Published in final edited form as:

Contemp Clin Trials. 2017 January ; 52: 46–53. doi:10.1016/j.cct.2016.11.002.

A Response-Adaptive Design of Initial Therapy for Emergency Department Patients with Heart Failure

Sijin Wen^a, Jing Ning^b, Sean Collins^c, and Donald Berry^{b,*}

^aDepartment of Biostatistics, School of Public Health, West Virginia University, Morgantown, WV 26506, USA

^bDepartment of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

^cDepartment of Emergency Medicine Vanderbilt University, Nashville, TN 37232, USA

Abstract

Finding safe and effective treatments for acute heart failure syndrome (AHFS) is a high priority. More than 80% of patients with AHFS who present to emergency departments are treated identically with intravenous diuretics, despite recognition of the syndrome's heterogeneity. We hypothesize that matching patient profiles with "personalized" AHFS treatments will improve outcomes. Matching multiple heterogeneous clinical profiles with a number of potentially effective treatments requires an adaptive trial design that can adjust for these complexities. We propose a Bayesian response-adaptive randomization trial design for AHFS patients. Baseline information collected for each patient with AHFS prior to randomization includes blood pressure, renal function, and dyspnea severity. The primary outcome is discharge readiness within 23 hours of presentation and no unplanned emergency visits or admissions for acute heart failure within 7 days of discharge. We use a Bayesian logistic regression model to characterize the association between primary outcome and patient profile. We adaptively randomize patients to one of five treatments, basing the randomization probability on the cumulative data from the ongoing trial and fitting results from the regression model. Simulations show high probability of selecting the best treatment corresponding to the patient's profile while allocating more patients to the efficacious treatments within the trial.

Keywords

Adaptive randomization; Bayesian; Clinical trial design; Urn design; Heart failure

*Corresponding author: dberry@mdanderson.org (Donald Berry).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Randomization ensures that the observed treatment effect is attributable to the treatment itself rather than to confounding elements, and is the hallmark of clinical trials assessing treatment effects. Recently, the response-adaptive (RA) randomization scheme has become popular in clinical research because of its flexibility and efficiency [1–7]. Based on the accruing history of patient responses to treatment, the RA randomization scheme adjusts the future allocation probabilities, thereby allowing more patients to be assigned to the superior treatment as the trial progresses. As a result, RA randomization can offer significant ethical and cost advantages over equal randomization. The Bayesian framework is particularly suitable for RA designs because it can incorporate historical information as a prior and allows for continually updating the model fitting based on cumulative data observed over time [8, 9].

When we have more than one experimental treatment to test against the traditional therapy, testing each treatment in separate trials is inefficient and wasteful of limited resources. To address this issue, many phase II designs for multiple treatments or combinations have been proposed for oncology trials [7, 10–12]. This paper is motivated by the challenge of designing a randomized trial with multiple treatment arms for patients with acute heart failure syndrome (AHFS). Heart failure affects nearly 6 million Americans and results in nearly one million annual hospital discharges [13]. It is estimated that by 2030, a 25% increase in the prevalence of AHFS will result in an additional 3 million people afflicted [13, 14]. The in-hospital mortality reported in major registries ranges from 4–12%, and may increase to 20–25% in high-risk subgroups [15–20]. Finding safe and effective treatments for patients with AHFS is a critical unmet need identified as a high priority by investigators of heart failure [21, 22]. Despite increasing recognition of disease heterogeneity and comorbidities in patients with AHFS, more than 80% are treated homogeneously with intravenous diuretics [15, 23]. An individualized approach that tailors therapy to improve symptoms, minimize adverse events, and promote earlier discharge has not been studied, largely due to significant limitations in current clinical trial design [24–27]. Such approaches study one therapy in a highly select group of patients, and have limited ability to test multiple treatment approaches within AHFS phenotypes, especially early in a patients course when symptoms are maximal. It is critical to explore novel approaches to improve clinical trial design.

Bayesian adaptive trial design is ideally suited for complex diseases and allows for the study of multiple treatment options in diverse patient phenotypes. After a burn-in period of equal treatment assignment, the randomization scheme is continuously adapted such that treatment allocation is directed toward a responding phenotype. This approach has successfully answered complex hypotheses in cancer trials, but represents an innovative and promising methodology for other complex diseases such as AHFS [2, 28]. Our design is further based on a significant body of research suggesting an association between initial AHFS therapy and near-term events [27, 29, 30]. This strongly suggests that AHFS patients with hypertension benefit from predominant vasodilator therapy, while those who are normotensive may improve with aggressive diuresis [31, 32]. We hypothesize that “personalizing” AHFS treatments by matching patient profiles with targeted treatment

options will more rapidly alleviate symptoms, improve outcomes and serve as a well-defined model for testing new therapies. We consider the following questions: For any treatment that does not work in all patients, is there a subset of patients in which it does work? Can we treat patients better during the trial based on each patient's profile? The Bayesian RA randomization design is ideally suited for facing these voluminous challenges. Bayesian RA designs use information existing at the initiation of the trial and combine it with data that accumulate during the trial to identify which treatments are most beneficial for which patients. Randomization is varied so that patients who are unlikely to benefit from a particular treatment are less likely to receive that treatment. Another benefit of this trial design is that ineffective treatments can be dropped and alternatives added as the trial is ongoing.

The remainder of this article is organized as follows. In Section 2, we specify our models and describe the proposed trial design. In Section 3, we evaluate the operating characteristics of the proposed design through simulation studies. We conclude with a brief discussion and comments in Section 4.

2. Methods

We implement a Bayesian RA randomization scheme in a clinical trial design to screen for effective treatments and identify which patient groups will and will not benefit from the treatments. To illustrate this design, we use a trial of initial therapy in patients who present to emergency departments with heart failure. The first line of emergency treatment for these patients consists of a diuretic (furosemide) and a vasodilator (nitroglycerin). The five targeted treatments (Figure 1) are low-dose diuretics plus intravenous (IV) vasodilator boluses (treatment 1), low-dose diuretics plus IV vasodilator infusion (treatment 2), daily-dose diuretics plus topical vasodilator (treatment 3), high-dose diuretics plus topical vasodilator (treatment 4), and daily-dose diuretics plus IV vasodilator infusion (treatment 5). A total of 1000 patients will be enrolled in the study. All enrolled and eligible patients are required to undergo baseline phenotype profile assessment before randomization. Specifically, three phenotypes are assessed before randomization for all enrolled patients: blood pressure ($> 160\text{mmHg}$ vs $< 160\text{mmHg}$), renal function (glomerular filtration rate [GFR] $< 60\text{ml/min}/1.73\text{m}^2$ vs $> 60\text{ml/min}/1.73\text{m}^2$) and dyspnea severity (initial respiration rate [RR] > 24 breaths/min vs < 24 breaths/min). Thus, this trial will sequentially enroll patients and assign each patient to receive one of five competing treatments based on his/her phenotype profile. Given five treatments and eight subgroups defined by patient phenotypes, there are forty subgroups in this trial.

