Real-time Adaptive Design for Clinical Trials

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 randomizations

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 Design for Clinical Trials

Abstract

Objective: To evaluate whether real-time (day-to-day) adaptation of clinical trials improves patient outcomes (lower mortality within the trial) and to identify the implied tradeoffs in confidence intervals, statistical power, and potential misidentification. The data and the analyses address the effect of delayed outcomes, e.g., a 30-day delay between treatment and observed outcomes.

Study Design and Setting: Using data from two large-scale random control test (RCT) clinical trials (30,732 patients in GUSTO-1 and 12,218 patients in EUROPA), we simulate the arm assignments and expected outcomes had the trials been run using real-time adaptation. Real-time adaptation is accomplished with a multi-arm bandit (MAB) in which blinded assignments are revised based on outcomes observed up to and including the previous day. We also investigate two variants, an η-variant that balances RCT and MAB assignments and a forward-looking multiday block-based MAB.

Results: Real-time assignment quickly learns which arm is superior. By the end of the trial, real-time assignment allocates more sample to the superior arm and less sample to the inferior arm(s) resulting in fewer mortalities over the course of the trial. Outcome probabilities are well within (frequentist) statistical confidence of RCT outcome probabilities, but with tighter confidence intervals on the superior arm and less-tight confidence intervals on the inferior arm(s). For the multi-arm trial in our data, statistical power is comparable (within a few percentage points) for pairwise comparisons of the superior arm and inferior arms but with a loss of statistical power for comparisons between inferior arms. The η-variant and the block-based MAB provide intermediate levels of benefits and costs while reducing the potential for misidentification of which arm is superior.

Conclusion: Real-time assignment has the potential to improve patient outcomes within the trial (beneficence) and the potential to reduce the confidence interval for the superior arm. These benefits are balanced with larger confidence intervals on inferior arms and less statistical power between inferior arms. Variants (η and block-based) provide intermediate benefits and costs.

Real-time Adaptive Design for Clinical Trials

1. Introduction

Significant breakthroughs in the medical treatment of patients have been achieved through the careful route from laboratory research to experimental animal studies, to clinical studies in subjects with established disease. Large Phase III trials on the effectiveness and safety of drug treatments form the important final phase of this development process. Often several thousands to tens of thousands of patients are studied before a drug is released to the market, and before a recommendation is included in the guideline. Typical examples include the introduction of renin angiotensin aldosterone system (RAAS) inhibitors in heart failure (~48-57,000 patients in Phase III trials)[1,2], and P2Y12-receptor inhibitors in patients with acute coronary syndrome (~145,000 patients) [3].

Increasingly, however, large Phase III drug trials are under pressure to reduce their length and cost while also reducing potential harm to patients. The COVID-19 pandemic made the balance between speed of new drug development and statistical robustness in trials ever more relevant. New legislation and regulations regarding drug trials exacerbate the administrative burden and may impede the practical implementation of clinical trials. This may encourage trialists to seek greater efficiency in trial designs.

We examine an alternative way to study clinical effects of drug treatments. We examine whether real-time adaptive (day-to-day) allocation can minimize the trial size (and the trial burden) without compromising the reliability of its final result. During the execution of a Phase III study, real-time adaptive allocation to the experimental drug and standard (or placebo) treatment might save healthy life-years if it becomes clear more quickly whether (and to what

extent) the innovation leads to a clinically more favorable result, or if the recommendation is identified successfully with fewer patients assigned to inferior arms. As uncertainty decreases during the conduct of a trial (RCT or adaptive), it may become unethical at some point to ignore the trend in outcomes when assigning patients. Trialists seek an optimal balance between (1) the need to learn the effectiveness and safety of the experimental treatment with (2) the benefits of treating patients with the treatment that is best within a reasonable statistical confidence based on current outcomes. Real-time adaptive allocation allows proportionally more assignments to the superior arm (as identified within the trial). Real-time adaptation might also allow stopping a trial early if necessary, based on futility or serious side effects.

In most Phase III drug trials, (frequentist) significance tests and control of the corresponding type I (false positive) and type II (false negative) errors are central. This focus often leads to a rigid design, with a specified number of patients based on the expected treatment effect and the chosen target confidence levels for the statistical hypothesis tests. Interim analyses of trial data are often included to understand the direction and magnitude of the effect of the experimental drug before the trial is completed. However, the number of interim analyses is usually limited and the bar to adapt the trial design is set high to avoid false positives that might result from multiple 'looks' at the data. We explore whether clinical trials can identify the effects of experimental drug treatments earlier (and thus be based on smaller samples or proportionally more sample to the superior arm) if *real-time* adaptation of clinical trials is implemented in combination with Bayesian thinking. Bayesian methods provide a more robust way of incorporating new information (such as a patient outcomes) about the parameter of interest (such as the efficacy of a drug). This means that Bayesian methods can update existing efficacy

estimates in real time, during the trial. Importantly, real-time adaptation can be implemented automatically, without jeopardizing blinded treatment allocation.

