The Comparison of ACI and MCB Methods for Choosing a Set that Contains the Optimal Dynamic Treatment Regime

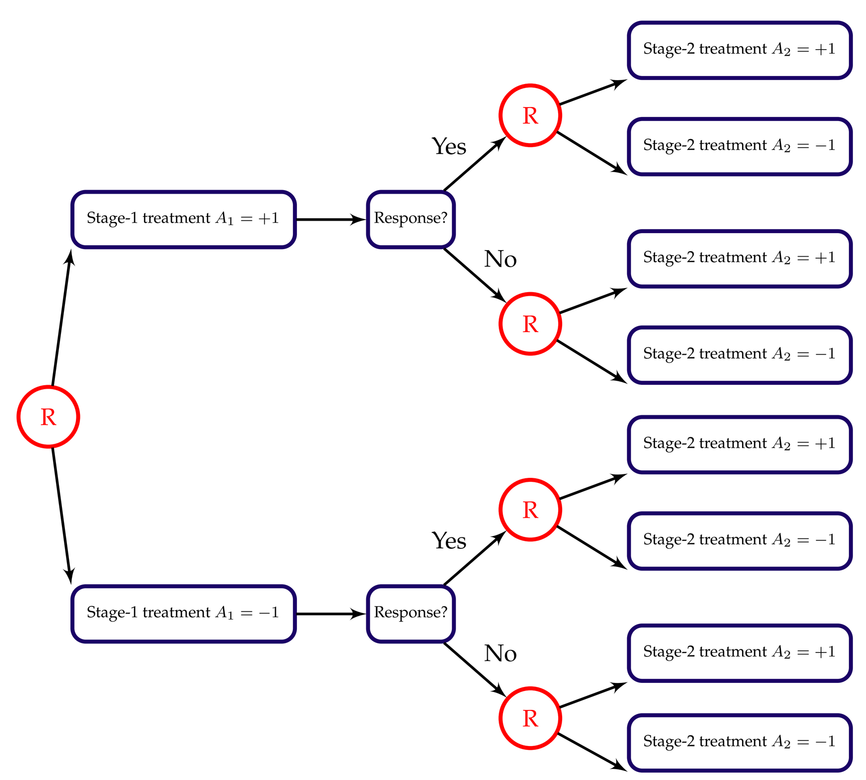
Rong Zhou[1*], Tianshuang Wu[1*]

1. Abstract

Dynamic treatment regimes (DTRs) have been used by clinicians for treatment decision-making at different stages for patient care. Sequential multiple assignment randomized trials (SMART) is a design to obtain data in order to find the optimal DTR that maximizes the expected cumulative outcome. Many methods have been developed to seek for the best DTR, including Q-learning methods and A-learning methods. In addition, a method called "Multiple Comparisons with the Best" is proposed by researchers to identify a set of DTRs that includes the optimal one. In this project, we focus on two methods: the modified ACI (Adaptive Confidence Intervals) method by Laber et al. and MCB (Multiple Comparisons with the Best) method by Ertefaie et al. We conduct simulation in four different scenarios for both methods and report the results. By comparing the average set size of each method in four different scenarios, we conclude that we will recommend the MCB method by Ertefaie et al. in general.

2. DTR and SMART

A **DTR** (Dynamic Treatment Regime) is a sequence of decision rules, which are based on patients' characteristics overtime and provide treatment decisions. The rules take information before the current decision-making as inputs and return a treatment decision at this stage. A **SMART** (Sequential, Multiple Assignment, Randomized Trial) starts with a randomization of individuals to available treatment options, then base on the intermediate outcome at each stage, individuals are re-randomized to receive the next stage's treatment.



We are considering a two-stage setting shown in the above graph. We use a time-ordered trajectory (X_1, A_1, X_2, A_2, Y) to record the process. X_1 is called baseline information, which contains the initial condition of the patient before the first stage treatment; A_1 is the treatment on the first stage, and is coded as binary values {-1,1}; X_2 is the response to the first stage treatment before the second stage treatment, and is coded as binary values {-1,1} with 1 indicating "Yes" and -1 indicating "No"; A_2 is the second stage treatment with binary values {-1,1}; Y is the final stage outcome of the regime and is coded such that a higher value is preferred.

3. ACI and MCB

So far, most methods proposed in this area are seeking for the optimal DTR under the SMART design. However, finding the sole best DTR should not always be the primary interest when two or more DTRs are "equally the best". In this study, the MCB method proposed by Ertefaie et al. identifies a set of DTRs that excludes relatively worse DTRs among several DTRs. And this provides more flexibility for clinicians to make decisions for patients. Here we introduce two of the methods.

