SMART Designs and Q-learning for Dynamic Treatment Regimens

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

bibhas.chakraborty@duke-nus.edu.sg

Victorian Centre for Biostatistics
Melbourne
May 21, 2015

DUKE NUS
GRADUATE MEDICAL SCHOOL SINGAPORE

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

- Overarching 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): An Overview
2. Sequential Multiple Assignment Randomized Trial (SMART) Design
3. Estimation of Optimal DTRs via Q-learning
4. Non-regular Inference for Parameters indexing Optimal DTRs
   - Adaptive m-out-of-n Bootstrap
   - Simulation Study
5. Analysis of Data from STAR*D, A SMART Study on Depression
6. Discussion

Outline

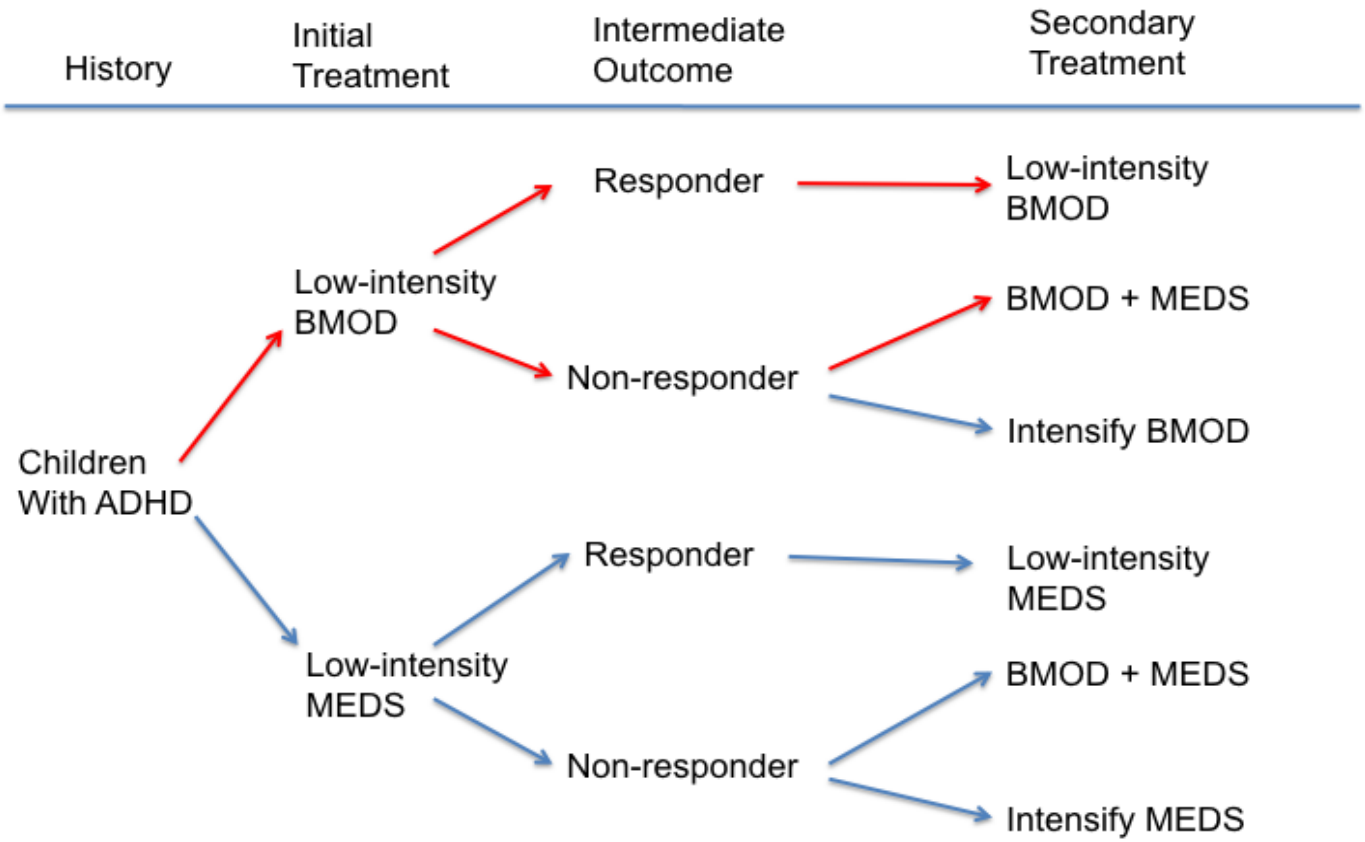
1. Dynamic Treatment Regimens (Regimes): An Overview
2. Sequential Multiple Assignment Randomized Trial (SMART) Design
3. Estimation of Optimal DTRs via Q-learning
4. Non-regular Inference for Parameters indexing Optimal DTRs
   - Adaptive m-out-of-n Bootstrap
   - Simulation Study
5. Analysis of Data from STAR*D, A SMART Study on Depression
6. Discussion

Dynamic Treatment Regimens (DTRs)

- DTRs offer a framework to operationalize personalized medicine in a time-varying setting
  - Clinical decision support systems for treating chronic diseases
- A DTR is a sequence of decision rules
  - Each decision rule takes a patient's treatment and covariate history as inputs, and outputs a recommended treatment
- A DTR is called optimal if it optimizes the long-term mean outcome (or some other suitable criterion)

ADHD Example: One Simple DTR

BMOD: Behavioral Modification Therapy; MEDS: Medication



"Give Low-intensity BMOD as initial treatment; if the subject responds, then continue BMOD, otherwise prescribe BMOD + MEDS"

ADHD Example: One Not-so-simple DTR

- Stage-1 Rule: "If the baseline level of impairment is greater than a threshold (say, ψ), prescribe MEDS; otherwise prescribe BMOD"
- Stage-2 Rule: "If the subject is a responder to initial treatment, continue the same treatment; if non-responder, prescribe BMOD + MEDS"

How to specify ψ?