"Continuously adaptative" trials seek to optimize the tradeoff between maximizing short-term outcomes of the trial, such as the best possible recovery of patients who must be treated now, and long-term outcomes, including the fastest path to discovery and dissemination of new treatments to all future patients. From a statistical point of view, this optimization underlying real-time adaptation is a multi-armed bandit (MAB) application.[5,6, 7, 8, 9, 10, 12]

The literature on adaptive designs for trials using between-interim points information is rarefied. Previously, an MAB strategy has been shown to improve outcomes in simulated data by assigning patients who arrive in blocks. The algorithm assumes that all outcomes are observed within the block such that there is no delay between block-based assignments and outcomes [11]

The goal of the current study is to assess the value, strengths, and caveats of continuously adapting trials. We focus on the trials where patients are randomized on a day-to-day basis (as opposed to arriving in multi-day blocks) and where there are substantial delays between randomization and outcomes. We implement real-time adaptation of clinical trials in an algorithm and test it on data from two large-scale trials, exploring its performance boundaries and comparing its performance with preplanned random assignment and with block-based approaches. Each comparison is based on individual patient data empirically-grounded from the large-scale clinical trials. The simulations represent would have happened had these trials been adapted in real-time (day-to-day).

In the next sections, we present how real-time adaptation, based solely on information available to the trialists during the trial, might have prevented cardiovascular events during the trial in both studies. Real-time adaptation reliably identifies the best treatment while using

observed outcomes to change the randomization rate, thus identifying and favoring the most effective treatment over the course of the trial.

The prevention of cardiovascular events does not come without ethical tradeoffs [11, 14,15] . Gains in patient beneficence and more (frequentist) statistical power for superior arms implies smaller samples on inferior arms and reduced statistical power on pairs of inferior arms (multi-arm trials). Fortunately, this does not affect the ability of real-time adaptation to identify the best treatment.

2. Methods: process/principle, calculation, software

2.1 The principle of real-time adaptation of trials

Real-time adaptation of trials requires a preplanned statistical algorithm to handle the continuous flow of information. An MAB algorithm assigns patients to one of the treatment arms, based on the knowledge of the outcomes of each arm as they appear during the course of the trial. Starting with a balanced 1:1 randomization ratio, MABs tend toward a ratio that favors randomization to the treatment arm with the best expected outcome, utilizing all information (on the occurrence of the trial endpoints) that has been gained within the trial up to the time that allocations are made. An MAB automatically randomizes a sufficient number of patients to all study arms, but not necessarily in a 1:1 ratio, until uncertainty in effect estimates is reduced to a level such that it is no longer of value to assign patients to the inferior arms. The algorithm optimizes patient outcomes, while avoiding endpoints that are 'unnecessary' for the learning process [13].[1] Because the randomization ratio changes dynamically as patients are assigned to

[1] Under reasonable assumptions about the independence of the arms, the algorithm has been proven to be the assignment strategy that optimally balances learning with maximizing patent outcomes. See, for example, Gittins, J.

arms, the MAB stops assigning patients to inferior arms automatically when reducing uncertainty in the inferior arm is no longer justified.

Figure 1 presents the conceptual framework of real-time adaptation of clinical trials, and its core two steps: assignments and updating.

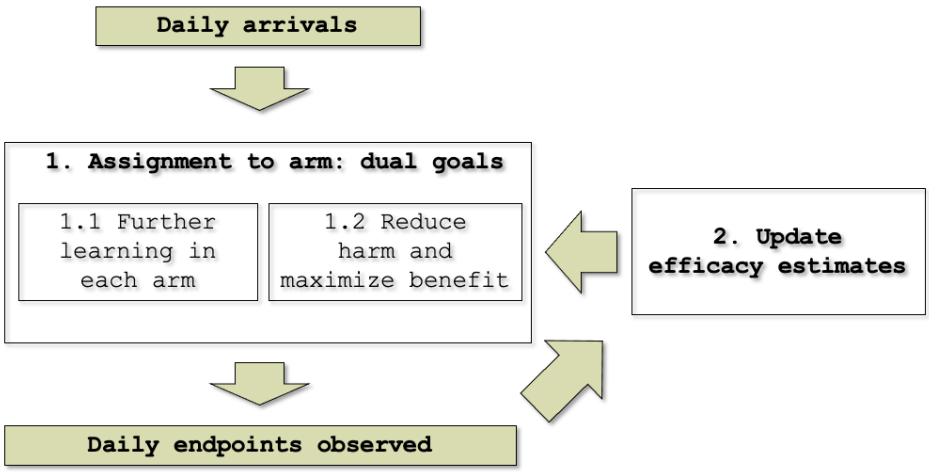


Fig. 1 - Real-time adaptation of clinical trials

2.2 The statistical details

Assignments to arms are based on the best estimates of the endpoint outcome probabilities, and the uncertainty about those probabilities, accrued up to and including each day of the trial that has occurred before patient assignment. Updating is the process of observing endpoints as they occur and automatically recalculating estimates and uncertainty about outcome probabilities based on all observations up to and including the observed endpoints. The typical MAB represents system knowledge about the endpoint probabilities using a beta probability distribution per arm [13]. The beta probability distribution has two parameters, α_a and β_a, for each arm a. Larger parameter values mean less uncertainty in beliefs about the arm's efficacy. The expected value of the beta distribution, $\alpha_a / (\alpha_a + \beta_a)$, gives an estimate of the endpoint

C. 1979. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical. Society*. Ser. B 41(2) 148–177, plus commentary.

probabilities for each arm (e.g., probability of a cardiological event). The parameters also enable the trialist to quantify uncertainty. The variance of knowledge about outcome probabilities is an analytic function of α_a and β_a. Technically, when a delayed outcome, say 30-day mortality, is observed, α_a is incremented by +1 for mortality and β_a is incremented by +1 for a survival. The updating rule is predetermined and can be blinded.

