates, it did not significantly impact the primary behavior of booking appointments. Generic reminders performed similarly. This suggests that, while personalization enhances intermediate engagement, structural barriers or lack of perceived need may limit behavior change. Future studies could investigate complementary interventions, such as incentives, to enhance follow-through. Additionally, future research should incorporate larger sample sizes to enhance statistical power, particularly for detecting small effects on low-frequency behavioral outcomes, such as healthcare visits.

# Appendix I: Example Email Messages

**Example Email to Placebo Group:**

Hello,

At Fia Care, we believe healthcare should come to you. We are a full-service virtual primary care clinic offering on-demand chat, video consultations, and home visits when needed. Our services include physicals, chronic disease management, some labs, prescription renewals, and referrals to specialists. We aim to make healthcare more accessible and cost-effective for individuals and companies alike.

To learn more about how we are redefining primary care, visit our About Us page.

Best regards,

The Fia Care Team
https://fia.care/about-us/

**Example Generic Email:**

Hello!

We are thrilled to have you as part of the Fia family! Are you ready to book your wellness visit? As a valued member of Fia, you have exclusive access to comprehensive benefits including health screenings, expert consultations, and much more.

Don't wait any longer to make the most of your Fia benefits. Schedule your visit today and ensure that you are keeping on top of your health. Your well-being is our priority, and we're here to support you every step of the way.

Kind regards,
The Fia Team
Connect with Fia Care

**Example Personalized Email:**

Hi [Patient Name]!

It's been a while since your last check-up on [Last Visit Date]. Let's get you back on track with your health goals by scheduling your next health screening. Staying current with regular screenings is key to maintaining good health.

Contact us today at [phone no.], and let's keep moving forward on your health journey. We are here to help you every step of the way.

Best regards,
The Fia Team
Connect with Fia Care

# Appendix II - R Code for Enhancing Patient Engagement in Preventative Care: A Field Experiment on the Effectiveness of Personalized and Generic Digital Outreach Strategies

Lynne Wang, Mridul Jain, Jesse Cox, and Tejasvi Kalakota

04/18/2025

```r
# Load and preview the first few rows of the data
# This gives a quick view of variables like variant_assignment, unsubscribed, opened, click, and visit
print(head(data))
```

```
##       id variant_assignment   last_seen unsubscribed opened click visit
##    <int>              <char>      <char>        <int>  <int> <int> <int>
## 1:    21             placebo   1/17/2025            1      1     0     0
## 2:    41  treatment_generic a while ago            1      1     0     0
## 3:    44             placebo   9/13/2023            1      1     0     0
## 4:    78  treatment_generic a while ago            1      1     0     0
## 5:   105  treatment_generic a while ago            1      1     0     0
## 6:   190             placebo a while ago            1      1     0     0
```

```r
d <- data
```

```r
# Count the number of patients assigned to each variant group
# Helps confirm balance of randomization across control and treatment groups
data %>%
  count(variant_assignment)
```

```
##        variant_assignment     n
##                     <char> <int>
## 1:                control   108
## 2:                placebo   132
## 3:      treatment_generic   120
## 4: treatment_personalized   109
```

