Sequential Multiple Assignment Randomized Trial with EnRichment Design (SMARTer) for Dynamic Treatment Regimes

Donglin Zeng

University of North Carolina

Joint work with Drs. Yuanjia Wang and Ying Liu

NUS, 2017

Outline

Introduction

SMART with EnRichment (SMARTer)

Inference for DTRs with SMARTer

Efficiency Analysis of SMARTer

Simulation Studies

SMARTer for the Autism Study

Discussion

OUTLINE

Introduction

SMART with EnRichment (SMARTer)

Inference for DTRs with SMARTer

Efficiency Analysis of SMARTer

Simulation Studies

SMARTer for the Autism Study

Discussion

PERSONALIZED MEDICINE

President's Council of Advisors on Science and Technology (PCAST):
Personalized medicine **"refers to the tailoring of medical treatment to the individual characteristics of each patient".**

Tailored medical treatments:

- Which drug? Which order? When/for whom to change drug or dose?

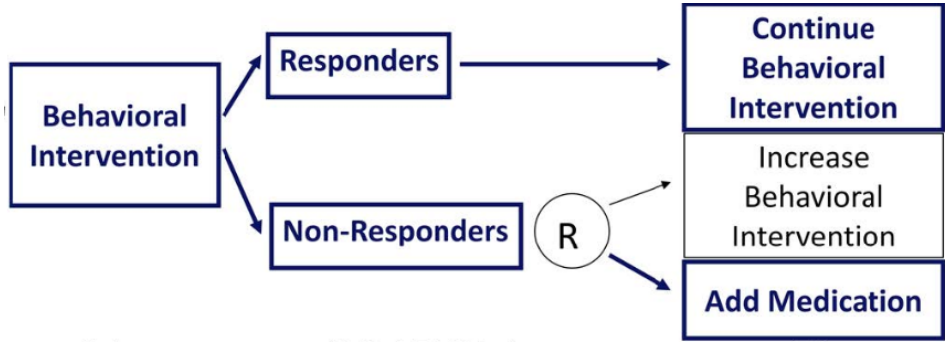
DYNAMIC TREATMENT REGIMES

- DTRs: sequential decision rules, tailored at each stage by patients' time-varying features and intermediate outcomes in previous stages (Lavori & Dawson 1998, Lavori et al. 2000, Murphy et al. 2001).
- Why DTRs?
  - Reflect clinical practice
  - Patients respond heterogeneously to treatments
  - Effect changes over time
  - Comorbidity conditions, relapses and side effects
  - High cost of intensive interventions (potential burden/side effects motivate intensity to be reduced when possible)
  - Improve adherence rate
- DTRs are often used in cancer, psychiatry, substance abuse research.

EXAMPLE OF DTRS

Adaptive Pharmacological Behavioral Treatments for Children with Attention Deficit Hyperactive Disorder (ADHD, Pelham 2002).

- DTR1: Prescribe medication (MED) as initial treatment; if a child responds then continue; if a child does not respond then augment with behavioral modification (BMOD).
- DTR2: Prescribe BMOD as initial treatment; if a child responds then continue; if a child does not respond then augment with MED.



SMART DESIGN

SMART: Sequential Multiple Assignment Randomized Trial
(Lavori & Dawson 2000, 2004; Murphy 2005)

- Patients are sequentially randomized at each critical decision stage.
- Randomization probability may depend on current states of patients.
- It enables causal comparisons among different DTRs due to randomization.
- – Adaptive Pharmacological and Behavioral treatments for ADHD (Pelham WE, 2002);
  – Sequenced Treatment Alternatives to Relieve Depression (STAR*D) (Rush, et al., 2004);
  – CATIE for schizophrenia (Schneider, et al., 2003);
  – ExTENd for alcohol dependence (Oslin, 2005);
  – Adaptive therapy for androgen independent prostate cancer (Thall et al. 2007)

A 2-STAGE SMART STUDY

- The study (Kasari et al., 2014) was designed to study communication intervention for minimally verbal children with autism.
- The study aimed to test the effect of SGD, each stage lasting 12 weeks.
- SGD: speech-generating device; (JASP+EMT): blended developmental/behavioral intervention
- The second stage had another 12 week follow-up.
- The study started with 61 eligible children and 46 completed both stages.