The algorithm, automatically and optimally, solves for the best assignment based on these parameters. Calculations are completed in real time and balance the short-term outcomes (best assignment for that patient) with long-term outcomes of the trial (best learning for all future patients). As data arrive, parameters increase in value and randomization rates change. When more than one patient arrives on a given day, we explore two strategies. In one strategy, all patients on a given day are assigned to one arm. In a second strategy, the algorithm provides randomization rates using methods inspired by previous research on block assignment [11]. We compare both strategies to the standard strategy of randomly assigning arms in a fixed ratio and passively observing the changes in the expected value of the efficacy of each arm. MABs differ from traditional adaptive trials in which Data and Safety Monitoring Boards (DSMBs) might recommend adjusting assignments a small number of times[16]. In a real-time adaptive trial, assignment ratios change automatically after each day (or, if practically feasible given the challenges of blinding, after each patient). Real-time adaptive trials may still decide to use DSMBs for safety checks and other tasks.

The algorithm used in our (simulation) analyses is based on Gittins' optimal solution to the dynamic-programming Bellman Equation for the MAB [17]. In Gittins's solution, patients are assigned to the arm with the largest index, where the index, which depends on the two parameters, α_a and β_a, and is computed based on all knowledge gained prior to the assignment.

The index also depends upon a parameter, called a discount rate [9, 11, 17] , chosen according to the length of the trial. For technical details and notation, please refer to the online Appendix. In the analyses used in this article, the initial parameters of the algorithm are set in advance so that we, as researchers, do not change the assignment rules based on intermediate outcomes. The "discount" parameter is chosen conservatively, and further analyses indicate that the algorithm is robust with respect to this parameter.

2.3 What if the GUSTO-1 trial and the EUROPA trial had been adapted in real time?

Patients

To study the potential performance of real-time adaptation of trials, we conducted simulations grounded by the data from the GUSTO-1 [18] and the EUROPA trials [19] 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[2] presenting with acute myocardial infarction to one of three thrombolytic strategies. (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 (mean) 4.2-year treatment with the angiotensin-converting-enzyme (ACE) inhibitor perindopril or to a matching placebo. The primary endpoint of was death by myocardial infarction or cardiac arrest and was lowest (8%) in those patients randomized to perindopril. In

[2] 30,752 patients after excluding observations with missing data.

2003, the investigators concluded that, “on top of other preventive medications, [perindopril] should be considered in all patients with coronary heart disease.”

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 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, myocardial infarction, 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
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)

2.4 Data and grounded simulations

The follow-up duration for the primary endpoint of GUSTO-1 was 30 days after randomization (Table 1)[18]. In contrast, the mean followup period in EUROPA was 4.2 years. We explore how these differences in trial design impact the application and effectiveness of the real-time adaptation. The detailed distribution of the number of Randomized Control Test (RCT) randomizations and endpoints per day in the original trials is presented in the online Appendix (eFigures 1 and 2). Both plots cover the entire duration of the trial, from the first randomization until the last primary endpoint was observed.

Based on the available trial data, we simulated what would have happened had the trial been adapted in real time by a MAB. We started with a uniform randomization ratio (1:1:1 for GUSTO-1; 1:1 for EUROPA). For each day (D) of the trial, the algorithm assigns all arriving patients at day D to the study arms, based on the informational state of the system up to that the beginning of D. Such assignments are done in the simulations by randomly drawing (with replacement, given exchangeability of patients[3]) a patient from the pool of patients in the chosen arm. To avoid a particularly favorable draw, we repeat the process with 100 replicates for each study. For example, 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. The grounded simulation continues until the number of patients in the simulation matches the number of patients in the original RCT trial. The assignment rule is automated and optimized according to pre-established rules as determined by automated use of the MAB. The MAB updates the information state of the system by observing the incidence of the primary endpoint for the patients in the trial using only information that is

[3] We are using the principle of exchangeability of patients to draw with replacement. Thus, the number of patients in an arm in the simulation can exceed the number of patients in an arm in the RCT. Drawing with replacement is a well-established statistical method (similar to bootstrapping for standard errors)

known at the time of the new patient assignment. Delayed outcomes that are not observed by day D are not used in the day D assignment. We repeat this procedure for all days of the trial.

Performance was studied, and is reported here, with respect to the number of assignments made to arms, and with respect to the primary study endpoints as estimated at the end of the empirically-grounded simulation. The confidence intervals reported are computed over 100 replicates, providing evidence of the uncertainty on the estimated incidence of the primary endpoint. Because most RCT trials are framed in classical frequentist interpretations, we report the odds ratios for all pairs of arms and corresponding Wald-based frequentist 95% confidence intervals using an univariate logistic regression in the online Appendix [20].

2.5 η-*Variation to Ensure a Target Minimum Power*

Empirically, the MAB identifies the superior arm quickly and assigns substantial sample to the superior arm. Less sample is assigned to inferior arms resulting in less frequentist statistical power for the inferior-arm outcome probabilities. This is an ethical dilemma. Trialists may wish to assure a minimum sample size (minimum statistical power) on the inferior arms [21].

To explore this issue while retaining some of the advantage of real-time adaptation, we report an η–variation of the MAB algorithm that explicitly guarantees a minimum level of statistical power to inferior arms. In the η-variation, we randomize patients across all arms that have not reached the minimum sample size with probability η, and assign patients with the Gittins solution with probability $1 - \eta k_D$, where k_D is the number of arms below the minimum sample size per arm at the start of day D. As an illustration of the η-varinant, we set the tuning parameters to $\eta = 0.25$ and the minimum sample size to 6,000 patients. These values of the tuning parameters are conservative making the MAB more like an RCT early in the trial. Depending upon the trialists' judgment with respect to the ethical dilemma, the trialist can set more-aggressive (favoring the

MAB) or less-aggressive (favoring the RCT) values of the parameters. See the last three rows of Table 2 for the effect of the illustrative parameters on the GUSTO-1 trial.

2.6 Patient blocking in the presence of outcome delays

The role of MABs was previously examined in trials where patients arrive in blocks and outcomes are observed before the next block is assigned [11]. The method provided "substantial improvements in terms of patient benefit, especially for small populations." Patients were assigned in blocks in part so that delayed outcomes could be observed within the time frame of the block. When comparing the block-based MAB to other trial strategies including fixed designs and Thompson sampling, the authors identified a tradeoff between improved patient outcomes and statistical power. For a scenario with response rates matching an empirical trial and for a block size large enough to observe all outcomes, the block-based MAB improved the expected number of patient successes by almost 50% but with a reduction in statistical power of approximately 70%. Results depended upon the block size, with better patient outcomes and lower power observed for smaller blocks. Other test designs produced intermediate outcomes and power relative to the block-based MAB and an RCT.

The block-based adaptive trial strategy 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 [11]. Outcomes 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 outcomes and Gittins-Index updates within a block. (*First-block* assignments are random draws based on the treatment priors.) The

simulation to identify assignment probabilities assumes (1) the first patient is assigned based on the Gittins Index calculated based on previous-block outcomes and (2) second and subsequent patients within a block are assigned, outcomes observed, and Gittins Indices updated based on simulated outcomes with priors based on all outcomes observed prior to the simulated assignment. Using patient interchangeability, the algorithm calculates the expected percentages of arm assignments over all possible patient orders. This is equivalent to considering the expected value of the assignments that the Gittins' algorithm would have made within a block were outcomes observed based on the history up to a patient. To make the algorithm feasible, the order of patient arrivals are sampled rather than exhaustively enumerated. This algorithm has become known in the literature as the forward-looking Gittins index (FLGI) algorithm [8, 11].

For the GUSTO-1 and EUROPA trials, we assume updating occurs only for those patients whose outcomes are observed within a block. This assures that the delayed-outcome-block-based algorithm is the same as a real-time assignment algorithm when the block size is one day. We use a block size consistent with the strategy in [11]; we set $b = 60$ days so that most 30-day outcomes are observed within a block. Our code is available for researchers who wish to evaluate other block sizes.

3. Results

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 sample size, number of primary endpoint events, and event rate for the original RCT (in the first three rows) along with confidence intervals [22]. In the last nine rows of Table 2, we present the results had these

trials used MABs to adapt in real time (day-to-day), to adapt in real-time with the η-variant, and to adapt with the block-based variant.

For both trials, the ranking of all arms in the simulation of real-time trial adaptation matches 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 the MAB are quite close to those estimated with the RCT and well within the confidence intervals (computed based on the 100 replicates). This is true for all tested MAB strategies. 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 real-time adaptation. But the tighter bound for the superior arm comes with a tradeoff: confidence intervals are not as tight on the (identified) inferior arms.

All tested MABs provide better patient outcomes (beneficence) with the lowest mortality for real-time adaptation (1,941 lives lost) and the highest mortality for the RCT (2,074 lives lost) – a net saving of 133 lives due to real-time adaptation. The net savings for the η-variant and the block-based MAB were 72 and 103 lives saved, respectively.

The gain in lives saved comes at the cost of making fewer assignments to the inferior arms. For real-time adaptation the assignments to the inferior arms were 2,259 as opposed to 10,268 in the SK + Heparin arm for GUSTO-1, 3,969 as opposed to 10,209 in the t-PA+SK + Heparin arm for GUSTO-1; and 5,692 as opposed to 6,108 in the placebo arm in EUROPA). The other variants allocated more sample to the inferior arms. Note that while the η-variant specifies a target minimum sample size, the actual assignments for inferior arms vary slightly from the target reflecting the tendency of the real-time adaptive portion of the η-variant to favor 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		
No.	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 (0.073, 0.087)	0.099 (0.092, 0.106)	
MAB[b] $\eta = 0$					MAB[b, d]		
No.	24,504	2,259	3,969	30,732	6,526	5,692	12,218
Events	1,509	162	271	1,941	524	560	1,084
Event rate	0.062 (0.061, 0.063)	0.076 (0.074, 0.077)	0.073 (0.070, 0.075)		0.080 (0.079 0.081)	0.098 (0.098.099)	
MAB[b,c] $\eta = 0.25$							
No.	18,479	5,683	6,570	30,732			
Events	1,136	413	453	2,002			
Event rate	0.062(0.061, 0.062)	0.073 (0.072,0.073)	0.069 (0.069, 0.070)				
MAB with blocks[b,d]							
No.	21,593	3,579	5,560	30,732			
Events	1,336	255	379	1,971			
Event rate	0.062 (0.062, 0.063)	0.072 (0.072, 0.073)	0.070 (0.069, 0.071)				

[a] Best arm in the trial

[b] Averaged over 100 replicates. MAB priors in GUSTO: $\alpha_0=6$, $\beta_0=390$. MAB priors in EUROPA: $\alpha_0=2$, $\beta_0=300$.

[c] Minimum n=6000.

[d] block size: 60. Monte Carlo draws: 100.

3.1 More lives would be saved for trials with larger differences in outcome probabilities

Every life is important, but one might ask whether 133 fewer deaths out of 2,074 mortalities 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, 0.063, 0.070, and 0.073. As a 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 with 30,732 patients, a real-time adaptive assignment would have saved 1,700 lives compared to a 1:1:1 RCT assignment. We kept the total patients the same for this comparison recognizing that if the trialist had strong priors on the mortality risk and required the same statistical power, the trialist would allocate less sample to both the MAB and to the RCT. However, the MAB would still save a substantial proportion of lives. The more that the true outcome probabilities differ for a given sample size, the greater the number of lives saved.

