Optimizing mHealth Interventions

Susan A Murphy
11.16.17
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

HeartSteps: 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)

HeartSteps: 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

**HeartSteps**: 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!

HeartSteps: Unavailable if sensors indicate that the individual may be operating a vehicle, is walking or has turned off the intervention.
4) Proximal Outcome $Y_{t+1}$

**HeartSteps**: 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?
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:

\[ Y_{t+1} \sim \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t \]

and then next,

\[ Y_{t+1} \sim \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t \]

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:

\[ Y_{t+1} \sim \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t \]

and then next,

\[ Y_{t+1} \sim \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t \]

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

\[ Y_{t+1} \sim \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t \]

\( \beta_0 \) is the effect, marginal over all observed and all unobserved variables, of the activity suggestion on subsequent activity.

\[ Y_{t+1} \sim \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t \]

\( \beta_0 + \beta_1 \) 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 $\beta$ coefficients

• Challenges:
  • Time-varying treatment ($A_t$, $t=1,...,T$)
  • “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 $\beta$’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

Heartsteps MRT to Promote Physical Activity Among Sedentary People

Each day of study
Observations are continuous (except self report.)
Randomizations to activity prompts occur 5x7 day at likely times for increasing physical activity

Next 30 minutes after intervention is delivered
Measured via accelerometer throughout study

Observations
- action (via GPS)
  - motion (via wrist band)
  - use of prompts (via user interaction)
  - self report of activity (via self in evening)

Start Intervention
Tailored prompt to become physically active

Proximal Outcome
Physical activity (steps taken)

Distal Outcome
Overall activity in the 60-day study

Following day

Proximal Outcome
Physical activity (steps taken)

No intervention
Average of 1x every other day

Start Intervention
Prompt planning of next day's activity

Driving?
YES
- No intervention

Walking?
YES
- Stop
- No intervention

Average 3x/day

Average 2x/day

Intervention tailored on
- weather
  - location
  - time of day
  - day of week

1x/day in evening
On each of $n=37$ participants:

a) Activity suggestion, $A_t$
   - **Provide a suggestion with probability $0.6$**
     - 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

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \beta_0 A_t \]
\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \alpha_2 d_t + \beta_0 A_t + \beta_1 A_t d_t \]

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

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \beta_0 A_t, \text{ and} \]

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \alpha_2 d_t + \beta_0 A_t + \beta_1 A_t d_t \]

<table>
<thead>
<tr>
<th>Causal Effect Term</th>
<th>Estimate</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 A_t )</td>
<td>( \hat{\beta}_0 = .13 )</td>
<td>(-0.01, 0.27)</td>
<td>.06</td>
</tr>
<tr>
<td>( \text{effect of an activity suggestion} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_0 A_t + \beta_1 A_t d_t )</td>
<td>( \hat{\beta}_0 = .51 )</td>
<td>(.20, .81)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>( \text{time trend in effect of an activity suggestion} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_1 = -.02 )</td>
<td>(-.03, -.01)</td>
<td>&lt;.01</td>
<td></td>
</tr>
</tbody>
</table>
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

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \beta_0 A_{1t} + \beta_1 A_{2t} \]

- \( A_{1t} = 1 \) if walking activity suggestion is delivered at the \( t^{th} \) decision point; \( A_{1t} = 0 \), otherwise,
- \( A_{2t} = 1 \) if sedentary-reducing activity suggestion is delivered at the \( t^{th} \) decision point; \( A_{2t} = 0 \), otherwise,

<table>
<thead>
<tr>
<th>Causal Effect</th>
<th>Estimate</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 A_{1t} + \beta_1 A_{2t} )</td>
<td>( \hat{\beta}_0 = .21 ) ( \hat{\beta}_1 &gt;0 )</td>
<td>(.04, .39)</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>
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 $0.5$**
  - Prompt using unstructured activity planning for following day with probability $0.25$
  - Prompt using structured activity planning for following day with probability $0.25$

- **Do nothing with probability $0.5$**

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

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \beta_0 A_t \]
\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \alpha_2 W_t + \beta_0 A_t W_t + \beta_1 A_t (1 - W_t) \]

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

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \beta_0 A_t, \text{ and} \]

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \alpha_2 W_t + \beta_0 A_t W_t + \beta_1 A_t (1-W_t) \]

<table>
<thead>
<tr>
<th>Causal Effect Term</th>
<th>Estimate</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 A_t ) (effect of planning)</td>
<td>( \hat{\beta}_0 = 1.7 )</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>( \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 ))</td>
<td>( \hat{\beta}_0 = 3.6 ) ( \hat{\beta}_1 &lt; 0 )</td>
<td>(.74, 6.4)</td>
<td>&lt;.02 ns</td>
</tr>
</tbody>
</table>
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

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \beta_0 A_{1t} + \beta_1 A_{2t} \]

- \( Y_{t+1} = \) square root-transformed step count on the day after the \( t^{th} \) day,
- \( A_{1t} = 1 \) if unstructured activity planning prompt on the evening of the \( t^{th} \) day; \( A_{1t} = 0 \), otherwise,
- \( A_{2t} = 1 \) if structured activity planning prompt on the evening of the \( t^{th} \) day; \( A_{2t} = 0 \), otherwise,
- \( Z_t = \) square-root step count on the \( t^{th} \) day,
HeartSteps Analysis

\[ Y_{t+1} \sim ^{\sim} \alpha_0 + \alpha_1 Z_t + \beta_0 A_{1t} + \beta_1 A_{2t} \quad t=0, \ldots, T=41 \]

- \( A_{1t} = 1 \) if unstructured activity planning prompt on the evening of the \( t \)th day; \( A_{1t} = 0 \), otherwise,
- \( A_{2t} = 1 \) if structured activity planning prompt on the evening of the \( t \)th day; \( A_{2t} = 0 \), otherwise,

<table>
<thead>
<tr>
<th>Causal Effect Term</th>
<th>Estimate</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 A_{1t} + \beta_1 A_{2t} )</td>
<td>( \hat{\beta}_0 = 3.1 ) ( \hat{\beta}_1 &gt;0 )</td>
<td>((-.22, 6.4)) ns</td>
<td>.07 ns</td>
</tr>
</tbody>
</table>
HeartSteps Analysis

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \alpha_2 W_t + \beta_0 A_{1t} W_t + \beta_1 A_{1t} (1-W_t) \]
\[ + \beta_2 A_{2t} W_t + \beta_3 A_{2t} (1-W_t) \]

\( A_{1t} = 1 \) if unstructured activity planning prompt on the evening of the \( t^{th} \) day; \( A_{1t} = 0 \), otherwise,

\( W_t = 1 \) if Sunday through Thursday; \( W_t = 0 \), otherwise

<table>
<thead>
<tr>
<th>Causal Effect</th>
<th>Estimate</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \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) )</td>
<td>( \hat{\beta}_0 = 5.3 ) ( \hat{\beta}_1 &lt; 0, \hat{\beta}_2 &gt; 0, \hat{\beta}_3 &lt; 0 )</td>
<td>(2.2, 8.5)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>all ns</td>
<td>all ns</td>
</tr>
</tbody>
</table>
Initial Conclusions

• The data indicates that there is a causal effect of planning the next day’s activity 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 outcome? *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” mHealth Intervention. *batch Reinforcement Learning*

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