From Past Abiding Behavior to Subsequent Intervention Compliance: A Mobile Phone Usage Case Study

Yang Meng   Rong Ji   Zeyue Zhang

Department of Biostatistics, University of Michigan, Ann Arbor, MI

1. Introduction

Background

- Mobile phone addiction is a crucial concern and excessive usage leads to various mental, psychological, and physical health issues[1];
- Many behavioral interventions have been developed but there is **a need to learn users' compliance and factors related with it;**

Study Design & Interventions

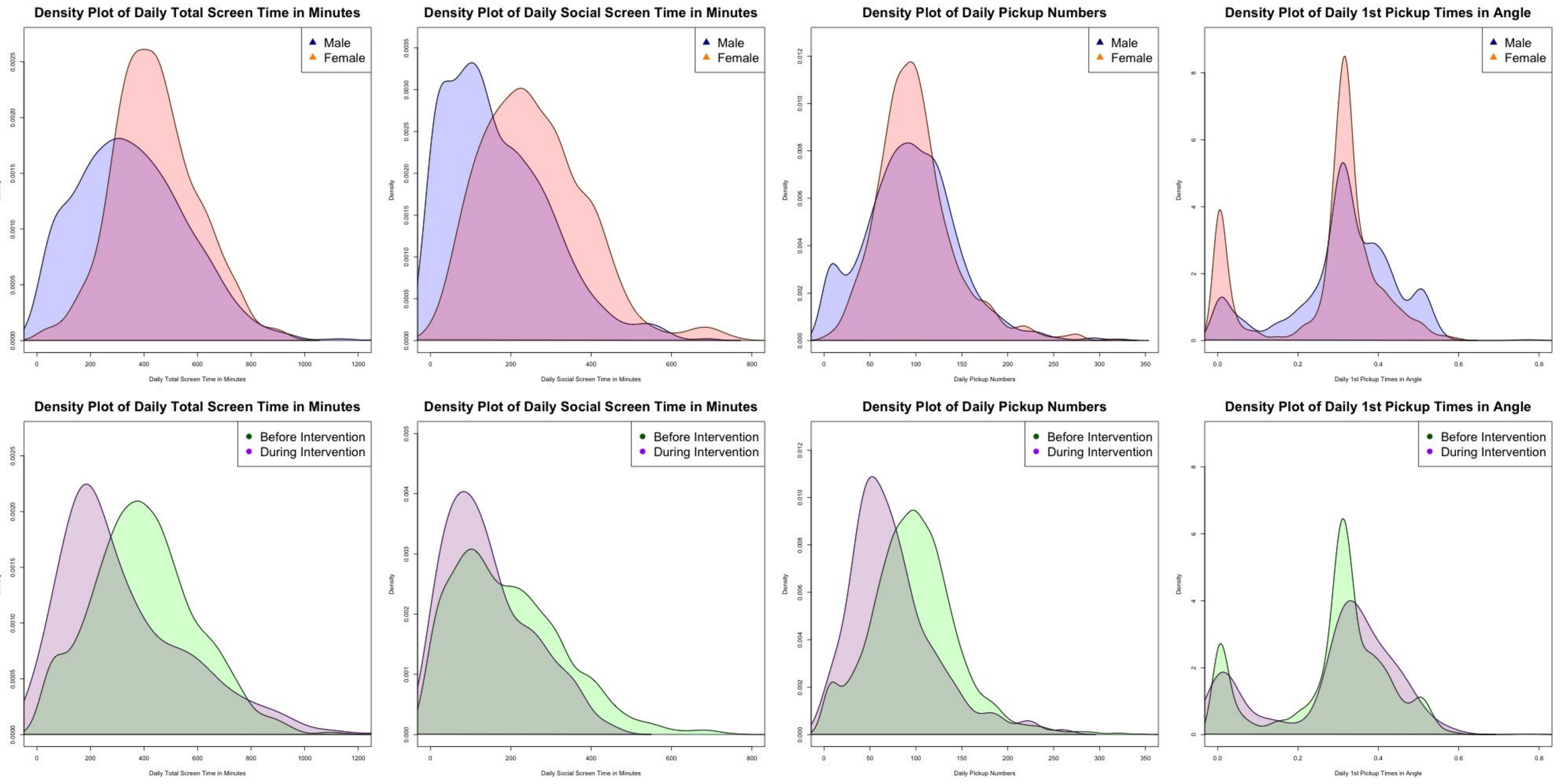
- A cohort of 34 users recorded their daily screen time data from mid-January to April 2;
- From March 27 to April 2 is the intervention treatment period which has **daily restricted allowance for phone usage** on users;
- Intervention assigned by randomization:
  - **Intervention A:** 200 screen time minutes;
  - **Intervention B:** 50 pick up numbers;

Statistical Aim

- **Are users' past committed behavior status associated with intervention adherence?**
- Deriving new variables from raw data to provide metrics for the past abiding behavior;

2. Data Description

- Original longitudinal data set consists of two sheets:
  - Sheet **"screentime"**: daily phone usage activity statistics for each participant;
  - Sheet **"baseline"**: individual characteristics by answering a questionnaire;
- **"compliance"**: binary outcome, whether complying with assigned intervention;
- **"early"**: new created main predictor, measuring users past punctuality and abiding tendencies to a 8 AM class; **hypothesized significantly associated with outcome;**



