“SMART” Clinical Trial Designs for Dynamic Treatment Regimes

Bibhas Chakraborty
Centre for Quantitative Medicine, Duke-NUS Graduate Medical School, Singapore

bibhas.chakraborty@duke-nus.edu.sg

MCP Conference
Hyderabad
September 4, 2015
Personalized Medicine

Believed by many as the future of medicine ...

Source: http://www.personalizedmedicine.com/

Often refers to tailoring by genetic profile, but it’s also common to personalize based on more “macro” level characteristics, some of which are time-varying
Personalized Medicine

- Paradigm shift from “one size fits all” to individualized, patient-centric care
  - Can address inherent heterogeneity across patients
  - Can also address variability within patient, over time
  - Can increase patient compliance, thus increasing the chance of treatment success
  - Likely to reduce the overall cost of health care

- Big methodological questions:
  - How to decide on the optimal treatment for an individual patient?
  - How to make these treatment decisions evidence-based or data-driven?
Outline

1. Dynamic Treatment Regimens (Regimes): What and Why?

2. Sequential Multiple Assignment Randomized Trial (SMART) Design
   - SMART Design: What and Why?
   - Primary Hypothesis and Sample Size


4. Discussion
Consider **personalized** management of **chronic** diseases

**Management of chronic diseases require ongoing interventions**

A **dynamic treatment regimen (DTR)** is a sequence of decision rules, one per stage of clinical intervention, that specify how to adapt the **type, dosage and timing of treatment** according to individual patient’s ongoing **response, adherence, burden, side effects, preference, and past treatments**

- Each decision rule takes a patient’s treatment and covariate history as inputs, and outputs a recommended treatment
ADHD Example: Treatment Scenarios


BMOD = Behavioral Modification Therapy
One Simple DTR

“Give **Low-intensity BMOD** as initial treatment; if the subject *responds*, then continue **BMOD**, otherwise prescribe **BMOD + MEDS**”
The Big Scientific Questions in DTR Research

- What would be the **mean outcome** if the population were to follow a particular pre-conceived DTR?

- How do the mean outcomes **compare** among two or more DTRs?

- What is the **optimal** DTR in terms of the mean outcome?
  - What is the best **sequencing** of treatments?
  - What are the best **timings** of alterations in treatments?
  - How do we best **personalize** the sequence of treatments? i.e. What **individual information (tailoring variables)** do we use to make these decisions?
The Big Statistical Questions

1. What is the right kind of data for comparing two or more DTRs, or estimating optimal DTRs? What is the appropriate study design?
   - Sequential Multiple Assignment Randomized Trial (SMART)

2. How can we compare pre-conceived, embedded DTRs?
   - primary analysis of SMART data

3. How can we estimate the “optimal” DTR for a given patient?
   - secondary analysis of SMART data
   - e.g. Q-learning, a stagewise regression-based approach
Outline

1 Dynamic Treatment Regimens (Regimes): What and Why?

2 Sequential Multiple Assignment Randomized Trial (SMART) Design
   - SMART Design: What and Why?
   - Primary Hypothesis and Sample Size

3 Secondary Analysis of SMART: Finding Optimal DTR

4 Discussion
Sequential Multiple Assignment Randomized Trial (SMART)

- Multi-stage trials with a goal to inform the development of DTRs (*Lavori and Dawson, 2000, 2004; Murphy, 2005*)

- Same subjects participate **throughout** (they are followed through stages of treatment)

- Each stage corresponds to a treatment decision

- At each stage the patient is **randomized** to one of the available treatment options

- Treatment options at randomization may be **restricted** on ethical grounds, depending on intermediate outcome and/or treatment history
Primary Outcome: Teacher-rated Impairment Rating Scale (TIRS)
Other Examples of SMART Studies

- **Schizophrenia:** CATIE (*Schneider et al.*, 2001)
- **Depression:** STAR*D (*Rush et al.*, 2003)
- **ADHD:** *Pellham et al.* (see, e.g., *Lei et al.*, 2012)
- **Prostate Cancer:** Trials at MD Anderson Cancer Center (e.g., *Thall et al.*, 2000)
- **Leukemia:** CALGB Protocol 8923 (see, e.g., *Wahed and Tsiatis*, 2004)
- **Smoking:** Project Quit (*Strecher et al.*, 2008)
- **Alcohol Dependence:** *Oslin et al.* (see, e.g., *Lei et al.*, 2012)

Recent examples at the Methodology Center, Pennsylvania State University website:  
http://methodology.psu.edu/ra/smart/projects
Can SMART be replaced by a series of RCTs?

Why not choose the best initial treatment on the basis of a standard RCT of initial treatments, and choose the best secondary treatment on the basis of another RCT of secondary treatments?

