Micro-Randomized Trials & mHealth

S.A. Murphy
ASA Biopharmaceutical Workshop
mHealth

MD2K Smoking Cessation Coach

- Wearable wrist/chest bands measure activity, stress, cigarette smoking…; phone records location, burden,…..
- Supportive stress-regulation interventions available on smartphone 24/7
- In which contexts should the wrist band provide supportive “cue” and smartphone activate to highlight associated support?
mHealth

HeartSteps Activity Coach

- Wearable bands measure activity and sleep quality; phone sensors measure busyness of calendar, location, weather, ….

- In which contexts should smartphone ping and deliver activity ideas?
Data from wearable devices that sense and provide treatments

On each individual:

\[ O_1, A_1, Y_2, \ldots, O_t, A_t, Y_{t+1}, \ldots \]

\( O_t \): Observations at \( t^{\text{th}} \) decision time (high dimensional)

\( A_t \): Action at \( t^{\text{th}} \) decision time (treatment)

\( Y_{t+1} \): Proximal Response (aka: Reward, Utility, Cost)
Examples

1) Decision Times  (Times at which a treatment can be provided.)
   1) Regular intervals in time (e.g. every 10 minutes)
   2) At user demand

HeartSteps: Approximately every 2-2.5 hours

Smoking Cessation: Every 1 minute during 10 hour day.
Examples

2) Observations $O_t$
   1) Passively collected (via sensors)
   2) Actively collected (via self-report)

HeartSteps: activity recognition, location, step count, usefulness ratings, adherence……..

Smoking Cessation: stress, smoking detection, mood, driving,…..
Examples

3) Actions $A_t$
   1) Treatments that can be provided at a decision time
   2) Whether to provide a treatment

HeartSteps: Activity recommendation on phone
Smoking Cessation: Cue on wrist band
Activity Recommendation

No Message or
Examples

4) Proximal Response (reward) \( Y_{t+1} \)

HeartSteps: Activity (step count) over next 60 minutes.
Smoking Cessation: Stress level over next x minutes
mHealth

HeartSteps Activity Coach

- Wearable bands measure activity and sleep quality; phone sensors measure busyness of calendar, location, weather, …..

- In which contexts should smartphone ping and deliver activity ideas?
Experimental Design: “Micro-Randomized Trial”

Randomize between actions at decision times → Each person may be randomized 100’s or 1000’s of times.

- These are sequential, “full factorial,” designs.
- Design trial to detect main effects.
Why Micro-Randomization?

• Treatment actions are often designed to have a near-time, proximal effect.
  – Randomization is the gold standard for providing data to assess the causal effect of a treatment
• Factorial designs are the gold standard when collecting data to build an intervention involving many treatment factors
• Sequential randomization will enhance quality of many interesting subsequent data analyses.
Micro-Randomized Trial Elements

1. **Record outcomes**
   - Distal (scientific/clinical goal) & Proximal Response
2. **Record** context (sensor & self-report data)
3. **Randomize** among treatment actions at decision points
4. **Use** data after study ends to assess treatment effects, learn warm-start treatment policy
Micro-Randomized Trial

How to justify the trial costs?

• Address a question that can be stated clearly across disciplinary boundaries and be able to provide guarantees.

• Design trial so that a variety of further interesting questions can be addressed.

First Question to Address: Do the treatment actions impact the proximal response? (aka, is there a main effect?)
Micro-Randomized Trial for HeartSteps

• 42 day trial
• Whether to provide an Activity recommendation? $A_t \in \{0, 1\}$
• Test for main effects on proximal response
• Randomization in HeartSteps

$$P[A_t = 1] = .4 \quad t = 1, \ldots, T$$
Time-varying Main Effects

Time varying potentially intensive/intrusive treatment actions $\rightarrow$ potential for accumulating habituation and burden

$\rightarrow$

In the test statistic allow the main effect of the treatment actions on proximal response to vary with time
Availability

• Treatment actions can only be delivered at a decision time if an individual is *available*

• Set $I_t=1$ if the individual is available at decision time $t$, otherwise, $I_t=0$
Potential Outcomes

• Define

\[ \bar{A}_t = \{A_1, A_2, \ldots, A_t\}, \quad \bar{a}_t = \{a_1, a_2, \ldots, a_t\} \]

• Define \( Y_{t+1}(\bar{a}_t) \) to be the observed response, \( Y_{t+1} \) if \( \bar{A}_t = \bar{a}_t \), e.g., \( Y_{t+1} = Y_{t+1}(\bar{A}_t) \)

• Define \( I_t(\bar{a}_{t-1}) \) to be the observed “available for treatment” indicator if \( \bar{A}_{t-1} = \bar{a}_{t-1} \)
Main Effect