DIAGRAM OF THE AUTISM STUDY

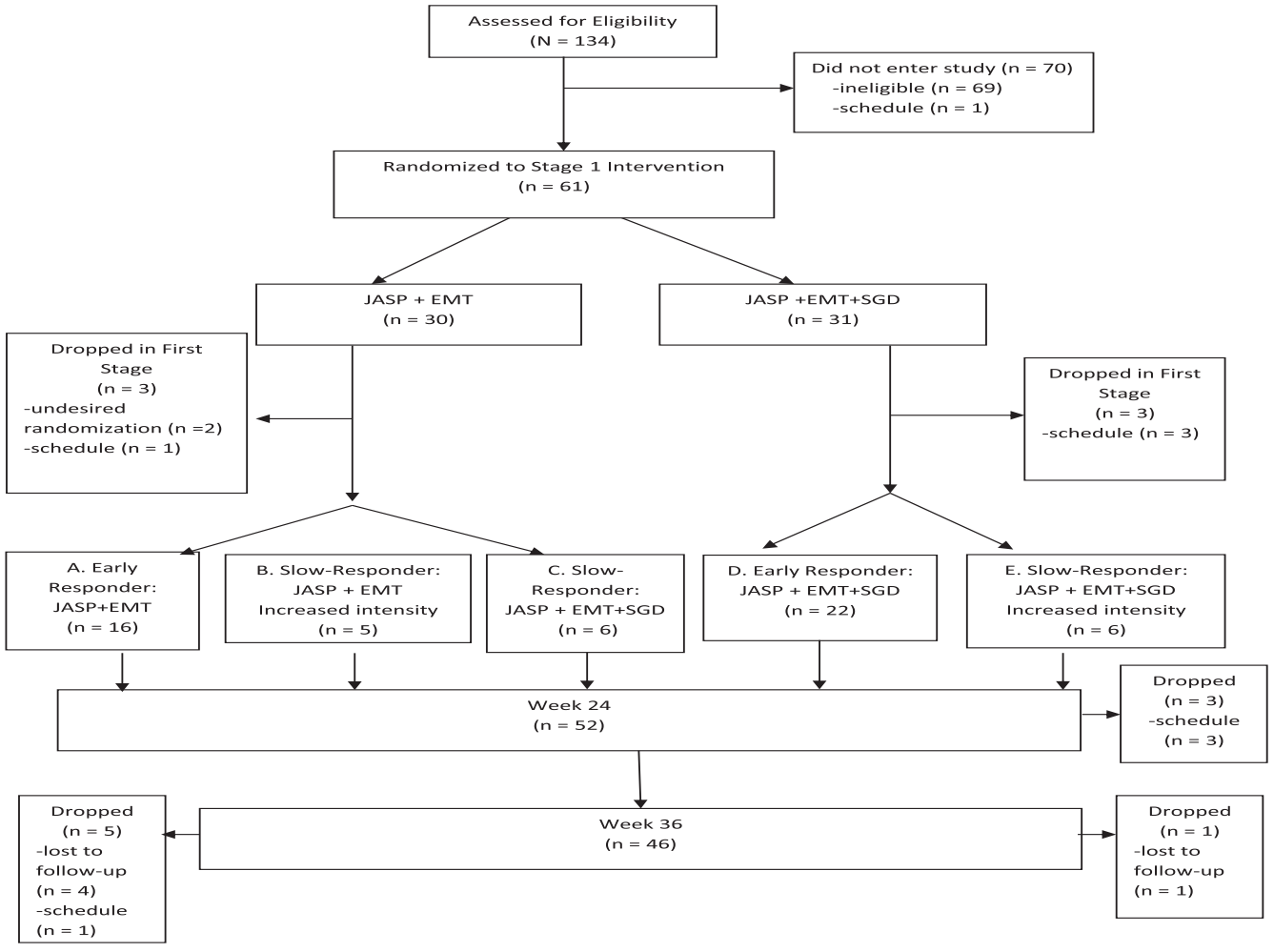


Figure: SMART Design of Autism Study (Kasari et al. 2014)

ADVANTAGE OF SMART

Research questions to be answered from a SMART

- **Main effects of treatments**
  - What is the better initial treatment, JASP+EMT or JASP+EMT+SGD?
  - What about the slow-responders: intensify or not?
- **Effects of embedded DTR**
  - JASP+EMT–>intensify vs JASP+EMT+SGD—> intensify vs JASP+EM–>JASP+EMT+SGD
- **Exploring optimal treatment strategy (deep tailoring)**
  - intensify or not in the second stage dependent on additional intermediate outcomes?

