Real-time Adaptive Randomization of Clinical Trials

Supplementary Online Content

Online Appendix

November 3, 2024

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix A. Notation and simulation procedures

eA.1. Notation and technical details

eA.2.1. Optimality (step IV.4.1 described below)

RTARs balance treating, providing the (likely) best treatment on day d, with learning, trying (currently) inferior arms to learn about endpoint rates so that better decisions can be made about all future patients, including those after the trial. For many reasons, positive endpoints and knowledge is better now than in the future, so we slightly discount future endpoints by a factor, δ, where $\delta \in (0, 1)$.

We represent knowledge about each arm by a (Bayesian) posterior distribution, a Beta distribution, with parameters α_{ad} and β_{ad}, for arm a based on endpoints observed up to day d. Gittins [e1] proved that the optimal balance between treating and learning is to compute a "Gittins index" for each arm and assign patients to the arm with the lowest index. (We are minimizing mortality, so lower is better.) We compute the Gittins index by comparing the "rewards" from a (currently) uncertain arm to the rewards from assigning patients to an arm where the endpoint rate is the Gittins index, $G(\alpha_{ad}, \beta_{ad})$. For the trials we analyze, "rewards" is mortality (GUSTO-1 and EUROPA) or another negative event (EUROPA). Fewer mortalities/negative-events are better. We write this value as G_{ad} for short. G_{ad} represents the comparative daily rewards for an arm in which the anticipated endpoint is known with certainty.

Let $R(\alpha_{ad}, \beta_{ad})$ be the expected discounted rewards for acting optimally on day d and all future days. We compare an uncertain arm to a certain arm to compute the Gittins index. In this comparison, the rewards for day d are related to G_{ad} and the rewards for day $d + 1$ by the Bellman equation. See derivation in [e2, e3].

$$R(\alpha_{ad}, \beta_{ad}) = min\left\{\frac{G_{ad}}{1-\delta}, \frac{\alpha_{ad}}{\alpha_{ad}+\beta_{ad}}[1 + aR(\alpha_{ad}+1, \beta_{ad})] + \frac{\beta_{ad}}{\alpha_{ad}+\beta_{ad}}\delta R(\alpha_{ad}, \beta_{ad}+1)\right\}$$

There is no analytical solution to this Bellman equation, but the $G(\alpha_{ad}, \beta_{ad})$'s are easy to compute numerically. We do the numerical calculations and store a table of Gittins indices by α and β.

The Gittins solution is provably optimal if one patient is assigned on day d and if the endpoint for patients assigned on day d is observed before assignments on day $d + 1$. The

Gittins solution becomes an heuristic algorithm, which we hope will reduce mortality relative to an RCT or the block-based FLGI, when more than one patient is assigned on day d and when there are delays in observed endpoints. This is an empirical question. The main paper tests whether the approximate optimal solution provides benefits relative to an RCT, the previously-proposed block-based FLGI solution, and an η-variant.

Future research might improve assignments with the solution to a fully optimal Bellman equation that accounts for multiple patients on day d and delays in endpoints. Thus, all results for the (hopefully) approximately-optimal RTAR in this paper are conservative relative to the solution to such a Bellman equation. Note that one way of handling delays is to change the discount rate, δ, to reflect a 30-day lag in updating. Fortunately, for our data, the Gittins solution appears to be robust to changes in the discount rate within a reasonable range suggesting that the loss of optimality due to delays may not be severe. Another heuristic might be to modify the RTAR to use the FLGI within each day d for which multiple patients are assigned. Such a variant, and many other variants, are readily explored by modifying our resampling simulation code. We did not explore all variations to avoid overfitting the empirical data.

eA.1.2. Learning (step IV.4.3)

The learning step is based on updating the Beta priors, α_{ad} and β_{ad}, with observations of the endpoints at day d. (The Beta priors are updated at the end of day d, patients assignments at the beginning of day d are based on all data up to, but not including, endpoints observed on day d.) Assuming the endpoints are observations from a Bernoulli process with stationary endpoint rates, the updating is simple and quick. When one endpoint is observed per day:

$$\alpha_{a,d+1} = \alpha_{ad} + 1, \beta_{a,d+1} = \beta_{ad} \text{ if the endpoint is a mortality}$$
$$\alpha_{a,d+1} = \alpha_{ad}, \beta_{a,d+1} = \beta_{ad} + 1 \text{ if the endpoint is survival}$$

If more than one endpoint is observed at the end of day d, say n_{md} mortalities and n_{sd} survivals, then we update using n_{md} and n_{sd}.

eA.2. Endpoints

In GUSTO-1, the endpoint is death or survival at 30-days since randomization. In EUROPA, the endpoint is a composite of cardiovascular mortality, non-fatal MI, and resuscitated cardiac arrest at any point in the trial.

eA.3. Simulation procedures

The RTAR multi-arm bandit (MAB) simulation can be run for any number of replicates. The number of replicates is set in the file "Parameters.R." All reported results in the manuscript are results averaged across 200 replicates. The RTAR MAB code is in the file "MAB.R." The η-variant follows the same procedure except that, with probability η, patients are assigned as in an RCT (equally likely) until the arm reaches a pre-defined minimum number of patients. With probability $1 - \eta k_d$, patients are assigned with the RTAR MAB. k_d is the number of arms that have not yet reached the minimum number of patients at the start of day d. The block-based forward-looking Gittins index (FLGI) algorithm is described in [e22]. In the block-based FLGI algorithm, patients are randomized in blocks. The code is available from the authors. We provide here the conceptual steps in the RTAR resampling simulations..