1. Department of Statistics, University of Michigan.

1. The Modified ACI (Adaptive Confidence Intervals) Method

The linear model for obtaining the means of the Q-learning functions and its estimator of each stage is as follows:

$$Q_i(h_i, a_i; \beta_i) = \beta_{i,0}^T h_{i,0} + \beta_{i,1}^T h_{i,1} 1_{a_i=1} \quad (1)$$

$$\hat{\beta}_i \triangleq arg \min_{\beta_i} \mathbb{P}_n (Y_i - Q_i(H_i, A_i; \beta_i))^2 \quad (2)$$

for $i = \{1, 2\}$. H_i is the joint expectation over the patient's history, and H_i includes information from the beginning up to the decision for ith stage treatment A_i. Correspondingly, $H_1 = X_1$, and H_2 contains X_2. Now Plug $\hat{\beta}_2$ into the predicted future reward, which is defined as:

$$\tilde{Y}_1 \triangleq Y_1 + \max_{a_2 \in \{1,-1\}} Q_2(H_2, A_2; \hat{\beta}_2) = Y_1 + H_{2,0}^T \hat{\beta}_{2,0} + [H_{2,1}^T \hat{\beta}_{2,1}]_+$$

where $[H_{2,1}^T \hat{\beta}_{2,1}]_+$ represents the positive part of $H_{2,1}^T \hat{\beta}_{2,1}$. By regressing $\tilde{Y}_1$ on H_1 and A_1, we can obtain $\hat{\beta}_1$. So the optimal DTR $\hat{a} = (\hat{a_1}, \hat{a_2})$ is:

$$\hat{a}_t(h_t) \triangleq arg \max_{a_t \in \{1,-1\}} Q_t(h_t, a_t; \hat{\beta}_t) \quad (3)$$

In this study, the ACI method was modified so that a set of DTRs is returned and the two methods are comparable with each other.

2. The MCB (Multiple Comparisons against the Best) Method

The goal of this method is to construct a set of indices (denoted as $\mathcal{B}$) such that $[K] := \arg\max_{k=1,\cdots,K} \theta_k$ belongs to the set with probability no smaller than a pre-specified level, say, $1-\alpha$. θ_k in the above formula represents the mean outcome of each DTR with k = 1 to K, and $\hat{\theta}_k$ is used to represent the estimation of each θ_k. A marginal structural model (MSM) is applied and the Inverse probability weighting (IPW) method is used to estimate the mean outcome $\hat{\theta}_k$ of each DTR with k = 1 to K. In the comparison process, an index will be included in the set if its estimated mean is not "too worse" than every other estimated mean. To be specific, an index i is included if and only if

$$\hat{\theta}_i \geq \max_{j \neq i} [\hat{\theta}_j - c_i \hat{\sigma}_{ij}]$$

where $c_i = c_i(R, \alpha)$ is a constant that depends on the correlation matrix R_θ of $\sqrt{n}\hat{\theta}$ as well as on $1-\alpha$; and $\hat{\sigma}_{ij}$ is the estimator of $\sigma_{ij} = \sqrt{var(\hat{\theta}_i - \hat{\theta}_j)}$. Hsu proved that if R_θ is known up to a constant, i.e., $R = \sigma^2 \Sigma$, where Σ is a known matrix and σ^2 is a constant that is either known or can be consistently estimated, then we have $P([K] \in \mathcal{B}) \geq 1 - \alpha$. But the problem is that we don't know the variance structure of $\sqrt{n}\hat{\theta}$. Thus an asymptotic theorem has been proved to allow that we still have the desired coverage probability as long as R itself can be consistently estimated.

4. The Simulation and Results

In this study, we focus on using simulation data and clinical data to:

1. Test if the two methods proposed will provide or include the optimal DTR(s) with the desired probability, and
2. Assess how good they are performing in excluding other DTRs

Simulations in this study are constructed based on a two-stage SMART that has eight DTRs in total and the generative model for final outcome Y is set as follows:

$$Y = \gamma_1 + \gamma_2 X_1 + \gamma_3 A_1 + \gamma_4 X_1 A_1 + \gamma_5 A_2 + \gamma_6 X_2 A_2 + \gamma_7 A_1 A_2 + \epsilon, \epsilon \sim N(0,1)$$

where Y is the final stage outcome; A_i = {1, -1} is the ith stage treatment decision for i = 1, 2; X_i = {1, -1} is the ith stage outcome for i = 1, 2.

This set-up allows different scenarios to be constructed by using different sets of γ vector $(\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7)$ and δ vector (δ_1, δ_2) that determines $X_2|X_1, A_1$ as $X_2|X_1, A_1 \sim$ Bernoulli(expit($\delta_1 X_1 + \delta_2 A_1$)), expit(x) = $e^x/(1+e^x)$.