GENERAL ADVANTAGES OF USING A SMART

- Valid comparisons of different treatment options at different stages due to the virtue of randomization.
- Discover adaptive treatment strategies that are embedded in the SMART trial.
- Inform the development of adaptive and deeply tailored treatments (using potentially high-dimensional biomarkers).

PRACTICAL LIMITATIONS OF SMART

- Operation cost of administrating multiple stage studies and multiple treatments is high.
- The length of trial period is long (March et al. 2010).
- Study dropout or compliance is common even in regular RCTs:
  - In the CATIE study, 705 of 1460 patients stayed for the entire 18 months of the study.
  - In ExTENd, the drop-out rate was 17% (52 out of 302) in the first-stage treatment and an additional 13% (41 out of 302) during the second stage.

OUTLINE

Introduction

SMART with EnRichment (SMARTer)

Inference for DTRs with SMARTer

Efficiency Analysis of SMARTer

Simulation Studies

SMARTer for the Autism Study

Discussion

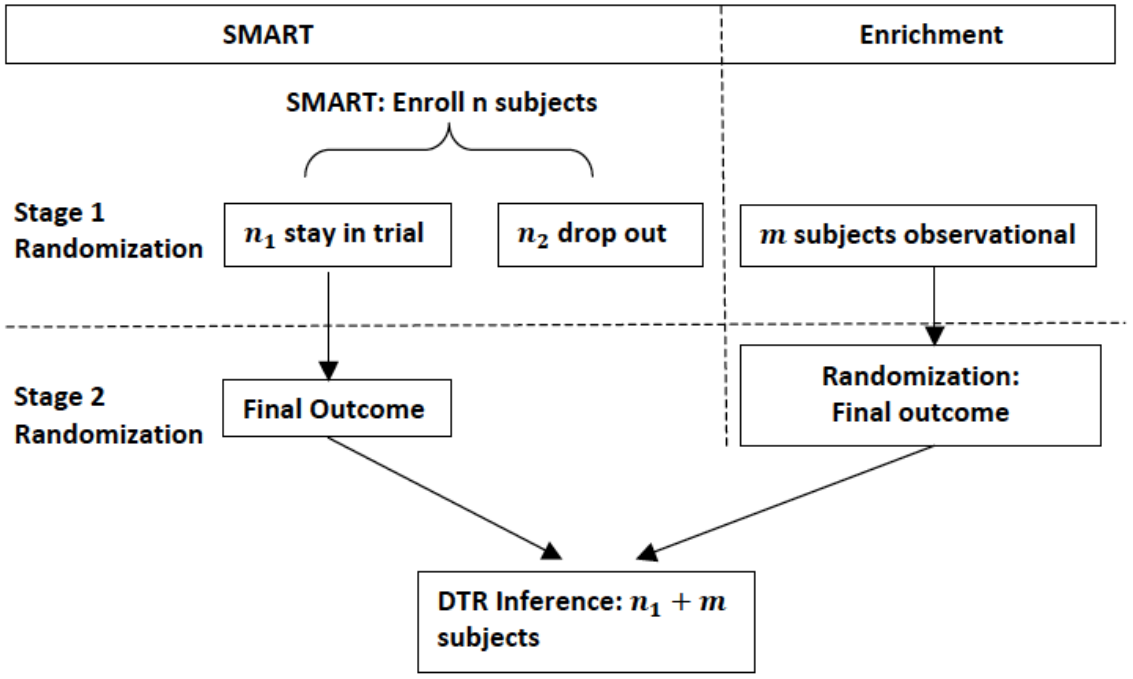
GENERAL IDEA

- SMARTer is a stage-wise enrichment design to improve efficiency of SMART.
- At each stage, an enrichment sample will be included in randomization even if they have received treatments in a naturalistic fashion in previous stages.
- SMARTer can be considered as a meta-analytic design to synthesize SMART with single-stage trials.

A GENERAL DIAGRAM OF SMARTER

- Central idea: we enrich the study at each stage randomization.

Figure: Diagram of SMART-EnRichment Trial (**SMARTER**)



- Key: natural treatment history of enrichment participants is collected.
- Not essential but useful: follow-up information of drop-out SMART participants.