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

Data Structure

K stages on a single patient:

$$O_1, A_1, \ldots, O_K, A_K, O_{K+1}$$

O_j	:	Observation (pre-treatment) at the j-th stage
A_j	:	Treatment (action) at the j-th stage, $A_j \in \mathcal{A}_j$
H_j	:	History at the j-th stage, $H_j = \{O_1, A_1, \ldots, O_{j-1}, A_{j-1}, O_j\}$
Y	:	Primary Outcome (larger is better)

A DTR is a sequence of decision rules:

$$d \equiv (d_1, \ldots, d_K) \text{ with } d_j(h_j) \in \mathcal{A}_j$$

For simplicity, restrict attention to $K = 2$ and $\mathcal{A}_j = \{-1, 1\}$

Data Sources

- Data from longitudinal observational studies have been widely used in the DTR context
  - This includes electronic medical records data
  - Usual concerns about observational data, e.g. confounding and other hidden biases (Rubin, 1974; Rosenbaum, 1991)
  - Need unverifiable assumptions to make causal inference about treatment effects
  - Analysis is more complex (Robins et al., 2008; Moodie, Chakraborty and Kramer, 2012)
- Better quality Data for estimating optimal DTRs can come from Sequential Multiple Assignment Randomized Trials (SMARTs) (*Lavori and Dawson, 2004; Murphy, 2005*)

In this talk, we will be dealing with SMART data only

Outline

1. Dynamic Treatment Regimens (Regimes): An Overview
2. Sequential Multiple Assignment Randomized Trial (SMART) Design
3. Estimation of Optimal DTRs via Q-learning
4. Non-regular Inference for Parameters indexing Optimal DTRs
   - Adaptive m-out-of-n Bootstrap
   - Simulation Study
5. Analysis of Data from STAR*D, A SMART Study on Depression
6. Discussion

Sequential Multiple Assignment Randomized Trial (SMART)

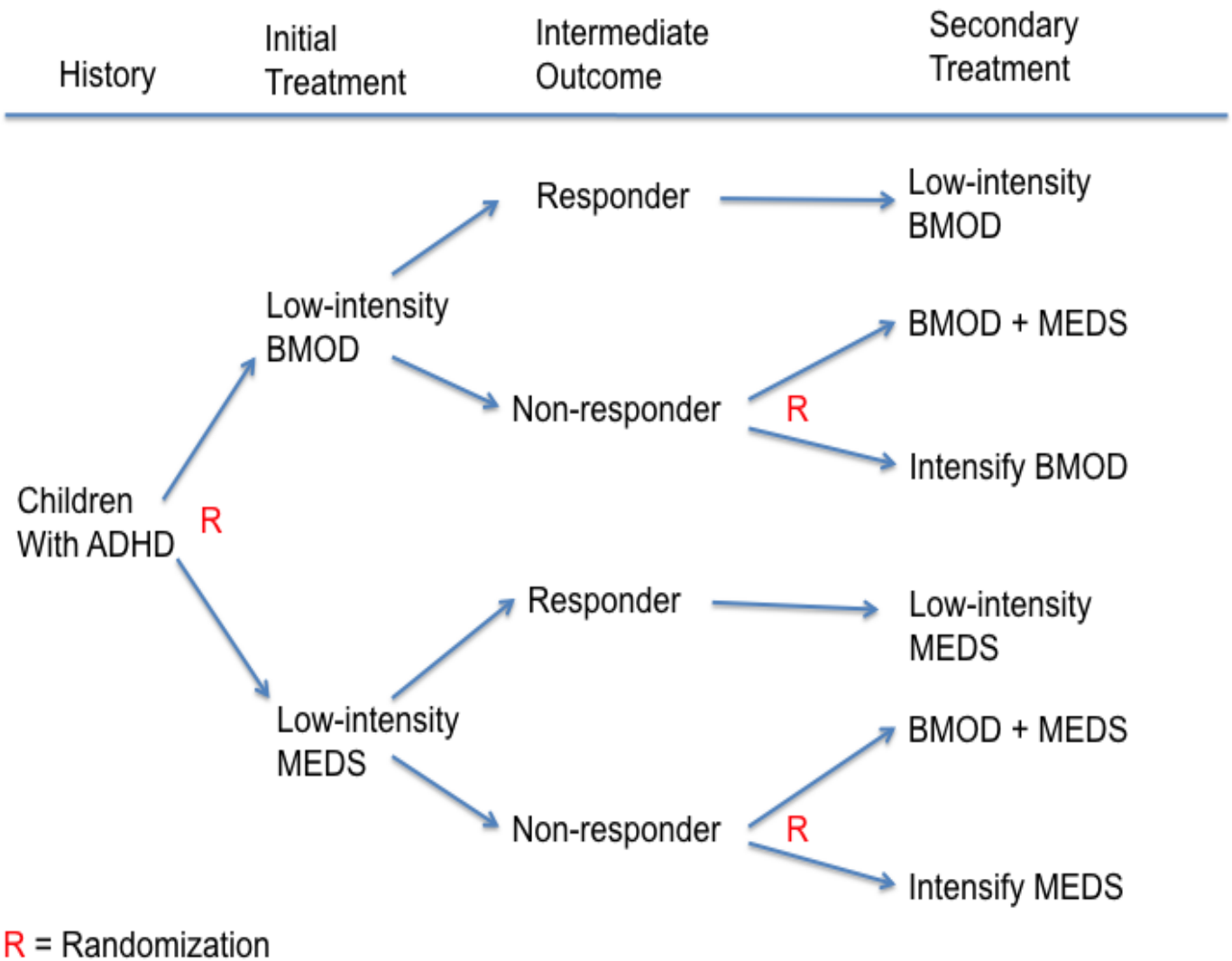
- Multi-stage trials with a goal to inform the development of DTRs
- Same subjects participate throughout (they are followed through stages of treatment)
- Each stage corresponds to a treatment decision
- At each stage the patient is randomized to one of the available treatment options
- Treatment options at randomization may be restricted on ethical grounds, depending on intermediate outcome and/or treatment history

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

A SMART Design in Children with ADHD



Primary Outcome: Teacher-rated Impairment Rating Scale (TIRS)

