Mobile Health Intervention Optimization

Susan Murphy

09.14.17

JITAIs

MD2K
Center of Excellence for Mobile Sensor Data-to-Knowledge

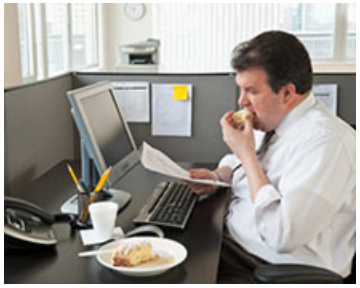
The Methodology Center
advancing methods, improving health

HeartSteps (PI Klasnja)

Goal: Develop an mobile activity coach for individuals who have coronary artery disease

Three iterative studies:

- 42 day micro-randomized pilot study with sedentary individuals,
- 90 day micro-randomized study,
- 365 day personalized study



HeartSteps

Context provided via data from:

Wearable band → minute level step count, bouts of activity and sleep quality;

Smartphone sensors → busyness of calendar, location at pre-specified time points during day, weather;

Self-report → nightly stress, user burden

In which contexts should the smartphone provide the user with a tailored activity suggestion?

Structure of a mobile health intervention that uses wearable devices to sense the context and deliver treatments

Structure of Mobile Health Intervention

1) Decision Points: times, *t*, at which a treatment might be delivered or "pushed"
   1) Regular intervals in time (e.g. every 10 minutes)
   2) At user demand

<u>HeartSteps</u>: approximately every 2-2.5 hours: pre-morning commute, mid-day, mid-afternoon, evening commute, after dinner

Structure of Mobile Health Intervention

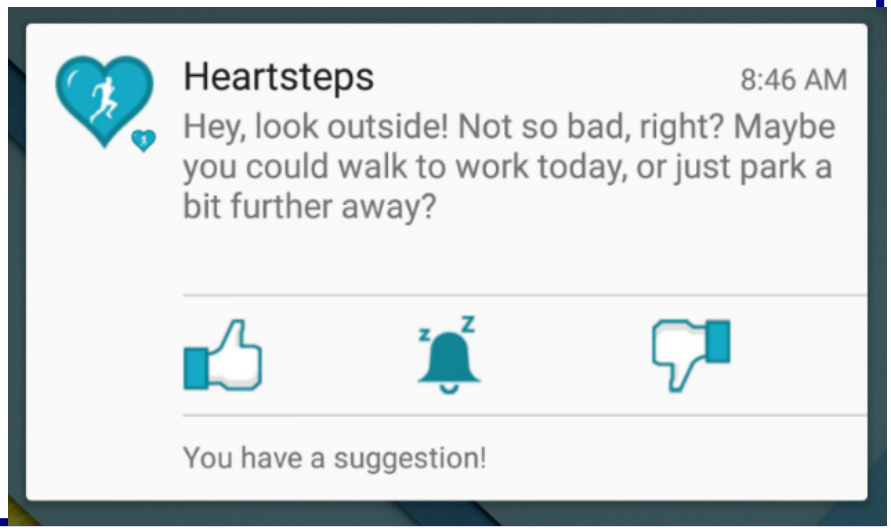
2) Observations at decision point *t*
   1) Passively collected (via sensors)
   2) Actively collected (via self-report)

<u>HeartSteps</u>: classifications of activity, location, weather, step count, busyness of calendar, usefulness ratings, adherence…….

Structure of Mobile Health Intervention

3) Actions A_t
   1) Types of treatments/engagement strategies that can be provided at a decision point, t
   2) Whether to provide a treatment

<u>HeartSteps</u>: tailored activity suggestion (yes/no)



Availability

Activity suggestions can only be delivered if the individual is currently *available*. -- I_t=1 if available, I_t=0 if not

- Unavailability is not the same as nonadherence!

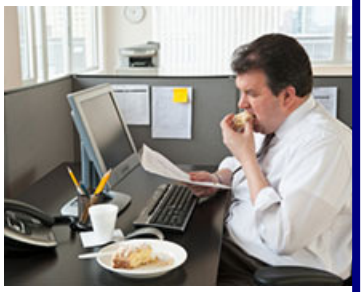
<u>HeartSteps</u>: Unavailable if sensors indicate that the individual may be operating a vehicle, is walking or has turned off the intervention.