Step I. Load parameters and set seed.

When the parameter file indicates that a single replicate is to be run, the system uses a fixed seed. When more than one replicate is to be run, the system uses different random seeds in each replicate. Results are averaged across replicates. In the case of confidence intervals, we note the values where 2.5% are below (lower) or 2.5% are above (upper) the confidence limits.

Step II. Load support functions.

Step III. Load data.

Step IV. Loop over all replicates (this the main part of the code)

Step IV.1. Build the pools of patients for this replicate.

The original RCTs (GUSTO-1 and EUROPA) assigned one set of patients to each arm (treatment). We refer to each of these sets of patients as a "pool of patients" for that arm. These pools will be used in step IV.4.2, when the RTAR MAB algorithm draws (with replacement) from these pools when making its assignments.

Step IV.2. Load priors for $\alpha_a(d)$ and $\beta_a(d)$ for this replicate.

These priors are set in the file "Parameters.R."

Step IV.3. Initialize intermediate data structures

See the file "Data dictionary.txt" for details on the intermediate variables.

Step IV.4. For each day d of the trial in this replicate, perform the following:

Step IV.4.1 DECIDE: select arm a_d^* to use on day d. This is the optimality step.

Select the optimal treatment arm, a_d^*, given current $\alpha_a(d)$ and $\beta_a(d)$ parameters of the Beta distribution over treatment-arm endpoint rates. The parameters are based on all endpoints observed at the start of day d. The RTAR algorithm chooses the arm with the largest Gittins index, G_{ad}. G_{ad} is a pre-computed tabled function $\alpha_a(d)$ and $\beta_a(d)$. The Gittins index optimally balances, on a daily basis, the amount of treating and learning the system does [e2, e3]. For details on the optimality step, please refer to §e3.2.

Step IV.4.2. RESAMPLING PATIENTS: The RTAR MAB draws patients from the pool of arm a_d^*.

Compute the number of patients N_d that were randomized by the RCT on day d.

Assign N_d patients to the optimal treatment arm, a_d^*, by drawing with replacement N_d patients from the pool of patients that were randomized by the original RCT to the treatment a_d^*.

Step IV.4.3. LEARN

For each treatment arm, learn from the endpoints observed for all the patients that had been assigned to that treatment arm and for whom endpoints have been observed by the start of day d. Update $\alpha_a(d)$ and $\beta_a(d)$ as described in §e3.2.

Step IV.5. Summarize results of the original RCT.

Step IV.6. Summarize the results of this RTAR replicate.

Step IV.7. Save all outputs to .csv files.

eAppendix B. Statistical concepts and potential adaptivity bias

eB.1. Statistical concepts and potential adaptivity bias

Learning in the Gittins framework is inherently Bayesian and, hence, interpretations are based on the posterior distribution (prior time likelihood normalized). Because the Bayesian prior does not depend on patient assignment, the posterior distribution of endpoint rates also does not depend on how the RTAR assigns patients. 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 (or parameters of the model) and can be removed from the likelihood [e4]. Thus, the Bayesian posterior likelihood and the posterior distribution do not depend explicitly on how the RAR assigns patients. All information about unequal sample sizes among arms is included in the likelihood function [e4, e5, e6].

Because the posterior 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 days), efficient estimators (no consistent estimator has a lower asymptotic mean-square error), and are asymptotically normal [e4, e7, e8, e9]. From a Bayesian perspective, MLEs are asymptotically equivalent to maximum posteriori estimation with weakly informative priors [e10]. The ability to factor the likelihood and the fact that the assignments are a known function of the data implies that MLEs can be reported and analyzed after the trial (or simulation) is completed, especially for large samples [e4, e7, e8, e9, e11].

MLEs are consistent, but they may be biased for small samples [e9, e11, e12, e13, e14]. 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 [e9, e12, e15, e16, e17]. Small-sample biases occur in many RARs and require advanced statistics or propensity scores [e4, e5, e8, e18]. Such biases are minimal for the large samples in GUSTO-1 or EUROPA as reported in this paper [e4, e11, e18].

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 [e9, e12,

e13, e14, e19, e20, e21, e22, e23] or resampling with replacement when patient-by-patient endpoints are observed [e2, e11]. When we have patient-by-patient end-point observations, resampling generates the distribution of observations from which we compute means, medians, 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 [e11, e23, e24, e25, e26]. 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.

eB.2. Anticipated performance of an RTAR relative to an RCT

As an RTAR learns endpoint rates for each arm, we expect the RTAR 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 endpoint rates.