3.2 When outcomes are severely delayed, an aggressive η-variant mitigates misidentification

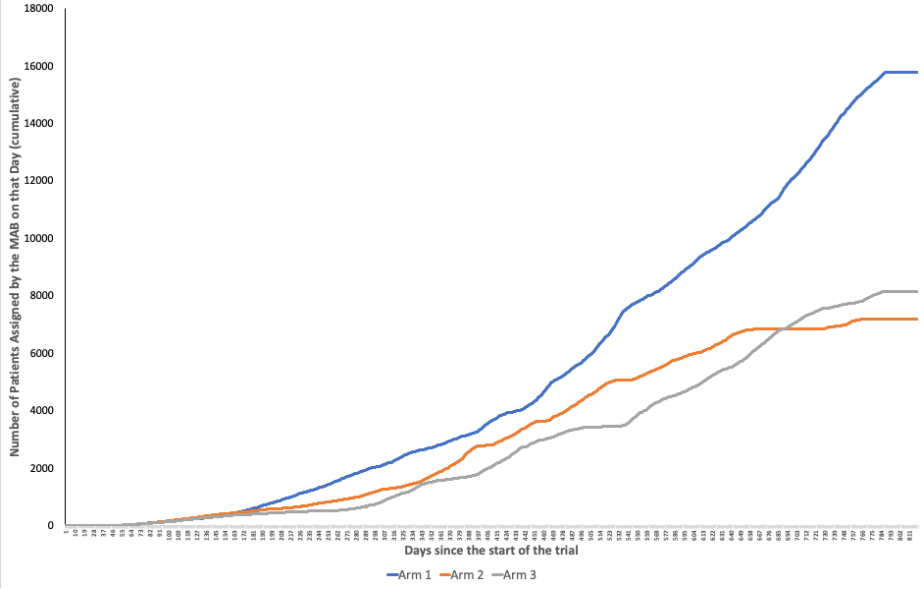
GUSTO-1 primary endpoints were observed at the 30th day after randomization, which is relatively fast. The EUROPA trial provides the opportunity to assess the effect of longer delays in outcomes. The longer duration in observing outcomes affects an MAB's ability to learn from the variation in endpoints during the early stages of the trial. Less rapid learning and adaptation affects the incremental advantage of optimal, rather than equally-likely, random assignment [23]. In the EUROPA-grounded simulation with a MAB, the treatment effect is statistically significant at $p \leq 0.05$ for 83% of of these replicates. If we apply an aggressive η-variant with $\eta = 0.50$ and sample size equal to 50% of each arm in the RCT trial, then all replicates had a p-

value that achieved significance, with only a slight change in the confidence intervals. (This η-variant is similar to a burn-in strategy as in [24].) As in Table 2, the lower likelihood of misidentification comes at the cost of smaller reduction in the incidence of cardiological events. For statistical details please refer to the online appendix D.

3.2 Visualizations of arm assignments and statistical power

Figure 2 summarizes arm assignments for (a) GUSTO-1 and (b) EUROPA. The blue, orange, and gray lines (GUSTO-1, blue and orange 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 real-time-adaptation MAB adapts as data on the patient endpoints become available because it learns the incidence rate of the primary endpoint (uncertainty decreases). As the trial progresses, the real-time MAB automatically assigns more patients to the superior arm (blue line). By the 200^{th} day of the 819-day GUSTO-1 trial, and the 350^{th} day of the 1,989-day EUROPA trial, the real-time MAB begins to assign all patients to the superior arm (blue line). Future research might explore the implications, especially ethical implications, of stopping the trial soon after assignments stabilize to the superior arm.

A. GUSTO – 1



B. EUROPA

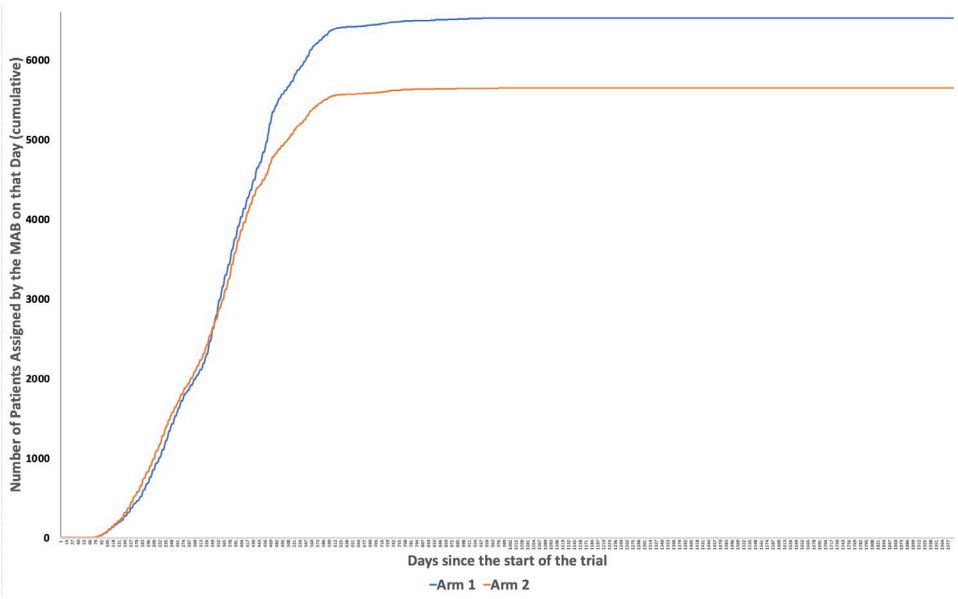


Fig. 2 – Assignments to Arms using the Day-to-Day Multi-arm Bandit Algorithm

Type I and II errors evolve throughout the trial as does the estimated odds ratio [20] because the MAB learns as more patients are allocated during both trials. By the end of the trial, the odds ratios based on the MAB converge to those based on the RCT. The mean and 95% bootstrapped confidence interval for the odds ratio of the superior arm, t-PA + Heparin, versus

SK + Heparin, at the end of the adaptive trials in GUSTO is 1.243 (1.214, 1.271) and versus tPA+SK+Heparin is 1.088 (1.07, 1.105). In EUROPA, the mean and 95% confidence interval of the odds ratio for Perindopril versus placebo is 1.288 (1.239, 1.337). The odds ratio for the superior to most sub-optimal are above 1.0 and tighter than the confidence interval comparing the two suboptimal treatments GUSTO (SK + Heparin and t-PA+SK + Heparin). The odds ratios decrease with sample size, as shown in the online Appendix (eFigures 3).

3.3 Effect of real-time trial adaptation on statistical power

Compared with the fixed-design RCT, real-time adaptation assigns a larger sample of patients to the superior arm. The increased sample for the superior arm does not come for free – it implies less sample for the inferior arms which, in turn, implies lower statistical power for the inferior arms. To focus on this tradeoff, we plot the change in statistical power over the duration of the GUSTO-1 trial enrollment (789 days until the last patient was enrolled).

Figure 3 presents the statistical power for an expected 15% reduction in mortality ($p_1 = 0.08$ and $p_2 = 0.068$). The type I error rate is set to $\alpha = 0.05$. The green line presents the statistical power for which the original RCT (all pairs of arms) was designed (90%). As expected, power increases smoothly to 90% as the RCT assigns sample randomly to the three arms. An MAB is likely to present the trialist with ethical tradeoff because, on the sole criterion of optimizing power for all pairs of arms, the RCT is best [11]. The blue, red, and purple lines represent the statistical power for all three pairs of arms given the daily randomization rate that would have been observed had a real-time adaptation MAB been used in this trial. The increase in power is less smooth and varies by arm-pair. At the end of the trial, the statistical power for comparisons of the superior arm (arm 1, determined endogenously within the trial) to one inferior arms (and 3) is slightly above

that of the RCT, but the comparison between the superior arm (arm 1) and the other inferior arm (arm 2) is slightly below that of the RCT. The comparison between the two inferior arms is well below that of the RCT. These power computations simply reflect that the MAB assigned fewer respondents to arm 3 than did the RCT, and even fewer respondents to arm 2.

The statistical power is computed as the power that the resulting sample size provides to detect the true mortality rates on that day. (The RCT endoutcome rates are: arm 1(t-PA, Heparin) = 0.062; arm 2(SK, Heparin) = 0.072, arm 3 (t-PA, SK, Heparin) = 0.069.