• Define the Main Effect at time $t$ as

$$E[Y_{t+1}(\bar{A}_{t-1}, 1) - Y_{t+1}(\bar{A}_{t-1}, 0) | I_t(\bar{A}_{t-1}) = 1]$$

• What does this estimand mean?
Main Effect

- The randomization implies that

\[
E\left[Y_{t+1}(\bar{A}_{t-1}, 1) - Y_{t+1}(\bar{A}_{t-1}, 0) \mid I_t(\bar{A}_{t-1}) = 1\right] = \\
E\left[Y_{t+1} \mid I_t = 1, A_t = 1\right] - E\left[Y_{t+1} \mid I_t = 1, A_t = 0\right]
\]

- Put

\[
\beta(t) = E\left[Y_{t+1} \mid I_t = 1, A_t = 1\right] - E\left[Y_{t+1} \mid I_t = 1, A_t = 0\right]
\]
Proposal

*Design and size micro-randomized trial to detect main effect of treatment on proximal response*

The main effect is a time-varying main effect $\beta(t), \ t=1,...,T$

The main effect is a causal effect!
Sample Size Calculation

• We calculate the number of subjects to test:
  \[ H_0 : \beta(t) = 0, \quad t = 1, 2, \ldots, T \]

• Size to detect a low dimensional, smooth alternate \( H_1 \).
  – Example: \( H_1 : \beta(t) \) quadratic with intercept, \( \beta_0 \), linear term, \( \beta_1 \), and quadratic term \( \beta_2 \) and test
    \[ \beta_0 = \beta_1 = \beta_2 = 0 \]
Test Statistic for Sample Size Calculation

• Test statistic based on fit of model:

\[ E[Y_{t+1}|I_t = 1, A_t] = \gamma(t) + \beta(t)(A_t - q_t) \]

where \( q_t \) is the randomization probability

• \( q_t = .4 \) in HeartSteps
Test Statistic for Sample Size Calculation

- Test statistic is based on least squares fit of

\[ E[Y_{t+1} | I_t = 1, A_t] = \gamma(t) + \beta(t)(A_t - q_t) \]

HeartSteps:

\[ \beta(t) = \beta_0 + \beta_1 \left\lfloor \frac{t-1}{5} \right\rfloor + \beta_2 \left\lfloor \frac{t-1}{5} \right\rfloor^2 \]

- You select parameterization of \( \gamma(t) \)
Alternative for Sample Size Calculation

• One calculates a sample size to detect a given alternative with a given power.

• Alternative:

$$H_1 : \beta_i = d_i \bar{\sigma}, \ i = 0, 1, 2$$

where $\bar{\sigma}^2$ is the average conditional variance.
Alternative for Sample Size Calculation

- Average conditional variance is

\[ \bar{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} E \left[ VAR(Y_{t+1} | I_t = 1, A_t) \right] \]
Sample Size Calculation

Alternative $H_1$ to a null, $H_0$, main effect is parameterized →

estimation uses both between person contrasts as well as within person contrasts →

sample sizes are small.
Specify Alternative for Sample Size Calculation

• Scientist indirectly specifies standardized $d_i$’s
  – initial proximal treatment effect: $d_0$,
  – average proximal effect over trial duration:
    \[
    \frac{1}{T} \sum_{t=1}^{T} \left( d_0 + d_1 \left\lfloor \frac{t-1}{5} \right\rfloor + d_2 \left\lfloor \frac{t-1}{5} \right\rfloor^2 \right),
    \]
  – and day of maximal proximal effect: $- \frac{d_1}{2d_2}$

• We solve for $d_0$, $d_1$, $d_2$
Test Statistic for Sample Size Calculation

• Put $Y_i = (Y_{i2}, \ldots, Y_{iT+1})^T$ for $i^{th}$ subject

$p$ is the total number of parameters ($p > 3$);
$X_i$ is the associated design matrix ($T$ by $p$)
$N$ is sample size

Last 3 columns of $X_i$ contain row entries:

$$I_{it}(A_{it} - q_t), I_{it}(A_{it} - q_t)\left[\frac{t-1}{5}\right],$$
$$I_{it}(A_{it} - q_t)\left[\frac{t-1}{5}\right]^2$$
Test Statistic for Sample Size Calculation

• “GEE” test statistic is

\[ N \hat{\beta}^T (K \hat{\Sigma} K^T)^{-1} \hat{\beta} = N \hat{\beta}^T (\hat{\Sigma}_\beta)^{-1} \hat{\beta} \]

where \( \hat{\Sigma} \) is the usual sandwich estimator of the variance-covariance and \( K \) is 3 by \( 3+q \) matrix picking out columns associated with coefficients \( \beta \)
Working Assumptions for Sample Size Calculation

1) \( E(\epsilon_{it}\epsilon_{it'} \mid I_{it} = 1, I_{it'} = 1, A_{it}, A_{it'}) \) is constant.