SMART Design Principles

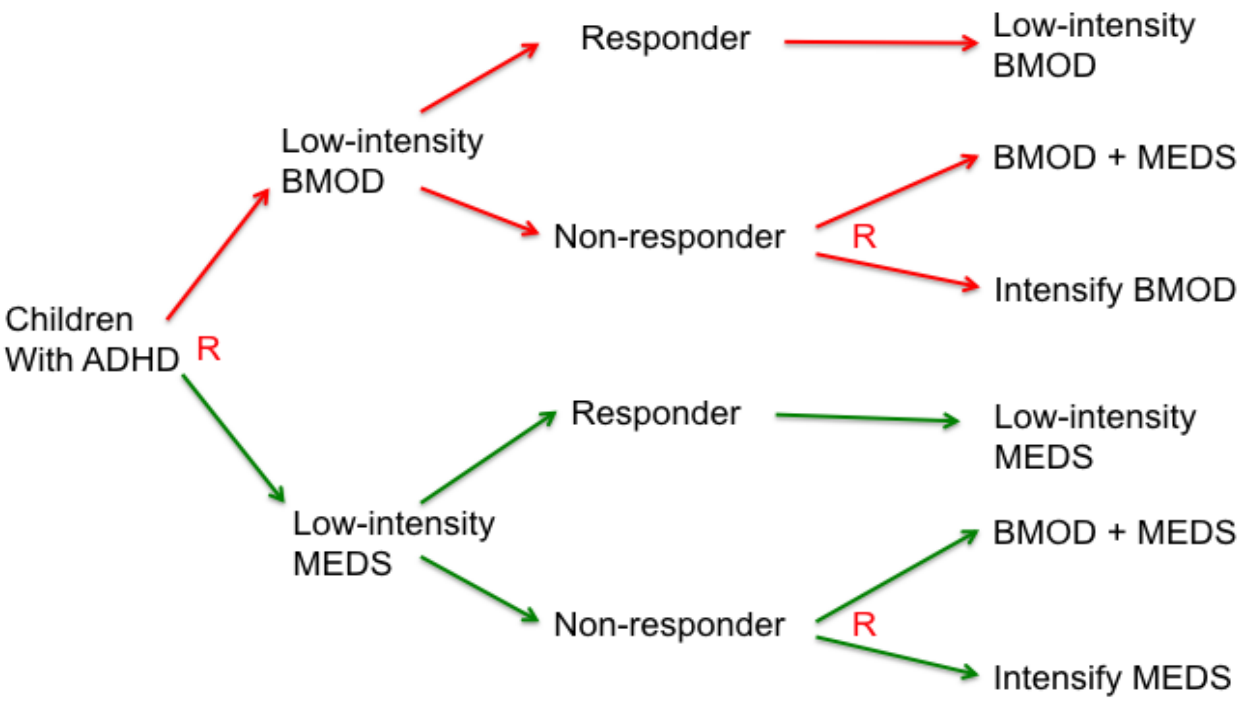
Primary and Secondary Hypotheses

- 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

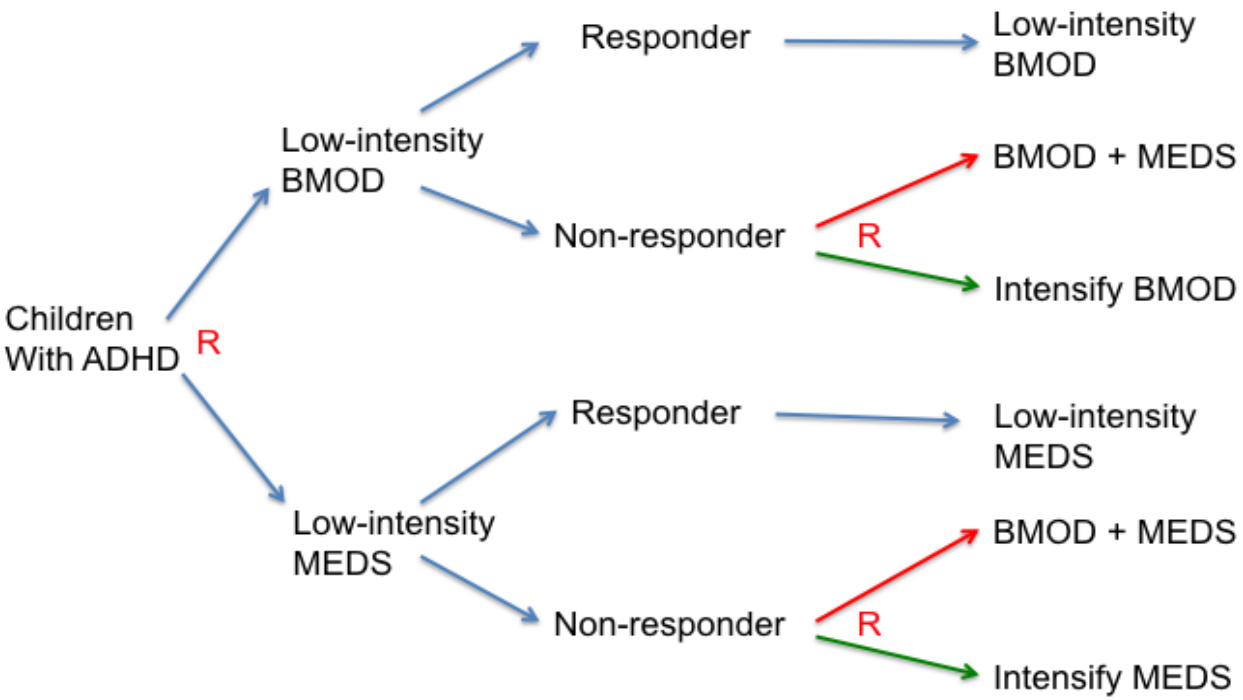
- *Sample size formula is same as that for a two group comparison*



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

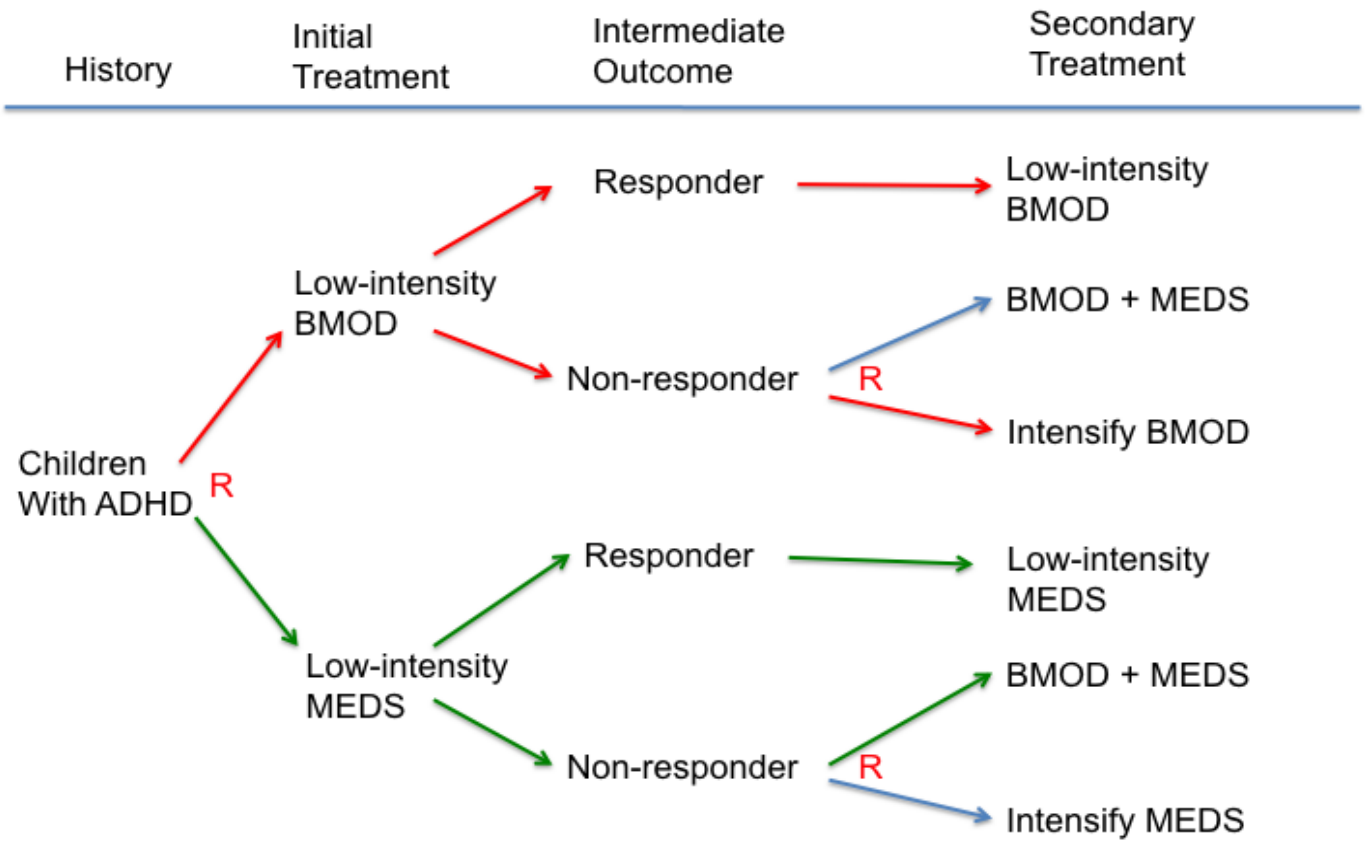
- *Sample size formula is same as that for a two group comparison of non-responders* (overall sample size depends on the presumed non-response rate)