For two arms, pairwise power will be maximal and odds-ratio confidence intervals are tightest 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 GUSTO-1 and EUROPA trials.

By design, an η-variant approaches an RTAR as $\eta \to 0$ and approaches an RCT as $\eta \to 1$, thus we expect the performance of an 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.

eAppendix C. Design and principle results of the GUSTO-1 and EUROPA large-scale randomized controlled trials

eTable 1. Summary of the GUSTO-1 and EUROPA trials (conducted as RCTs)

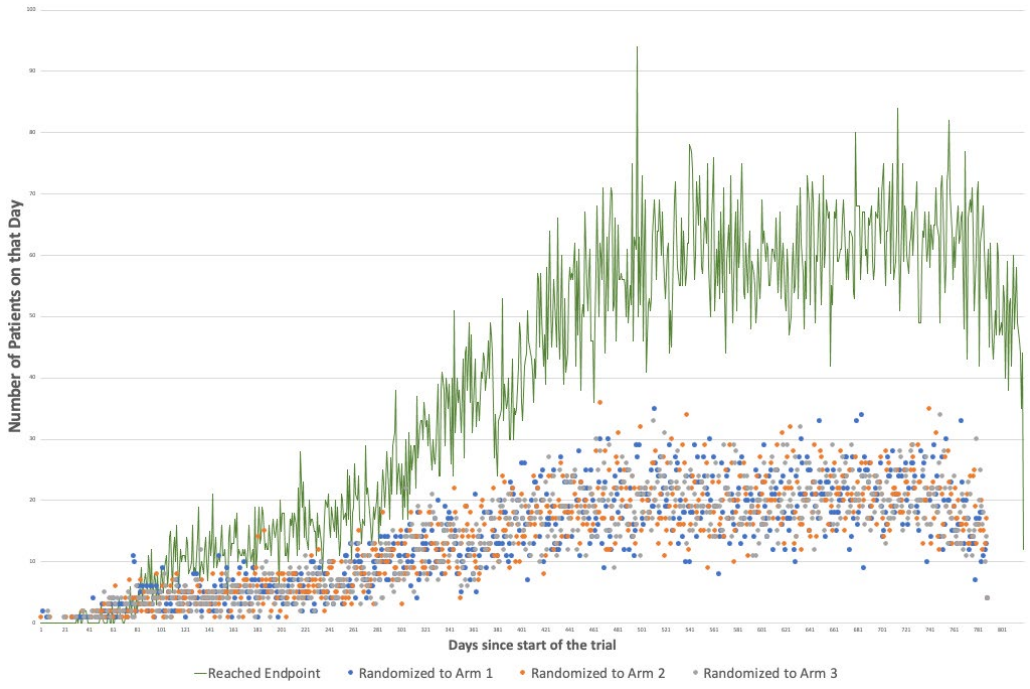
Trial details	GUSTO-1	EUROPA
Goal	Compare streptokinase and tissue plasminogen activator thrombolytic strategies in the treatment of acute myocardial infarction	Assess the effect of perindopril versus placebo on the combined endpoint of cardiovascular death, non-fatal MI, and resuscitated cardiac arrest 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)

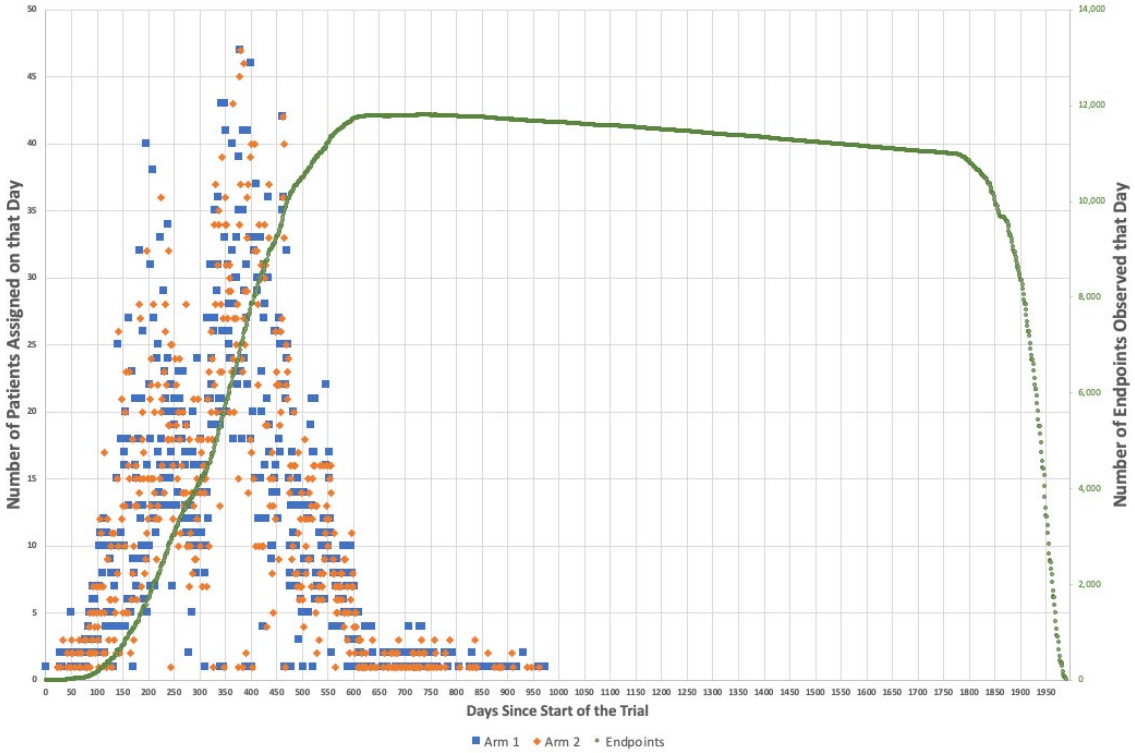
eAppendix D. Empirical details on the original RCTs and the RTAR simulations

eD.1. Randomization and endpoints for GUSTO-1 and EUROPA RCTs

eFigure 1 (for GUSTO-1) and eFigure 2 (for EUROPA) present the randomization of patients and the endpoints observed in both studies. The dots at the bottom of eFigure 1 and to the left of eFigure 2 correspond to the number of patients that were randomized by the RCT in each day of the GUSTO-1 trial (eFigure 1) and EUROPA trial (eFigure 2). This information is shown separately per treatment, using a color code. For GUSTO-1, blue corresponds to RCT randomizations to t-PA+Heparin (arm 1). Red corresponds to RCT randomizations to SK+Heparin (arm 2). Gray corresponds to RCT randomizations to t-PA+SK+Heparin (arm 3). For EUROPA, blue corresponds to RCT randomizations to Perindopril and orange corresponds to RCT randomizations to placebo. eFigures 1 and 2 also present, in the green solid line on the top of the figures, the number of endpoints that were observed in each day of the trial. This is the total number of daily endpoints summed over all arms in each study (GUSTO-1 had three arms; EUROPA had two arms).



eFig. 1: RCT Randomizations (in blue, orange and gray) and endpoints (in green) in GUSTO-1

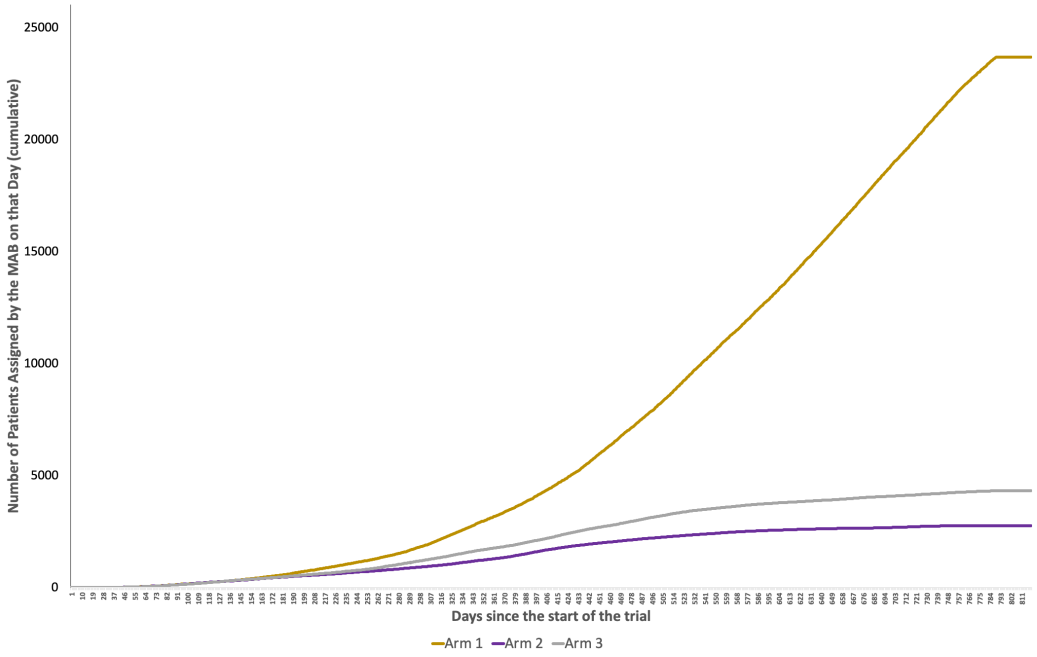


eFig. 2: RCT Randomizations (in blue and orange) and follow-up study's endpoints (in green) in EUROPA

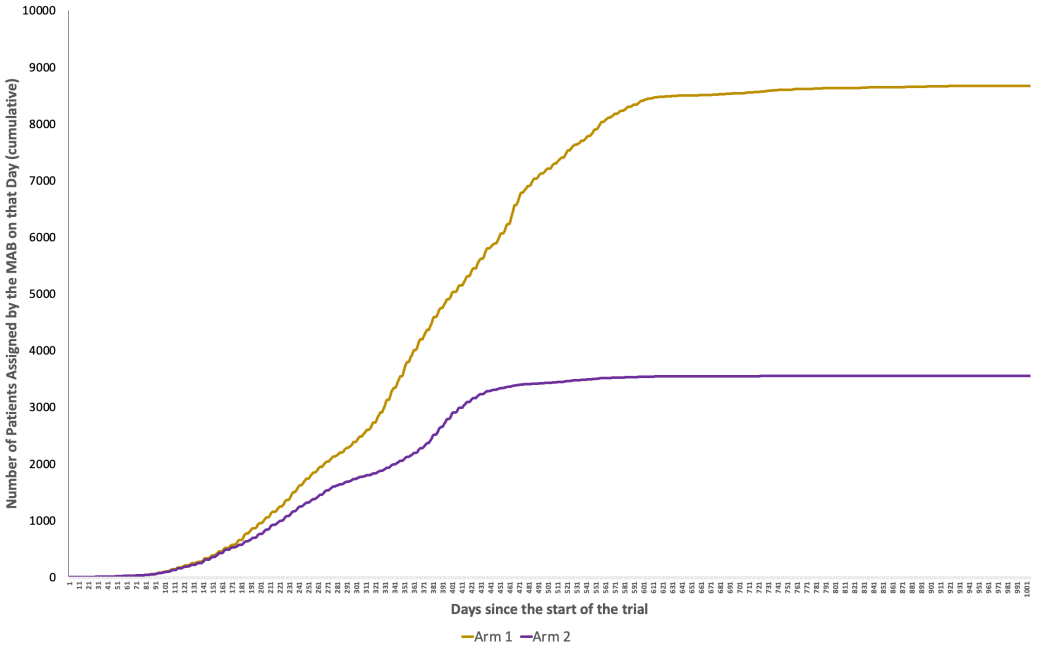
eD.2. Evolution of RTAR arm assignments for GUSTO-1 and EUROPA

eFigure 3summarizes arm assignments for (a) GUSTO-1 and (b) EUROPA. The purple, gold, and gray lines (GUSTO-1) or purple and gold lines (EUROPA) and the left vertical axis present the cumulative number of assignments over the duration of the trials. The horizontal axis represents the days of the trial. The RTAR adapts as data on patient endpoints become available. As the trial progresses, the RTAR automatically assigns more patients to the superior arm (gold line). By roughly the 500th day of the 819-day GUSTO-1 trial, and the 500th day of the 1,989-day EUROPA trial, the RTAR begins to assign almost all patients to the superior arm (gold line). In theory, a particularly adverse, but random, run of negative endpoints might lead an MAB to explore inferior arms after stabilization, but that probability is low. Future research might explore optimal stopping rules which could save even more lives than the RTAR studied in this paper.

A. GUSTO – 1



B. EUROPA



eFig. 3. Assignments to arms using the day-to-day RTAR

eAppendix E. Temporal Changes in Endpoint Rates

Temporal changes in endpoint rates are a known issue with RARs and are potentially an issue with RTARs [e5, e9, e14, e21]. 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 [e14, e21]. 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 [e9, e21].

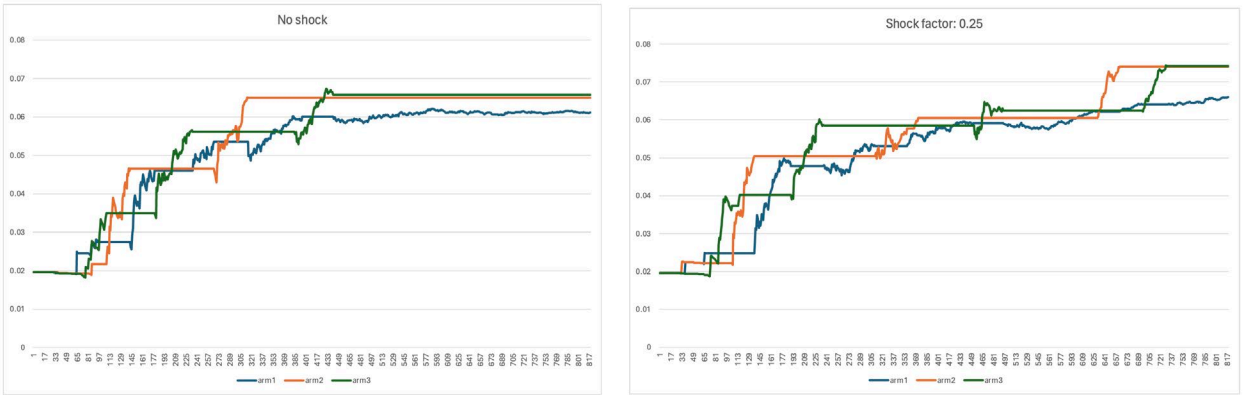
RTARs are based on Gittins indices which react to observed endpoint rates [e27]. If there are sufficient post-shock observations, the Gittins indices will evolve after the shock causing the MAB to reexplore the inferior arms. As a result, RTARs might be better able to react to shocks than many RARs.