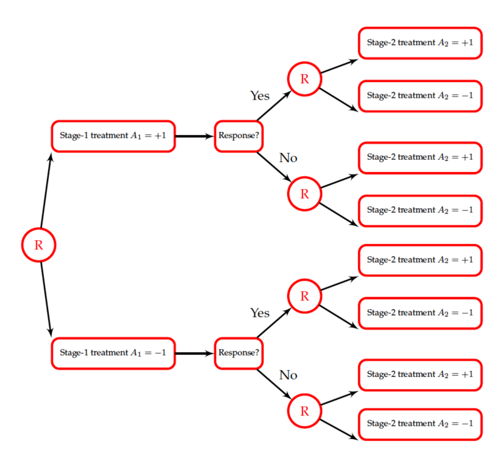
The way to determine whether ACI or MCB is a better method in different scenarios is by comparing the average set size from each method, which is the sum of probabilities of including each of the DTRs within each method. The result tables are shown below.

The labels in the following tables are defined as:
$N_{iter}, SSize$ = "Number of iteration and Sample size", $BestIdx$ = "Best DTR index",
P_{ACI} = "Best DTR Probabilities using ACI", P_{MCB} = "Best DTR Probabilities using MCB",
S_{ACI} = "Average set size ACI", S_{MCB} = "Average set size MCB"

(a) Scenario 1: MCB performs better than ACI.

N_{iter}, $SSize$	$BestIdx$	P_{ACI}	P_{MCB}	S_{ACI}	S_{MCB}
10000, 150	1-8	0.9638	0.9393	7.6990	7.5162
		0.9655	0.9416		
		0.9594	0.9442		
		0.9646	0.9410		
		0.9634	0.9381		
		0.9611	0.9376		
		0.9604	0.9366		
		0.9608	0.9378		

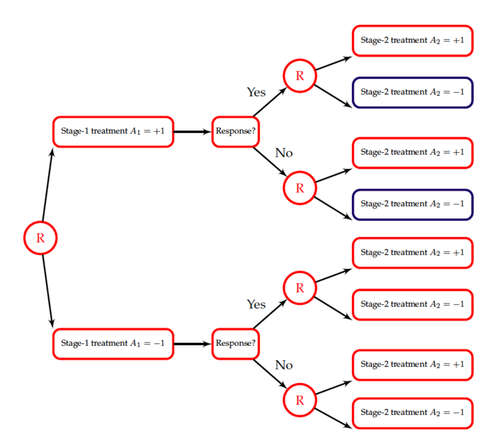
(b)



(c) Scenario 2: MCB performs similar to ACI.

N_{iter}, $SSize$	$BestIdx$	P_{ACI}	P_{MCB}	S_{ACI}	S_{MCB}
10000 150	1,5,6,7,8	0.9842	0.9666	4.8283	4.8174
		0.9630	0.9628		
		0.9608	0.9642		
		0.9599	0.9612		
		0.9604	0.9617		

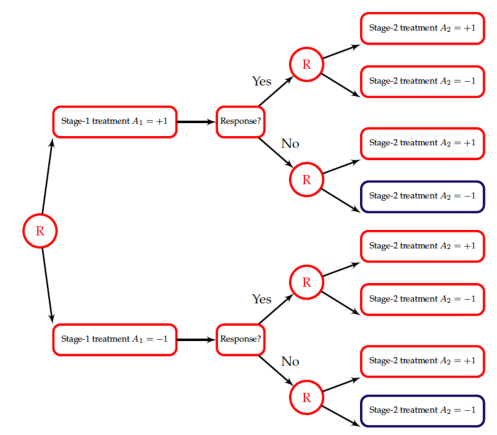
(d)



(e) Scenario 3: MCB performs similar to ACI.

N_{iter}, $SSize$	$BestIdx$	P_{ACI}	P_{MCB}	S_{ACI}	S_{MCB}
2000 150	1,3,5,7	0.9815	0.9745	3.8925	3.8815
		0.9705	0.9690		
		0.9730	0.9642		
		0.9675	0.9620		

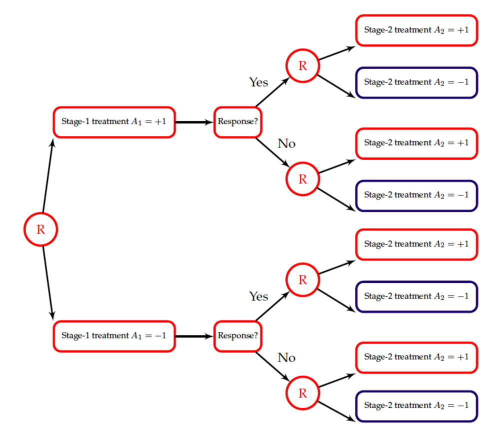
(f)



(g) Scenario 4: ACI performs better than MCB.

N_{iter}, $SSize$	$BestIdx$	P_{ACI}	P_{MCB}	S_{ACI}	S_{MCB}
2000, 150	4,8	0.9820	0.9905	1.9585	1.9800
		0.9765	0.9890		

(h)



5. Conclusion and Future Study

We recommend the MCB method in general because:
1. On average, MCB performs better than ACI.
2. In real clinical trails, Scenario One is the more common case.
The next step is to use data from Extending Treatment Effectiveness of Naltrexone (EXTEND) study to further test the performance of the two methods.