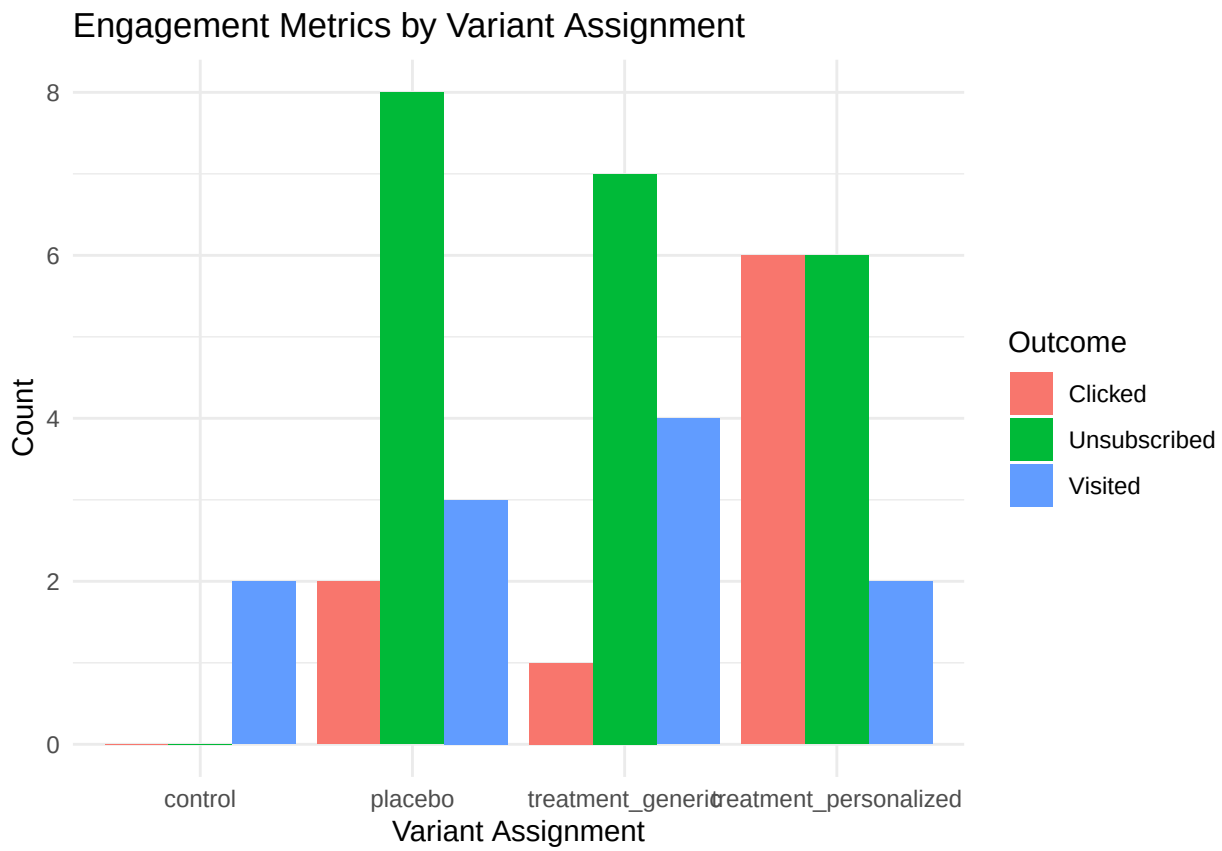
```r
# Summarize click behavior by treatment arm
# Outputs total clicks and average click rate per group
d %>%
  group_by(variant_assignment) %>%
  summarise(n = n(), clicks = sum(click, na.rm = TRUE), click_rate = mean(click, na.rm = TRUE))
```

```
## # A tibble: 4 x 4
##   variant_assignment         n clicks click_rate
##   <chr>                  <int>  <int>      <dbl>
```

```
## 1 control                 108      0    0
## 2 placebo                 132      2    0.0152
## 3 treatment_generic       120      1    0.00833
## 4 treatment_personalized  109      6    0.0550
```

```r
# Summarize key engagement outcomes by treatment group
# Includes counts of clicks, visits, and unsubscribes for plotting
summary_counts <- d %>%
  group_by(variant_assignment) %>%
  summarise(
    Clicked = sum(click),
    Visited = sum(visit),
    Unsubscribed = sum(unsubscribed)
  ) %>%
  pivot_longer(
    cols = c(Clicked, Visited, Unsubscribed),
    names_to = "Outcome",
    values_to = "Count"
  )

# Plot bar chart comparing counts of engagement metrics by variant group
# Visualizes Clicked, Visited, and Unsubscribed across Control, Placebo, and Treatment groups
ggplot(summary_counts, aes(x = variant_assignment, y = Count, fill = Outcome)) +
  geom_bar(stat = "identity", position = position_dodge()) +
  labs(title = "Engagement Metrics by Variant Assignment", x = "Variant Assignment", y = "Count") +
  theme_minimal()
```



```r
# Filter data to include only treatment-related groups (excluding control)
# Used for comparing email open rates across different outreach strategies
data_subset_1 <- data %>%
  filter(variant_assignment %in% c("placebo", "treatment_generic", "treatment_personalized"))

# Estimate linear model to compare open rates across variant_assignment
# Robust standard errors account for heteroskedasticity
model_opened <- lm(opened ~ variant_assignment, data = data_subset_1)
robust_se <- coeftest(model_opened, vcov = vcovHC(model_opened, type = "HC1"))
print(robust_se)
```

```
##
## t test of coefficients:
##
##                                          Estimate Std. Error  t value Pr(>|t|)
## (Intercept)                             0.9924242  0.0075786 130.9515   <2e-16
## variant_assignmenttreatment_generic     0.0075758  0.0075786   0.9996   0.3182
## variant_assignmenttreatment_personalized 0.0075758  0.0075786   0.9996   0.3182
##
## (Intercept)                             ***
## variant_assignmenttreatment_generic
## variant_assignmenttreatment_personalized
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Intention-to-treat analysis: calculate open rate and sample size per variant
# Used to assess compliance and attrition
itt_open <- data_subset_1 %>%
  group_by(variant_assignment) %>%
  summarise(open_rate = mean(opened == 1), count = n())
print(itt_open)
```

```
## # A tibble: 3 x 3
##   variant_assignment     open_rate count
##   <chr>                      <dbl> <int>
## 1 placebo                    0.992   132
## 2 treatment_generic          1       120
## 3 treatment_personalized     1       109
```

```r
# Filter to just treatment_generic and treatment_personalized groups
# Fit model to compare unsubscribe rates between personalized and generic emails
data_subset_2 <- data %>%
  filter(variant_assignment %in% c("treatment_generic", "treatment_personalized"))

model_unsubscribed <- lm(unsubscribed ~ variant_assignment, data = data_subset_2)

# Output summary with robust standard errors (HC1)
coeftest(model_unsubscribed, vcov = vcovHC(model_unsubscribed, type = "HC1"))
```

```
##
## t test of coefficients:
##
```

```