2.1. Probability model

The limited sample size and low prevalence of some phenotype subgroups make it difficult to draw inferences for the low prevalence subgroups when considered separately. We address this challenge by using covariate analysis to borrow information across subgroups to enhance the precision of the inference. We use a logistic regression model to evaluate the main effects of the five treatments and eight patient phenotypes (covariates) and their interactions. Our formulation of the model, which includes the treatment effects (β^T),

covariate effects (β^Z), and treatment-covariate interactions (β^J), is shown hereafter. For patient i ,

$$\text{logit}(Pr(Y_i=1)) = \beta_0 + \sum_{j=1}^3 \beta_j^Z Z_{ij} + \sum_{k=1}^5 \beta_k^T T_{ik} + \sum_{j=1}^3 \sum_{k=1}^5 \beta_{jk}^J Z_{ij} T_{ik}, \quad (1)$$

where Y_i is the response indicator that takes a value of 1 if the i th patient is discharged within 23 hours of presentation and has no unplanned emergency visits or admissions for acute heart failure within 7 days of discharge, Z_{ij} is the phenotype covariate j , and T_{ik} is the indicator for the k th treatment. To complete the Bayesian model formulation, we assume that the parameters follow a multivariate normal distribution with independent components. Historical data or elicited information may be used to construct informative priors on non-treatment parameters, such as the effects of phenotype covariates in the present setting. However, to ensure that the trial design is both ethical and widely acceptable, we take the prior of each component to be $Normal(\hat{\beta}_{MLE}, \sigma^2)$, where $\hat{\beta}_{MLE}$ is the maximum likelihood estimators of the regression coefficients from the run-in phase in this study (the first 200 patients) and $\sigma^2 \sim Gamma(0.001, 0.001)$. This Bayesian structure via logistic regression in the RA randomization design allows for borrowing strength of information across patients who are receiving the same treatment but have different phenotypes. Hence, the information obtained in one subset of patients provides some information that is useful for patients with similar phenotype profiles.

2.2. Trial design

We propose an RA randomization scheme for the conduct of the trial. We assume that a total of 1000 patients will be enrolled in this study. Since the discharge rate for each of the treatments by phenotype combination is unknown at the beginning of the trial, we equally randomize the first 200 patients into the five competing treatment arms. After that, we randomize the patients based on the estimated posterior means of the discharge rates under each of the five treatments. Let P_{ik} denote the discharge rate of patient i undergoing treatment k . Utilizing the aforementioned Bayesian logistic model and all available data, we continuously update the posterior distributions of the probability P_{ik} during the trial. With the posterior distribution of P_{ik} , $k = 1, \dots, 5$, we calculate the posterior probability that the k th treatment is best for subject i ,

$$P_{ik}^M = Pr(p_{ik} = \max\{p_{ij}, j=1, \dots, 5\} | data).$$

An intuitive way to perform RA randomization is to assign patients to treatment k with the posterior probability P_{ik}^M . However, this leads to some undesirable properties [33], such as a substantial risk of unbalancing the samples in favor of the inferior treatment and a very low probability of selecting a superior treatment due to the fact that the probability P_{ik}^M is quite variable. To deal with this issue, we assign patients to the treatment k with the following stabilized probability

$$R_{ik}(c) = \frac{\left(P_{ik}^M\right)^c}{\sum_{j=1}^5 \left(P_{ij}^M\right)^c},$$

where c is a positive tuning parameter. The value of $c = 0$ gives us conventional equal randomization, and $c=1$ gives us $R_{ik}(1) = P_{ik}^M$. In practice, we should use a value of c between 0 and 1. According to Thall et al. [34] and Rosenberger et al. [35], the value of $c=1/2$ gives a good performance in that it maintains the balance between treating patients efficiently and collecting reasonable information on all treatment arms.

At the end of the trial, treatment K will be declared effective for some phenotype subgroup if the probability that the k th treatment is best for this subgroup is greater than a pre-specified control rate θ_U . The value of this control rate is selected via simulations to ensure that the design has desirable operating characteristics (see the section on simulations for details).

2.3. Adaptive blocked randomization

Due to the large number of subgroups created from the possible combinations of phenotype profile and treatment, there is a possibility of having a severe imbalance in the numbers of patients assigned to the various treatment arms for some phenotype subgroups. As a result, inference about treatment effects and their interactions with phenotypes may suffer from a loss in precision. To ensure that each subgroup has a certain number of patients assigned to each of the five treatment arms, we use the adaptive blocked randomization approach to balance the conflict between collective information and individual ethics.

We start with eight hypothetical urns corresponding to the eight subgroups classified by the three phenotype covariates. For convenience, we denote the balls in the urns by different types that correspond with the five treatments: type 1, type 2, type 3, type 4 or type 5. At the start of the trial, there are 10 balls in each urn, 2 balls for each type. Patients are assigned to a treatment by randomly selecting a ball from the urn corresponding to the patient's phenotype. The balls are drawn completely at random; hence the unequal (adaptive) treatment allocation is only through the unequal allocation of balls in each urn. The allocation of balls within an urn is only updated when the number of balls in an urn equals 5, in which 10 balls are added to the urn before patient randomization. The 10 additional balls will consist of equal numbers of each type during the burn-in phase, and will generally consist of greater numbers for the effective treatments for a given phenotype during the adaptive phase, which results in greater probability of randomization to the successful treatments within a given phenotype. Meanwhile, for each urn update, we keep the unused balls in the urn so that the urn builds up more balls representing treatments with fewer patients assigned. Mathematical details for the allocation of the balls to the urns are provided below.