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:

Effect Size	Scenario 1	Scenario 2	Scenario 3
0.3	$N_1 = 350$	$N_2 = \frac{N_1}{(1-\gamma)} = 700$	$N_3 = N_1 \times (2-\gamma) = 525$
0.5	$N_1 = 128$	$N_2 = \frac{N_1}{(1-\gamma)} = 256$	$N_3 = N_1 \times (2-\gamma) = 192$
0.8	$N_1 = 52$	$N_2 = \frac{N_1}{(1-\gamma)} = 104$	$N_3 = N_1 \times (2-\gamma) = 78$

Outline

1. Dynamic Treatment Regimens (Regimes): An Overview
2. Sequential Multiple Assignment Randomized Trial (SMART) Design
3. Estimation of Optimal DTRs via Q-learning
4. Non-regular Inference for Parameters indexing Optimal DTRs
   - Adaptive m-out-of-n Bootstrap
   - Simulation Study
5. Analysis of Data from STAR*D, A SMART Study on Depression
6. Discussion

Q-learning: A Secondary Analysis of SMART Data

How to estimate the optimal DTR for an individual patient?

- Q-learning (*Watkins, 1989*)
  - A popular method from Reinforcement (Machine) Learning
  - A generalization of least squares regression to multistage decision problems (*Murphy, 2005*)
  - Implemented in the DTR context with several variations (*Zhao et al., 2009; Chakraborty et al., 2010; Schulte et al., 2012; Song et al., 2014*)
  - We developed an R package called qLearn (*Xin et al., 2012*) that conducts Q-learning (Freely available at CRAN):

    http://cran.r-project.org/web/packages/qLearn/

- The intuition comes from dynamic programming (*Bellman, 1957*) in case the multivariate distribution of the data is known
  - Q-learning is an approximate dynamic programming approach

Motivation for Q-learning

- Move backward in time to take care of the delayed effects
- Define the "Quality of treatment", Q-functions:

$$
\begin{aligned}
Q_2(h_2, a_2) &= \mathbb{E}\Big[Y \Big| H_2 = h_2, A_2 = a_2\Big] \\
Q_1(h_1, a_1) &= \mathbb{E}\Big[\underbrace{\max_{a_2} Q_2(H_2, a_2)}_{\text{delayed effect}} \Big| H_1 = h_1, A_1 = a_1\Big]
\end{aligned}
$$

- Optimal DTR:

$$
d_j(h_j) = \arg\max_{a_j} Q_j(h_j, a_j),\ j = 1, 2
$$

When the true Q-functions are not known, one needs to estimate them from data, using regression models ...

DUKE NUS GRADUATE MEDICAL SCHOOL SINGAPORE

Q-learning with Linear Regression ($K = 2$)

- Regression models for Q-functions:

$$Q_j(H_j, A_j; \beta_j, \psi_j) = \beta_j^T H_j + (\psi_j^T H_j) A_j, \; j = 1, 2,$$

- At stage 2, regress Y on $(H_2, \; H_2 A_2)$ to obtain $(\hat{\beta}_2, \hat{\psi}_2)$

- Construct stage-1 Pseudo-outcome:

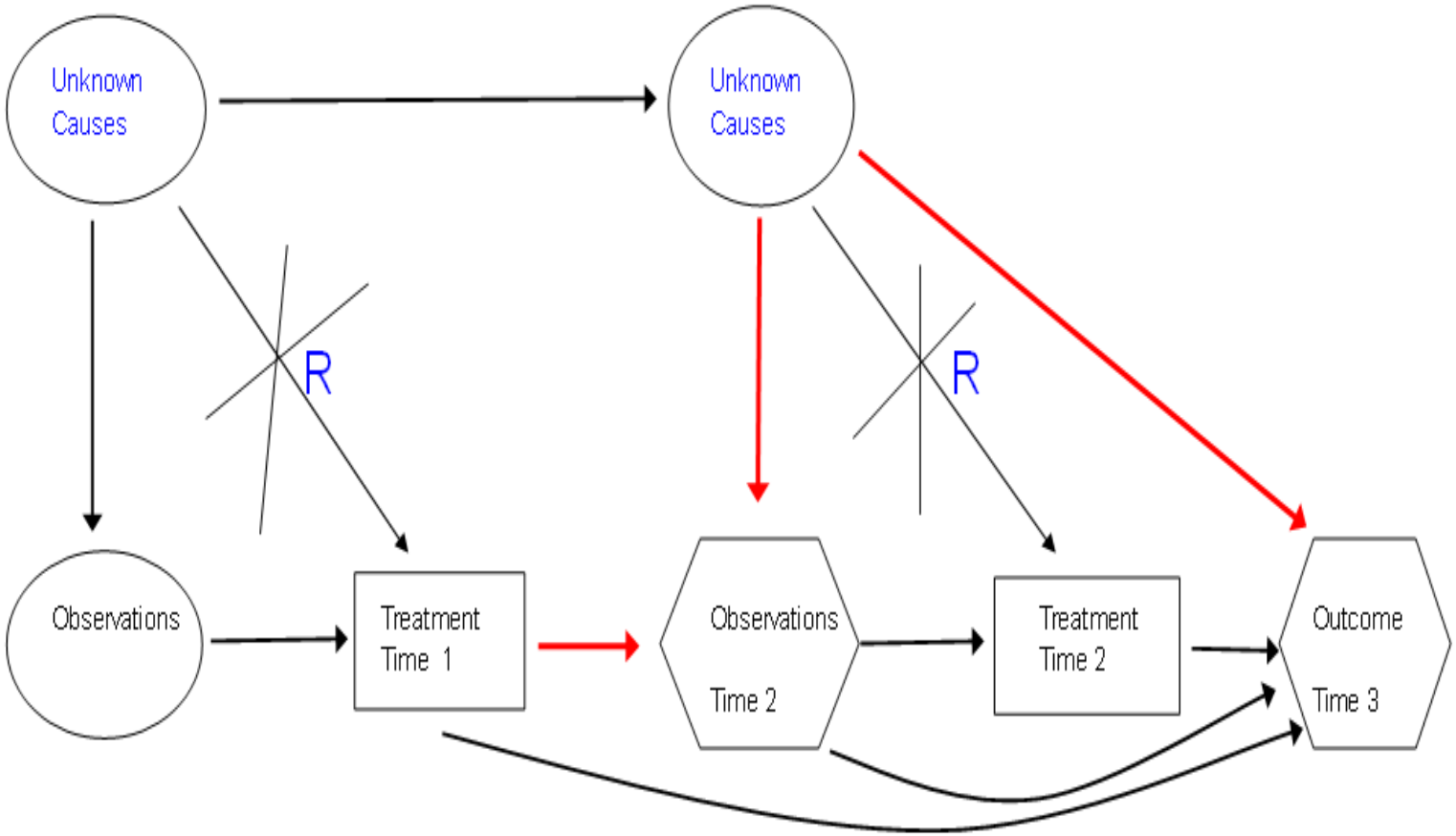
$$\tilde{Y}_{1i} = \max_{a_2} Q_2(H_{2i}, a_2; \hat{\beta}_2, \hat{\psi}_2), \; i = 1, \ldots, n$$

- At stage 1, regress $\tilde{Y}_1$ on $(H_1, \; H_1 A_1)$ to obtain $(\hat{\beta}_1, \hat{\psi}_1)$

- Estimated Optimal DTR:

$$\hat{d}_j(h_j) = \arg\max_{a_j} \; Q_j(h_j, a_j; \hat{\beta}_j, \hat{\psi}_j) = sign(\hat{\psi}_j^T h_j)$$

Why move through stages as in Q-learning? Why not run an "all-at-once" multivariable regression?



Berkson's Paradox or Collider-stratification Bias: There may be non-causal association(s) even with randomized data, leading to biased stage-1 effects (Berkson, 1946; Greenland, 2003; Murphy, 2005; Chakraborty, 2011)

Outline

1. Dynamic Treatment Regimens (Regimes): An Overview
2. Sequential Multiple Assignment Randomized Trial (SMART) Design
3. Estimation of Optimal DTRs via Q-learning
4. Non-regular Inference for Parameters indexing Optimal DTRs
   - Adaptive m-out-of-n Bootstrap
   - Simulation Study
5. Analysis of Data from STAR*D, A SMART Study on Depression
6. Discussion

Inference for Optimal Regimen Parameters

$$d_j(h_j) = sign(\psi_j^T h_j)$$

- "Regimen parameters" ψ_j – parameters that index the decision rules
  - Reduce the number of variables on which data must be collected for future implementations of the DTR
  - Know when there is insufficient evidence in the data to recommend one treatment over another – choose treatment based on cost, familiarity, preference etc.
- Inference for the optimal regimen parameters based on Q-learning has been a topic of active research for last 10 years (*Robins, 2004; Moodie and Richardson, 2010; Chakraborty et al., 2010; 2013; Laber et al., 2014; Song et al., 2014*)

Non-regularity in Inference for ψ_1 ($K = 2$)

$$\tilde{Y}_{1i} = \max_{a_2} Q_2(H_{2i}, a_2; \hat{\beta}_2, \hat{\psi}_2) = \hat{\beta}_2^T H_{2i} + |\hat{\psi}_2^T H_{2i}|$$

- Due to the non-differentiability of $\tilde{Y}_{1i}$, the asymptotic distribution of $\hat{\psi}_1$ does not converge uniformly over the parameter space – non-regular *(Robins, 2004; Laber et al., 2014)*
  - It is problematic if $p > 0$, where $p \stackrel{\text{def}}{=} P[\psi_2^T H_2 = 0]$
  - The problem persists even when $|\psi_2^T H_2|$ is "small" with non-zero probability ("local asymptotics"; *Laber et al., 2011, 2014*)
- Practical consequence: Both Wald type CIs and standard bootstrap CIs perform poorly (*Robins, 2004; Moodie and Richardson, 2010; Chakraborty et al., 2010*)
- In a K-stage setting, the same issues arise for all ψ_k, $k = K-1, \ldots, 1$

DUKE NUS GRADUATE MEDICAL SCHOOL SINGAPORE

m-out-of-n Bootstrap: A Feasible Solution

- m-out-of-n bootstrap is a tool for remedying bootstrap inconsistency due to non-smoothness (*Shao, 1994; Bickel et al., 1997*)
- Efron's nonparametric bootstrap with a smaller resample size, $m = o(n)$
- Choice of m has always been difficult – resulting in a historical lack of popularity of the approach
- We developed a choice of m for the regime parameters in the context of Q-learning – adaptive to the degree of non-regularity present in the data[1]

[1] Chakraborty B, Laber EB, and Zhao Y (2013). Inference for optimal dynamic treatment regimes using an adaptive m-out-of-n bootstrap scheme. *Biometrics*, 69: 714 - 723.