- Particularly attractive since potential initial treatment may have been evaluated in prior trials – so the current investigator has to only propose a responder study or a non-responder study

Combining best treatment for each stage, obtained from separate randomized trials, does not necessarily give the optimal DTR, due to delayed effects!
Example: Treating MDD

- Suppose we wish to compare both front-line and second-line treatment of major depressive disorder (MDD):
  - Front-line options: citalopram (Cit) or cognitive behavioral therapy (CBT)
  - Second-line options: treatment switch to Cit, CBT, or Lithium (Li)
  - All responders to first-line therapy will continue with maintenance and follow-up
Example: Treating MDD

Sequential Multiple Assignment Randomized Trial (SMART) Design

SMART Design: What and Why?

Maintenance dose +
telephone monitoring

Responder

Non-responder

CBT
Cit
R

Telephone monitoring

Responder

Non-responder

CBT
Li
R

Cit
Li
R

CBT
Cit
R

Li
Example: Treating MDD

Sequential Multiple Assignment Randomized Trial (SMART) Design

SMART Design: What and Why?

Maintenance dose + telephone monitoring

Responder

Non-responder

Telephone monitoring

Responder

Non-responder

CBT

Cit

R

Li

R

CBT

Li

R

Cit

Li
Front-line Treatment of MDD

- Suppose we observe 60% response with Cit, and only 50% with CBT
- Conclude: Cit is the best front-line therapy
- Now run another one-stage trial amongst Cit non-responders
Second-line Treatment of MDD

- We now observe 40% response to **CBT** and 20% to **Li**

- Conclude: **CBT** is the best second-line therapy

- Final treatment sequence: “**Cit**, followed by **CBT** for non-responders”
  - Under this regimen, we expect to see 76% of patients respond
**Delayed Effect**

- What if initial treatment with **CBT** increases treatment adherence ⇒ subsequent therapies more successful?

![Diagram showing treatment outcomes]

- Optimal DTR: “**CBT, followed by Cit for non-responders**”; 80% response expected
Primary and Secondary Hypotheses

- Choose scientifically important primary hypotheses that also aid in developing DTRs
  - Power trial to address these hypotheses

- Depending on the research question, the primary analysis can be a comparison of two or more means (or, proportions) corresponding to two or more DTRs embedded in the SMART, or components thereof

- Choose secondary hypotheses that further develop the DTR, and use randomization to eliminate confounding
  - Trial is not necessarily powered to address these hypotheses
  - Still better than post hoc observational analyses
  - Underpowered randomizations can be viewed as pilot studies for future full-blown comparisons
Primary Hypothesis and Sample Size: Scenario 1

Hypothesize that averaging over the secondary treatments, the initial treatment **BMOD** is as good as the initial treatment **MEDS**

– *Sample size formula is same as that for a two group comparison*
Primary Hypothesis and Sample Size: Scenario 2

Hypothesize that among non-responders a treatment augmentation (BMOD+MEDS) is as good as an intensification of treatment

– *Sample size formula is same as that for a two group comparison of non-responders* (overall sample size depends on the presumed non-response rate)
Primary Hypothesis and Sample Size: Scenario 3

Hypothesize that the “red” DTR is as good as the “green” DTR

- Sample size formula involves a two group comparison of “weighted” means (overall sample size depends on the presumed non-response rate)
Sample Size Requirements

Assume continuous outcome, e.g., TIRS in case of ADHD

Key Parameters:

- Effect Size = \( \frac{\Delta \mu}{\sigma} \) (Cohen’s \( d \))
- Type I Error Rate = \( \alpha = 0.05 \)
- Desired Power = \( 1 - \beta = 0.8 \)
- Initial Response Rate = \( \gamma = 0.5 \)

Trial Size:

<table>
<thead>
<tr>
<th>Effect Size</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>( N_1 = 350 )</td>
<td>( N_2 = \frac{N_1}{(1-\gamma)} = 700 )</td>
<td>( N_3 = N_1 \times (2 - \gamma) = 525 )</td>
</tr>
<tr>
<td>0.5</td>
<td>( N_1 = 128 )</td>
<td>( N_2 = \frac{N_1}{(1-\gamma)} = 256 )</td>
<td>( N_3 = N_1 \times (2 - \gamma) = 192 )</td>
</tr>
<tr>
<td>0.8</td>
<td>( N_1 = 52 )</td>
<td>( N_2 = \frac{N_1}{(1-\gamma)} = 104 )</td>
<td>( N_3 = N_1 \times (2 - \gamma) = 78 )</td>
</tr>
</tbody>
</table>
Outline

1. Dynamic Treatment Regimens (Regimes): What and Why?

2. Sequential Multiple Assignment Randomized Trial (SMART) Design
   - SMART Design: What and Why?
   - Primary Hypothesis and Sample Size


4. Discussion
Secondary Analysis

- **Goal:** To find *optimal* treatment sequence for each individual patient by *deeply tailoring* on their time-varying covariates and intermediate outcomes