Let n_{ik}^j be the number of current balls of type k in the i th urn immediately preceding the j th update, and let m_{ik}^j denote the number of balls of type k to be added to the i th urn for the urn's j th update. The total number of balls of type k in the i th urn immediately following the

update but before randomization of the patient is given by $n_{ik}^j + m_{ik}^j$. When a new patient is recruited and is waiting for treatment assignment, we first decide the corresponding urn based on the patient's phenotype profile and then check the number of balls left in the urn. If the number of balls remaining is greater than 5, a ball is randomly drawn and then the treatment assignment for this patient is determined by the type of the drawn ball: if the ball is type i , the patient is assigned to treatment i . On the other hand, if the number of balls remaining in the urn is equal to 5 and we have few patients enrolled at that moment, such as fewer than 200 patients, we add 2 balls to the urn for each type and then draw a ball randomly from the urn. Otherwise, we fit the logistic model (described in Section 2.2) by using all available data and calculate the stabilized randomization probabilities $R_{ik}^j(c)$, $k = 1, \dots, 5$. The number of balls to be added to the urn (m_{ik}^j) of type k is a function of the probabilities $R_{ik}^j(c)$, $k = 1, \dots, 5$, and the observed residuals from the previous update, $e_{ik}^{j-1} = 10R_{ik}^{j-1}(c) - m_{ik}^{j-1}$, which represent the differences in the expected number of additional balls and the actual number of additional balls for each type in the previous update ($j-1$). Specifically, for the i th urn, let $a_{ik}^j = \text{round}(10R_{ik}^j(c) + e_{ik}^{j-1})$ be the candidate number of balls of type k to be added to the i th urn for the j th update, where $\text{round}(\cdot)$ is a function to round to the nearest integer. If the total number of candidate balls equals 10, i.e. $\sum_{k=1}^5 a_{ik}^j = 10$, we set the number of additional balls actually added (m_{ik}^j) for each type equal to the candidate number, $m_{ik}^j = a_{ik}^j$. Given the constraint of adding exactly 10 balls, $\sum_{k=1}^5 m_{ik}^j = 10$, we will either add or subtract balls from the candidate number of balls (a_{ik}^j) to obtain the actual number of balls to be added (m_{ik}^j). For example, if the total number of candidate balls equals 9, i.e. $\sum_{k=1}^5 a_{ik}^j = 9$, we add one more ball of type k^* that has the biggest value of candidate residuals to the candidate number (a_{ik}^j) to obtain the actual number to be added, i.e. $m_{ik^*}^j = (a_{ik^*}^j + 1)$ and $m_{ik'}^j = a_{ik'}^j$, $k \neq k^*$, where the candidate residuals are defined as $e_{a,ik}^j = (10R_{ik}^j(c) - a_{ik}^j)$. On the other hand, if $\sum_{k=1}^5 a_{ik}^j = 11$, we compare the five candidate residuals $e_{a,ik}^j$ for the j th update and delete one ball from the candidate balls of type k^* with the smallest residual, i.e. $m_{ik^*}^j = (a_{ik^*}^j - 1)$ and $m_{ik'}^j = a_{ik'}^j$, $k \neq k^*$. If the number of candidate balls is less than 9 or greater than 11, we repeat the iteration of adding or removing balls to the candidate numbers (a_{ik}^j) by the same principle until we have exactly 10 balls to be added. If there are ties for the smallest or largest residuals, the choice of the ball to be removed or added will be randomly selected among the tied types.

For each update, we keep the unused balls in each urn to improve the balance between treatment arms within each phenotype subgroup. In this way, each urn builds up more balls representing the more successful treatments, while we control for a severe imbalance among the treatment arms.

3. Simulations

We used simulations to evaluate the operating characteristics of the proposed Bayesian RA randomization trial design and compare it to the performance of an equal randomization design under several scenarios.

3.1. Data generation

In the simulations, patients were sequentially enrolled and assigned to receive one of five competing treatments. Based on an accrual rate of 40 patients per month, the trial design had a maximum sample size of 1000 patients for a maximum trial duration of roughly 25 months. To ensure that the adaptive criterion was reasonably informative, the trial was begun by equally randomizing 200 patients to the five treatment arms. The patient population was partitioned into eight subgroups: two subgroups based on blood pressure ($> 160\text{mmHg}$ vs $< 160\text{mmHg}$), two subgroups based on renal function ($\text{GFR} < 60\text{ml/min/1.73m}^2$ vs $\text{GFR} > 60\text{ml/min/1.73m}^2$) and two subgroups based on dyspnea severity ($\text{RR} > 24$ breaths/min vs $\text{RR} < 24$ breaths/min), which are denoted as Z_1 , Z_2 and Z_3 , respectively. Three binary phenotype covariates were generated independently for each patient, and the prevalence of each subgroup in the simulation study was based on clinical practice from the literature and historical data (Table 1). Patient responses (binary) were based on the logistic regression model (1) given in Section 2.1.

As seen in Table 1, some subgroups classified by the patient's phenotype covariates have very low prevalence. Hence, we focused on "signatures" that are combinations of some subgroups. Table 2 shows the ten signatures based on (Z_1, Z_2, Z_3) and addressed in the simulation study. Table 3 lists the discharge rates of patients within the phenotype signature groups defined in Table 2 under four different scenarios:

1. We set all coefficients in model (1) equal to 0 except for $\beta_0 = -0.84$ which yielded no treatment-covariate interactions. This is a scenario of the null hypothesis, in which the discharge rate is 0.30 for all patients with signature 1 in Table 2.
2. We set all coefficients in model (1) equal to 0 except for $\beta_0 = -0.84$, $\beta_{21}^I = 2$. This results in success rates of 0.76 on treatment 1 for patients with signature 4 in Table 2: $(-, +, -)$, $(-, +, +)$, $(+, +, -)$, or $(+, +, +)$.
3. We set all coefficients in model (1) equal to 0, except for $\beta_0 = -0.84$, $\beta_{21}^I = 1.7$, and $\beta_{11}^I = 1.7$. This results in success rates of 0.70 on treatment 1 for patients with signature 8 in Table 2: $(-, +, -)$, $(-, +, +)$, $(+, -, -)$, or $(+, -, +)$; and success rates of 0.93 for patients with signature 8 in Table 2: $(+, +, -)$, or $(+, +, +)$.
4. We set all coefficients in model (1) equal to 0, except for $\beta_0 = -0.84$, $\beta_{31}^I = 1.25$, and $\beta_{22}^I = 2.25$. This results in success rates of 0.60 on treatment 1 for patients with signature 6: $(-, -, +)$, $(-, +, +)$, $(+, -, +)$, or $(+, +, +)$; and success rates of 0.80 on treatment 2 for patients with signature 4: $(-, +, -)$, or $(-, +, +)$ $(+, +, -)$, or $(+, +, +)$.

We simulated 1000 data sets for each scenario given in Table 3. For each simulated data set, the starting value of the regression parameter β for the Markov chain Monte Carlo (MCMC) algorithm was the maximum likelihood estimate from the first 200 patients. The MCMC algorithm run was based on 5,000 iterations. The first 2,500 iterations were discarded as a burn-in period for the algorithm to converge.