Our Approach

- Key idea: Since non-regularity arises when $p > 0$, an adaptive choice of m should depend on an estimate of p
- Consider a class of resample sizes: $m = n^{\frac{1+\alpha(1-p)}{1+\alpha}}$, where $\alpha > 0$ is a tuning parameter
- Estimate p by "pre-test" of $\psi_2^T H_2 = 0$ for fixed H_2 over the training data set:

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\left\{ \frac{n(\hat{\psi}_2^T H_{2,i})^2}{H_{2,i}^T \hat{\Sigma}_2 H_{2,i}} \leq \chi^2_{1,1-\nu} \right\}$$

- Plug in $\hat{p}$ for p in the above formula for m to get: $\hat{m} = n^{\frac{1+\alpha(1-\hat{p})}{1+\alpha}}$

DUKE NUS GRADUATE MEDICAL SCHOOL SINGAPORE

Implementation

- α can be chosen in a data-driven way via double-bootstrapping (Davison and Hinkley, 1997)
- R package qLearn: http://cran.r-project.org/web/packages/qLearn/
- Constructing one CI via double bootstrap takes about 3 minutes on a machine with dual core 2.53 GHz processor and 4GB RAM

Inference for ψ_{10}: Simulation Design

- A Simple Class of Generative Models

$$
\begin{aligned}
O_1, A_1, A_2 &\in \{-1, 1\} \quad \text{with probability } 0.5 \\
O_2 &\in \{-1, 1\} \quad \text{with } P[O_2 = 1 | O_1, A_1] = \frac{\exp(\delta_1 O_1 + \delta_2 A_1)}{1 + \exp(\delta_1 O_1 + \delta_2 A_1)} \\
Y|\cdot &\sim N(\gamma_1 + \gamma_2 O_1 + \gamma_3 A_1 + \gamma_4 O_1 A_1 + \gamma_5 A_2 + \gamma_6 O_2 A_2 + \gamma_7 A_1 A_2, 1)
\end{aligned}
$$

- Analysis Model:

$$
\begin{aligned}
Q_2 &= \beta_{20} + \beta_{21} O_1 + \beta_{22} A_1 + \beta_{23} O_1 A_1 + \underbrace{(\psi_{20} + \psi_{21} O_2 + \psi_{22} A_1)}_{\psi_2^T S_2} A_2 \\
Q_1 &= \beta_{10} + \beta_{11} O_1 + (\psi_{10} + \psi_{11} O_1) A_1
\end{aligned}
$$

- The size of the stage-2 treatment effect $\psi_2^T S_2$ determines the extent of nonregularity, e.g. $p = P[\psi_2^T S_2 = 0]$

Inference for ψ_{10}: Simulation Design

Example Generative Models[2]

Example	γ^T	δ^T	Type	p
1	$(0, 0, 0, 0, 0, 0, 0)$	$(0.5, 0.5)$	NR	1
2	$(0, 0, 0, 0, 0.01, 0, 0)$	$(0.5, 0.5)$	NNR	0
3	$(0, 0, -0.5, 0, 0.5, 0, 0.5)$	$(0.5, 0.5)$	NR	0.5
4	$(0, 0, -0.5, 0, 0.5, 0, 0.49)$	$(0.5, 0.5)$	NNR	0
5	$(0, 0, -0.5, 0, 1.0, 0.5, 0.5)$	$(1.0, 0.0)$	NR	0.25
6	$(0, 0, -0.5, 0, 0.25, 0.5, 0.5)$	$(0.1, 0.1)$	R	0
7	$(0, 0, -0.25, 0, 0.75, 0.5, 0.5)$	$(0.1, 0.1)$	R	0
8	$(0, 0, 0, 0, 0.25, 0, 0.25)$	$(0, 0)$	NR	0.5
9	$(0, 0, 0, 0, 0.25, 0, 0.24)$	$(0, 0)$	NNR	0

[2] Ex. 1 – 6 taken from Chakraborty et al. (2010), and Ex. 7 – 9 taken from Laber et al. (2014)

Inference for ψ_{10}: Simulation Design

- Focus on the 95% nominal CI for the stage-1 treatment effect parameter ψ_{10}
- Compare Monte Carlo estimates of coverage and mean width of
  - n-out-of-n bootstrap (usual)
  - m-out-of-n bootstrap
- 1000 simulated data sets, each of size $n = 300$
- 1000 bootstrap replications to construct CIs

Coverage and Mean Width of the 95% nominal CI for ψ_{10}

Table : Coverage Rates (color-coded as under-coverage, nominal coverage)

	Ex. 1 NR	Ex. 2 NNR	Ex. 3 NR	Ex. 4 NNR	Ex. 5 NR	Ex. 6 R	Ex. 7 R	Ex. 8 NR	Ex. 9 NNR
n-out-of-*n*	0.936	0.932	0.928	0.921	0.933	0.931	0.944	0.925	0.922
m-out-of-*n*	0.964	0.964	0.953	0.950	0.939	0.947	0.944	0.955	0.960

Table : Mean Width of CIs

	Ex. 1 NR	Ex. 2 NNR	Ex. 3 NR	Ex. 4 NNR	Ex. 5 NR	Ex. 6 R	Ex. 7 R	Ex. 8 NR	Ex. 9 NNR
n-out-of-*n*	0.269	0.269	0.300	0.300	0.320	0.309	0.314	0.299	0.299
m-out-of-*n*	0.331	0.331	0.321	0.323	0.330	0.336	0.322	0.328	0.328

DUKE NUS GRADUATE MEDICAL SCHOOL SINGAPORE

Outline

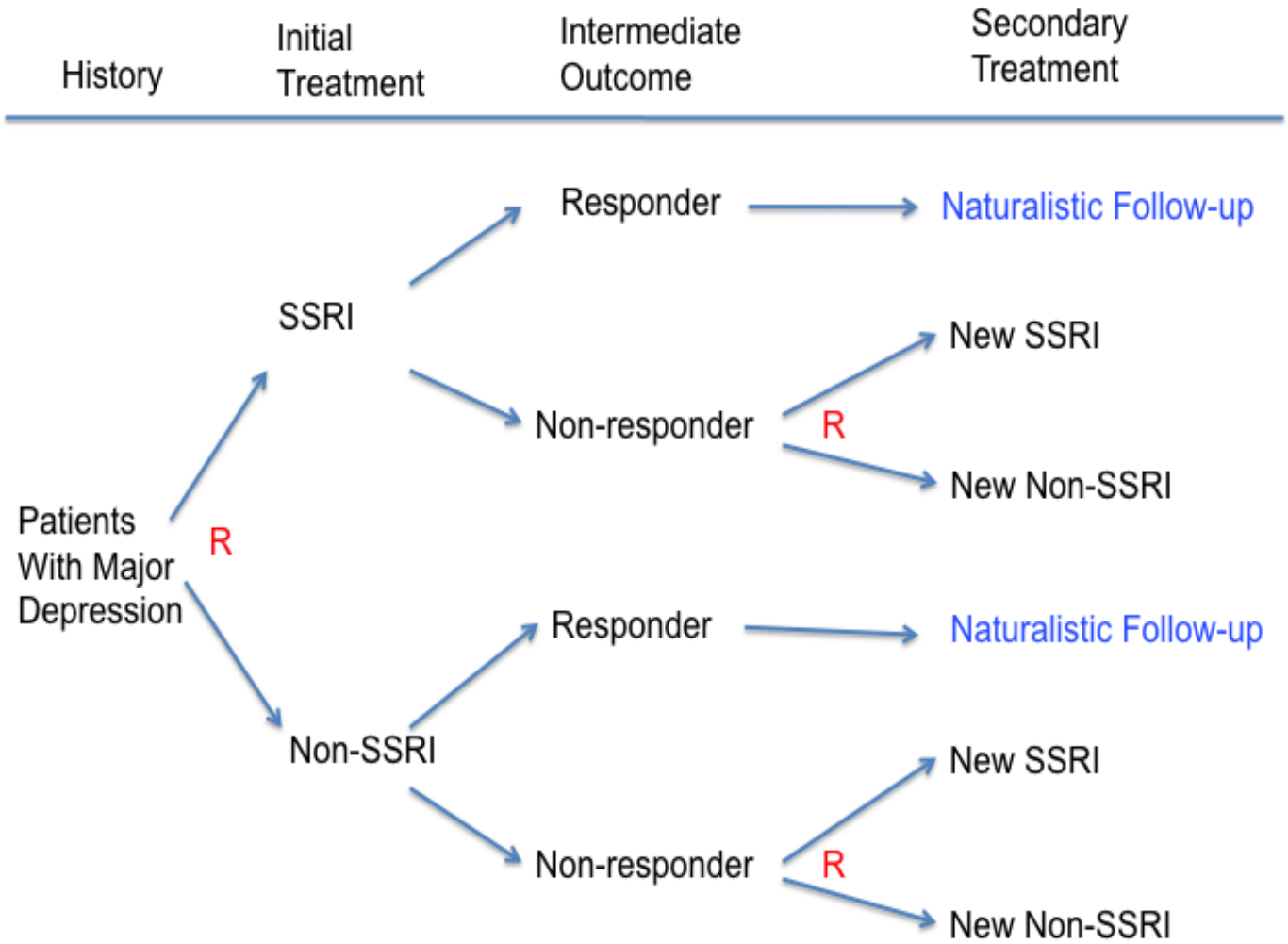
1. Dynamic Treatment Regimens (Regimes): An Overview
2. Sequential Multiple Assignment Randomized Trial (SMART) Design
3. Estimation of Optimal DTRs via Q-learning
4. Non-regular Inference for Parameters indexing Optimal DTRs
   - Adaptive m-out-of-n Bootstrap
   - Simulation Study
5. Analysis of Data from STAR*D, A SMART Study on Depression
6. Discussion

STAR*D Study (Vastly Simplified Version)

- Sequenced Treatment Alternatives to Relieve Depression (STAR*D) (*Fava et al., 2003; Rush et al., 2004*) – one of the earliest SMART designs
- Only non-responders move to the next stage and get re-randomized, but the responders move to a naturalistic follow-up phase with no new treatment (exit study)
- At each stage, treatment is binarized, SSRI (+1) or non-SSRI (-1)[3]
- Symptom severity was measured by Quick Inventory of Depressive Symptomatology (QIDS) score
- We consider $-$QIDS as the outcome (goal is to maximize)
- Covariates and/or tailoring variables (as in *Pineau et al., 2007*): preference (switch vs. augment), QIDS.start, QIDS.slope

[3] SSRI = Selective Serotonin Reuptake Inhibitor

STAR*D Design (Simplified)



STAR*D Study: Clinical Research Questions

- Based on the data from STAR*D study, how can we recommend optimal treatment sequences (in terms of SSRI vs. non-SSRI) for a future patient with known values of preference (switch vs. augment), QIDS.start and QIDS.slope, so as to achieve greatest reduction in symptom severity (e.g. QIDS score)?
  - This is about point estimation of the optimal DTR
- What measures of uncertainty, if any, can we attach to the treatment recommendations?
  - This is about inference on the the optimal DTR

STAR*D Study: Simpler Analysis

- The two Q-functions are of the form:

$$\begin{aligned} Q_2 &= \beta_{02} + \beta_{12}\text{QIDS.start}_2 + \beta_{22}\text{QIDS.slope}_2 + \beta_{32}\text{Preference}_2 + \beta_{42}A_1 \\ &\quad + \left(\psi_{02} + \psi_{12}\text{QIDS.start}_2 + \psi_{22}\text{QIDS.slope}_2\right)A_2 \\ Q_1 &= \beta_{01} + \beta_{11}\text{QIDS.start}_1 + \beta_{21}\text{QIDS.slope}_1 + \beta_{31}\text{Preference}_1 \\ &\quad + \left(\psi_{01} + \psi_{11}\text{QIDS.start}_1 + \psi_{21}\text{QIDS.slope}_1 + \psi_{31}\text{Preference}_1\right)A_1 \end{aligned}$$

- Thus the optimal decision rules are of the form:

$$\begin{aligned} d_2(H_2) &= sign(\psi_{02} + \psi_{12}\text{QIDS.start}_2 + \psi_{22}\text{QIDS.slope}_2) \\ d_1(H_1) &= sign(\psi_{01} + \psi_{11}\text{QIDS.start}_1 + \psi_{21}\text{QIDS.slope}_1 + \psi_{31}\text{Preference}_1) \end{aligned}$$

STAR*D Analysis Results

Parameter	Variable	Estimate	90% m-out-of-n bootstrap CI
Stage 2 ($n = 327; m = n$)			
β_{02}	Intercept$_2$	-1.36	$(-3.41,\ 0.65)$
β_{12}	QIDS.start$_2$	$-0.73*$	$(-0.88,\ -0.57)$
β_{22}	QIDS.slope$_2$	0.88	$(-0.04,\ 1.84)$
β_{32}	Preference$_2$	$0.66*$	$(0.12,\ 1.25)$
β_{42}	Treatment$_1$	0.20	$(-0.29,\ 0.75)$
ψ_{02}	Treatment$_2$	-0.51	$(-2.58,\ 1.50)$
ψ_{12}	Treatment$_2$ $\times$ QIDS.start$_2$	0.02	$(-0.14,\ 0.18)$
ψ_{22}	Treatment$_2$ $\times$ QIDS.slope$_2$	-0.30	$(-1.17,\ 0.64)$
Stage 1 ($n = 1260; m = \hat{m} = 910$)			
β_{01}	Intercept$_1$	-0.93	$(-4.76,\ 1.64)$
β_{11}	QIDS.start$_1$	$-1.12*$	$(-1.32,\ -0.93)$
β_{21}	QIDS.slope$_1$	0.34	$(-0.55,\ 1.20)$
β_{31}	Preference$_1$	$1.65*$	$(0.63,\ 2.60)$
ψ_{01}	Treatment$_1$	-0.93	$(-3.22,\ 1.48)$
ψ_{11}	Treatment$_1$ $\times$ QIDS.start$_1$	0.01	$(-0.14,\ 0.15)$
ψ_{21}	Treatment$_1$ $\times$ QIDS.slope$_1$	0.04	$(-0.92,\ 0.89)$
ψ_{31}	Treatment$_1$ $\times$ Preference$_1$	$-1.23*$	$(-2.17,\ -0.29)$

Outline

1. Dynamic Treatment Regimens (Regimes): An Overview
2. Sequential Multiple Assignment Randomized Trial (SMART) Design
3. Estimation of Optimal DTRs via Q-learning
4. Non-regular Inference for Parameters indexing Optimal DTRs
   - Adaptive m-out-of-n Bootstrap
   - Simulation Study
5. Analysis of Data from STAR*D, A SMART Study on Depression
6. Discussion

From SMART to SMART-AR

- 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)[4] has been developed recently
  - In general, how best to do this is not known yet

[4] Cheung YK, Chakraborty B, and Davidson K (2014). Sequential multiple assignment randomized trial (SMART) with adaptive randomization for quality improvement in depression treatment program. *Biometrics*, DOI: 10.1111/biom.12258.

Summary

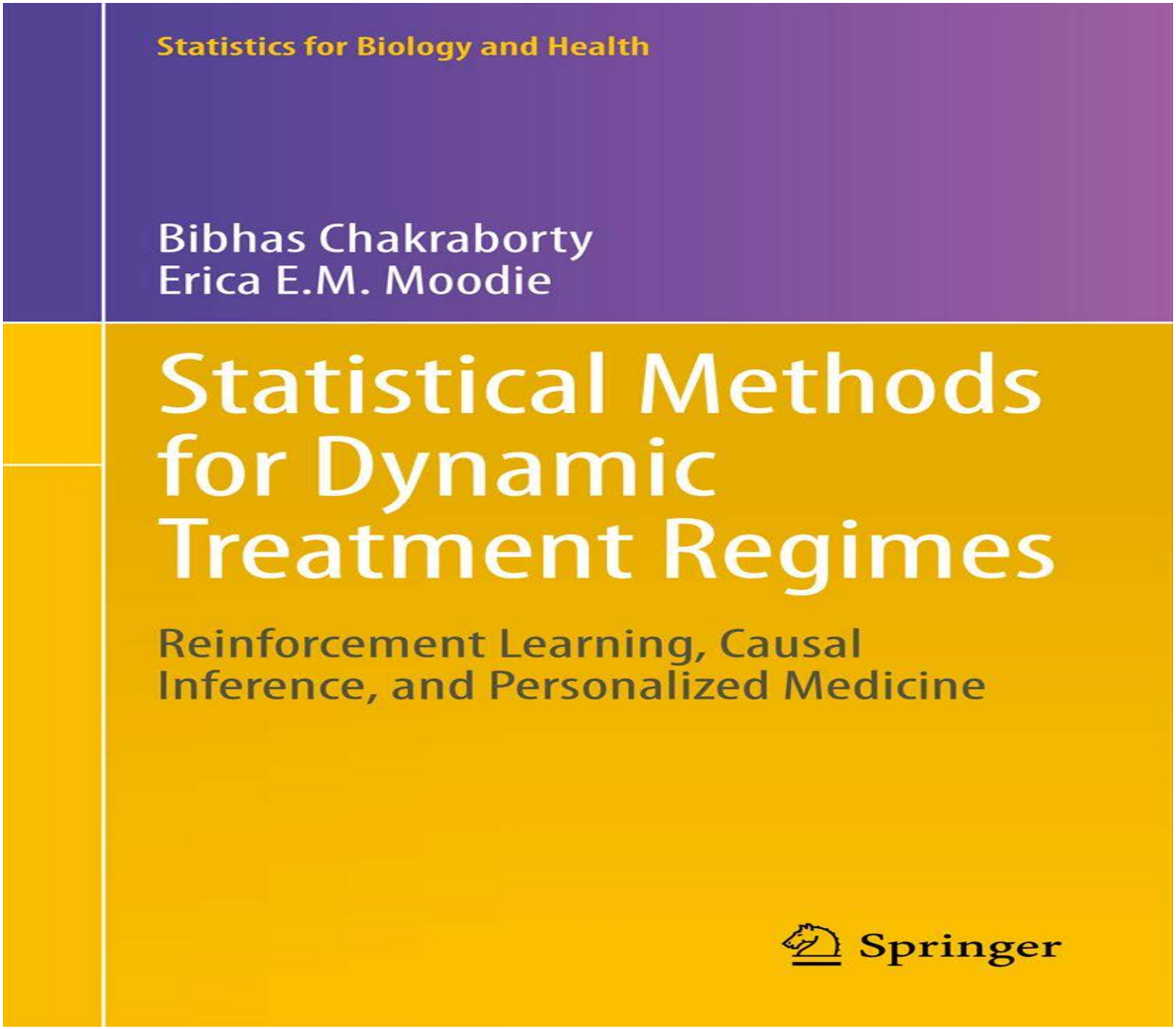
- DTRs offer a framework for operationalizing, and thus potentially improving, adaptive clinical practice for chronic diseases
- SMART designs are useful for comparing pre-conceived DTRs, as well as generating high quality data that can aid in constructing optimal DTRs
  - Sample size formulae are available 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*)
- 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

Discussion

- At least in case of SMARTs, regular settings (in which treatment effects are "too different") are much less likely to occur than non-regular settings, due to clinical equipoise (*Freedman, 1987*)
  - Hence any method of inference in the DTR context should deal with non-regularity seriously

- We have proposed an adaptive m-out-of-n bootstrap scheme for constructing CIs for the optimal regimen parameters
  - The procedure is consistent, and successfully adapts to the degree of non-regularity present in the data
  - It is conceptually simple, likely to be palatable to practitioners
  - We have developed an R package to facilitate wide dissemination

- Extending the m-out-of-n bootstrap procedure to settings with more stages and more treatment choices per stage is conceptually not too problematic, but can be operationally messy


Statistics for Biology and Health
Bibhas Chakraborty
Erica E.M. Moodie
Statistical Methods for Dynamic Treatment Regimes
Reinforcement Learning, Causal Inference, and Personalized Medicine
Springer

- Shoot your questions, comments, criticisms, and collaboration request to: bibhas.chakraborty@duke-nus.edu.sg