- **Methodologically challenging – custom-made analytic tools necessary**
  - One popular approach is **Q-learning**, a stage-wise regression-based method *(Murphy, 2005; Zhao et al., 2009; Chakraborty et al., 2010)*
  - **Nonstandard bootstrap** is necessary for inference *(Chakraborty et al., 2013, 2014; Laber et al., 2015)*
  - We developed an **R package** called **qLearn** *(Xin et al., 2012)* that conducts Q-learning and associated inference (Freely available at CRAN):
    
    `http://cran.r-project.org/web/packages/qLearn/`
Project Quit: A Smoking Cessation Trial (Simplified)

Two-stage Web-based (*eHealth*) Behavioral Intervention Trial for Smoking Cessation conducted at the University of Michigan

**Stage-1 Covariate**: education (≤ high school vs. > high school)

**Stage-1 Intervention**: tailoring of success story, low vs. high
(in addition to free nicotine patch)

**Stage-2 Covariates**: quit status at 6 months (1 = quit, 0 = not quit), months non-smoking over 6 months

**Stage-2 Intervention**: booster prevention vs. control

**Primary Outcome**: months non-smoking over 12 months

---

SMART Design Schematic (Simplified)
Secondary Research Questions

- **Stage-1 question:** (In future) How should a web-based smoking cessation intervention be designed so as to maximize each individual’s chance of quitting over the two stages? Should this intervention be adapted to the smoker’s baseline education?

- **Stage-2 question:** Should the stage-2 intervention be adapted to either the stage-1 intervention the smoker has already received and/or the smoker’s intermediate outcome (e.g., stage-1 quit status)?
Secondary Analysis of SMART: Finding Optimal DTR

Results from Q-learning Analysis

No significant stage-2 treatment effect \( (n = 479) \)

Stage-1 Analysis Summary \( (n = 1848) \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>95% Bootstrap CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>education</td>
<td>0.01</td>
<td>(-0.18, 0.20)</td>
</tr>
<tr>
<td>high vs. low tailoring</td>
<td>0.07</td>
<td>(-0.01, 0.11)</td>
</tr>
<tr>
<td>tailoring:education</td>
<td>-0.11</td>
<td>(-0.24, -0.00)*</td>
</tr>
</tbody>
</table>

- The “highly individually tailored” level of story appears more effective for subjects with less education \( \leq \) high school
- This finding is consistent with that of Strecher et al. (2008) – a logistic regression analysis of the stage-1 quit status (binary) data
Outline

1. Dynamic Treatment Regimens (Regimes): What and Why?

2. Sequential Multiple Assignment Randomized Trial (SMART) Design
   - SMART Design: What and Why?
   - Primary Hypothesis and Sample Size


4. Discussion
SMART vs. Crossover Trial Designs

- Operationally, SMARTs look very similar to crossover trials.

- However, conceptually they are very different.
  - Unlike SMART, Crossover trials aim to evaluate stand-alone treatments, not DTRs.
  - Unlike a crossover trial, treatment allocation in a SMART is typically adaptive to a subject’s intermediate outcome.
  - Crossover trials attempt to “wash out” the “carry-over” effect while SMARTs attempt to capture it.
SMART vs. Usual Adaptive Trial Designs

- SMART is different from usual adaptive trial wherein the design elements (e.g., randomization probabilities) can change during the course of the trial
  - Within-subject vs. between-subject adaptation

- Combination of the two concepts is a topic of current research
  - SMARTs can be made more ethically appealing by incorporating adaptive randomization or sequential elimination
  - In certain modern contexts (e.g., implementation research and mHealth), SMART with Adaptive Randomization (SMART-AR)\(^3\) has been developed recently
  - In general, how best to do this is not known yet

---

Practical Appeal of SMART Designs

- SMARTs can be best viewed as developmental trials
  - Run a SMART to find the optimal DTR
  - Run a confirmatory RCT to compare the optimal DTR with suitable control

- It can be argued that by virtue of the option to alter a non-functioning treatment, SMARTs should enjoy, in principle, greater recruitment success and lower patient dropout compared to standard RCT

- Simple, intuitive primary analysis – less complex than one might think

- A stage-wise regression-based approach called Q-learning can be used for secondary analysis of SMART data to construct evidence-based optimal DTRs for specific patient subgroups
Summary

- DTRs offer a framework for operationalizing, and thus potentially improving, adaptive clinical practice for chronic diseases

- SMART designs are useful for comparing pre-conceived DTRs, as well as generating high quality data that can aid in constructing optimal evidence-based DTRs
  
  - Sample size formulae are available for hypotheses involving components of DTR, as well as entire DTRs, for continuous (and binary) outcomes, as illustrated (Oetting et al., 2011)
  
  - Sample size formulae are also available for survival outcomes (Li and Murphy, 2011)
  
  - SMARTs do not necessarily require larger sample sizes – e.g., the ADHD trial recruited only 155 patients!
Shoot your questions, comments, criticisms, and collaboration request on trial design to: bibhas.chakraborty@duke-nus.edu.sg

Thank you