Structure of Mobile Health Intervention

4) Proximal Outcome Y_{t+1}

<u>HeartSteps</u>: Step count over 30 minutes following decision point, t

Treatment Effects



The tailored activity suggestions are designed to be near-term actionable: to impact activity in the near term.

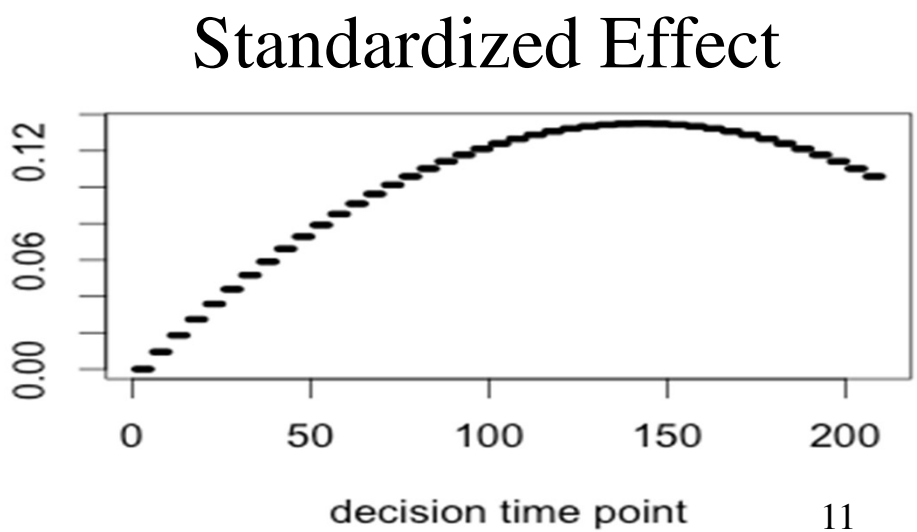
Does the tailored activity suggestion influence step count in the subsequent 30 minutes?

- Does this effect deteriorate over time?

Treatment Effect

Effect of activity suggestion on step count is likely time-varying

What does this effect mean?



HeartSteps
Micro-Randomized Trial

On each participant, randomize delivery of a mobile intervention component (activity suggestion) each time that component *may be* delivered.

Activity suggestion (210 randomizations)

- If available, provide an activity suggestion with probability .6; do nothing with probability .4

Data analyses following the micro-randomized mobile health trial

Conceptual Models

Generally data analysts fit a series of increasingly more complex models:

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t$$

and then next,

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t$$

and so on…

- Y_{t+1} is subsequent activity over next 30 min.
- $A_t = 1$ if activity suggestion and 0 otherwise
- Z_t summaries formed from t and past/present observations
- S_t potential moderator (e.g., current weather is good or not)

Conceptual Models

Generally data analysts fit a series of increasingly more complex models:

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t$$

and then next,

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t$$

and so on…

$\alpha_0 + \alpha_1^T Z_t$ is used to reduce the noise variance in Y_{t+1} (Z_t is sometimes called a vector of control variables)

Causal Effects

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t$$

β_0 is the effect, marginal over all observed and all unobserved variables, of the activity suggestion on subsequent activity

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t$$

$\beta_0 + \beta_1$ is the effect when the weather is good (S_t=1), marginal over other observed and all unobserved variables, of the activity suggestion on subsequent activity.