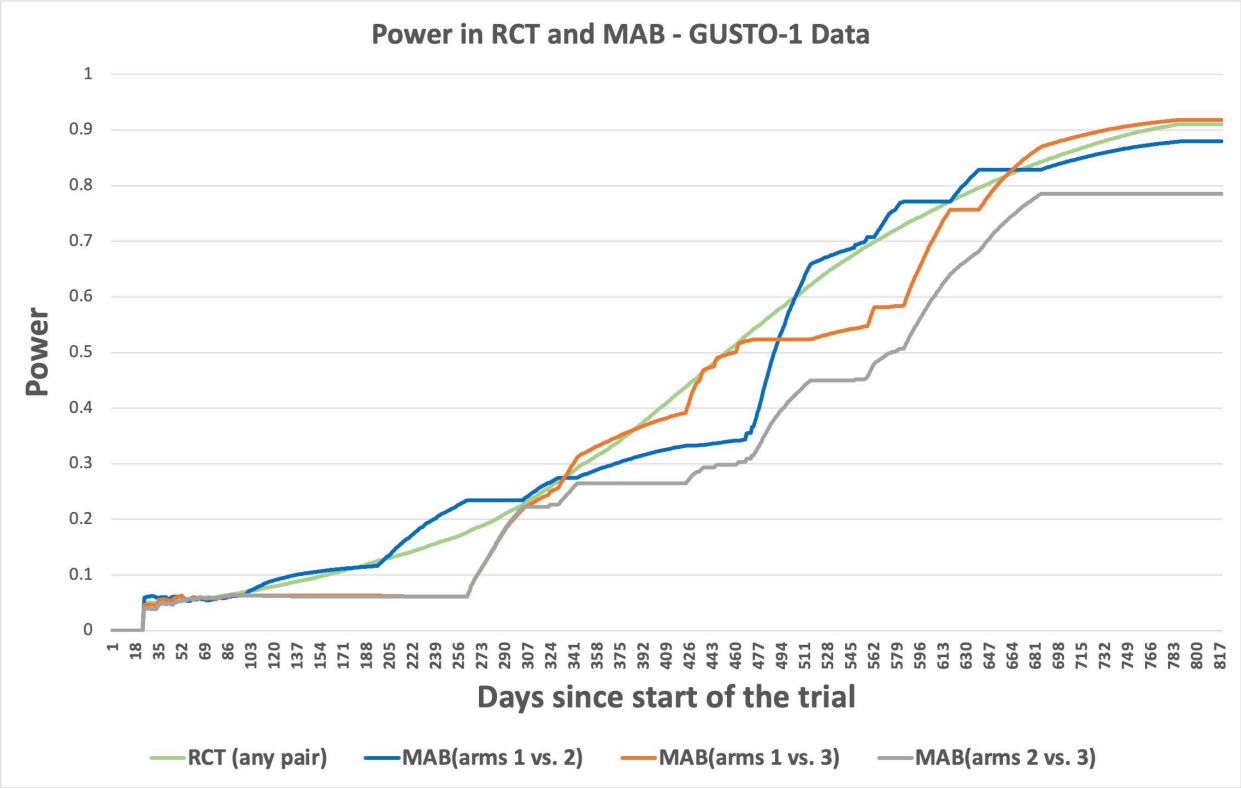


Fig. 3 - Power in RCT and MAB GUSTO-1 simulations

Figure 3 illustrates the ethical decisions that are implied by the use of real-time adaptation. Patient risk is reduced, and patient benefit increased when more sample is allocated to the superior arm. This also implies tighter confidence intervals on the mortality rates of the

superior arm. The implied samples sizes in our simulations suggest slightly more power when comparing the superior arm to one inferior arm and slightly less power to the other inferior arm. However, real-time adaptation implies less statistical power to distinguish between the two inferior arms.

The η-variant of real-time adaptation and the block-based MAB allocate more sample to the superior arm than the RCT, but less sample than real-time adaptation. Relative to real-time (day-to-day) adaptation, these MAB variants provide more power between the superior arm and each inferior arm, and reduce the risk of misidentification (as discussed in an earlier section), but to achieve these gains the MAB variants trade off slightly less patient benefit.

4. Discussion

Our analyses suggest that the use of real-time adaptation in the GUSTO-1 and EUROPA trials would likely have saved lives and avoided cardiovascular events relative to classical RCTs. Real-time adaptive trials enhance the ethical principle of beneficence in the sense of the Belmont report—"maximize possible benefits and minimize possible harm [16].” Real-time adaptive trials also respect persons and justice because *a priori* arm assignment depends upon outcomes not knowable in advance. The algorithm does not depend upon demographic indicators. However, there is an ethical issue because the likelihood of receiving the best treatment changes over time. In real-time adaptive trials, 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.). The use of real-time adaptation implies tighter confidence intervals involving the superior arm, but at the cost of less tight confidence intervals for the inferior arms. For the multi-arm trial,

statistical power for real-time adaptation for pairwise comparisons to the superior arm were comparable to those in the RCT – a sometimes above and sometimes below the RCT in our simulations. Statistical power for comparisons between inferior arms was much below those observed in the RCT. These ethical issues are beyond the scope of the present article. We seek to highlight the issues for the MAB and its variants. The η-variant and the previously proposed block-based MAB provide the trialist with a means to balance the ethical issues and achieve benefits and costs that are intermediate between an RCT and real-time adaptation..

Our analyses highlight when real-time adaptation is best and highlight tradeoffs that must be made. Lives would have been saved in the GUSTO-1 and EUROPA trials, but the number of lives saved was not dramatic because the arms were close in outcomes. In trials where the outcomes are likely to vary more, the number of lives saved could be more dramatic. Even more lives could be saved with stopping rules, but at the cost of substantially less sample for the inferior arm(s) implying larger confidence intervals and less statistical power for those arms.

Our discussion attempted to balance Bayesian and frequentist perspectives. Real-time adaptation is inherently Bayesian because the MAB updates posterior distributions of outcomes and assigns patients to arms by solving a Bayesian optimization problem. However, RCT sample sizes are traditionally based on frequentist power calculations and analyzed in that domain. Because trialists who use RCT are more likely to be familiar with classical statistics, we chose to present results such as confidence intervals, odds ratios, and statistical power within the domain that is used most often to analyze RCT trials. There is nothing incompatible with analyzing the resulting outcomes from a classical perspective. After the data are obtained, the endogeneity of assignments does not affect the calculations. We could have chosen to present the results from a purely Bayesian perspective and used Bayesian concepts such as Bayes factors or Bayesian

confidence intervals to analyze the trial results. The conceptual issues raised in this paper would have been the same.

In the GUSTO-1 and EUROPA trials, our analysis of real-time adaptation and the variants are limited to the starting arms in the RCT, but real-time adaptation (and variants) are not limited to the starting arms. Additional arms can be added during the trial. Arms are eliminated automatically if the MAB no longer allocates sample – such elimination is blind to the researchers (and optimal). We encourage further research to test stopping rules and to test real-time adaptation with the inclusion of arms that were not present in the original planning. When a new arm is introduced after convergence of the original arms, Bayesian priors are necessary for the new arm—priors must be sufficiently optimistic so that the algorithm explores the new arms, but not so optimistic that it over-explores the new arms. Setting such priors is practical because, ethically, the new arm would not have been introduced unless it had a reasonable probability of improving outcomes.

4.1 Study Limitations

Real-time adaptive trials outperform fixed-design RCTs with respect to patient beneficence. We evaluated this hypothesis by *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 simulations assume patient exchangeability. We have no reason to question patient exchangeability in the GUSTO-1 and EUROPA trials. However, when secular time trends independent of the trial are present, patients are not exchangeable and any time trends would need to be modeled [28].

Our data are time-stamped at the daily level, so our MAB assignments and learning occur once a day and are conservative. Within-day data, when available, would allow for even more gains in patient beneficence.