3.2. Simulation results

In the simulation study, treatment K was declared effective for some phenotype subgroup if the probability that the k th treatment is best for this subgroup is greater than a pre-specified control rate θ_U . We calibrated the value of θ_U via simulations to ensure that the design has desirable operating characteristics. Table 4 lists the selection probabilities under the proposed RA randomization design as well as under equal randomization in scenario 1. By choosing $\theta_U = 0.84$ for the equal randomization design and $\theta_U = 0.82$ for the proposed RA randomization design, the empirical probability of choosing one treatment as the best treatment in 1000 runs by the two designs was around 0.05 for each signature, and corresponded to a two-sided type I error rate of 0.10 in a frequentist design. We used the same values of θ_U for the two designs in all of the scenarios, so that the selection probability for a treatment arm in scenarios 2 through 4 can be interpreted as the power in a frequentist design.

In Table 5, we list the average number of patients (with standard deviation) assigned to each treatment arm. Because there was no effective treatment in the null scenario (scenario 1), on average, both the proposed design and the equal randomization design assigned roughly equal numbers of patients to each treatment arm within each subgroup. However, the standard deviations are quite different under the two types of randomizations. For example, for the subgroup with phenotype profile $(-, -, -)$, the standard deviation of the number of patients assigned to treatment 1 when using the proposed design was almost six times higher than that under equal randomization. Under scenarios 2 through 4, as expected, the average sample sizes for the five treatments were still approximately the same within each subgroup under the equal randomization design (not listed due to space limitations); although there was clear evidence that some treatments outperformed others for some subgroups. In contrast to equal randomization, the proposed RA randomization design assigned more patients to the treatment arms that were better for the patients. Under scenario 2, on average in each of 4 phenotype subgroups with signature 4, the larger proportion of patients was assigned to treatment 1. For example, of the 197.8 patients in the subgroup $(-, +, -)$, 47.9% (94.8/197.8) of the patients were randomized to treatment 1, but only 12.8% (25.4/197.8), 13.1% (25.9/197.8), 12.9% (25.6/197.8), and 13.2% (26.1/197.8) of the patients were randomized to treatments 2 through 5, respectively. On the other hand, no allocation preference was observed for all subgroups that had normal renal function, since none of the treatments were effective for that subgroup. In addition, on average, 181 (94.8+44.2+31.1+10.6) patients with signature 4 were assigned to treatment 1 by the RA randomization design, and the overall response rate was 36.2% (362/1000), in comparison with a response rate of 32.2% (322/1000) for the equal randomization design (Table 6). The unbalanced allocation probability that favored the superior treatment arm resulted in a higher number of patients achieving treatment success under the RA randomization design than

under the traditional equal randomization design. Similarly, under scenarios 3 and 4, the proposed RA randomization design assigned more patients to the better treatments that corresponded with their phenotype signatures.

In scenarios 2, 3 and 4 with treatment-covariate interactions, we evaluated the statistical power under the two randomization designs. Table 6 lists the probabilities of declaring effective treatments for each phenotype signature and the overall number of respondents (i.e., discharged within 23 hours without unplanned emergency visits or admissions for acute heart failure within 7 days of discharge). In scenario 2, treatment 1 outperformed the others for patients with high renal function. The probability of selecting treatment 1 as the best treatment for patients with the signature “renal function +” was about 95 % under the proposed design and 94 % under the equal randomization design. The probability of declaring treatment 1 as the best treatment for patients with the signature “renal function –” was 2% under the proposed design and 3 % under the equal randomization design. When there was one best treatment (treatment 1) that was effective for patients with high blood pressure and/or high renal function (scenario 3), the probability of selecting treatment 1 as the best treatment for patients with high blood pressure and/or high renal function was 90% (65%) under the proposed design, and 90% (71%) under the equal randomization design.

In summary, considering the numbers of patients assigned to better treatment arms, the advantage of the proposed RA randomization design over the traditional equal randomization design is clear. With its ability to adjust the unbalance in patient allocation among treatment arms via the urn model, the power of the proposed RA design is comparable to that of the equal randomization design.

4. Discussion

We have proposed a Bayesian response-adaptive randomization trial design, with the added consideration of balancing patients among treatment arms via the urn model, for a clinical trial of early treatment for patients presenting with acute heart failure syndrome. This design uses information existing at the time of the trial’s initiation, combined with data that accumulates during the trial, to identify the most beneficial treatments for newly enrolled patients. The trial continues to update the inferences about the effects of competing treatments, patients’ phenotype profiles and the treatment-phenotype interactions, and accordingly adjusts the randomization probabilities. Based on this learning process, we adaptively allocate more patients to more effective treatment arms. Simulation studies can be used to calibrate the design parameters to control the type I error rate. Although the imbalanced allocation among treatment arms can result in a mild loss of statistical power for comparing treatments, the proposed design allocates more patients to more effective treatments and substantially outperforms the traditional equal randomization design in terms of efficacy. The sensitivity analysis on the response profiles over time shows that the type I error rates are quite robust against the changes of response rates over time, although they are inflated slightly in the large drift case (Table S1 in the supplemental material).

The RA randomization design generally requires that the response to treatment be quickly ascertainable, and is thus particularly suitable for the proposed trial with a primary outcome

within one week. However, in some applications, the response assessment will require a long follow-up time and consequently the responses of previously enrolled patients will not be observed when a new patient is enrolled and ready for treatment assignment. One possible approach to this issue is to suspend the accrual and wait until the responses of all patients treated in the trial are fully observed before enrolling a new patient. As a result of frequent suspensions of patient accrual, which waste resources and add administrative inconvenience, such an approach may lead to an unfeasibly long trial. Another possible approach to trials involving delayed treatment response is to determine the adaptive randomization probabilities based on the data observed thus far. Although this observed-data-based approach may not fully use all information from the enrolled patients to determine current treatment assignments, it does not require suspending accrual and thus dramatically shortens the trial duration.

As pointed out by Thall [36], response-adaptive randomization designs can result in larger variations in the sample size of each arm and more bias in the estimates of the treatment effect. To reduce the potential for extremely unbalanced arms, we have included the run-in phase, the tuning parameter in the randomization probability and the urn model in the proposed design. Our simulations illustrate that our design ensures that each subgroup has a certain number of patients assigned to each of the five treatment arms, although our design produces a larger variation in the estimates compared with that of designs that use equal randomization. In addition, we have examined the chance that a trial heads off in the wrong direction and ends up allocating more patients to inferior arms in the simulation study. Overall speaking, the chance to have the imbalance in the wrong direction is small, and decreases with the increasing sample size. For example, under Scenario 2 in Table 5, the chance to have the imbalance in the wrong direction is 1% for subgroup $(-, +, -)$, 2% for subgroup $(-, +, +)$, 4% for subgroup $(+, +, -)$ and 6% for subgroup $(+, +, +)$. The higher chances for subgroups $(+, +, -)$ and $(+, +, +)$ are due to the low prevalence of the two subgroups. In terms of estimation precision, the empirical biases of the proposed design are larger than those of equal randomization designs, but all are within a reasonable range (Table S2 in the supplemental material).