RATIONALE BEHIND A 2-STAGE SMARTER

- At Stage 2, the continuing participants from SMART and the enrichment participants provide unbiased prediction of stage 2 treatment effects given history at Stage 1, due to RANDOMIZATION.
- This prediction provides unbiased predicted future outcomes for the participants at Stage 1–we "recover" what outcome would be for those drop-out participants.
- At Stage 1, the predicted and observed outcomes from SMART can be used to infer unbiased treatment effects, again due to RANDOMIZATION.
- Therefore, SMARTER protects against bias due to sequential randomization; SMARTER improves efficiency due to enrichment.

SYNTHETIC MULTIPLE-STAGE TRIALS: AN EXTREME CASE OF SMARTER

- Two independent trials are run in stage 1 and stage 2 respectively.
- This is equivalent to an extreme case of SMARTer: $n_1 = 0$.
- The trial stage 2 can be used to predict what would be for the subjects in the first trial at stage 1.
- Thus, the predicted outcomes can be combined with the actual stage 1 trial to mimick a 2-stage SMART.
- Note that this is completely different from simply piecing together trial results (Collins et al. 2014)–here, we piece trial data together.

OUTLINE

Introduction

SMART with EnRichment (SMARTer)

Inference for DTRs with SMARTer

Efficiency Analysis of SMARTer

Simulation Studies

SMARTer for the Autism Study

Discussion

DATA FROM SMARTER

- Notation. Stage 1: S_1, A_1; Stage 2: $S_2 = (S_1, A_1)$, A_2; Final outcome: Y.
- The goal is to evaluate the expected outcome for any given treatment strategy: $a_1 = d_1(S_1), a_2 = d_2(S_2)$.
- Data from the SMART sample: $S_{1i}, A_{1i}, Z_i A_{2i}, Z_i Y_i, i = 1, ..., n$. ($Z_i$: Stage 2 continuation status)
- Data from the enrichment sample: $S_{1j}, A_{1j}, A_{2j}, Y_j, j = 1, ..., m$.
- Note that the distributions of (S_1, A_1) may be different between the SMART group and the enrichment group!

KEY ASSUMPTIONS

- Assumption (a): the dropout in the SMART group only depends on observed (A_1, S_1) (non-informative dropout).
- Assumption (b): the conditional distribution of Y given (A_1, A_2, S_1) is the same between the SMART population and the enrichment population (equivalent outcome distributions).
- Remark. Assumption (a) is needed only if we want to use data from stage 2 subjects in the SMART sample; this is minimal for any RCTs. Assumption (b) is natural to ensure utility of the enrichment sample.

INFERENCE FROM SMARTER

- First, we estimate the predicted outcome using the stage 2 data:

$$\widehat{Y}(a_1, a_2, s) =$$

$$\frac{\sum_{i=1}^{n} Z_i Y_i I(A_{1i} = a_1, A_{2i} = a_2, S_{1i} = s) + \sum_{j=1}^{m} Y_j I(A_{1j} = a_1, A_{2j} = a_2, S_{1j} = s)}{\sum_{i=1}^{n} Z_i I(A_{1i} = a_1, A_{2i} = a_2, S_{1i} = s) + \sum_{j=1}^{m} I(A_{1j} = a_1, A_{2j} = a_2, S_{1j} = s)}.$$

INFERENCE FROM SMARTER

- For any given treatment regime (d_1, d_2), the estimator of its value using SMARTer is given as a weighted average of the outcomes from the SMART participants who were assigned to such treatment regime:

	outcome	weight
$Z_i = 1$	Y_i	$\frac{I(A_{1i}=d_1(S_{1i}),A_{2i}=d_2(S_{1i},A_{1i}))}{p(A_{1i}\|S_{1i})p(A_{2i}\|S_{1i},A_{1i})}$
$Z_i = 0$	$\widehat{Y}(A_{1i}, d_2(S_{1i}, A_{1i}), S_{1i})$	$\frac{I(A_{1i}=d_1(S_{1i}))}{p(A_{1i}\|S_{1i})}$

- We show that this estimator is unbiased for $E[Y(d_1, d_2)]$.
- This is because $\widehat{Y}(a_1, a_2, s)$ is unbiased for $E[Y(a_1, a_2)|A_1 = a_1, S_1 = s]$.