- Differences between males and females in terms of daily phone usage;
- **Less time on phones for users during intervention, compared with prior data;**

3. Data Preprocessing

Quality Control

- **Common Categories:**
  - **Invalid:** valid ranges of variables' measurements;
  - **Inconsistency:** variables not marked in the same unit;
  - **Accuracy:** under-reporting of variable information;
  - **Precision:** recording lower values than actual amount;
- **Outliers Detection:**
  - With-and-without outliers analysis on mean calculation;
  - **Minor concern if results show trivial change;**

Variable Creation

- A common choice in practice to derive new variables from raw data to **provide other types of information;**
- **"lag_y1"** or **"lag_y2":**
  - Yesterday's phone usage data (minutes or pickup);
- **"early":**
  - Past proportion of daily 1st pickup time prior 8 AM;
- **"base_mean":**
  - Last 3 weeks' (ahead March 27) mean usage data;

Missing Data

- R-package MICE for multiple imputation procedures;
  - **mice()** function in the MICE package;
  - **"method=pmm"** for continuous variables;
  - **"m=5"** indicates generating five imputed datasets;
- **Separate original data by intervention category;**
  - Producing 5 imputed sets for Intervention A data and Intervention B data respectively using **mice()**;
- **After-imputation analysis:** estimation and inference;

4. Data Analysis

Model Setup

- Response **"compliance"** is binary ↔ logistic regression model
- **"score"**: a continuous variable of a procrastination score value from an online test;
- **Intervention A Model:**
- Model A: $\mathbf{logit(\pi) = \beta_0 + \beta_1 \cdot (lag\_y1) + \beta_2 \cdot (base\_mean) + \beta_3 \cdot (score) + \beta_4 \cdot (early)}$
- **Intervention B Model:**
- Model B: $\mathbf{logit(\pi) = \beta_0 + \beta_1 \cdot (lag\_y2) + \beta_2 \cdot (base\_mean) + \beta_3 \cdot (score) + \beta_4 \cdot (early)}$
- **Hypothesis Testing:**

  Null Hypothesis: $\beta_4 = 0$ & Alternative Hypothesis: $\beta_4 \neq 0$
- **Focus on observations in intervention program period** (March 27 – April 2);

Model Fitting Steps

- Fitted Model A on 5 generated imputed sets from Intervention A data;
  - Obtained 5 fitted GLMs' summary statistics;
- Fitted Model B on 5 generated imputed sets from Intervention B data;
  - Obtained 5 fitted GLMs' summary statistics;
- Summary statistics — e.g., estimates, standard errors —- could used to **manually compute the key metrics of missing data imputation diagnostics;**
  - **pool()** function is an alternative option to deliver key metrics;
  - also returning model summary by merging all the fitted GLMs;
- **Results:**

# of Model	β_0	β_1	β_2	β_3	β_4
Model A P-Values	0.00011	$1.15544e^{-9}$	0.43219	0.28683	0.08939
Model B P-Values	0.02222	$3.11952e^{-9}$	0.70035	0.84552	0.03543

Key Findings

- **$\beta_1 < 0.05$ in both Model A & B, $\beta_4 > 0.05$ in Model A**, and **$\beta_4 < 0.05$ in Model B**;
  - **Lag-1 effects significantly associated with both intervention compliance;**
  - **Prior abiding habit significant associated with Intervention B compliance;**

Imputation Diagnostics

- Show **missing data imputation effectiveness** after generating imputed datasets;
- Using **Rubin's rule**[2] can find out key metrics regarding imputation diagnostics;
  - **RI:** represents **a proportional increase in total sampling variance** compared with its sampling variance would have had the missing values been complete;
  - **FMI:** exhibits **a proportion of total sampling variance** due to a missing variable and **how correlated** this variable with others in imputation model;
  - **RE:** provides **an accuracy of estimating true population parameters** and **an estimate of efficiency compared to infinite number of imputations;**

# of Model's β	Relative Increase(RI)	Fraction Missing Info(FMI)	Relative Efficiency(RE)
Model A: `lag_y1`	0.03872	0.03795	0.99247
Model A: `base_mean`	0.08989	0.08559	0.98317
Model A: `score`	0.02696	0.02659	0.99471
Model A: `early`	0.23682	0.20591	0.96045
Model B: `lag_y2`	0.05541	0.05381	0.98935
Model B: `base_mean`	0.01973	0.01954	0.99611
Model B: `score`	0.08590	0.08198	0.98387
Model B: `early`	0.33809	0.27543	0.94779

5. Future Directions

- Longitudinal analysis given **longitudinal nature of repeated-measures data;**
- Expectation–Maximization **(EM) algorithm for handling missing values;**
- Dynamic treatment regimes to **design individualized intervention policies;**

References

[1] I. O. Adam and M. D. Alhassan. *The effect of mobile phone penetration on the quality of life.* Telecommunications Policy, vol. 45, no. 4, p. 102109, 2021.

[2] D. Rubin. *Multiple Imputation for Nonresponse in Surveys,*. New York: John Wiley and Sons., 1987.