A number of Bayesian response-adaptive randomization designs have been implemented in real clinical trials including two adaptive clinical trials conducted at MD Anderson Cancer Center: I-SPY 2 for breast cancer [28] and BATTLE for non-small cell lung cancer [37]. Both of these trials used an adaptive randomization approach for treatment allocation, which utilized information from patients previously enrolled in the trial to guide decisions about which treatments might be more beneficial to future patients. The success of these response-adaptive randomization designs confirmed the feasibility of conducting a Bayesian adaptive randomization design for large and multi-center clinical trials. Such an adaptive randomization scheme provides a scientific basis for researchers to eliminate ineffective treatments, allow effective treatments to graduate more quickly, and assign more patients to the treatments that will likely benefit them the most. This approach has successfully answered complex hypotheses in cancer trials, and represents an innovative and promising methodology for other complex diseases. In this paper, we have proposed a Bayesian response-adaptive randomization trial design for patients presenting with acute heart failure syndrome, with the added consideration of balancing patients among treatment arms via the

urn model. Although the urn model is embedded in the response-adaptive randomization schema to improve the statistical performance, it will not increase the operational burden of implementing the design, which can be easily conducted via a software package with user-friendly instructions. Further, many contemporary clinical trials have required frequent contact with the patient, providing real-time updating of events. There is a nearly a decade of emergency-department-based clinical trial experience in patients with AHFS [38–40]. Many of these studies require continuous monitoring of patients through the first 30 days to accurately capture events. A Bayesian RA design would be welcomed and fit within the construct of our AHFS clinical trial network [41]. With 24-hour interactive voice response systems (IVRS), adjusting randomization blocks and schema is relatively easy and not considered a hurdle from a clinical trial implementation perspective.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The work was supported in part by the U.S. National Institute of Health (NIH) grants U54GM104942, CA193878 and CA016672. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

1. Berry DA. The brave new world of clinical cancer research: Adaptive biomarker-driven trials integrating clinical practice with clinical research. *Molecular oncology*. 2015; 9(5):951–959. [PubMed: 25888066]
2. Berry DA. Adaptive clinical trials in oncology. *Nature Reviews Clinical Oncology*. 2011; 8:199–207.
3. Cheung Y, Inoue L, Wathen J, Thall P. Continuous bayesian adaptive randomization based on event times with covariates. *Statistics in Medicine*. 2006; 25:55–70. [PubMed: 16025549]
4. Hu F, Rosenberger W. Optimality, variability, power: Evaluating response-adaptive randomization procedures for treatment comparisons. *Journal of American Statistical Association*. 2003; 98:671–678.
5. Hu, F., Rosenberger, W. *The Theory of Response-Adaptive Randomization in Clinical Trials*. Wiley; New York: 2006.
6. Hu F, Rosenberger W, Zhang L. Asymptotically best response-adaptive randomization procedures. *Journal of Statistical Planning and Inference*. 2006; 136:1911–22.
7. Zhou X, Liu S, Kim E, Lee J. Bayesian adaptive design for targeted therapy development in lung cancer – a step toward personalized medicine. *Clinical Trials*. 2008; 5:181–93. [PubMed: 18559407]
8. Berry DA. Bayesian statistics and the efficiency and ethics of clinical trials. *Statistical Science*. 2004; 19:175–187.
9. Berry DA. Bayesian clinical trials. *Nature Reviews Drug Discovery*. 2006; 5:27–36. [PubMed: 16485344]
10. Thall PF, Simon R, Ellenberg SS. Two-stage selection and testing designs for comparative clinical trials. *Biometrika*. 1988; 75:303–310.
11. Schaid D, Wieand S, Therneau T. Optimal two-stage screening designs for survival comparisons. *Biometrika*. 1990; 77:507–513.
12. Ji Y, Bekele N. Adaptive randomization for multiarm comparative clinical trials based on joint efficacy/toxicity outcomes. *Biometrics*. 2009; 65:876–884. [PubMed: 19173694]