##                                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)                             0.0583333  0.0214892  2.7145 0.007147
## variant_assignmenttreatment_personalized -0.0032875  0.0307116 -0.1070 0.914849
##
## (Intercept)                             **
## variant_assignmenttreatment_personalized
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Filter to just treatment_generic and treatment_personalized groups
# Fit model to compare click rates between personalized and generic emails
data_subset_3 <- data %>%
  filter(variant_assignment %in% c("treatment_generic", "treatment_personalized"))

model_click <- lm(click ~ variant_assignment, data = data_subset_3)

# Output summary with robust standard errors (HC1)
coeftest(model_click, vcov = vcovHC(model_click, type = "HC1"))
```

```
##
## t test of coefficients:
##
##                                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)                             0.0083333  0.0083350  0.9998  0.31847
## variant_assignmenttreatment_personalized 0.0467125  0.0234710  1.9902  0.04777
##
## (Intercept)
## variant_assignmenttreatment_personalized *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Filter to just control, treatment_generic and treatment_personalized groups
# Fit model to compare visit rates between personalized and generic emails

data_subset_4 <- data %>%
  filter(variant_assignment %in% c("control", "treatment_generic", "treatment_personalized"))

model_visit <- lm(visit ~ variant_assignment, data = data_subset_4)

# Output summary with robust standard errors (HC1)
coeftest(model_visit, vcov = vcovHC(model_visit, type = "HC1"))
```

```
##
## t test of coefficients:
##
##                                           Estimate  Std. Error t value
## (Intercept)                              0.01851852  0.01303089  1.4211
## variant_assignmenttreatment_generic      0.01481481  0.02099368  0.7057
## variant_assignmenttreatment_personalized -0.00016989  0.01834491 -0.0093
##                                          Pr(>|t|)
## (Intercept)                                0.1562
## variant_assignmenttreatment_generic        0.4809
## variant_assignmenttreatment_personalized   0.9926
```

```r
# Add covariate for recency of healthcare interaction, and generate interaction term with treatment.
# Enables HTE analysis to see if treatment works differently for patients not recently seen.

today <- as.Date("2025-04-15")

d <- d %>%
  mutate(last_seen = ifelse(last_seen == "a while ago", '04/01/2024', last_seen))

d <- d %>%
  mutate(
    # Step 2: Convert from m/d/Y to Date
    last_seen_clean = mdy(last_seen),

    # Step 3: Compute days since seen
    days_since_seen = as.numeric(Sys.Date() - last_seen_clean),

    # Step 4: Define seen_recently as  365 days
    seen_recently = ifelse(days_since_seen <= 365, 1, 0)
  )
```

```r
# Fit a model including an interaction between treatment assignment and not_seen_recently.
# Helps test whether timing of last engagement modifies treatment effects.

# Set 'not seen recently' (0) as the reference level
# Create a new variable: not_seen_recently
d$not_seen_recently <- as.numeric(d$seen_recently == 0)

# Filter out the placebo group
d_filtered <- d %>% filter(variant_assignment %in% c("control", "treatment_generic", "treatment_personal

# Create a new variable: not_seen_recently
d_filtered$not_seen_recently <- as.numeric(d_filtered$seen_recently == 0)

# Fit the model using the new variable
model_HTE <- lm(visit ~ variant_assignment * not_seen_recently, data = d_filtered)
robust_se <- vcovHC(model_HTE, type = "HC1")
robust_results <- coeftest(model_HTE, vcov = robust_se)
print(robust_results)
```

```