We focused on within-trial optimization. For across-trials multiple-population settings, researchers might explore methods to merge real-time adaptive methods and platform-trials methods. Adaptive platform trials provide an alternative that compares multiple interventions, generates subgroup estimates, and minimizes downtime between trials [26].

Blinding is an important matter in clinical trials. Without blinding, participants in an adaptive trial may be tempted to guess the superior arm. Real-time adaptive trials such as those using MAB assignment can be blinded because assignment is based on pre-defined rules such that neither the experimenter nor the patient knows which patient is assigned to which arm. Careful protocols need to be developed, and independent Data and Safety Monitoring Committees should be used, to check on the adaptive process and identify any undesirable patterns of adverse event occurrence [27]. Finally, we might improve assignments further with the use of biomarkers as surrogate measures of outcomes [8, 12, 26].

4.2 Conclusion

In this article, we explicitly addressed and discussed the trade-off implicit to real-time adaptive trials: real-time adaptive trials increase patient beneficence (e.g., fewer cardiovascular events) by allocating less sample to inferior arms. While real-time adaptation saves lives and reduces confidence intervals for the superior arm, it reduces confidence intervals in pairs of inferior arms. The trialist must also compare the benefits to changes in statistical power, especially among inferior arms.

References

[1] Cheng J, Zhang W, Zhang X, et al. Effect of Angiotensin-Converting Enzyme Inhibitors and Angiotensin II Receptor Blockers on All-Cause Mortality, Cardiovascular Deaths, and Cardiovascular Events in Patients With Diabetes Mellitus: A Meta-analysis. JAMA Intern Med. 2014;174(5):773–785. doi:10.1001/jamainternmed.2014.348

[2] Tai, C., Gan, T., Zou, L. et al. Effect of angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers on cardiovascular events in patients with heart failure: a meta-analysis of randomized controlled trials. BMC Cardiovasc Disord 17, 257 (2017). https://doi.org/10.1186/s12872-017-0686-z

[3] Baldetti L, Melillo F, Moroni F, Gallone G, Pagnesi M, Venuti A, Beneduce A, Calvo F, Gramegna M, Godino C, D'Ascenzo F, De Ferrari GM, Capodanno D, Cappelletti AM. Meta-Analysis Comparing P2Y12 Inhibitors in Acute Coronary Syndrome. Am J Cardiol. 2020 Jun 15;125(12):1815-1822. doi: 10.1016/j.amjcard.2020.03.019. Epub 2020 Apr 2. PMID: 32305225.

[4] Berry DA. A case for Bayesianism in clinical trials. Stat Med. 1993 Aug;12(15-16):1377-93; discussion 1395-1404. doi: 10.1002/sim.4780121504. PMID: 8248653.

[5] Katehakis M and Veinott. A. The multi-armed bandit problem: decomposition and computation, Mathematics of Operations Research, 1987; 12 (2): 262–268, doi:10.1287/moor.12.2.262.

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

[7] Connor, Jason T, Jordan J Elm, Kristine R. Broglio (2013) Bayesian adaptive trials offer advantages in comparative effectiveness trials: an example in status epilepticus. Journal of Clinical Epidemiology 66:S130-S137.

[8] Barnett, Helen Yvette, Sofia S. Villar, Helena Geys, Thomas Jaki (2020) A novel statistical test for treatment differences in clinical trials using a response-adaptive forward-looking Gittins Index Rule. Biometrics 79:86-97.

[9] Chick, Stephen E., Noah Gans, Özge Yapar (2022) Bayesian sequential learning for clinical trials of multiple correlated medical interventions. Management Science 68(7):4919-4938.

[10] Robertson, David S., Kim May Lee, Boryana C. Lopez-Kolkovska, Sofia S. Villar (2023) Response-adaptive randomization in clinical Trials: From myths to practical considerations. Statistical Science, 38(2):185-208.

[11] Villar, Sofia S. James Wason, Jack Bowden (2015) Response-adaptive randomization for multi-arm clinical trials using the forward-looking Gittins Index rule. Biometrics 71:969-978.

[12] Hauser JR, Urban G, Liberali G and Braun M. Website Morphing. Marketing Science. 2009; 28, 2, (March-April), 202-224.

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

[14] FDA (2019) Adaptive designs for clinical trials of drugs and biologics: Guidance for industry. Biostatistics (November)

[15] Pallmann, Philip , Alun W. Bedding, Babak Choodari-Oskooei, Munyaradzi Dimairo, Laura Flight, Lisa V. Hampson, Jane Holmes, Adrian P. Mander, Lang'o Odondi, Matthew R. Sydes, Sofia S. Villar, James M. S. Wason, Christopher J. Weir, Graham M. Wheeler, Christina Yap, Thomas Jaki (2018) Adaptive designs in clinical trials:

why use them, and how to run and report them. MBC Medicine 16(29):1-15.

[16] 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.

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

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

[19] 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

[20] Galbraith R. A note on graphical presentation of estimated odds ratios from several clinical trials. Stat Med. 1988;7: 889–94.

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

[22] Pocock S. Clinical Trials. A Practical Approach. 1983; London: Wiley.

[23] Hadad, Vitor, David A. Hirshberg, Ruohan Zhan, Stefan Wager, Susan Athey (2021) Confidence intervals for policy evaluation in adaptive experiments. PNAS 118(15):1-10.

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

[26] 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.

[27] Villar S, Bowden J and Wason J. Multi-armed bandit models for the optimal design of clinical trials: benefits and challenges. Statistical science: a review" Journal of the Institute of Mathematical Statistics. 2015; 30(2):199–215.

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