To explore the impact of temporal changes in the endpoint rates on the performance of an RTAR, we use sample enrichment to simulate the effect of a change in endpoint rates midway through the trial (after half of the patients have been assigned). Patient enrichment is an accepted way to model temporal changes [e13] and provides equivalent implications to changing endpoint rates in simulations. With patient enrichment, we add sufficiently many patients to the pool for each arm such that the endpoint rate in each arm is the endpoint rate we seek to simulate.

For example, the RCT endpoint rate (mortality) for arm 1 in GUSTO-1 is 0.0615 and we wish to simulate a shock of 5%. For arm a in the first period, we draw from GUSTO-1 patients

for the days corresponding to the first $N_{a,total}/2$ patients in GUSTO-1 arm a. We modify the pool of patients from which we draw patients for the days corresponding to the second $N_{a,total}/2$ patients in the GUSTO-1 trial. A 5% increase in 0.0615 is 0.0646, an increase of 0.0031 in the mortality rate. To maintain consistency with the literature, we increase all arms by 0.0031 resulting in a vector of endpoint rates of [0.0646, 0.0753, 0.0717] for arms 1, 2, and 3, respectively.

eFigure 4 illustrates how the Gittins index changes during the trial for two separate GUSTO-1 simulations. A change in the Gittins index indicates that the arm is being used – when an arm is used, the Gittins index is updated with the end-outcome. For example, the left pane of eFigure 4 shows that *when there are no shocks*, the RTAR algorithm for this replicate stabilizes slightly before the 500th day of the trial, i.e., the RTAR stops assigning the two worst arms (arms 2 and 3). After the 500th day, the Gittins indices for arms 2 and 3 do not change as represented by the flat green and red lines. Note also that the ranking of the three arms also does not change.



eFig. 4. Gittins indices indicating assignments to arms using the day-to-day RTAR for one replicate of the GUSTO-1 data in the absence (left) and presence (right) of a 25% shock.

Separately, in another replicate, we introduced a 25% shock after 50% of the patients are assigned (day 530). The results, shown in the right pane of eFigure 4, indicates that the Gittins indices driving RTAR assignments have stabilized for this replicate around the 450th day (represented by the flat red and green lines for the inferior arms). Around day 560 (when the first 30-day after-shock GUSTO endpoint-outcomes were observed), the index detects changes in mortality rates and the RTAR returns to using all arms (trying first the red arm and then the green arm, both inferior). This is represented by changes in the green, blue and red lines. By day 730, the RTAR algorithm has sufficiently explored and learned about the arms' endpoint rates given the new mortality rates (post shock). Assignments again stabilize, and RTAR assigns

patients only to the arm it has automatically determined is the best arm (arm 1, blue line). The green and red lines for arms 2 and 3 become flat, representing no new patients being assigned to arms 2 and 3. The data for shock levels from 5% to 25% are provided in eTable 2. For exploratory research on MABs under temporal changes see [e28, e29, e30, e31].

eTable 2: The effect of temporal changes (temporal shock) for an RTAR and an RCT using sample enrichment of the GUSTO-1 trial

Shock		Arm 1	Std Dev Arm 1	Arm 2	Std Dev Arm 2	Arm 3	Std Dev Arm 2	Arm 1:2	Arm 1:3	Arm 2:3
5%	Mortality rate at start of trial	0.0620	0.0034	0.0720	0.0036	0.0690	0.0035	0.0100	0.0070	0.0030
	Mortality rate at end of trial	0.0651	0.0034	0.0751	0.0037	0.0721	0.0036	0.0100	0.0070	0.0030
	Mortality rate RCT	0.0636	0.0017	0.0736	0.0048	0.0706	0.0040	0.0100	0.0070	0.0030
	Mortality rate RTAR	0.0637	0.0017	0.0747	0.0049	0.0720	0.0040	0.0111	0.0083	0.0028
	Percent bias vs. RCT	0.16%		1.62%		2.02%		10.89%	18.88%	-7.76%
	Patients pre shock, RTAR	8,068		2,337		3,068				
	Patients post shock, RTAR	11,913		574		1,046				
	Patients to arm, RTAR	19,981		2,911		4,114				
10%	Mortality rate at start of trial	0.0620	0.0034	0.0720	0.0036	0.0690	0.0035	0.0100	0.0070	0.0030
	Mortality rate at end of trial	0.0682	0.0035	0.0782	0.0037	0.0752	0.0037	0.0100	0.0070	0.0030
	Mortality rate RCT	0.0651	0.0018	0.0751	0.0047	0.0721	0.0037	0.0100	0.0070	0.0030
	Mortality rate RTAR	0.0655	0.0018	0.0766	0.0047	0.0725	0.0037	0.0111	0.0070	0.0041
	Percent bias vs. RCT	0.57%		1.97%		0.55%		11.13%	0.41%	36.12%
	Patients pre shock, RTAR	8,171		2,217		3,085				
	Patients post shock, RTAR	10,702		983		1,848				
	Patients to arm, RTAR	18,873		3,200		4,933				
15%	Mortality rate at start of trial	0.0620	0.0034	0.0720	0.0036	0.0690	0.0035	0.0100	0.0070	0.0030
	Mortality rate at end of trial	0.0713	0.0036	0.0813	0.0038	0.0783	0.0038	0.0100	0.0070	0.0030
	Mortality rate RCT	0.0667	0.0018	0.0767	0.0045	0.0737	0.0037	0.0100	0.0070	0.0030
	Mortality rate RTAR	0.0671	0.0018	0.0781	0.0045	0.0742	0.0037	0.0109	0.0071	0.0038
	Percent bias vs. RCT	0.71%		1.85%		0.79%		9.46%	1.57%	27.88%
	Patients pre shock, RTAR	8,157		2,194		3,122				
	Patients post shock, RTAR	10,252		1,314		1,967				
	Patients to arm, RTAR	18,409		3,508		5,089				
20%	Mortality rate at start of trial	0.0620	0.0034	0.0720	0.0036	0.0690	0.0035	0.0100	0.0070	0.0030
	Mortality rate at end of trial	0.0744	0.0037	0.0844	0.0039	0.0814	0.0038	0.0100	0.0070	0.0030
	Mortality rate RCT	0.0682	0.0019	0.0782	0.0043	0.0752	0.0036	0.0100	0.0070	0.0030
	Mortality rate RTAR	0.0684	0.0019	0.0791	0.0043	0.0756	0.0036	0.0106	0.0072	0.0035
	Percent bias vs. RCT	0.35%		1.11%		0.52%		6.32%	2.17%	16.00%
	Patients pre shock, RTAR	8,132		2,254		3,087				
	Patients post shock, RTAR	9,491		1,729		2,312				
	Patients to arm, RTAR	17,623		3,984		5,399				
25%	Mortality rate at start of trial	0.0620	0.0034	0.0720	0.0036	0.0690	0.0035	0.0100	0.0070	0.0030
	Mortality rate at end of trial	0.0775	0.0037	0.0875	0.0039	0.0845	0.0039	0.0100	0.0070	0.0030
	Mortality rate RCT	0.0698	0.0019	0.0798	0.0042	0.0768	0.0037	0.0100	0.0070	0.0030
	Mortality rate RTAR	0.0703	0.0019	0.0812	0.0042	0.0775	0.0037	0.0109	0.0072	0.0037
	Percent bias vs. RCT	0.86%		1.83%		1.02%		8.56%	2.65%	22.37%
	Patients pre shock, RTAR	8,100		2,377		2,996				
	Patients post shock, RTAR	9,359		1,866		2,308				
	Patients to arm, RTAR	17,459		4,243		5,303				

eAppendix F. Early stopping, biomarkers, and discount rates

eF.1. Early stopping

An RTAR method adapts during the trial (increasing sample in the superior arm and reducing sample in inferior arms). By design, without early-stopping the trial the total sample size in the overall trial is not reduced. RTARs might be further be improved by addressing stopping early for futility or for compelling results. Optimal stopping in MABs is a challenging problem. One approach might be to adopt alpha-spending considerations [e32].

eF.2. Biomarkers and earlier adaptivity

An RTAR, with or without early stopping, saves lives because. relative to an RCT, the RTAR learns quickly to assign more sample to the superior arm. The GUSTO-1 and EUROPA simulations suggest that an RTAR has advantages even when there are delays in observing endpoints. If endpoints could be observed sooner, the advantages of an RTAR might improve. One potential practical solution to earlier observation might be to observe biomarkers and learn how (whether) the biomarkers predict endpoints. Adapting arm assignments based on biomarkers complicates the technical problem because, when we observe a biomarker, we observe a probabilistic prediction of the endpoint outcome. Adapting on biomarkers transforms the MAB to a partially-observable Markov dynamic program (POMDP). Such POMDPs have been studied in other contexts and provide promising results [e33].

eF.3. Sensitivity to "discount" rates

A "discount" rate enables an MAB to value current (positive) outcomes more than future (positive) outcomes. By using a discount rate, the MAB bases the treating-versus-learning decisions on both trial-based and post-trial patients. For RTARs and for MABs in general, unlike in financial decisions, the amount by which future outcomes are discounted is an ethical decision. For our simulations we chose a discount rate ($\delta = 0.9999$) that values future endpoint outcomes almost as much as current endpoint outcomes. A trialist might want to discount future endpoint outcomes more or less. To aid the trialist, we ran simulations with 200 replicates for discount rates of $\delta = 0.95, 0.99, 0.999, 0.9999,$ and 0.99995. The RTAR that we simulated appears to be robust to these changes in the discount rate. We hope future research will explore this sensitivity further.