13. Roger VL, Go AS, Lloyd-Jones DM, Adams RJ, Berry JD, Brown TM, Carnethon MR, Dai S, de Simone G, Ford ES, et al. Heart disease and stroke statistics 2011 update a report from the american heart association. *Circulation*. 2011; 123(4):e18–e209. [PubMed: 21160056]
14. Heidenreich PA, Trogon JG, Khavjou OA, Butler J, Dracup K, Ezekowitz MD, Finkelstein EA, Hong Y, Johnston SC, Khera A, et al. Forecasting the future of cardiovascular disease in the united states a policy statement from the american heart association. *Circulation*. 2011; 123(8):933–944. [PubMed: 21262990]
15. Adams KF, Fonarow GC, Emerman CL, LeJemtel TH, Costanzo MR, Abraham WT, Berkowitz RL, Galvao M, Horton DP, A. S. A. Committee, Investigators. et al. Characteristics and outcomes of patients hospitalized for heart failure in the united states: rationale, design, and preliminary observations from the first 100,000 cases in the acute decompensated heart failure national registry (adhere). *American heart journal*. 2005; 149(2):209–216. [PubMed: 15846257]
16. Cleland J, Swedberg K, Follath F, Komajda M, Cohen-Solal A, Aguilar JC, Dietz R, Gavazzi A, Hobbs R, Korewicki J, et al. The euroheart failure survey programme a survey on the quality of care among patients with heart failure in europe part 1: patient characteristics and diagnosis. *European heart journal*. 2003; 24(5):442–463. [PubMed: 12633546]
17. Follath F, Yilmaz M, Delgado J, Parissis J, Porcher R, Gayat E, Burrows N, McLean A, Vilas-Boas F, Mebazaa A. Clinical presentation, management and outcomes in the acute heart failure global survey of standard treatment (alarm-hf). *Intensive care medicine*. 2011; 37(4):619–626. [PubMed: 21210078]
18. Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, O'Connor CM, Pieper K, Sun JL, Yancy CW, et al. Influence of a performance-improvement initiative on quality of care for patients hospitalized with heart failure: results of the organized program to initiate lifesaving treatment in hospitalized patients with heart failure (optimize-hf). *Archives of internal medicine*. 2007; 167(14):1493–1502. [PubMed: 17646603]
19. Komajda M, Follath F, Swedberg K, Cleland J, Aguilar J, Cohen-Solal A, Dietz R, Gavazzi A, Van Gilst W, Hobbs R, et al. The euroheart failure survey programme a survey on the quality of care among patients with heart failure in europe part 2: treatment. *European Heart Journal*. 2003; 24(5): 464–474. [PubMed: 12633547]
20. Nieminen MS, Brutsaert D, Dickstein K, Drexler H, Follath F, Harjola VP, Hochadel M, Komajda M, Lassus J, Lopez-Sendon JL, et al. Euroheart failure survey ii (ehfs ii): a survey on hospitalized acute heart failure patients: description of population. *European heart journal*. 2006; 27(22):2725–2736. [PubMed: 17000631]
21. Peacock WF, Braunwald E, Abraham W, Albert N, Burnett J, Christenson R, Collins S, Diercks D, Fonarow G, Hollander J, et al. National heart, lung, and blood institute working group on emergency department management of acute heart failure: research challenges and opportunities. *Journal of the American College of Cardiology*. 2010; 56(5):343–351. [PubMed: 20650354]
22. Weintraub NL, Collins SP, Pang PS, Levy PD, Anderson AS, Arslanian-Engoren C, Gibler WB, McCord JK, Parshall MB, Francis GS, et al. Acute heart failure syndromes: Emergency department presentation, treatment, and disposition: Current approaches and future aims a scientific statement from the american heart association. *Circulation*. 2010; 122(19):1975–1996. [PubMed: 20937981]
23. Collins SP, Pang PS, Lindsell CJ, Kyriacou DN, Storrow AB, Hollander JE, Kirk J Douglas, Miller CD, Nowak R, Peacock W Frank, et al. International variations in the clinical, diagnostic, and treatment characteristics of emergency department patients with acute heart failure syndromes. *European journal of heart failure*. 2010; 12(11):1253–1260. [PubMed: 20797987]
24. Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De Simone G, Ferguson TB, Ford E, Furie K, Gillespie C, et al. Heart disease and stroke statistics 2010 update a report from the american heart association. *Circulation*. 2010; 121(7):e46–e215. [PubMed: 20019324]
25. Cotter G, Metzko E, Kaluski E, Faigenberg Z, Miller R, Simovitz A, Shaham O, Marghitay D, Koren M, Blatt A, et al. Randomised trial of high-dose isosorbide dinitrate plus low-dose furosemide versus high-dose furosemide plus low-dose isosorbide dinitrate in severe pulmonary oedema. *The Lancet*. 1998; 351(9100):389–393.

26. Hasselblad V, Stough WG, Shah MR, Lokhnygina Y, O'Connor CM, Califf RM, Adams KF. Relation between dose of loop diuretics and outcomes in a heart failure population: results of the escape trial. *European journal of heart failure*. 2007; 9(10):1064–1069. [PubMed: 17719273]
27. Gottlieb SS, Abraham W, Butler J, Forman DE, Loh E, Massie BM, O'Connor CM, Rich MW, Stevenson LW, Young J, et al. The prognostic importance of different definitions of worsening renal function in congestive heart failure. *Journal of cardiac failure*. 2002; 8(3):136–141. [PubMed: 12140805]
28. Barker A, Sigman C, Kelloff G, Hylton N, Berry D, Esserman L. I-spy 2: An adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clinical Pharmacology & Therapeutics*. 2009; 86(1):97–100. [PubMed: 19440188]
29. Aronson D, Burger AJ. The relationship between transient and persistent worsening renal function and mortality in patients with acute decompensated heart failure. *Journal of cardiac failure*. 2010; 16(7):541–547. [PubMed: 20610229]
30. Mebazaa A, Nieminen MS, Packer M, Cohen-Solal A, Kleber FX, Pocock SJ, Thakkar R, Padley RJ, Pöder P, Kivikko M, et al. Levosimendan vs dobutamine for patients with acute decompensated heart failure: the survive randomized trial. *Jama*. 2007; 297(17):1883–1891. [PubMed: 17473298]
31. Levy P, Compton S, Welch R, Delgado G, Jennett A, Penugonda N, Dunne R, Zalenski R. Treatment of severe decompensated heart failure with high-dose intravenous nitroglycerin: a feasibility and outcome analysis. *Annals of emergency medicine*. 2007; 50(2):144–152. [PubMed: 17509731]
32. Felker GM, Lee KL, Bull DA, Redfield MM, Stevenson LW, Goldsmith SR, LeWinter MM, Deswal A, Rouleau JL, Ofili EO, et al. Diuretic strategies in patients with acute decompensated heart failure. *New England Journal of Medicine*. 2011; 364(9):797–805. [PubMed: 21366472]
33. Thall PF, Wathen JK. Practical bayesian adaptive randomization in clinical trials. *Eur J Cancer*. 2007; 43:859–866. [PubMed: 17306975]
34. Thall PF, Inoue LYT, Martin TG. Adaptive decision making in a lymphocyte infusion trial. *Biometrics*. 2002; 58:560–568. [PubMed: 12229990]
35. Rosenberger W, Stallard N, Ivanova A, Harper C, Ricks M. Optimal adaptive designs for binary response trials. *Biometrics*. 2001; 57:909–913. [PubMed: 11550944]
36. Thall P, Fox P, Wathen J. Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials. *Annals of Oncology*. 2015; 26(8):1621–8. [PubMed: 25979922]
37. Kim ES, Herbst RS, Wistuba II, Lee JJ, Blumenschein GR, Tsao A, Stewart DJ, Hicks ME, Erasmus J, Gupta S, et al. The battle trial: personalizing therapy for lung cancer. *Cancer discovery*. 2011; 1(1):44–53. [PubMed: 22586319]
38. McMurray JJ, Teerlink JR, Cotter G, Bourge RC, Cleland JG, Jondeau G, Krum H, Metra M, O'Connor CM, Parker JD, et al. Effects of tezoseptan on symptoms and clinical outcomes in patients with acute heart failure: the veritas randomized controlled trials. *Jama*. 2007; 298(17):2009–2019. [PubMed: 17986694]
39. O'Connor C, Starling R, Hernandez A, Armstrong P, Dickstein K, Hasselblad V, Heizer G, Komajda M, Massie B, McMurray J, et al. Effect of nesiritide in patients with acute decompensated heart failure. *New England Journal of Medicine*. 2011; 365(1):32–43. [PubMed: 21732835]
40. Teerlink JR, Cotter G, Davison BA, Felker GM, Filippatos G, Greenberg BH, Ponikowski P, Unemori E, Voors AA, Adams KF, et al. Serelaxin, recombinant human relaxin-2, for treatment of acute heart failure (relax-ahf): a randomised, placebo-controlled trial. *The Lancet*. 2013; 381(9860):29–39.
41. Collins SP, Levy PD, Lindsell CJ, Pang PS, Storrow AB, Miller CD, Naftilan AJ, Thohan V, Abraham WT, Hiestand B, et al. The rationale for an acute heart failure syndromes clinical trials network. *Journal of cardiac failure*. 2009; 15(6):467–474. [PubMed: 19643356]

Figure 1. Schematic of treatment combinations of nitroglycerin and furosemide dosing and route (B = intravenous vasodilator bolus; I = intravenous vasodilator infusion; T = topical vasodilator; R = regular/daily-dose diuretic; L = low-dose diuretic; H = high-dose diuretic). The open circles are potential combinations not considered in this trial.