VARIANCE COMPUTATION

- For any given treatment regime (d_1, d_2), the variance of the value estimator is the variance of

$$
\begin{aligned}
& Var \quad \Bigg(Z \frac{I(A_1 = d_1(S_1), A_2 = d_2(S_1, A_1))}{p(A_1|S_1)p(A_2|S_1, A_1)} \\
& \qquad \times \left\{ (Y - \mu(d_1, d_2)) + \frac{1 - \alpha(A_1, S_1)}{\alpha(A_1, S_1) + \beta r(A_1, S_1)} (Y - E[Y|A_1, A_2, S_1]) \right\} \\
& \qquad + (1 - Z) \frac{I(A_1 = d_1(S_1))}{p(A_1|S_1)} E[Y - \mu(d_1, d_2)|A_1, A_2 = d_2(S_1, A_2), S_1] \Bigg) \\
+\beta \quad & Var_e \quad \Bigg(\frac{(1 - \alpha(A_1, S_1))(Y - E[Y|A_1, A_2, S_1])}{\alpha(A_1, S_1) + \beta r(A_1, S_1)} \\
& \qquad \times \frac{I(A_1 = d_1(S_1), A_2 = d_2(S_1, A_1))}{p(A_1|S_1)p(A_2|S_1, A_1)} \Bigg) .
\end{aligned}
$$

- $\beta = m/n$, $\alpha(a_1, s_1)$ is the non-drop-out probability, and $r(a_1, s)$ is the probability ratios of stage 1 data between the enrichment sample and SMART sample.

MORE GENERAL INFERENCE

- The proposed method can be easily generalized to compare two different DTRs:

$$E[Y(d_1, d_2)] - E[Y(\widetilde{d}_1, \widetilde{d}_2)].$$

- If intermediate outcomes (s_2) between two stages are observed for all subjects in both SMART sample and enrichment sample, such outcomes can be incorporated into estimating predicted outcomes at stage 2, $\widehat{Y}(a_1, a_2, s_2, s_1)$, and DTRs can allow d_2 to depend on s_2.

OUTLINE

Introduction

SMART with EnRichment (SMARTer)

Inference for DTRs with SMARTer

Efficiency Analysis of SMARTer

Simulation Studies

SMARTer for the Autism Study

Discussion

EFFICIENCY ANALYSIS

- Simplifications:
  (1) pure randomization: $p_1(S) = p_1, p_2(S) = p_2$;
  (2) completely random drop-out: $P(Z = 1|A_1, S) = \alpha$
  (3) the stage 1 distributions between two samples are similar.
- Relative efficiency of SMARTER to SMART with no drop-out:

$$\rho \approx \frac{1+\gamma}{1-(1-\alpha)(1-p_2)+\gamma\frac{\alpha(1+\beta)^2+\beta(1-\alpha)^2}{(\alpha+\beta)^2}},$$

  γ is the ratio of the within-subgroup variance versus the between-subgroup variance.

RELATIVE EFFICIENCY IN SOME SIMPLE CASES

- $\alpha = 1$, that is, none drops out of the first stage. Under this case, $\rho \approx (1+\gamma)/(1+\gamma/(1+\beta)^2)$. Thus, the enrichment with $\beta > 0$ always increases efficiency.
- $\alpha = 0$, that is, all subjects drop out of the first stage. Under this case, $\rho \approx (1+\gamma)/(p_2+\gamma/\beta)$ so clearly, $\beta \geq 1$, we always gain efficiency. Indeed, $\beta > \gamma/(1+\gamma-p_2)$ ensures efficiency gain.
- In fact, for any $0 \leq \alpha < 1$, SMARTER is more efficient if $\beta \geq 1$.