eTable 3: Sensitivity to the discount parameter

Discount parameter	Arm	Mean Endpoint Rate	CI Lower	CI Upper	Sample Allocated	Events
0.95000	Arm 1	0.062	0.059	0.065	23,253	1,431
	Arm 2	0.072	0.066	0.089	3,086	221
	Arm 3	0.069	0.064	0.080	4,393	302
0.99000	Arm 1	0.062	0.059	0.068	22,947	1,413
	Arm 2	0.072	0.066	0.090	3,002	216
	Arm 3	0.068	0.063	0.081	4,783	327
0.99900	Arm 1	0.062	0.059	0.066	23,420	1,440
	Arm 2	0.072	0.066	0.097	3,021	218
	Arm 3	0.069	0.065	0.086	4,291	2966
0.99990	Arm 1	0.061	0.058	0.066	23,538	1,446
	Arm 2	0.072	0.066	0.089	2,922	210
	Arm 3	0.069	0.064	0.089	4,272	294
0.99995	Arm 1	0.061	0.059	0.065	23,442	1.441
	Arm 2	0.072	0.066	0.094	2,864	207
	Arm 3	0.069	0.064	0.085	4,426	304

eReferences (repeated and renumbered from the text for convenience)

[e1] Gittins JC 1979. Bandit processes and dynamic allocation indices. Journal of the Royal Statistical. Society 1979; (Ser B)41(2):148–177, plus commentary

[e2] Gittins J, Glazebrook K and Weber R. Multi-armed bandit allocation indices. London: Wiley. 2011.

[e3] Hauser JR, Urban G, Liberali G and Braun M. Website morphing. Marketing Science 2019;28(2):202-224.

[e4] Bowden J, Trippa L. Unbiased estimation for response adaptive clinical trials. Statistical Methods in Medical Research 2017;26(5):2376-2388.

[e5] Proschan M, Evans S. Resist the temptation of response-adaptive randomization. Clinical Infectious Diseases 2020;71:3002-3004.

[e6] Carlin BP, Louis TA Bayes and empirical Bayes methods for data analysis. Chapman & Hall/CRC, New York NY. 2000.

[e7] Hu F, Rosenberger WF. The theory of response-adaptive randomization in clinical trials. Wiley Series in Probability and Statistics. Wiley Interscience. Hoboken NJ. 2006.

[e8] Melfi VF, Page C. Estimation after adaptive allocation. Journal of Statistical Planning and Inference 2000;87(2):353-363.

[e9] Robertson DS, Lee KM, Lopez-Kolkovska BC, Villar SS. Response-adaptive randomization in clinical Trials: From myths to practical considerations. Statistical Science 2023;38(2):185-208.

[e10] Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis, 2E. New York NY: Chapman & Hall/CRC 2004.

[e11] 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. MBC Medicine 2018;16(29):1-15.

[e12] 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.

[e13] 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.

[e14] 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.

[e15] Pocock SJ. Group sequential methods in the design and analysis of clinical trials. Biometrika 1977;64(2):191-199.

[e16] Pocock SJ (2006) Current controversies in data monitoring for clinical trials. Clinical Trials 2006;3:513-521.

[e17] Todd S, Whitehead A, Stallard N, Whitehead J. Interim analyses and sequential designs in phase III studies. Clin. Phama. 2001;51:394-399.

[e18] Hadad V, Hirshberg DA, Zhan R, Wager S, Athey S. Confidence intervals for policy evaluation in adaptive experiments. PNAS 2021;118(15):1-10.

[e19] 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.

[e20] Wathen JK, Thall PF A simulation study of outcome adaptive randomization in multi-arm clinical trials. Clinical Trials 2017;14(5):432-440.

[e21] Villar SS, Bowden J and Wason J. Response-adaptive designs for binary responses: how to offer patient benefit while being robust to time trends? Pharm Stat 2018;17:182–19

[e22] Villar SS, Wason J, Bowden J. Response-adaptive randomization for multi-arm clinical trials using the forward-looking Gittins Index rule. Biometrics 2015;71:969-978.

[e23] 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.

[e24] Du Y, Wang X, Lee JJ. Simulation study for evaluating the performance of response-adaptive randomization. Contemporary Clinical Trials 2015;40:15-25.

[e25] 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. Journal of Clinical Epidemiology 2023;153:45-54.

[e26] Efron B, Tibshirani RJ. An introduction to the bootstrap. New York NY :Chapman & Hall/CRC 1994.

[e27] The GUSTO Investigators. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. NEJM 1993;329(10):673:682

[e28] Chen Q, Golrezaei N, Bouneffouf D. Non-stationary bandits with auto-regressive temporal dependency. 37th Conference on Neural Information Processing Systems 2023:1-22.

[e29] Liu, Y, Kuang, X, van Roy B. Non-stationary bandit learning via predictive sampling [arXiv:2205.01970]. arXiv preprint arXiv:2205.01970. 2023

[e30] Simchi-Levi D, Wang C, Zheng Z. Non-stationary Experimental Design under Structured Trends (July 18, 2023). Available at SSRN: https://ssrn.com/abstract=4514568 or http://dx.doi.org/10.2139/ssrn.4514568

[e31] Liu Y, Kuang X, van Roy B. A definition of non-stationary bandits [arXiv:2302.12202]. arXiv preprint arXiv:2302.12202. 2023.

[e32] DeMets DL, Lan KKG. Interim analysis: The alpha spending function approach. Statistics in Medicine 1994;13:1341-1352.

[e33] Hauser JR, Urban G, Liberali G, Braun M. Website morphing. Marketing Science 2019;28(2):202-224.