##
## t test of coefficients:
##
##                                                          Estimate Std. Error
## (Intercept)                                              0.083333   0.056926
## variant_assignmenttreatment_generic                     -0.018817   0.072269
## variant_assignmenttreatment_personalized                -0.011905   0.075182
## not_seen_recently                                       -0.083333   0.056926
## variant_assignmenttreatment_generic:not_seen_recently    0.041289   0.073987
## variant_assignmenttreatment_personalized:not_seen_recently  0.011905   0.075182
##                                                         t value Pr(>|t|)
## (Intercept)                                              1.4639   0.1442
## variant_assignmenttreatment_generic                     -0.2604   0.7947
## variant_assignmenttreatment_personalized                -0.1583   0.8743
```

```
## not_seen_recently                                          -1.4639   0.1442
## variant_assignmenttreatment_generic:not_seen_recently       0.5581   0.5772
## variant_assignmenttreatment_personalized:not_seen_recently  0.1583   0.8743
```

```r
# Simulate statistical power at different sample sizes for three effect sizes: 0.005, 0.01, 0.02.
# Helps visualize sensitivity of the study to detect effects under realistic assumptions.

set.seed(3)

power_test_t <- function(
  mean_control = 0.02,
  mean_treat = 0.03,
  sd_control = sqrt(0.02 * 0.98),
  sd_treat = sqrt(0.02 * 0.98),
  number_per_condition = 40,
  power_loops = 1000,
  verbose = FALSE) {
    p_values <- numeric(power_loops)
    for (i in 1:power_loops) {
        control_group <- rnorm(number_per_condition, mean = mean_control, sd = sd_control)
        treatment_group <- rnorm(number_per_condition, mean = mean_treat, sd = sd_treat)

        test_result <- t.test(control_group, treatment_group, var.equal = FALSE)
        p_values[i] <- test_result$p.value
    }

    power_estimate <- mean(p_values < 0.05)
    return(list('p_values' = p_values, 'power' = power_estimate))
}
```

```r
# Generate line plot showing how sample size impacts power for each effect size.
# A horizontal line at 0.8 indicates the standard threshold for acceptable power.

# Sample sizes
samples_per_condition <- c(100, 200, 300, 300, 400, 500, 600, 700, 800, 900, 1000, 1200, 1500, 1750, 20

# Calculate power for treatment effects of +0.005, +0.01, +0.02
power_005 <- numeric(length(samples_per_condition))
power_01 <- numeric(length(samples_per_condition))
power_02 <- numeric(length(samples_per_condition))

for(i in seq_along(samples_per_condition)) {
  power_005[i] <- power_test_t(mean_control = 0.02, mean_treat = 0.025,
                               number_per_condition = samples_per_condition[i])$power
  power_01[i] <- power_test_t(mean_control = 0.02, mean_treat = 0.03,
                              number_per_condition = samples_per_condition[i])$power
  power_02[i] <- power_test_t(mean_control = 0.02, mean_treat = 0.04,
                              number_per_condition = samples_per_condition[i])$power
}

# Assemble power table
power_table <- data.frame(
  "Sample Size" = samples_per_condition,
  "ATE 0.005" = power_005,
```

```
  "ATE 0.01" = power_01,
  "ATE 0.02" = power_02
)

# Plot results
plot(samples_per_condition, power_005, type = "o", col = "red", lwd = 2, pch = 16, ylim = c(0, 1),
     xlab = "Sample Size per Condition", ylab = "Power", main = "Power Curves for ATEs of 0.005, 0.01, 0
lines(samples_per_condition, power_01, col = "blue", lwd = 2, type = "o", pch = 17)
lines(samples_per_condition, power_02, col = "green", lwd = 2, type = "o", pch = 18)
abline(h = 0.8, col = "gray", lty = 2, lwd = 2)
legend("bottomright", legend = c("ATE = 0.005", "ATE = 0.01", "ATE = 0.02"),
       col = c("red", "blue", "green"), lwd = 2, pch = c(16, 17, 18))
```