Goal

- Develop data analytic methods that are consistent with our scientific understanding of the meaning of the β coefficients

- Challenges:
  - Time-varying treatment (A_t, $t=1,\ldots T$)
  - "Independent" variables: Z_t, S_t, I_t that may be affected by prior treatment
- Robustly facilitate noise reduction via use of controls, Z_t

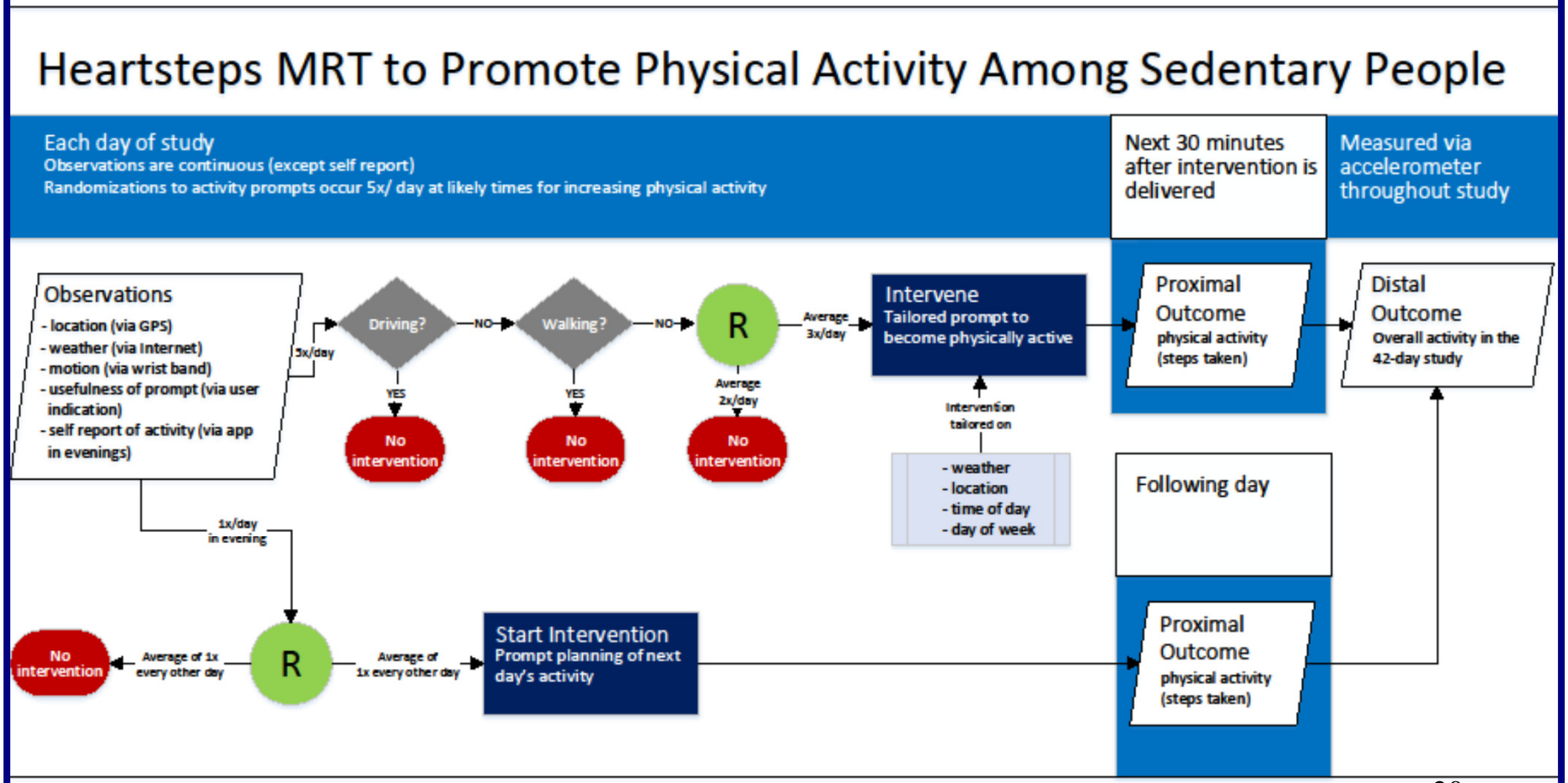
"Centered and Weighted Least Squares Estimation"

- Simple method for complex data!
- Enables unbiased inference for a causal, marginal, treatment effect (the β's)
- Inference for treatment effect is not biased by how we use the controls to reduce the noise variance in Y_{t+1}

https://arxiv.org/abs/1601.00237

Application of the “Centered and Weighted Least Squares Estimation” method in an initial analysis of HeartSteps