Table 1

Prevalence of the eight subgroups in the simulations

Blood Pressure	Dyspnea GFR < 60	(RR < 24) GFR > 60	Dyspnea GFR < 60	(RR > 24) GFR > 60
< 160mmHg	0.3	0.2	0.15	0.10
> 160mmHg	0.1	0.07	0.05	0.03

Table 2

Candidate phenotype signatures for the study

Signature	Possible subgroups based on phenotype (Z_1, Z_2, Z_3) [†]									
	(-, -, -)	(-, -, +)	(-, +, -)	(-, +, +)	(+, -, -)	(+, -, +)	(+, +, -)	(+, +, +)	(+, -, +)	(+, +, +)
1 (All)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2 (Blood pressure +)	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
3 (Blood pressure -)	Yes	Yes	Yes	Yes	No	No	No	No	No	No
4 (Renal function +)	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes
5 (Renal function -)	Yes	No	No	No	Yes	No	No	Yes	No	No
6 (Dyspnea severity +)	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes
7 (Dyspnea severity -)	Yes	Yes	Yes	No	No	No	Yes	No	No	No
8 (Blood pressure +, Renal function +)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9 (Blood pressure +, Dyspnea severity +)	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10 (Renal function +, Dyspnea severity +)	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Estimated prevalence	0.30	0.20	0.15	0.10	0.10	0.10	0.07	0.05	0.03	0.03

[†]Yes indicates that the respective subtype is included in the signature, no indicates that it is not

Response rates in the four simulation scenarios where response is defined as the patient being discharged within 23 hours of presentation and had no unplanned emergency visits or admissions for acute heart failure within 7 days of discharge.

Table 3

(Z_1, Z_2, Z_3)	$(-, -, -)$	$(-, -, +)$	$(-, +, -)$	$(-, +, +)$	$(+, -, -)$	$(+, -, +)$	$(+, +, -)$	$(+, +, +)$
Scenario 1								
Treatment 1	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 2	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 3	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 4	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 5	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Scenario 2								
Treatment 1	0.30	0.76	0.30	0.76	0.30	0.76	0.30	0.76
Treatment 2	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 3	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 4	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 5	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Scenario 3								
Treatment 1	0.30	0.70	0.30	0.70	0.7	0.93	0.70	0.93
Treatment 2	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 3	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 4	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 5	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Scenario 4								
Treatment 1	0.30	0.30	0.30	0.60	0.3	0.30	0.60	0.60
Treatment 2	0.30	0.80	0.30	0.80	0.30	0.80	0.30	0.80
Treatment 3	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 4	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Treatment 5	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30

Selection proportions by equal randomization and RA randomization designs under the scenario without treatment-covariate interactions (Scenario 1)

Table 4

Treatment	1	2	3	4	5	Overall
Equal randomization design:						
Blood pressure +	0.006	0.014	0.011	0.014	0.008	0.046
Blood pressure –	0.005	0.008	0.015	0.012	0.014	0.053
Renal function +	0.001	0.012	0.010	0.014	0.012	0.050
Renal function –	0.010	0.009	0.012	0.008	0.009	0.048
Dyspnea severity +	0.005	0.012	0.010	0.012	0.011	0.049
Dyspnea severity –	0.006	0.010	0.013	0.011	0.010	0.050
Proposed RA randomization design						
Blood pressure +	0.002	0.011	0.020	0.010	0.011	0.054
Blood pressure –	0.002	0.012	0.011	0.005	0.010	0.041
Renal function +	0.000	0.009	0.015	0.009	0.014	0.046
Renal function –	0.005	0.005	0.022	0.006	0.007	0.046
Dyspnea severity +	0.001	0.007	0.020	0.007	0.016	0.052
Dyspnea severity –	0.004	0.006	0.018	0.007	0.005	0.040

Table 5

Mean numbers of patients (standard deviation) assigned to each treatment arm from a total of 1000 patients