2) \( E(\epsilon_{it} \mid I_{it} = 1, A_{it}) = 0 \)

3) \( Var(\epsilon_{i,t} \mid I_{it} = 1, A_{it}) \) is constant.

\[
\epsilon_{it} = Y_{i,t+1} - \left( \gamma(t) + \beta(t)(A_{it} - q_t) \right)
\]
Sample Size Calculation

- Under the working assumptions, $\Sigma_{\beta}$ only depends on polynomials in $[\frac{t-1}{5}]$, the marginal distribution of $I_t$ and on the randomization probabilities.

- $\Sigma_{\beta}$ does not depend on the form of $\gamma(t)$
Sample Size Calculation

- Under standard moment assumptions, the asymptotic distribution of the “GEE test statistic” is a Chi-Squared on 3 degrees of freedom with non-centrality parameter:

\[ N d^T (\Sigma_\beta)^{-1} d \]

- Instead of a Chi-Squared on 3 degrees we use

\[ \frac{3(N-q-1)}{N-q-3} F_{3,N-q-3} \]

with the same non-centrality parameter.
HeartSteps Example

- Standardized $d_i$’s
  - initial proximal effect: $d_0 = 0$
  - output average proximal effect
  - day of maximal proximal effect: $- \frac{d_1}{2d_2} = 28$

- Model for test statistic:

$$
\gamma(t) + \beta(t)(A_{it} - .4), \ t = 1, \ldots, 210
$$

where

$$
\gamma(t) = \gamma_0 + \gamma_1 \left[ \frac{t-1}{5} \right] + \gamma_2 \left[ \frac{t-1}{5} \right]^2
$$
<table>
<thead>
<tr>
<th>Standardized Average Main Effect over 42 Days</th>
<th>Sample Size For 70% availability or 50% availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.06 standard deviations</td>
<td>81 or 112</td>
</tr>
<tr>
<td>0.08 standard deviations</td>
<td>48 or 65</td>
</tr>
<tr>
<td>0.10 standard deviations</td>
<td>33 or 43</td>
</tr>
</tbody>
</table>
Same Test Statistic for Analysis

- “GEE” test statistic is
  
  $$N\hat{\beta}^T (K\hat{\Sigma}K^T)^{-1}\hat{\beta}$$

  where $K$ is 3 by $3+q$ matrix picking out columns associated with $\beta$ coefficients

- No working assumptions
Small Sample Adjustment

- $\hat{e}_{it}$ is the $i^{th}$ subject, $t^{th}$ time point residual and
  $\hat{e}_i = (\hat{e}_{i1}, \ldots, \hat{e}_{iT})^T$

- Adjusted sandwich estimator:
  $\hat{\Sigma} = \hat{\sigma}^2 N \left( \sum_{i=1}^{N} X_i^T X_i \right)^{-1} \left\{ \sum_{i=1}^{N} X_i^T B_i \hat{e}_i \hat{e}_i^T B_i X_i \right\} \left( \sum_{i=1}^{N} X_i^T X_i \right)^{-1}$

  $B_i = (I - H_{i\cdot})^{-1}$
### Simulation Results

#### Type 2 Error Rate (2000 data sets)

<table>
<thead>
<tr>
<th>Average Proximal Effect (Sample Size)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05(115)</td>
<td>0.790</td>
</tr>
<tr>
<td>0.06(81)</td>
<td>0.794</td>
</tr>
<tr>
<td>0.07(61)</td>
<td>0.800</td>
</tr>
<tr>
<td>0.08(48)</td>
<td>0.801</td>
</tr>
<tr>
<td>0.09(39)</td>
<td>0.798</td>
</tr>
<tr>
<td>0.10(33)</td>
<td>0.803</td>
</tr>
</tbody>
</table>
The micro-randomized trial is a sequential factorial trial with multiple factors, e.g. in HeartSteps:

Factor 1: Activity recommendation is randomized 5 times per day (yes/no)

Factor 2: Daily activity planning is randomized each evening (yes/no)
Experimental Design Challenges

Micro-randomized trials are a new type of factorial design

i. Time varying factors → time varying main effects, time-varying two-way interactions, different delayed effects

ii. Better trial designs?

iii. Design studies specifically to detect interactions between factors or delayed effects.
Collaborators