HeartSteps Micro-Randomized Trial



On each of $n=37$ participants:

a) Activity suggestion, A_t

- **Provide a suggestion with probability .6**
  - a tailored sedentary-reducing activity suggestion (probability=.3)
  - a tailored walking activity suggestion (probability=.3)
- **Do nothing (probability=.4)**

- 5 times per day * 42 days= 210 decision points

Conceptual Models

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_t$$

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \alpha_2 d_t + \beta_0 A_t + \beta_1 A_t d_t$$

- $t=1,\ldots T=210$
- Y_{t+1} = log-transformed step count in the 30 minutes *after* the tth decision point,
- $A_t = 1$ if an activity suggestion is delivered at the tth decision point; $A_t = 0$, otherwise,
- Z_t = log-transformed step count in the 30 minutes *prior* to the tth decision point,
- d_t =days in study; takes values in (0,1,….,41)

HeartSteps Analysis

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_t \text{, and}$$

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \alpha_2 d_t + \beta_0 A_t + \beta_1 A_t d_t$$

Causal Effect Term	Estimate	95% CI	p-value
$\beta_0 A_t$ *(effect of an activity suggestion)*	$\hat{\beta}_0$=.13	(-0.01, 0.27)	.06
$\beta_0 A_t + \beta_1 A_t d_t$ *(time trend in effect of an activity suggestion)*	$\hat{\beta}_0$ = .51	(.20, .81)	<.01
	$\hat{\beta}_1$ = -.02	(-.03, -.01)	<.01

Heartsteps

On each of *n=37* participants:

a) Activity suggestion

- Provide a suggestion with probability .6
  - **a tailored walking activity suggestion (probability=.3)**
  - **a tailored sedentary-reducing activity suggestion (probability=.3)**
- **Do nothing (probability=.4)**

- 5 times per day * 42 days= 210 decision points

HeartSteps Analysis

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_{1t} + \beta_1 A_{2t}$$

- $A_{1t} = 1$ if walking activity suggestion is delivered at the tth decision point; $A_{1t} = 0$, otherwise,
- $A_{2t} = 1$ if sedentary-reducing activity suggestion is delivered at the tth decision point; $A_{2t} = 0$, otherwise,

Causal Effect	Estimate	95% CI	p-value
$\beta_0 A_{1t} + \beta_1 A_{2t}$	$\hat{\beta}_0 = .21$ $\hat{\beta}_1 > 0$	(.04, .39) ns	.02 ns

Initial Conclusions

- The data indicates that there is a causal effect of the activity suggestion on step count in the succeeding 30 minutes.
  - This effect is primarily due to the walking activity suggestions.
  - This effect deteriorates with time
  - The walking activity suggestion initially increases step count over succeeding 30 minutes by approximately 271 steps but by day 20 this increase is only approximately 65 steps.

On each of $n=37$ participants:

b) Evening planning prompt, A_t

- **Provide a prompt with probability .5**
  - Prompt using unstructured activity planning for following day with probability=.25
  - Prompt using structured activity planning for following day with probability=.25
- **Do nothing with probability=.5**

- 1 time per day * 42 days= 42 decision points

Conceptual Models

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_t$$

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \alpha_2 W_t + \beta_0 A_t W_t + \beta_1 A_t (1-W_t)$$

- $t=1,\ldots T=42$
- Y_{t+1} = square root-transformed step count on the day *after* the tth day,
- $A_t = 1$ if activity planning prompt on the evening of the tth day; $A_t = 0$, otherwise,
- Z_t = square-root step count on the tth day,
- $W_t = 1$ if Sunday through Thursday; $W_t = 0$, otherwise

HeartSteps Analysis

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_t \text{, and}$$

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \alpha_2 W_t + \beta_0 A_t W_t + \beta_1 A_t (1-W_t)$$

Causal Effect Term	Estimate	95% CI	p-value
$\beta_0 A_t$ *(effect of planning)*	$\hat{\beta}_0 = 1.7$	ns	ns
$\beta_0 A_t W_t + \beta_1 A_t (1-W_t)$ *(effect of planning for weekday ($W_t = 1$) and for weekend($W_t = 0$)*	$\hat{\beta}_0 = 3.6$ $\hat{\beta}_1 < 0$	(.74, 6.4) ns	<.02 ns

On each of *n=37* participants:

b) Evening planning prompt

- Provide a prompt with probability .5
  - **Prompt using unstructured activity planning for following day with probability=.25**
  - **Prompt using structured activity planning for following day with probability=.25**
- **Do nothing with probability=.5**

- 1 time per day * 42 days= 42 decision points

Conceptual Model

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_{1t} + \beta_1 A_{2t}$$

- Y_{t+1} = square root-transformed step count on the day *after* the tth day,
- $A_{1t} = 1$ if *unstructured* activity planning prompt on the evening of the tth day; $A_{1t} = 0$, otherwise,
- $A_{2t} = 1$ if *structured* activity planning prompt on the evening of the tth day; $A_{2t} = 0$, otherwise,
- Z_t = square-root step count on the tth day,

HeartSteps Analysis

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_{1t} + \beta_1 A_{2t} \quad t=0,\ldots T=41$$

- $A_{1t} = 1$ if *unstructured* activity planning prompt on the evening of the tth day; $A_{1t} = 0$, otherwise,
- $A_{2t} = 1$ if *structured* activity planning prompt on the evening of the tth day; $A_{2t} = 0$, otherwise,

Causal Effect Term	Estimate	95% CI	p-value
$\beta_0 A_{1t} + \beta_1 A_{2t}$	$\hat{\beta}_0 = 3.1$ $\hat{\beta}_1 > 0$	(-.22, 6.4) ns	.07 ns

HeartSteps Analysis

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \alpha_2 W_t + \beta_0 A_{1t} W_t + \beta_1 A_{1t}(1-W_t) + \beta_2 A_{2t} W_t + \beta_3 A_{2t}(1-W_t)$$

$A_{1t} = 1$ if *unstructured* activity planning prompt on the evening of the tth day; $A_{1t} = 0$, otherwise,

$W_t = 1$ if Sunday through Thursday; $W_t = 0$, otherwise

Causal Effect	Estimate	95% CI	p-value
$\beta_0 A_{1t} W_t + \beta_1 A_{1t}(1-W_t) + \beta_2 A_{2t} W_t + \beta_3 A_{2t}(1-W_t)$	$\hat{\beta}_0 = 5.3$ $\hat{\beta}_1 < 0$, $\hat{\beta}_2 > 0$, $\hat{\beta}_3 < 0$	(2.2, 8.5) all ns	<.01 all ns

Initial Conclusions

- The data indicates that there is a causal effect of planning the next day's activity on on the following day's step count
  - This effect is due to the unstructured planning prompts.
  - This effect occurs primarily on weekdays
  - On weekdays the effect of an unstructured planning prompt is to increase step count on the following day by approximately 780 steps.

Discussion

Problematic Analyses

- GLM & GEE analyses
- Random effects models & analyses
- Machine Learning Generalizations:
  - Partially linear, single index models & analysis
  - Varying coefficient models & analysis

--These analyses do not take advantage of the micro-randomization. Can accidentally eliminate the advantages of randomization for estimating causal effects--

Discussion

- Randomization enhances:
  - Causal inference based on minimal structural assumptions
- Challenge:
  - How to include random effects which reflect scientific understanding ("person-specific" effects) yet not destroy causal inference?

Continually Learning Mobile Health Intervention

1) Trial Designs: Are there effects of the actions on the proximal response? *experimental design*

2) Data Analytics for use with trial data: Do effects vary by the user's internal/external context,? Are there delayed effects of the actions? *causal inference*

3) Learning Algorithms for use with trial data: Construct a "warm-start" treatment policy. *batch Reinforcement Learning*

4) Online Algorithms that personalize and continually update the mHealth Intervention. *online Reinforcement Learning*

Collaborators!