(Z_1, Z_2, Z_3)	$(-, -, -)$	$(-, -, -)$	$(-, +, -)$	$(-, +, +)$	$(+, -, -)$	$(+, -, +)$	$(+, +, -)$	$(+, +, +)$
Equal randomization design								
Treatment 1	60.1(3.1)	40.1(2.6)	29.8(2.3)	20.1(2.0)	19.8(1.9)	14.1(1.9)	10.0(1.6)	6.0(1.4)
Treatment 2	60.1(3.1)	40.0(2.6)	29.9(2.3)	20.0(2.1)	19.9(1.9)	14.1(1.8)	10.1(1.7)	6.0(1.5)
Treatment 3	60.0(3.1)	39.9(2.6)	29.9(2.2)	20.0(2.0)	19.9(2.0)	13.9(1.9)	10.0(1.7)	6.1(1.4)
Treatment 4	60.0(3.2)	40.0(2.6)	29.9(2.3)	20.1(2.0)	20.0(2.0)	14.0(1.8)	10.1(1.6)	6.0(1.4)
Treatment 5	60.0(3.0)	40.1(2.6)	30.0(2.3)	20.1(2.1)	19.9(1.9)	13.9(1.9)	10.1(1.7)	6.0(1.4)
RA randomization design under scenario 1								
Treatment 1	65.2(24.5)	37.3(14.4)	32.5(10.7)	18.6(5.9)	21.0(6.9)	12.8(4.3)	10.1(3.6)	5.6(1.9)
Treatment 2	55.9(17.8)	39.6(13.6)	29.4(11.0)	21.6(7.1)	18.2(5.7)	13.5(4.6)	9.90(3.3)	6.1(2.2)
Treatment 3	59.9(21.4)	40.4(14.1)	28.2(11.3)	19.2(6.8)	22.0(7.3)	14.7(4.8)	10.3(3.4)	5.9(1.9)
Treatment 4	57.7(19.0)	40.0(13.4)	30.1(10.1)	19.6(6.2)	20.3(6.8)	14.8(5.2)	10.3(3.4)	6.1(2.0)
Treatment 5	60.3(20.5)	42.5(15.0)	30.3(10.3)	21.3(7.4)	19.5(6.6)	13.5(4.8)	10.0(3.4)	6.1(2.1)
RA randomization design under scenario 2								
Treatment 1	66.6(26.2)	94.8(10.4)	32.6(10.2)	44.2(7.0)	20.0(7.3)	31.1(4.6)	10.5(4.0)	10.6(3.4)
Treatment 2	56.4(20.7)	25.4(2.2)	30.4(8.9)	14.3(2.8)	19.0(7.1)	10.3(1.9)	10.1(3.4)	5.2(1.7)
Treatment 3	59.0(19.6)	25.9(3.2)	26.7(8.1)	13.2(2.2)	20.2(6.1)	10.1(2.0)	9.80(3.5)	4.7(1.4)
Treatment 4	56.7(21.2)	25.6(2.1)	30.0(10.2)	13.6(1.8)	21.0(7.8)	11.0(2.4)	10.5(3.3)	4.8(1.2)
Treatment 5	64.6(23.7)	26.1(3.5)	27.7(8.9)	13.6(2.0)	18.7(6.9)	10.4(1.8)	9.60(3.2)	4.9(1.4)
RA randomization design under scenario 3								
Treatment 1	64.2(21.8)	65.7(16.4)	32.4(11.6)	29.9(10.7)	42.2(7.8)	32.1(4.8)	18.5(4.8)	11.3(3.9)
Treatment 2	53.3(17.4)	31.8(9.3)	30.3(11.2)	18.3(5.1)	13.8(1.9)	10.0(1.6)	7.6(2.3)	4.8(1.4)
Treatment 3	60.6(22.5)	32.8(8.8)	26.0(8.1)	15.9(4.3)	15.1(3.3)	9.90(1.5)	8.2(2.9)	4.5(1.3)
Treatment 4	60.0(23.2)	32.8(10.7)	32.2(10.9)	17.4(6.1)	14.2(2.7)	10.3(1.9)	8.1(1.8)	4.7(1.3)
Treatment 5	65.0(21.5)	35.5(11.2)	27.5(8.1)	17.2(5.4)	13.7(2.0)	10.0(1.6)	7.7(1.8)	4.6(1.3)
RA randomization design under scenario 4								
Treatment 1	64.8(23.2)	32.4(4.5)	37.4(8.9)	35.9(5.8)	21.3(4.3)	11.6(4.1)	18.8(4.0)	7.2(1.9)
Treatment 2	63.5(18.5)	96.2(9.1)	31.1(8.0)	36.4(5.2)	21.5(4.5)	30.1(8.7)	8.0(2.1)	9.5(2.3)
Treatment 3	54.9(18.3)	23.5(3.7)	27.2(6.9)	9.5(3.1)	18.4(4.1)	10.3(1.9)	7.7(2.1)	4.5(1.4)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

(Z_1, Z_2, Z_3)	$(+, -, +)$	$(-, -, +)$	$(-, +, +)$	$(+, -, -)$	$(+, +, -)$	$(-, -, -)$	$(-, +, -)$	$(+, -, -)$
Treatment 4	4.1(1.2)	7.6(1.9)	9.9(1.9)	19.1(4.0)	9.1(2.8)	27.3(7.1)	24.1(4.0)	60.1(23.9)
Treatment 5	4.2(1.3)	7.9(2.0)	8.7(1.6)	19.5(4.3)	9.2(2.9)	26.8(6.8)	23.7(3.8)	56.5(18.5)

Table 6

Selection proportions under equal randomization and RA randomization designs under scenarios with treatment-covariate interactions

Selection proportions under scenario 2:										
Treatment	Equal randomization ($r^{\dagger}=322$)					RA randomization ($r^{\dagger}=362$)				
	1	2	3	4	5	1	2	3	4	5
Blood pressure +	0.47	0.00	0.00	0.00	0.00	0.48	0.00	0.00	0.01	0.00
Blood pressure –	0.50	0.00	0.00	0.00	0.00	0.51	0.01	0.00	0.00	0.00
Renal function +	0.94	0.00	0.00	0.00	0.00	0.95	0.00	0.00	0.00	0.00
Renal function –	0.03	0.01	0.01	0.01	0.01	0.02	0.01	0.00	0.01	0.00
Dyspnea severity +	0.47	0.00	0.00	0.00	0.00	0.48	0.00	0.00	0.01	0.00
Dyspnea severity –	0.50	0.00	0.00	0.00	0.00	0.51	0.00	0.00	0.0	0.00

Selection proportions under scenario 3:										
Treatment	Equal randomization ($r^{\dagger}=359$)					RA randomization ($r^{\dagger}=449$)				
	1	2	3	4	5	1	2	3	4	5
Blood pressure +	0.90	0.01	0.00	0.01	0.01	0.90	0.00	0.00	0.00	0.00
Blood pressure –	0.26	0.00	0.01	0.00	0.00	0.17	0.01	0.01	0.00	0.00
Renal function +	0.71	0.00	0.00	0.00	0.00	0.65	0.00	0.00	0.00	0.00
Renal function –	0.44	0.01	0.00	0.00	0.01	0.43	0.01	0.01	0.00	0.00
Dyspnea severity +	0.52	0.01	0.00	0.00	0.01	0.48	0.00	0.00	0.00	0.00
Dyspnea severity –	0.64	0.00	0.01	0.00	0.00	0.60	0.00	0.01	0.00	0.00

Selection proportions under scenario 4:										
Treatment	Equal randomization ($r^{\dagger}=366$)					RA randomization ($r^{\dagger}=433$)				
	1	2	3	4	5	1	2	3	4	5
Blood pressure +	0.26	0.38	0.00	0.00	0.00	0.28	0.38	0.00	0.00	0.00
Blood pressure –	0.26	0.39	0.00	0.00	0.00	0.27	0.38	0.00	0.00	0.00
Renal function +	0.02	0.80	0.00	0.00	0.00	0.04	0.79	0.00	0.00	0.00
Renal function –	0.50	0.00	0.00	0.01	0.01	0.52	0.00	0.00	0.00	0.00
Dyspnea severity +	0.53	0.26	0.00	0.00	0.00	0.53	0.25	0.00	0.00	0.00
Dyspnea severity –	0.01	0.50	0.00	0.01	0.01	0.02	0.50	0.00	0.00	0.00

$\bar{x}_j$ = mean numbers of respondents (i.e., discharged within 23 hours and without visits or admissions within 7 days) out of 1000 patients.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript