Real-time Adaptive Randomization of Clinical Trials

May 21, 2024

Prof. Gui Liberali,[a,c] Prof. Eric Boersma,[b] Hester Lingsma,[c] Jasper Brugts,[b] Prof. Diederik Dippel,[d] Prof. Jan Tijssen[e], Prof. John Hauser[f]

a. Erasmus University, Rotterdam School of Management, Rotterdam
b. Erasmus MC, University Medical Center, Department of Cardiology, Rotterdam
c. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam
d. Erasmus MC, University Medical Center, Department of Neurology, Rotterdam
e. Amsterdam UMC, University of Amsterdam, Department of Cardiology, Amsterdam
f. Massachusetts Institute of Technology (MIT), MIT Sloan School of Management, Cambridge, U.S.

Contact author: Gui Liberali, liberali@rsm.nl. Address for correspondence: P.O. box 1738. 3000 DR Rotterdam. Netherlands.

Keywords: adaptive clinical trials, multi-armed bandits, response-adaptive randomization

Acknowledgements: The authors are very thankful for comments and support from Eric Langhamer, Diana Toli, Gaelle Saint Hilary, and Laura Velasco, and for the clinical trial data from Servier and Duke University School of Medicine

Real-time Adaptive Randomization of Clinical Trials

Abstract

Objective: To evaluate real-time (day-to-day) adaptation of randomized controlled clinical trials (RCTs) with delayed endpoints – a "forward-looking optimal-experimentation" form of response-adaptive randomization (RAR). To identify the implied tradeoffs between lowered mortality, confidence intervals, statistical power, and potential misidentification based on empirically-grounded data and on simulations of temporal changes in endpoint rates.

Study Design and Setting: Using data from RCTs in acute myocardial infarction (30,732 patients in GUSTO-1) and coronary heart disease (12,218 patients in EUROPA), we resample treatment-arm assignments and expected endpoints to simulate (1) real-time assignment, (2) forward-looking multiday-block-based assignment, and (3) an η-variant that balances RCT and real-time assignments. (Stationarity, which enables resampling, is tested.) Blinded day-to-day arm assignments are adjusted by optimizing the tradeoff between assigning the (likely) best treatment and learning about endpoint rates for future random assignments.

Results: Despite delays in endpoints (in EUROPA), real-time assignment quickly learns which arm is superior. In the simulations, by the end of the trials, real-time assignment allocated more patients to the superior arm and fewer patients to the inferior arm(s) resulting in fewer mortalities over the course of the trial. Endpoint probabilities and odds ratios were well within (resampling) confidence intervals of the actual RCTs, but with tighter confidence intervals on the superior arm and less-tight confidence intervals on the inferior arm(s) and the odds ratios. The η-variant and block-based assignments provide intermediate levels of benefits and costs. With temporal changes, real-time assignment improves estimation of the end-of-trial superior-arm endpoint rates, but exaggerates differences relative to inferior arms. Unlike most RARs, real-time assignment automatically adjusts to reduce biases when temporal changes are large.

Conclusion: Real-time assignment would have improved patient outcomes within the trial (beneficence) and reduced the confidence interval for the superior arm. Benefits are balanced with larger confidence intervals on inferior arms and odds ratios. Variants and block-based assignments provide intermediate benefits and costs.

Real-time Adaptive Randomization of Clinical Trials

1. Adaptive trials, response-adaptive-randomization, and real-time adaptive randomization

The gold standard for large-scale Phase III trials is randomized controlled trials (RCT) in which patients are assigned randomly (and usually in equal proportions) to different treatments, or "arms." However, patient lives might be saved (and non-fatal endpoints prevented) if blinded information, gained from patients within the trial, is used to automatically assign more patients to the best treatment (endogenously-identified superior arm) and fewer patients to other treatments (inferior arms). Such beneficence and equipoise might come at the cost of higher-variance odds-ratios, a change in the ability to identify the best arm, or statistical confidence that the superior arm is indeed best.

Adaptive designs reallocate patients based on observations throughout the trial. For example, a Data Safety Monitor Board (DSMB) might periodically review results and reallocate the next batch of patients among arms [1, 2]. Alternatively, trialists might use prospectively-planned algorithms to reassign patients more often [3]. Such designs are often called response-adaptive randomization (RAR) methods [4, 5, 6, 7, 8].

RAR methods include Thompson Sampling (assign patients proportional to the probability that an arm is best), modifications of Thompson sampling, play the winner, sequential maximum likelihood, sequential posterior mean, and various other methods based on Bayes updating [4, 5, 6, 7, 8]. RARs vary in how they choose to use information in their adaptive-sampling strategies, leading to different trade-offs between patient beneficence and uncertainty reduction (e.g., statistical power for endpoint rates or for odds ratios). For example, an RAR might increase the likelihood of positive endpoints within a trial, but at the expense of

less power to distinguish among pairs of inferior arms. Because RARs often allocate more patients to the superior arm as the trial progresses, RARs may or may not be robust to temporal changes in (true) endpoint rates [12, 15, 16]. Furthermore, RAR analysis must use all available information and account for small-sample biases [2, 9, 11, 15, 16, 28, 30]. When RARs are Bayesian in nature, reported statistics must be justified as appropriate for the data-generating process [15, 25, 26].

Most RARs tend to be myopic, require randomization within periods (blocks of patients), and use fixed overall sample sizes [9]. Recently, researchers have proposed variations of forward-looking patient-by-patient optimal experimentation. Forward-looking patient assignment balances the benefit of learning the endpoint rates of the arms to make better patient assignments during the remainder of trial (and post-trial) against the immediate expected best-arm assignments in the current period [10, 11, 12, 13, 14].

Forward-looking methods show much promise in terms of positive endpoints, such as lives saved, relative to myopic RARs and fixed randomization [13], but forward-looking optimization faces theoretical challenges when endpoints are delayed. The proof of optimality for forward-looking experimentation assumes that a trialist observes an endpoint before the next patient is assigned. However, in our first empirical example, mortality is observed 30 days after arm assignment and in our second empirical example the last primary endpoints (a composite of cardiovascular mortality, myocardial infarction and cardiac arrest) are observed in the follow-up study with a mean of 4.2 years after treatment. To address delayed outcomes and to maintain randomization within periods, most previously-proposed forward-looking algorithms group patients into blocks and sample from all potential patient orders within the block. Simulated optimal assignments within block provide (unequal) randomization probabilities.

Assigning patients within blocks is a creative and effective strategy, but does not fully address delayed endpoints. Unless the trialist plans no-assignment periods between successive blocks, delayed patient endpoints for assignments late in the block period are not observed in time for the next-block patient assignments. One solution is to use large blocks so that the percentage of such patients is small, but large blocks decrease the advantages of optimal experimentation [13].

We examine an alternative forward-looking optimal-experimentation algorithm which assigns patients on a real-time (day-to-day) basis based on all data observed up to the day of patient assignment – real-time adaptive randomization (RTAR). Our analyses suggest that, had RTAR assignments been used in two large-scale cardiovascular trials, then lives would have been saved and non-fatal cardiovascular events prevented. We evaluate advantages and disadvantages of RTARs and highlight the ethical issues raised by RTARs. In §5.2, we examine the impact of temporal changes in endpoint rates, a known issue with RARs [12, 15, 16].

1.1. Multi-arm bandit algorithm when there are no delayed outcomes

RTARs use a preplanned statistical algorithm to assign patients to arms based on the endpoints observed up to the time of assignment. For ease of exposition, we first describe a real-time adaptive design in which exactly one patient arrives each day and endpoints are observed the day of assignment. We next extend the discussion to the more-realistic situation where more than one patient arrives each day and endpoints are delayed. Based on tests of stationarity (§5.1), we assume that patients are interchangeable – the true endpoint rates (e.g., mortality) per arm do not change throughout the trial.

Optimal experimentation is based on a multi-arm bandit (MAB) [17]. Conceptually, the trialist seeks to optimize endpoints over all current and future patients, including those after the

trial. To best assign arms to patients, the trialist must balance learning about the endpoint probabilities ("learning") and assigning the treatment most likely to be best ("earning"). See Figure 1.

We use Bayesian thinking. Let $p_a(d)$ be our current beliefs at the start of day d about the endpoint probability for arm a. If we assign arm a for the day-d patient, we get an expected reward of $p_a(d)$ for that patient, but we also learn more about arm a. On day $d+1$, we have a better estimate of the endpoint probability, $p_a(d+1)$, which enables the trialist to make a better decision on day $d+1$. Mathematically, the trialist has posterior beliefs about the distributions of $p_a(d)$ and $p_a(d+1)$ for all arms in the trial.

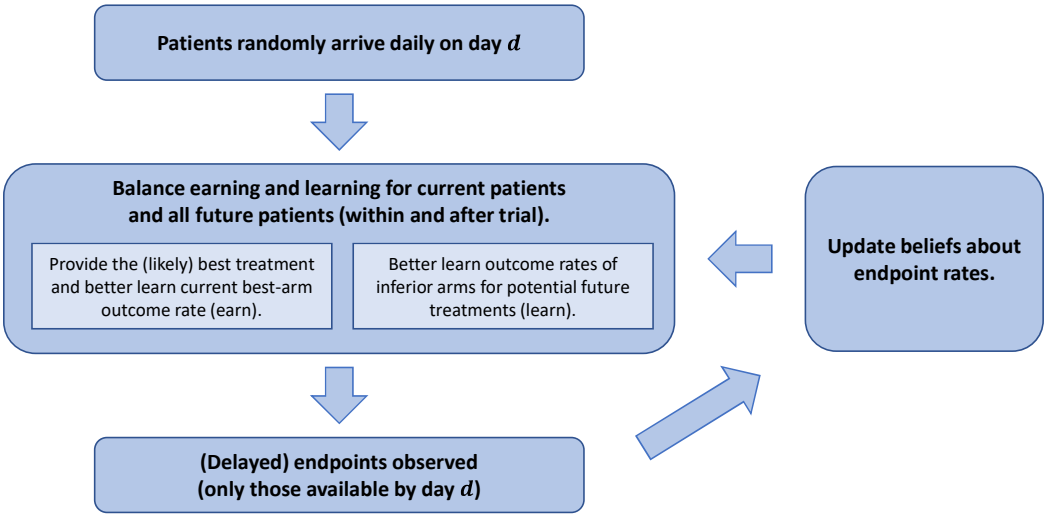

Fig. 1. Real-time adaptive randomization

We would like to learn the best treatment earlier than later. For example, if we know the best treatment today rather than a year from now, we can save lives and prevent non-fatal cardiovascular events while we are learning. To capture that concept, optimal experimentation introduces a "discount" parameter to value endpoints today slightly more than endpoints tomorrow [11, 13, 17, 18, 19]. The "discount" parameter is chosen conservatively; analyses for our data suggest that the RTAR for the two empirical trials is robust with respect to this parameter. We seek an algorithm that minimizes discounted negative outcomes, such as

mortality or cardiovascular events, for all future periods recognizing that we value saving lives earlier rather than later.

An RTAR represents system knowledge about the endpoint probabilities using a beta probability distribution per arm [11, 12, 21]. The beta probability distribution has two parameters, $\alpha_a(d)$ and $\beta_a(d)$, for each arm a. Larger parameter values mean less uncertainty in beliefs about the arm's efficacy. The expected value of the beta distribution, $\alpha_a(d)/(\alpha_a(d) + \beta_a(d))$, gives an estimate of the endpoint probabilities for each arm. Suppose the endpoint is mortality, then, when an endpoint is observed at the end of day d, $\alpha_a(d+1)$ is incremented by +1 for mortality and $\beta_a(d+1)$ is incremented by +1 for survival.

Gittins proved that the optimal dynamic-programming solution for the Bellman Equation, that represents the tradeoff between "learning" and "earning, is to compute a "Gittins index" that is a function of $\alpha_a(d)$ and $\beta_a(d)$ [11, 19]. Details in eAppendix A. The optimal allocation is to assign to the arriving patient the arm with the lowest Gittins index (ties broken randomly). (Lowest because the endpoint is mortality; highest if the endpoint were survival.) The first patient is assigned based on prior beliefs, which may be 1:1. The calculations necessary to compute Gittins indices are easily completed and tabled before (blinded) patients are assigned to an arm.

The RTAR algorithm differs from many RARs because the patient is assigned to an arm deterministically based on the endpoints observed up to that day. (Variants introduced in §1.3-1.4 allow randomization within day.) When patients arrive randomly and when endpoint probabilities do not change over the course of the trial, the random arrival of patients assures that the RTAR algorithm is a randomization procedure –an optimal randomization procedure. In RTARs, the data and safety monitoring boards still maintain their role as independent oversight and monitoring of the trial progress, data integrity, and participant safety [2, 20].

RTARs optimize discounted patient endpoints, while avoiding arm assignments that are "unnecessary" for the learning process [11, 21]. RTARs automatically assign a sufficient number of patients to all study arms, but not necessarily in a 1:1 ratio, until uncertainty in endpoint rates is reduced enough that assigning patients to the inferior arms no longer provides value.

Stabilization to the superior arm occurs when the trial be sufficiently long. In our primary application, arm assignments almost always stabilize across replicates. In our experience, a trial size that is sufficiently powered as if run as an RCT is sufficiently long for real-time assignments to stabilize. In a new trial, trial size can be planned with simulation [2].

1.2. Multi-arm bandit algorithm with delays and multiple patients per day

As an approximation to optimality, if more than one patient arrives on a given day, we assign all patients to the arm with the lowest index. When there are delays, we use only endpoints that have been observed by day d. The Gittins algorithm is no longer provably optimal, but we expect the algorithm to be close to optimal if (1) the number of patients that arrive on each day is small compared to the total patients in the trial and (2) the delay is small compared to the length of the trial. The first condition is met in both trials that we analyze, but it remains an empirical question whether an RTAR saves lives and minimize other adverse end outcomes. The second condition is met in the first trial we analyze, but not necessarily the second trial, thus enabling us to examine the impact of substantial endpoint delays.

1.3. Block-based MAB based on Gittins' solution

RTARs assign patients deterministically, relying on random arrival and stationarity. An alternative MAB-based algorithm assigns patients to blocks and randomizes within blocks [12, 13]. If the block is sufficiently large relative to the endpoint delay, then most, but not all, endpoints can be observed before assignments are made in the next block.

The block-based adaptive design enrolls patients in J blocks of size b, assigning patients in block j using the information gathered up to and including the $j-1^{st}$ block [13]. The design assumes endpoints are observed immediately at the end of a block and used for assignments in the next block. Because, theoretically, learning could happen within a block, the block-based algorithm looks forward through the block by simulating the expect endpoints and Gittins-Index updates within a block. The simulation to identify assignment probabilities assumes (1) the first patient is assigned based on the Gittins index calculated based on previous-block endpoints and (2) second and subsequent patients within a block are assigned, endpoints observed, and Gittins indices updated based on simulated endpoints within the block. Assuming stationarity and random arrival, the algorithm calculates the expected percentages of arm assignments over all possible patient orders. To make the algorithm feasible, the order of patient arrivals is sampled rather than exhaustively enumerated. This algorithm is known as the forward-looking Gittins index algorithm [FLGI, 10, 12, 13]. We refer to it as the block-based MAB.

In simulations grounded to a breast-cancer-treatment RCT, the block-based MAB algorithm provided "substantial improvements in terms of patient benefit" relative to other trial strategies including RCTs and other RARs. The block-based MAB improved the expected number of positive endpoints by almost 50%, but with a reduction in statistical power of approximately 70%. Results depended upon the block size, with more positive endpoints and lower power observed for smaller blocks. Other RARs produced intermediate patient successes and power relative to the block-based MAB and an RCT.

1.4. RTAR η-variant to ensure a target minimum power

Theoretically, an RTAR identifies the superior arm quickly and assigns substantial sample to the superior arm. Less sample is assigned to inferior arms resulting in less statistical power for

the inferior-arm endpoint rates. This is an ethical dilemma. Trialists may wish to assure a minimum sample size (minimum statistical power) on the inferior arms or on odds ratios [12, 22].

To explore this issue while retaining many advantages of RTARs, we examine an η–variant of an MAB algorithm that seeks a minimum level of statistical power to inferior arms. In the η-variation, with probability η we randomize patients in equal proportions to all arms that have not yet reached a target minimum number of patients and, with probability $1 - \eta k_d$, we assign patients with Gittins indices. k_d is the number of arms that have not yet reached the minimum number of patients at the start of day d. The η-variant is an alternative means to achieve burn-in [5, 8, 23].

RTARs, the block-based MAB, and the η-variants are all MAB-based algorithms. An MAB-based algorithm is a type of RAR and an RAR is a type of adaptive design.

2. Statistical concepts, adaptivity bias, and expected performance

2.1. Statistical concepts and potential adaptivity bias

Learning in the Gittins framework is inherently Bayesian. Bayesian interpretations are based on the posterior likelihood. For RARs, the posterior likelihood can be factored into a term based on the observed endpoint conditioned on the assignment and a term based on the probability of assignment given the data from prior endpoints. Because the latter is a known function of the data, the second term does not depend upon the unknown endpoint probabilities and can be removed from the likelihood [25]. Thus, the Bayesian posterior likelihood does not depend explicitly on how the RAR assigns patients. All information about unequal sample sizes among arms is included in the likelihood function [15, 25, 26].

Because the likelihood does not depend upon how the RAR assigns patients, neither do typically-used maximum-likelihood estimators (MLE). MLEs are consistent estimators

(asymptotically unbiased for large numbers of patients), efficient estimators (no consistent estimator has a lower asymptotic mean-square error), and are asymptotically normal [9, 25, 27, 28]. From a Bayesian perspective, MLEs are asymptotically equivalent to maximum posteriori estimation with weakly informative priors [29]. The factored likelihood implies that MLEs can be reported and analyzed after the trial is completed, especially for large samples [9, 25, 27, 28, 30].

MLEs are consistent, but they may be biased for small samples [2, 11, 16, 28, 30]. When adaptive designs are based on a small number of intermediate analyses, trialists use standard corrections for estimation biases and especially for Type 1 error inflation [2, 28, 32, 33, 34]. Small-sample biases occur in many RARs and require advanced statistics or propensity scores [15, 25, 27, 35]. Such biases are minimal for the large samples in this paper [25, 30, 35].

Researchers estimate the distributions of statistics to evaluate RARs, such as the percent of times the superior arm is identified as superior, by sampling from a known model [2, 7, 8, 11, 12, 13, 16, 24, 28] or resampling with replacement when patient-by-patient endpoints are observed [30, 21]. When we have patient-by-patient observations, resampling generates the distribution of observations from which we compute means and confidence intervals for statistics such as odds ratios and endpoint rates. Resampling also provides the percent assignments to arms, the probability an arm is identified as superior, Type 2 error (the probability of declaring a trial inconclusive when it is not), and the percent of adverse imbalance in arm assignments [5, 23, 24, 30, 31]. Resampling statistics are consistent with the (Bayesian) likelihood principle. When the number of patients is sufficiently large, they are consistent with commonly-reported post-trial statistics.

2.2. Anticipated performance of an RTAR relative to an RCT

As an RTAR learns endpoint rates for each arm, we expect it to allocate more patients to the (endogenously-identified) superior arm. With more patients allocated to the superior arm and fewer patients allocated to inferior arms, we expect the negative endpoints to be fewer for RTARs relative to an RCT. When the endpoint is mortality, RTARs will lead to greater patient beneficence.

When more patients are allocated to the superior arm, we expect that the (resampling) confidence intervals, relative to RCT confidence intervals, will be tighter for the superior arm at the expense of less-tight confidence intervals for the inferior arms. We also expect there will be more power to estimate superior-arm endpoint rates and less power for inferior arms.

For two arms, pairwise power will be maximal and odds-ratio confidence intervals are tighter for equal allocation. With three (or more) arms, predictions are less clear. With three arms, we expect that the RTAR will allocate fewer than $N/3$ patients to the worst inferior arm, resulting in more than $2N/3$ patients split between the superior arm and the second-best arm. Depending on the specific allocation, the superior-arm-to-inferior pairwise power may either increase or decrease relative to the corresponding RCT. Similarly, the confidence intervals for the odds ratios may be tighter or less-tight depending upon the specific allocation of patients to arms. We resolve this ambiguity empirically for the two trials analyzed in this paper.

By design, the η-variant approaches an RTAR as $\eta \to 0$ and approaches an RCT as $\eta \to 1$, thus we expect the performance of the η-variant to be between that of an RTAR and an RCT. By choosing η between 0 and 1, the trialist can finetune emphasis on patient beneficence, estimating the endpoint rates for the superior arm, estimating the endpoint rates for the inferior arms, power for endpoint rates, and power for odds ratios.

We examine the performance in terms of endpoint outcomes and statistical power of an RTAR and variants on data from two large-scale trials. We also examine the percent of times the superior arm (and inferior arms) are identified as superior, and how quickly the real-time algorithm converges to the superior arm. In §5.1, we test stationarity of endpoint rates. In §5.2, we examine the impact of temporal changes in endpoint rates.

3. What if the GUSTO-1 and EUROPA trials had been adapted in real time?

3.1. The GUSTO-1 and EUROPA RCT trials

To study the potential performance of RTARs, we use resampling simulations grounded by the data from the GUSTO-1 and the EUROPA trials [36, 37]. The design and principal results of both trials have been published and are summarized in Table 1. Briefly, GUSTO-1 randomized a total of 31,180 patients presenting with acute myocardial infarction to one of three thrombolytic strategies. (30,732 patients after excluding observations with missing data. A fourth strategy was added later into the trial.) The primary endpoint was 30-day all-cause mortality and was lowest in the patients randomized to accelerated tissue plasminogen activator (t-PA) with intravenous heparin, 6.3%. The GUSTO-1 investigators concluded that this combination "is the best thrombolytic strategy to date (i.e., 1993) for patients with acute myocardial infarction."

The EUROPA investigators randomly assigned 12,218 patients with stable coronary heart disease to either a treatment with the angiotensin-converting-enzyme (ACE) inhibitor perindopril or to a matching placebo. The primary endpoint was a composite of cardiovascular death, non-fatal myocardial infarction and cardiac arrest with successful resuscitation, and was lowest (8%) in those patients randomized to perindopril. In 2003, the investigators concluded that, "on top of other preventive medications, [perindopril] should be considered in all patients with coronary heart disease." The mean time-to-the-last observation of outcomes across all patients in

EUROPA was 4.2 years since the start of the trial.

The GUSTO-1 and EUROPA trials were conducted according to the prevailing ethical regulations at the time, which included approval of the protocol by the institutional review board at the participating hospitals, and informed consent by the study participants. Our analyses are based on the individual (anonymized) patient data from the trials, which we obtained by courtesy of Duke University School of Medicine and Servier.

Table 1. GUSTO-1 and EUROPA Trials (conducted as RCTs)

Trial details	**GUSTO-1** [ii]	**EUROPA** [iii]
Goal	Compare streptokinase and tissue plasminogen activator thrombolytic strategies in the treatment of acute myocardial infarction	Assess the effect of perindopril versus placebo on the combined endpoint of cardiovascular death, non-fatal MI, and resuscitated cardiac arrest in patients with stable coronary heart disease
1st Enrollment	December 27, 1990	27 October, 1997
Termination	February 22, 1993	20 March, 2003
Arms at the start of the trial	Arm 1: t-PA, IV Heparin Arm 2: SK, IV Heparin Arm 3: t-PA+ SK, IV Heparin	Arm 1: Perindopril Arm 2: Placebo
Patients per randomly allocated treatment [a]	t-PA, IV Heparin: 10,396 SK, IV Heparin: 10,410 t-PA+ SK, IV Heparin: 10,374	Perindopril: 6,110 Placebo: 6,108
Primary endpoint	Death from any cause at 30 days of follow-up	Composite of cardiovascular mortality, non-fatal MI, and resuscitated cardiac arrest during (mean) 4.2 year follow-up
Incidence of the primary efficacy endpoints [a]	t-PA, IV Heparin: 653 (6.3 %) SK, IV Heparin: 763 (7.3 %) t-PA+ SK, IV Heparin: 723 (7.0 %)	Perindopril: 488 (8.0%) Placebo: 603 (9.9%)
Eligibility	Patients presenting to a participating hospital less < 6 hours after symptoms, with chest pain lasting at least 20 minutes and accompanied by electrocardiographic signs of ≥0.1mV of ST-segment elevation in two or more limb leads or ≥ 0.2 mV in two or more contiguous precordial leads	Men and women ≥ 18 years with evidence of coronary heart disease per MI, percutaneous or surgical coronary revascularization, angiographic evidence ≥ 70% narrowing of at least one major coronary artery, or a history of typical chest pain in male patients with an abnormal stress test
Exclusion	Previous stroke, active bleeding, previous treatment with streptokinase or anistreplase, recent trauma or major surgery, previous participation in the trial, or non-compressible vascular punctures	Clinically evident heart failure, planned revascularization procedure, hypotension, uncontrolled hypertension, use of ACE-inhibitors or angiotensin-2 receptor blockers in the last month, renal insufficiency, and serum potassium

[a] Before removing observations with missing data. GUSTO-1 sample sizes after removing missing data are: 10,255 (arm 1),10,268 (arm 2); 10,209 (arm 3)

3.2. Data and grounded simulations

The detailed distribution of the RCT randomizations and endpoints per day in the GUSTO-1 and EUROPA trials are presented in the online eAppendix B (eFigures 1 and 2). The figures cover the entire duration of the trial, from the first randomization until the last primary endpoint was observed.

Using the empirical trial data, we resampled patients to simulate what would have happened had the trial been based on an RTAR. Priors were weakly informative and equal for all arms, thus starting with an equally-likely ratio (1:1:1 for GUSTO-1; 1:1 for EUROPA). For each day of the trial, the RTAR automatically assigns patients arriving on day d to the study arms, based on observed endpoints up to that the beginning of day d. (Only endpoints observed before day-d assignments are used.) Patients for each arm are drawn randomly (with replacement, given stationarity and exchangeability of patients) from the pool of RCT patients in the chosen arm. To avoid a particularly favorable draw and to compute confidence intervals for all statistics, we repeat the process with 200 replicates for each study. In GUSTO-1, these pools have 10,255 patients in arm 1, 10,268 patients in arm 2, and 10,209 patients in arm 3. In EUROPA, these pools have 6,100 patients in Perindopril and 6,108 in the placebo. The empirically-grounded simulations continue until the final day of the original RCT trials. Mean endpoint rates, confidence intervals, power, pairwise odds ratios, and other statistics of interest are based on the distributions over replicates.

3.3. Odds ratios

As more patients are assigned to arms, the estimated odds ratios evolve and the odds-ratio confidence intervals become tighter. Figures 2a to 2c plot the evolution of the mean and the confidence intervals for the odds ratios of all pairs of arms (averaged over replicates). We observe a

tighter confidence interval for the arm-1-to-arm-3 odds ratio (the two arms with lowest mortality rates) with a less-tight confidence interval for the arm-1-to-arm-2 odds ratio (superior to third best) and for the arm-2-to-arm-3 odds ratio (second best to third best).

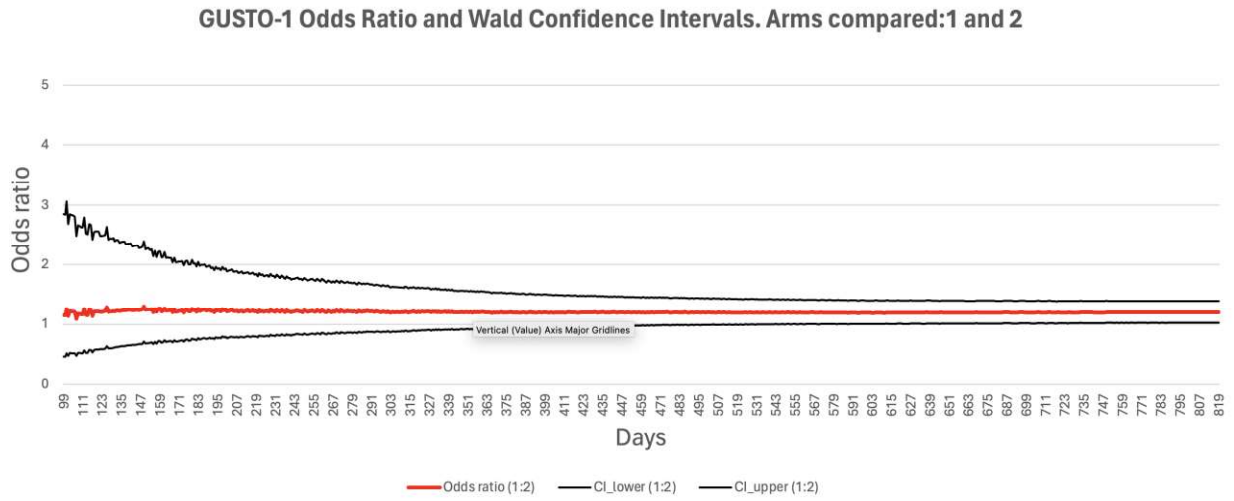

Fig. 2a. Changes in real-time adaptive trial odds ratios during the trial for arms 1 and 2, averaged over 200 replicates

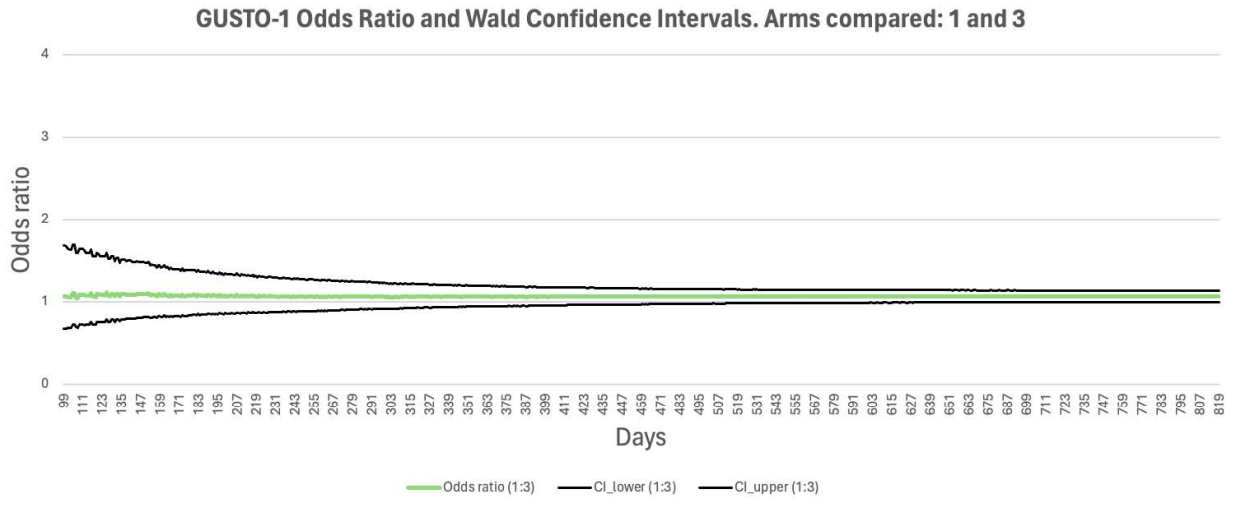

Fig. 2b. Changes in real-time adaptive trial odds ratios during the trial for arms 1 and 3, averaged over 200 replicates

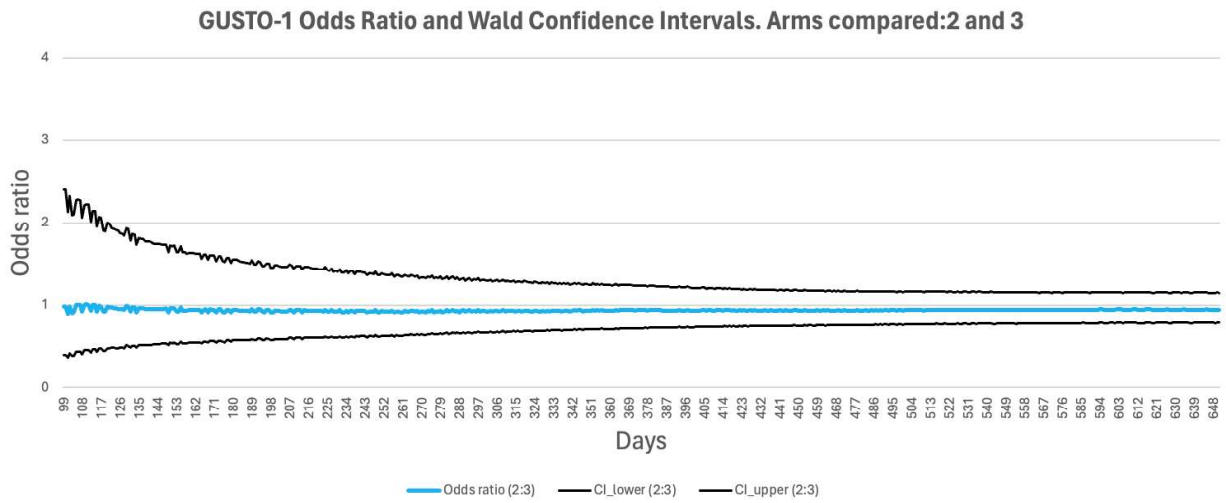

Fig. 2c. Changes in real-time adaptive trial odds ratios during the trial for arms 2 and 3, averaged over 200 replicates

For the RCT and the RTAR, we compute 2.5%, 50%, and 97.5% confidence intervals over 200 replicates, each based on patient resampling. For GUSTO-1, the resampling median RCT arm 1:2 odds ratio is 1.182 (1.075, 1.297), the arms 1:3 odds ratio is 1.116 (0.997, 1.240), and the arms 2:3 odds ratio is 0.944 (0.852, 1.052). (In the RCT, the observed means were 1.184, 1.118, and 0.945, respectively.) The median RTAR odds ratio estimates are 1.204 (1.073, 1.632) for arms 1:2, 1.150 (1.032, 1.414) for arms 1:3, and 0.946 (0.721, 1.199) for arms 2:3. The RCT medians are within the RTAR confidence intervals and the RTAR medians are within the RCT confidence intervals for all pairs of arms. Nonetheless, the confidence intervals for the RTAR odds ratios are larger than those for the RCT.

We get similar results for EUROPA. For example, the median odds ratio for Perindopril versus a placebo is 1.246 (1.118, 1.412) and the median RTAR odds ratio is 1.248 (1.080, 1.455), $\eta = 0.25$.

3.4. Number of Patients Assigned, Mortalities, and Mortality Rates

The first four columns of Table 2 present the results of GUSTO-1 trial. The last three columns present the results of the EUROPA trial. For each trial, we present the number of assigned patients, the number of primary endpoint events, and the endpoint rates for the original RCT (in the first three rows) along with confidence intervals. In the last nine rows of Table 2, we present the results had these trials used an RTAR, an η-variant, or a block-based MAB to assign patients in real time or by block.

For both trials, the ranking of all arms in the simulations by the RTAR, the η-variant, and the block-based MAB match the RCT ranking (t-PA with IV Heparin is the best, SK with IV Heparin is the worst in GUSTO; Perindopril is the best, placebo is the worst in EUROPA). The primary endpoint rates estimated with all three adaptive algorithms are quite close to those estimated with the RCT and well within the confidence intervals. Relative to the RCT, all MAB variants provided tighter confidence intervals on the mortality rate for the (identified-within-the-trial) superior arm, with the tightest confidence interval provided for by the RTAR. As expected, the tighter bound for the superior arm comes with a tradeoff: confidence intervals are not as tight for the (endogenously-identified) inferior arms.

The lowest mortality (greatest beneficence), in GUSTO-1, was observed for the RTAR (1,952 lives lost) and the highest mortality for the RCT (2,074 lives lost) – a net saving of 122 lives due to real-time adaptation. The net savings for the η-variant and the block-based MAB were 72 and 102 lives saved, respectively.

Resampling suggests that the RCT would have identified the best arm in 98% of the replicates, comparable to the 99% achieved by the RTAR. There were no cases, for either the RTAR or the RCT, where an inferior arm (arm 2 or arm 3) was identified as statistically significantly better than the superior arm (arm 1). There were only 2% cases of arm imbalance where one of

the inferior arms (arm 2 or arm 3) was assigned more patients than the superior arm (arm 1). We obtain similar results for EUROPA despite the substantial delays in observing outcomes in the follow-up.

The gain in the reduction of negative endpoint outcomes comes at the cost of making fewer assignments to the inferior arms. For the RTAR, the assignments to the inferior arms averaged 2,755 as opposed to 10,268 in GUSTO-1 arm 2, 4,311 as opposed to 10,209 in GUSTO-1 arm 3; and 1,859 as opposed to 6,108 in the placebo arm in EUROPA). The η-variant and the block-based MAB allocated fewer patients to the superior arm and more patients to the inferior arms than the RTAR. While the η-variant specifies a minimum patient target for each arm, the observed minimums for inferior arms vary slightly because the real-time portion of the η-variant favors the superior arm.

Table 2. Benchmark simulation results for GUSTO-1 and EUROPA

	GUSTO-1				EUROPA		
	Arm 1[a]: t-PA, IV Heparin	Arm 2: SK, IV Heparin	Arm 3: t-PA+ SK, IV Heparin	Total Mortality	Arm 1 [a]: Perindopril	Arm 2: Placebo	Totals
RCT					RCT		
Patients	10,255	10,268	10,209	30,732	6,110	6,108	12,218
Events	631	742	701	2,074	489	603	1,092
Event rate	0.062 (0.055, 0.068)	0.072 (0.068, 0.077)	0.069 (0.068, 0.069)		0.080 (0.074, 0.087)	0.099 (0.091, 0.106)	
Real-time adaptive randomization (RTAR)[b]					Real-time adaptive randomization (RTAR)[b, d]		
Patients	23,666	2,755	4,311	30,732	10,359	1,859	12,218
Events	1,455	200	297	1,952	828	183	1,012
Event rate	0.061 (0.059, 0.065)	0.073 (0.066, 0.096)	0.069 (0.064, 0.085)		0.080 (0.075 0.153)	0.098 (0.088, 0.227)	
Real-time adaptive randomization variant, $\eta = 0.25$ [b, c]					Real-time adaptive randomization variant, $\eta = 0.25$ [b,d]		
Patients	18,479	5,683	6,570	30,732	9,768	2,450	12,218
Events	1,136	413	453	2,002	782	241	1,023
Event rate	0.062 (0.059, 0.066)	0.073 (0.067,0.079)	0.069 (0.067, 0.076)		0.080 (0.075 0.086)	0.099 (0.088, 0.110)	
Block-based MAB [b, d]							
Patients	21,593	3,579	5,560	30,732			
Events	1,336	255	379	1,971			
Event rate	0.062 (0.059, 0.068)	0.071 (0.066, 0.081)	0.068 (0.066, 0.083)				

[a] Best arm in the trial

[b] Averaged over 200 replicates. Priors in GUSTO: $\alpha_o = 6$, $\beta_o = 390$ for all arms. Priors in EUROPA: $\alpha_o = 40$, $\beta_o = 1{,}800$ for both arms

[c] Minimum equally-likely allocation = 6,000.

[d] Block size = 60. Monte Carlo draws = 100.

3.5. The trade-off between (odds-ratio) statistical power and patient beneficence

More patients assigned to the superior arms implies greater power for the superior-arm endpoint rates. Fewer patients assigned to inferior arms implies less power for inferior-arms' endpoint rates. However, the deviation from equal allocation of patients to arms implies lower power for the odds ratios [9, 10]. To examine this tradeoff further, we plot the change in odds-ratio statistical power (solid lines, left vertical axis) and the number of patient exposed to the superior and inferior arms (dotted lines, right vertical axis) had different values of the tuning parameter, η, been used in the GUSTO-1 trial ($\eta = 0$ implies an RTAR) .

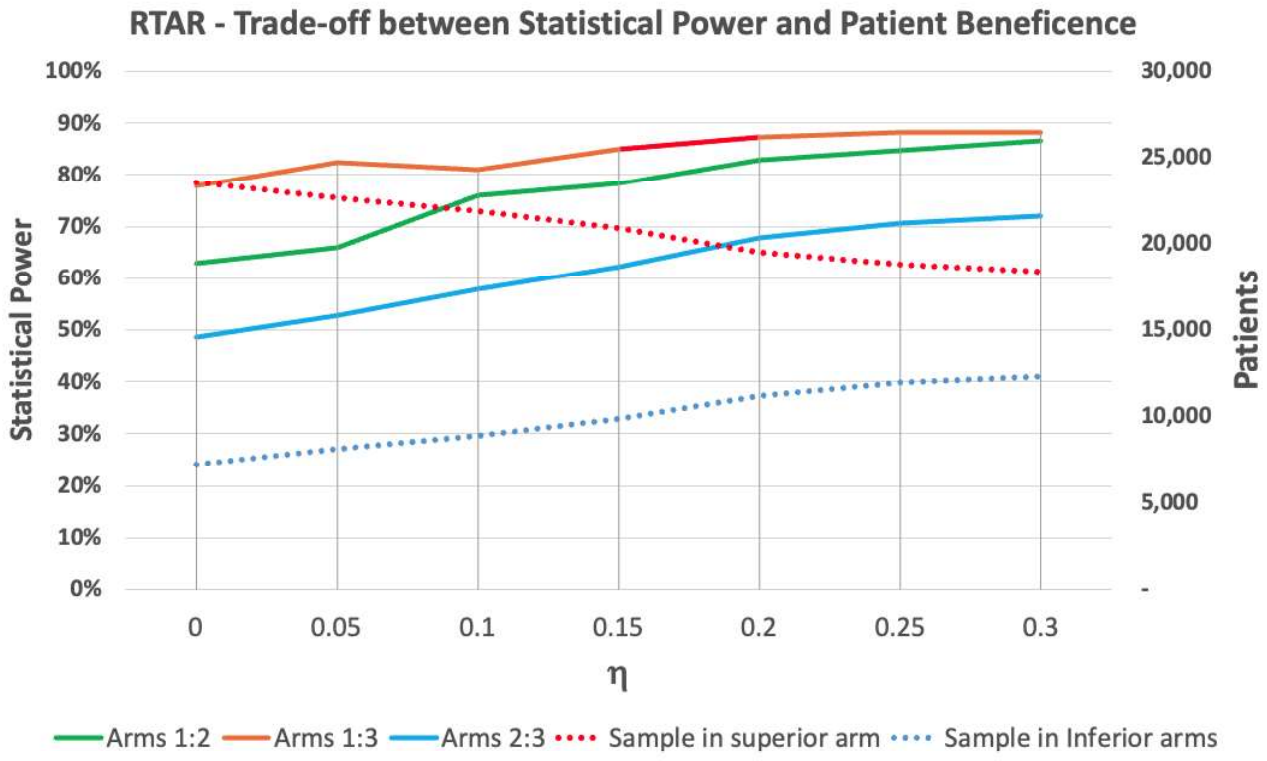

Fig. 3. Trade-off between statistical power and patient beneficence in GUSTO-1

More aggressive η's (right side of Figure 3) provide higher statistical power for the odds ratios, but also lower patient beneficence (spread from the orange to the blue dotted lines). At $\eta = 0.30$, the power of the η-variant is almost indistinguishable from the power of the RCT

(90%). The trialist can choose η to make ethical judgments between beneficence and odds-ratio power. (The code is provided as open-source.)

3.6. Greater patient beneficence with larger differences in endpoint probabilities

Every life is important, but one might ask whether, for example, 122 fewer deaths out of 2,074 mortalities in GUSTO-1 justifies the use of a new method. This is an ethical issue beyond the scope of this article. However, if we examine the GUSTO-1 trial, we see that the three arms are close in mortality risk, 0.062, 0.069, and 0.072. As an hypothetical, we examine more substantial differences—mortality rates of 0.063, 0.126, and 0.189 for the three arms, unknown before the trial. In this hypothetical world with 30,732 patients, a real-time adaptive assignment would have saved 1,700 lives compared to an RCT assignment. We kept the total patients the same for a clear comparison. If the trialist had strong priors on the mortality risk and required the same statistical power, the trialist would allocate fewer patients to both the RTAR or the RCT. Even in this case, the RTAR would lead to substantial greater patient beneficence.

3.7. Summary of the GUSTO-1 and EUROPA empirically-grounded simulations

The GUSTO-1 and EUROPA empirically-grounded resampling simulations illustrate the ethical decisions when choosing between an RTAR and an RCT. Patient risk is reduced, and patient beneficence increased, when more patients are automatically allocated to the (endogenously identified) superior arm. RTARs result in tighter confidence intervals and more power for the endpoint rates of the superior arm. In GUSTO-1, the allocations implied slightly more pairwise power when comparing the superior arm to one inferior arm and slightly less power relative to the other inferior arm. There was less statistical power to distinguish between the two inferior arms. However, power for the odds ratios was reduced.

The η-variant and the block-based MAB algorithm provide the trialist with the ability to

balance the benefits and costs of an RTARs versus an RCT. The η-variant and the block-based MAB allocate more patients to the superior arm than an RCT, but fewer patients than RTARs. Confidence intervals for the superior arm are tighter than an RCT, but less tight than an RTAR. Power to compare superior-to-inferior arms is between that of an RCT and RTARs and odds-ratio power is comparable to an RCT for $\eta = 0.30$ or higher. Misidentification of the inferior arm as superior is not a problem. Neither the RTAR, the η-variant, nor the block-based MAB identify an inferior arm as statistically superior.

4. Stationarity and temporal changes in endpoint rates

4.1. Tests of stationarity

To test stationarity, we split the RCT trial by quantiles on the date of assignment and examine whether endpoint rates vary significantly by quantile. For deciles, the null hypothesis of stationarity was not rejected for all GUSTO-1 arms (arm 1 $p = 0.47$, arm 2 $p = 0.45$, arm 3 $p = 0.80$) and for both EUROPA arms (Perindopril $p = 0.37$, placebo $p = 0.39$). Particularly relevant for the analyses in the next section, two-way quantile splits were also not significantly different. Other stationarity tests (available from the authors) do not reject stationarity.

4.2. Temporal changes in endpoint rates

Temporal changes in endpoint rates are a known issue with RARs and are potentially an issue with RTARs [12, 15, 16, 28]. Suppose that there is a shock to the system, perhaps due to a mutation in a virus, a change in the demographics of patients, environmental changes, or the advent of auxiliary treatments. Such a temporal change might imply that the mortality rate is higher for later patients than for earlier patients.

RARs tend to allocate relatively more patients to superior arms and fewer patients to inferior arms as the trial progresses. To visualize the effect, assume the temporal change happens

midway through the number of patients in the trial. For an RCT, the estimated endpoint rates will be the average of the endpoint rates in the two periods. For RARs, because the superior-arm sample grows relative to the inferior-arm sample, the endpoint rate for the superior arm will be closer to the endpoint rate at the end of the trial and the endpoint rate for the inferior arm will be closer to the endpoint rate at the beginning of the trial. The net result will be that, relative to an RCT, the difference in endpoint rates between the superior and inferior arms will be overestimated. The logic generalizes, for example, we would observe similar effects when endpoint rates drift throughout the trial [12, 16]. Prior research suggests that RARs are robust to drift as long as the drift is less than 25%, that the block-based MAB is less sensitive to drift than Thompson sampling, and that it is important to distinguish RAR biases from biases induced by early stopping [12, 28].

RTARs are based on Gittins indices which react to observed endpoint rates [36]. If there are sufficient post-shock observations, the Gittins indices will evolve causing the MAB to reexplore the inferior arms. The MAB will automatically begin to learn the new endpoint rates.

To examine temporal changes, we simulate midpoint shocks of 5%, 10%, 15%, 20%, and 25%. See eAppendix C. As expected, the RTAR estimates are closer to the end-of-trial mortalities than the RCT estimates. Consistent with the literature, shock leads to an upward bias in the differences between arms relative to an RCT. For arms 1:3, the biases are approximately 19% for a 5% shock, but smaller (1-3%) for higher shocks because the MAB explores more with higher shocks. For arms 1:2 the bias is approximately 11% for low shocks but decreases to 6-9% for higher shocks. For comparisons among the inferior arms 2:3, the bias is usually higher due to fewer patients being allocated to inferior arms. RTAR improves end-of-trial estimates, but at the cost of a modest bias between arms. The effect of shock and drift on MABs is still under

development. Researchers are exploring MABs that allocate patients while anticipating temporal changes such as shock and drift [40, 41]. These developments are promising.

5. Discussion

Our goal is to evaluate empirically an RTAR relative to an RCT when there are delays in observing endpoints. The RTAR would likely have saved lives (in GUSTO-1) and avoided cardiovascular events (in EUROPA) relative to an RCT while providing estimates of endpoint rates and odds ratios within statistical confidence of the RCT – even though endpoints are delayed by 30 days or more. The tradeoff is that confidence intervals for the odds ratios and inferior-arm event rates increase. Confidence intervals are tighter for superior-arm event rates. Power is lower for odds ratios but higher for superior-arm event rates. The η-varient enables the trialist to finetune the power/beneficence tradeoff. For example, for $\eta = 0.30$ we obtain substantial beneficence with little loss of power. A trialist can balance ethical considerations for a planned trial by resampling using priors on endpoint rates for that trial.

Ethically, RTARs enhance the principle of beneficence in the sense of the Belmont report—"maximize possible benefits and minimize possible harm [2].” RTARs also respect persons and justice because *a priori* arm assignments depend upon endpoints not knowable in advance. The algorithm does not depend upon demographic indicators. On the other hand, the likelihood of receiving the best treatment changes over time (violating the ethical principle of equality). In RTARs, patients who enter the trial late or after the trial has ended are more likely to receive the best treatment than patients who enter the trial early. (This is also true to a lesser extent when trial assignment ratios are adapted due to a small number of interim reviews and is always true when comparing patients in a trial to those who receive treatment after a trial.)

5.1 Study limitations

Our simulations are empirically-grounded and their implications are as predicted by theory, but all of our simulations are *post hoc* analyses of the GUSTO-1 and EUROPA trials. There is nothing in our analyses that used knowledge that was not available at the time of RCT patient assignment. Nonetheless, any *post hoc* analyses must be treated with caution.

Our data are time-stamped at the daily level, hence our real-time assignments and end-point-rate learning are conservative relative to an MAB that adapts assignments within days (if feasible from a blinding standpoint). For multiple-trial multiple-population settings, researchers can merge RTARs and platform-trials. Adaptive platform trials provide a means to compare multiple interventions, generate subgroup estimates, and minimize downtime between trials [36]. Finally, we might improve assignments further with the use of biomarkers as surrogate measures of endpoints [10, 38, 39].

5.2 Conclusion

We used conceptual arguments and empirically-grounded simulations to examine the trade-offs trialists face when using RTARs, the η-variant, and the block-based MAB. RTARs increase patient beneficence (e.g., fewer cardiovascular events) and enable better estimates of endpoint rates for superior arms.

References

[1] Ryan, EG, Bruce J, Metcalfe AJ, Stallard N, Lamb SE, Viele K, Young D, Gates S. Using Bayesian adaptive designs to improve phase III trials: a respiratory care example. BMC Medical Research Methodology 2019;19(99):1-10.

[2] FDA. Adaptive designs for clinical trials of drugs and biologics: Guidance for industry. Biostatistics. Food and Drug Administration, Center for Drug Evaluation and Research, November; 2019; FDA-2018-D-3124.

[3] Tricco, A, Tovey, D. Flexible approaches to clinical trials. Journal of Clinical Epidemiology, 2023; 154: A1-A2.

[4] Berry DA, Eick SG. Adaptive assignment versus balanced randomization in clinical trials: a decision analysis. Statistics in Medicine 1995;14:231-246.

[5] Du Y, Wang X, Lee JJ. Simulation study for evaluating the performance of response-adaptive randomization. Contemporary Clinical Trials 2015;40:15-25.

[6] Thall PF, Wathen JK. Practical Bayesian adaptive randomization in clinical trials. European Journal of Cancer 2007;43:589-866.

[7] Trippa L, Le EQ, Wen PW, Batchelor TT, Cloughesy T, Parmigaiani G, Alexander BM. Bayesian adaptive randomized trial design for patients with recurrent glioblastoma. Journal of Clinical Oncology 2012;30(2):3258-3263.

[8] Wathen JK, Thall PF (2017) A simulation study of outcome adaptive randomization in multi-arm clinical trials. Clinical Trials 14(5):432-440.

[9] Hu F, Rosenberger WF. The theory of response-adaptive randomization in clinical trials. Wiley Series in Probability and Statistics. Wiley Interscience. Hoboken NJ. 2006.

[10] Barnett HY, Villar SS, Geys H, Jaki T. A novel statistical test for treatment differences in clinical trials using a response-adaptive forward-looking Gittins Index Rule. 2020; Biometrics 79:86-97.

[11] Villar SS, Bowden J, Wason J. Multi-armed bandit models for the optimal design of clinical trials: Benefits and challenges. Stat. Sci. 2015;30(2):199-215.

[12] Villar SS, Bowden J and Wason J. Response-adaptive designs for binary responses: how to offer patient benefit while being robust to time trends? 2018;Pharm Stat; 17: 182–19

[13] Villar SS, Wason J, Bowden J. Response-adaptive randomization for multi-arm clinical trials using the forward-looking Gittins Index rule. 2015;Biometrics 71:969-978.

[14] Williamson SF, Villar SS. A response-adaptive randomization procedure for multi-armed clinical trials with normally distributed outcomes. 2020;Biometrics 76:197–209.

[15] Proschan M, Evans S. Resist the temptation of response-adaptive randomization. Clinical Infectious Diseases 2020;71:3002-3004.

[16] 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;26:1621-1628.

[17] Berry D and Fristedt B. Bandit Problems – Sequential Allocation of Experiments; London: Chapman and Hall. 1985.

[18] Chick SE, Gans N, Yapar Ö. Bayesian sequential learning for clinical trials of multiple correlated medical interventions. Management Science 2022;68(7):4919-4938.

[19] Gittins JC 1979. Bandit processes and dynamic allocation indices. Journal of the Royal Statistical. Society.1979; (Ser B)41(2):148–177, plus commentary

[20] Clemons F, Elbourne D, Darbyshire J, Pocock S, Damocles Group. Data monitoring in randomized controlled trials: Surveys of recent practice and policies. Clinical Trials 2005;2:22-33.

[21] Gittins J, Glazebrook K and Weber R. Multi-armed bandit allocation indices. London: Wiley. 2011.

[22] Flehinger BJ, Louis TA, Robbins H, Singer BH. Reducing the number of inferior treatments in clinical trials, PNAS 1972;69(10):2993-2994.

[23] Granholm A, Kaas-Hansen BS, Lange T, Schjørring OL, Andersen LW, Perner A, Jensen AKG, Hylander Møller MH, An overview of methodological considerations regarding adaptive stopping, arm dropping, and randomization in clinical trials. 2023;Journal of Clinical Epidemiology 153:45-54.

[24] Connor JT, Elm JJ, Broglio KR. Bayesian adaptive trials offer advantages in comparative effectiveness trials: an example in status epilepticus. Journal of Clinical Epidemiology 2013;66:S130-S137.

[25] Bowden J, Trippa L. Unbiased estimation for response adaptive clinical trials. Statistical Methods in Medical Research 2017;26(5):2376-2388.

[26] Carlin BP, Louis TA Bayes and empirical Bayes methods for data analysis. Chapman & Hall/CRC, New York NY. 2000.

[27] Melfi VF, Page C. Estimation after adaptive allocation. Journal of Statistical Planning and Inference 2000;87(2):353-363.

[28] Robertson DS, Lee KM, Lopez-Kolkovska BC, Villar SS. Response-adaptive randomization in clinical Trials: From myths to practical considerations. 2023; Statistical Science, 38(2):185-208.

[29] Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis, 2E. New York NY: Chapman & Hall/CRC 2004.

[30] Pallmann P, Bedding AW, Choodari-Oskooei B, Dimairo M, Flight L, Hampson LV, Holmes J, Mander AP, Odondi L, Sydes MR, Villar SS, Wason JMS, Weir CJ, Wheeler GM, Yap C, Jaki T. Adaptive designs in clinical trials: Why use them, and how to run and report them. 2018;MBC Medicine 16(29):1-15.

[31] Efron B, Tibshirani RJ. An introduction to the bootstrap. New York NY :Chapman & Hall/CRC 1994.

[32] Pocock SJ. Group sequential methods in the design and analysis of clinical trials. Biometrika 1977;64(2):191-199.

[33] Pocock SJ (2006) Current controversies in data monitoring for clinical trials. 2006;Clinical Trials 3:513-521.

[34] Todd S, Whitehead A, Stallard N, Whitehead J. Interim analyses and sequential designs in phase III studies. 2001; Clin. Phama. 51:394-399.

[35] Hadad V, Hirshberg DA, Zhan R, Wager S, Athey S. Confidence intervals for policy evaluation in adaptive experiments. PNAS 2021;118(15):1-10.

[36] The GUSTO Investigators. An International Randomized Trial Comparing Four Thrombolytic Strategies for Acute Myocardial Infarction. NEJM 1993;329(10): 673:682

[37] The European trial on reduction of cardiac events with perindopril in stable coronary artery disease investigators. Efficacy of perindopril in reduction of cardiovascular events among patients with stable coronary artery disease: randomized, double-blind, placebo-controlled, multicenter trial (the EUROPA study). Lancet. 2003; 362:782-788

[38] Hauser JR, Urban G, Liberali G, Braun M. Website Morphing. Marketing Science 2019;28(2):202-224.

[39] The Adaptive Platform Trials Coalition., Angus, D.C., Alexander, B.M. et al. Author Correction: Adaptive platform trials: definition, design, conduct and reporting considerations. Nat Rev Drug Discov 18, 808. 2019. https://doi.org/10.1038/s41573-019-0045-0.

[40] Chen Q, Golrezaei N, Bouneffouf N. Non-stationary bandits with auto-regressive temporal dependency. 37th Conference on Neural Information Processing Systems. 2023:1-22.

[41] Liu Y, Xu K, Van Roy B. Non-stationary bandit learning via predictive sampling. arXiv:2205.01970v6. 2023:1-42.