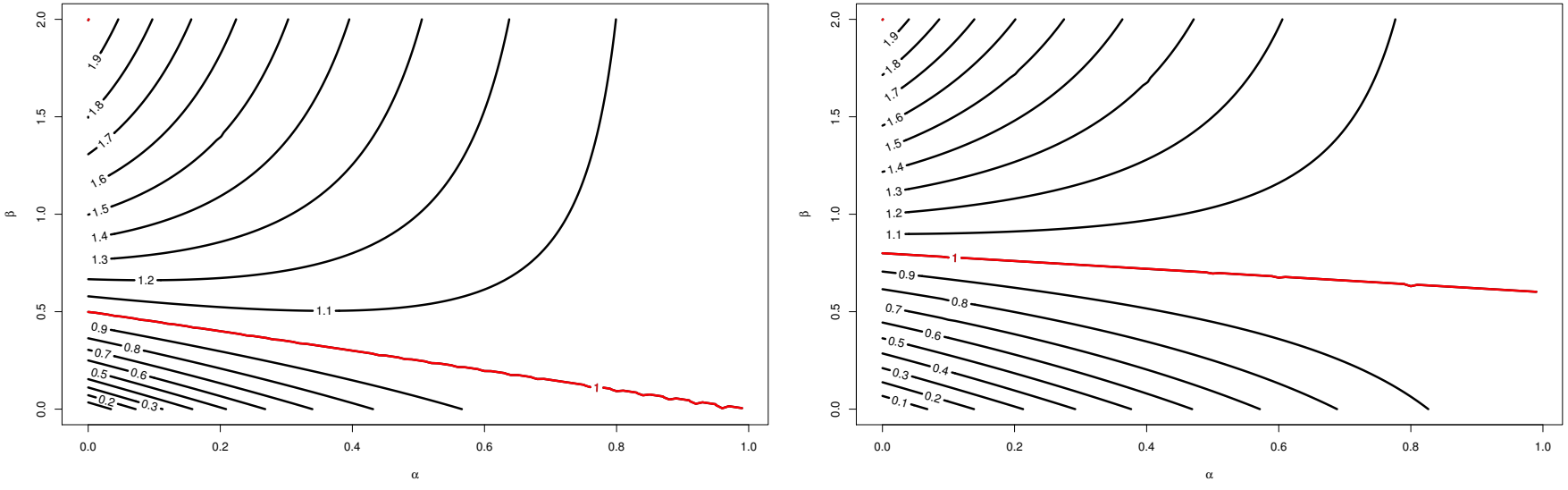
NUMERICAL EXAMPLE

Table: Comparison of relative efficiency (RE, variance ratio) of SMART and SMARTER

α	β	Theoretical RE	Empirical RE
0.5	0.5	0.870	0.899
0.5	1	1.053	1.075
0.5	2	1.266	1.301
0	0.5	0.588	0.592
0	1	1.111	1.163
0	2	2.000	2.205

CONTOUR PLOTS OF REs

Figure: Relative efficiencies of SMARTER compared to SMART; γ ratio of within and between stratum variance; $\gamma = 0.5$ (left); $\gamma = 2$ (right)



Outline

Introduction

SMART with EnRichment (SMARTer)

Inference for DTRs with SMARTer

Efficiency Analysis of SMARTer

Simulation Studies

SMARTer for the Autism Study

Discussion

SIMULATION SETTING 1 FOR SMARTER

Simulation setting 1 (no intermediate outcome)

- In the SMART sample ($n = 800$), consider two stages and randomization probabilities are both 0.5 ($A_1, A_2 \in \{-1, 1\}$) A baseline state S_1 takes values $(0, 1, 2)$ with equal probabilities.
- In the enrichment sample, $P(A_1 = 1|S_1) = 1/(1 + \exp(-I(S_1 < 2) + 0.5))$ and S_1 takes values $(0, 1, 2)$ with probabilities (0.5, 0.25, 0.25).
- Final outcome model

$$Y = (A_1+A_2)(1-S_1)+I(S_1 = 1, A_1 = 1, A_2 = -1)+N(0, 1).$$

- We estimate the value for DTR

$$d_1(s_1) = 2I(s_1 < 2) - 1, \quad d_2(s_1, a_1) = 2I(s_1 < 1) - 1$$

and compare with one-size-fits-all $d_1 = d_2 = 1$.

SIMULATION SETTING 2 FOR SMARTER

Simulation setting 2 (with intermediate outcome)

- An intermediate binary outcome S_2 is generated from

$$logit(P(S_2 = 1|S_1, A_1)) = A_1(1 - S_1).$$

- Final outcome model:

$$Y = S_1 + A_2(1 - S_1) + I(S_1 = 1)A_2(2S_2 - 1) + N(0, 1).$$

- We consider DTR

$$d_1(s_1, a_1) = 2I(s_1 = 1) - 1,$$

$$d_2(s_1, a_1, s_2) = I(s_1 \neq 1)(2I(s_1 = 0) - 1) + I(s_1 = 1)sign(2s_2 - 1)$$

and compare with one-size-fit-all $d_1 = d_2 = 1$.

SIMULATION 1 RESULTS FROM 1,000 REPLICATES

α	β	Est	ESE	SD	CI	$\widehat{\rho}$
Value estimation of one DTR (true value 1.667)						
0.0	0.5	1.668	0.100	0.100	0.943	0.647
0.0	1.0	1.663	0.073	0.072	0.955	1.230
0.0	2.0	1.669	0.054	0.053	0.957	2.316
0.5	0.5	1.663	0.082	0.082	0.944	0.946
0.5	1.0	1.664	0.075	0.074	0.948	1.141
0.5	2.0	1.665	0.069	0.068	0.948	1.353
Comparing two different DTRs† (truth = 1.667)						
0.0	0	1.672	0.147	0.147	0.946	0.557
0.0	1	1.665	0.106	0.108	0.946	1.040
0.0	2	1.667	0.077	0.075	0.953	2.157
0.5	0	1.667	0.115	0.116	0.946	0.845
0.5	1	1.668	0.104	0.104	0.954	1.040
0.5	2	1.667	0.095	0.094	0.944	1.269

SIMULATION 2 RESULTS FROM 1,000 REPLICATES

α	β	Est	ESE	SD	CI	$\widehat{\rho}$
Value estimation of one DTR (true value 1.654)						
0.0	0.5	1.653	0.105	0.114	0.924	0.504
0.0	1.0	1.650	0.078	0.079	0.941	1.011
0.0	2.0	1.658	0.058	0.059	0.948	1.812
0.5	0.5	1.653	0.084	0.084	0.942	0.901
0.5	1.0	1.656	0.077	0.074	0.955	1.150
0.5	2.0	1.653	0.070	0.071	0.939	1.243
Comparing two different DTRs (truth= 1.154)						
0.0	0	1.160	0.174	0.191	0.921	0.435
0.0	1	1.147	0.127	0.138	0.929	0.778
0.0	2	1.157	0.095	0.096	0.950	1.719
0.5	0	1.159	0.135	0.134	0.953	0.876
0.5	1	1.156	0.123	0.121	0.956	1.016
0.5	2	1.153	0.112	0.112	0.950	1.206

SIMULATIONS FOR ESTIMATING **OPTIMAL** DTRS

Simulation setting

- $R_1 = 1 + A_1 * S_1 + N(0, 2)$; $R_2 = A_2 * R_1 + N(0, 2)$.
- $S_1 \sim N(0, 1)$ plus 4 additional noise baseline covariates.
- In SMART group, A_1 and A_2 are purely randomized.
- In the enrichment group of the same size, A_1 is observational and depends on R_1 and R_2; only A_2 is purely randomized.
- We vary the drop-out rates of subjects in SMART component.

SIMULATIONS FOR EXPLORING **OPTIMAL** DTRS

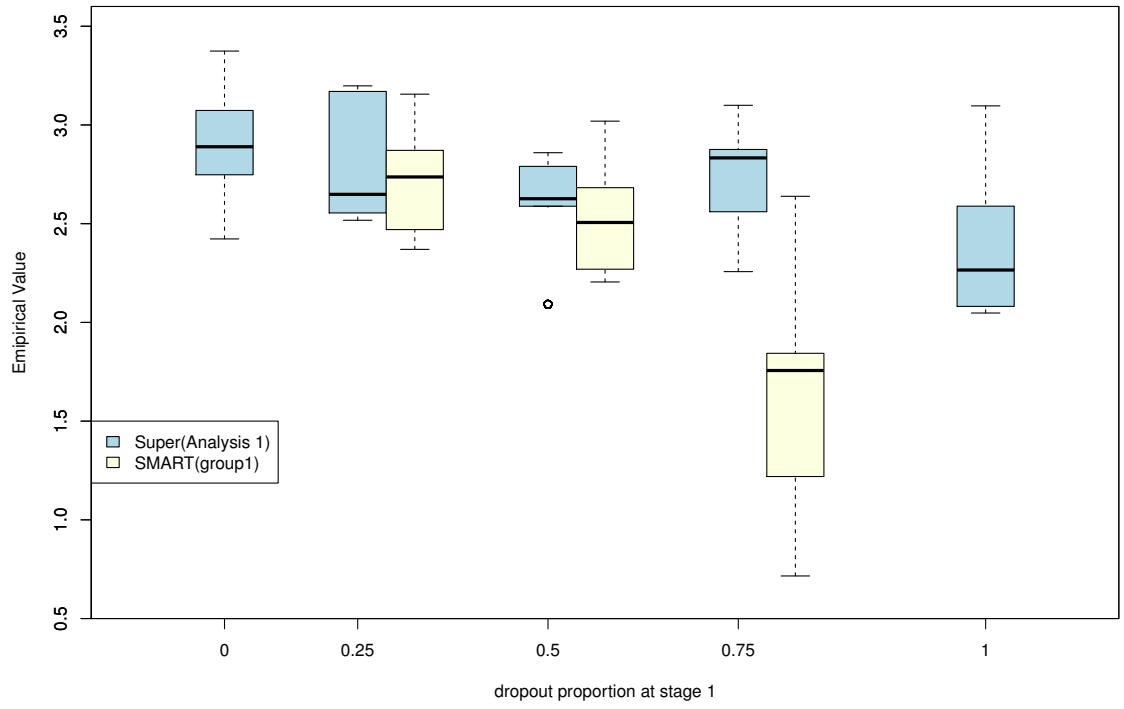


Figure: Estimates of the value functions using the complete SMART subjects (yellow) and the SMARTer (blue)

OUTLINE

Introduction

SMART with EnRichment (SMARTer)

Inference for DTRs with SMARTer

Efficiency Analysis of SMARTer

Simulation Studies

SMARTer for the Autism Study

Discussion

SAMPLE SIZE CALCULATION

- We use the autism study as an illustration of SMARTer example.
- The study (Kasari et al. 2014) finds that SGD (JASP+EMT+SGD) had a better treatment effect compared with spoken words alone (JASP+EMT) and that the adaptive intervention (JASP+EMT+SGD)—> intensify for slow responders led to better post-treatment outcomes.
- We compare two DTRs in this example:

$$(JASP + EMT + SGD) - -- > \text{intensify for slow responders}$$

vs

$$(JASP + EMT) - -- > \text{ add SGD for slow responders.}$$

- We exam the sample sizes needed for SMARTer for powers of 90, 85, 80% using the original study effect sizes.

SAMPLE SIZE RESULTS

γ	d%	Power 90%			Power 85%			Power 80%		
		SMART	SMARTER		SMART	SMARTER		SMART	SMARTER	
			n	m		n	m		n	m
0.2	0%	202	202	0	173	173	0	151	151	0
	15%	238	202	27	203	173	15	178	151	18
	40%	337	202	54	288	173	43	252	151	40
0.5	0%	202	202	0	173	173	0	151	151	0
	15%	238	202	104	203	173	81	178	151	76
	40%	337	202	120	288	173	100	252	151	89
1	0%	202	202	0	173	173	0	151	151	0
	15%	238	202	144	203	173	117	178	151	106
	40%	337	202	155	288	173	130	252	151	115

d% is the dropout rate; γ is ratio of within- and between-stratum variance between responding status.

CONCLUSION FROM THE EXAMPLE

- SMARTer requires smaller total number of patients than SMART if within-responding status outcomes are more homogeneous.
- The benefit of SMARTer is more obvious when dropout rate is higher.
- This benefit is even more significant, considering operating cost in maintaining SMART as compared to single-randomization in enrichment sample.

OUTLINE

Introduction

SMART with EnRichment (SMARTer)

Inference for DTRs with SMARTer

Efficiency Analysis of SMARTer

Simulation Studies

SMARTer for the Autism Study

Discussion

CONCLUDING REMARKS

- SMARTER supplements SMART to improve efficiency to salvage potential high drop-out in SMART.
- By enrichment, it ensures sufficient sample size at each stage.
- More efficiency gain if intermediate outcomes are available for all participants.
- We may not even need SMART component to infer DTRs by synthesizing independent trials.

CAUTIONS USING SMARTER

- Potential difference between enrichment population and SMART population. The latter tends to have more stringent eligibility requirement.
- Quality of the first stage (naturalistic) treatment delivery in the enrichment sample.
- Feasibility of retrieving treatment history of the enrichment sample and matching with the SMART sample.
- The enrichment sample may not be useful for the first stage treatment if its a novel treatment not immediately available in communities.

THANK YOU!