

The Winner’s Curse as an e-Process Compensator with Applications to Safe Anytime-Valid Testing for Dose-Ranging Trials

Victor K. de la Peña^{1,2,†} Fangyuan Lin^{1,†} Demissie Alemayehu¹ Victor H. de la Peña¹

¹Department of Statistics, Columbia University ²Inference LLC

[†]These authors contributed equally.

June 2026

Abstract

Phase II dose-ranging trials often report the largest observed dose–control effect while inspecting accumulating data repeatedly. This creates two coupled distortions: selection optimism from choosing the empirical winner, known as the winner’s curse, and Type-I error inflation from multiplicity across doses and interim looks. We develop an anytime-valid procedure for testing whether the best true dose effect exceeds a clinically meaningful margin. The mathematical starting point is a recent selection-premium identity for the running maximum: for dose–control scores, the expected gain from re-selecting the current leader becomes a predictable selection charge. Subtracting this charge gives a residual with nonpositive drift under the composite null; applying a one-sided mixture-exponential construction then yields an e-process and hence an anytime-valid global test. The resulting rule has a transparent ledger form: raw best effect minus selection charge minus monitoring margin, and a “go” decision is made only when the remaining evidence still exceeds the clinical margin. We give plug-in implementations for Gaussian and binary outcomes, prove finite-sample Type-I control and anytime lower confidence bounds for the best dose effect, and illustrate the method through worked examples and simulations.

Keywords: dose-ranging, selection premium, winner’s curse, e-process, confidence sequence, Gaussian endpoint, binary endpoint.

1 Introduction

1.1 The decision problem

Suppose K active doses are randomized against one concurrent control. We use the following convention throughout:

Symbol	Meaning
t	index for sequential information update or analysis look
n_t	cumulative number of evaluable subjects per arm at look t
$X_{t,k}$	newly observed outcome from dose k at update t ; $X_{t,0}$ is control
$\varepsilon_{t,k}$	centered noise, $X_{t,k} - \mathbb{E}(X_{t,k})$

The balanced unit-block theory observes one outcome per arm and control at each update, so $n_t = t$ in that case. $X_{t,k}$ and $X_{t',k'}$ are assumed to be independent whenever $(t, k) \neq (t', k')$. Let

$$\mu_k := \mathbb{E}(X_{t,k}), \quad \mu_0 := \mathbb{E}(X_{t,0}), \quad \delta_k := \mu_k - \mu_0.$$

The best program-level efficacy effect is

$$\delta_* := \max_{1 \leq k \leq K} \delta_k. \quad (1)$$

A protocol specifies a clinically meaningful margin δ_0 and asks

$$H_0(\delta_0) : \delta_* \leq \delta_0 \quad \text{versus} \quad H_1(\delta_0) : \delta_* > \delta_0. \quad (2)$$

This is a global proof-of-concept question: does the tested dose range contain at least one dose whose mean effect clears the clinical requirement?

At look t , define cumulative outcomes

$$U_{t,k} := \sum_{s=1}^t X_{s,k}, \quad C_t := \sum_{s=1}^t X_{s,0},$$

and pairwise sample-mean differences

$$\hat{\delta}_{t,k} := \frac{U_{t,k} - C_t}{n_t}, \quad \hat{\delta}_{\max,t} := \max_k \hat{\delta}_{t,k} = \frac{\max_k U_{t,k} - C_t}{n_t}.$$

The maximum is operationally natural to look at because a global go decision is often driven by whichever pairwise comparison currently looks strongest. It, however, creates two logically separate problems.

1. **Winner's curse/selection optimism.** Even when every $\hat{\delta}_{t,k}$ is unbiased before selection, the largest one is optimistic because it was selected for being large.
2. **Multiplicity over doses and time.** Looking across K doses and repeatedly over time creates multiple opportunities for a false global go decision.

A multiplicity-adjusted test can solve the second problem while leaving the reported largest effect optimistic. On the other hand, subtracting a bias correction does not license repeated peeking. The paper therefore keeps bias correction and Type-I control as separate guarantees, even though the same selection-premium quantity drives both.

1.2 Statistical guarantee

For the present paper,

$$\sup_{\delta_* \leq \delta_0} \mathbb{P}\{\text{a global go occurs at some look } t\} \leq \alpha \quad (3)$$

is the relevant Type-I requirement. It says that, when no tested dose truly clears the clinical margin, the probability of ever declaring that the program contains an efficacious dose is at most α .

Strong familywise error rate (FWER) is a different and stronger armwise requirement: under every configuration of true and false individual hypotheses, the probability of rejecting at least one true named-dose null is controlled. Our procedure makes no individual dose rejection, but authorizes a program-level efficacy conclusion; named-dose localization requires a another layer.

1.3 Related work

Fixed-look many-to-one comparisons. Dunnett's procedure uses the correlation induced by a shared control to provide simultaneous many-to-one inference and strong FWER control at a fixed analysis [Dunnett, 1955]. It is usually more efficient than a Bonferroni analysis that ignores this correlation. This solves dose multiplicity, but it does not by itself license arbitrary repeated looks or explain how optimistic the raw observed winner is.

Pre-planned group-sequential and MAMS designs. Treatment-selection group-sequential designs and generalized Dunnett/MAMS procedures extend shared-control testing to a finite, pre-specified stage structure, including efficacy and futility boundaries and interim treatment selection [Stallard and Todd, 2003, Magirr et al., 2012, Wason and Jaki, 2012, Wason et al., 2017]. These methods can provide strong FWER across arms and stages and may be the preferable primary design when stage times, selection rules, and endpoints can be fully specified in advance. The MAMS software described by Jaki et al. [2019] implements several such designs and can also use a conditional-error calculation for unexpected modifications.

Combination tests and adaptive Dunnett tests. Adaptive combination tests combine valid stagewise p -values from disjoint increments. Adaptive Dunnett procedures use the conditional error of the original many-to-one test, while closure can extend these local tests to strong FWER control for a family of named-dose hypotheses [Bauer and Köhne, 1994, Koenig et al., 2008, Friede and Stallard, 2008, Abery and Todd, 2019]. These methods are designed around stagewise tests and individual hypotheses.

Dose-response modelling. MCP-Mod combines a pre-specified family of dose-response contrasts with model fitting to establish a dose-response signal and estimate a target dose under model uncertainty [Bretz et al., 2005]. It exploits dose ordering and shape; our maximum-based global test deliberately does not. The two can be used in sequence or side by side, but they answer different scientific questions.

Inference after choosing a winner. Winner-inference methods target the effect of the selected named option and correct confidence intervals or point estimates for the selection event [Andrews et al., 2024]. Our inferential target is instead δ_* , the best true effect available in the tested dose range, uniformly over analysis time. It does not identify the best dose or give an exactly unbiased effect for the empirical leader.

Anytime-valid inference and betting. Confidence sequences and e-processes already solve optional stopping in broad settings [Howard et al., 2021, Ramdas et al., 2023], and multi-armed betting tests provide anytime-valid global testing under other observation models [Sandoval et al., 2026]. The new object in this paper is the state-dependent selection premium, which has the same units as the clinical effect and simultaneously serves as (i) an interpretable winner-optimism ledger and (ii) the predictable conditional-drift charge needed before safe exponentiation.

2 The selection premium identity and the equal-mean reference

2.1 The selection premium identity

Let $X_t = (X_{t,1}, \dots, X_{t,K})$ denote the observed dose outcomes at update t and write $\varepsilon_t := X_t - \mathbb{E}(X_t)$ for their centered noise. Let

$$S_{t,k} := \sum_{s=1}^t \varepsilon_{s,k}, \quad M_t := \max_k S_{t,k}.$$

The selection premium function (see [de la Pena et al., 2026] for its properties) is a function of a state and an increment law. Let $\mathcal{L}$ denote the law of one fresh centered null update, and let

$\eta = (\eta_1, \dots, \eta_K) \sim \mathcal{L}$ be an independent copy of that next update. Define

$$\text{sp}_{\mathcal{L}}(u) := \mathbb{E}_{\eta \sim \mathcal{L}} \left[\max_k (u_k + \eta_k) \right] - \max_k u_k, \quad u \in \mathbb{R}^K. \quad (4)$$

The symbol η is not an observed residual and not a second data set. It is the generic fresh noise used to compute the expectation of the next update. If the increments are i.i.d., the same function $\text{sp}_{\mathcal{L}}$ is used at every update.

Let $\mathcal{L}_t$ denote the incremental law for the t -th trial and $A_t = \sum_{s \leq t} a_s$.

If one must commit now, the value is $\max_k u_k$. If one waits for one more update and may then choose the leader, the expected value is the first term in (4). Thus $\text{sp}_{\mathcal{L}}(u)$ is the expected one-step value of retaining the option to switch leaders.

The foundation paper [de la Pena et al., 2026] proves, in the i.i.d. case,

$$\mathbb{E}M_t = \sum_{s=1}^t \mathbb{E}\{\text{sp}_{\mathcal{L}}(S_{s-1})\} = \mathbb{E} \sum_{s=1}^t a_s, \quad (5)$$

with a stopping-time analogue and a heterogeneous-mean extension [de la Pena et al., 2026, Thms. 3.1, 3.5 and 4.3]. It also proves translation invariance $\text{sp}_{\mathcal{L}}$, maximality at the tie configuration, domination of heterogeneous-mean premiums by the equal-mean premium, and decay after a true leader emerges.

2.2 Equal-mean configuration is a calibration reference

For a candidate clinical effect δ , the *equal-margin reference* is the hypothetical model in which every dose has mean $\mu_0 + \delta$, so every pairwise dose–control effect equals δ . The actual trial can (and most likely) have arbitrary heterogeneous effects.

The reference is useful for two operational reasons.

1. **It isolates pure selection.** When dose means are equal, deterministic mean differences cannot explain the maximum. Its excess is entirely the value of being able to choose the noisiest leader.
2. **It gives the largest unknown-gap premium.** [de la Pena et al., 2026] proves that true mean gaps only reduce the premium. When the gaps between the true dose effects are unknown, the equal-mean premium is therefore a conservative envelope for the winner’s curse.

The equal-mean calculation is therefore directly useful even when no investigator believes the true effects are exactly equal: it provides an exact benchmark for pure winner optimism, a worst-case charge when gaps between the true dose effects are unknown, and a state-dependent diagnostic of whether dose competition is still active.

Figure 1 illustrates the behavior of the selection premium function $\text{sp}_{\mathcal{L}}$.

3 From selection premium to e-process

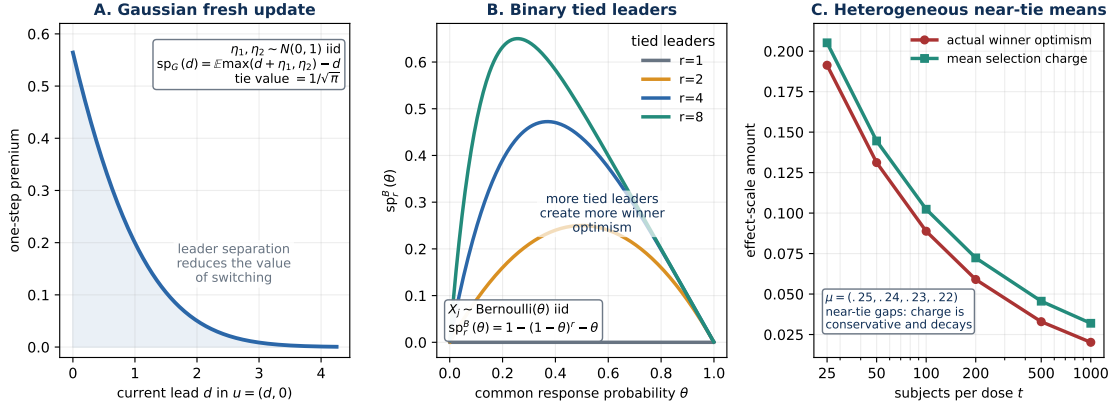
3.1 Cumulative score and corrected residual

For any candidate effect δ , define the cumulative global score at look t by

$$G_t(\delta) := \max_{1 \leq k \leq K} U_{t,k} - C_t - n_t \delta = n_t \{\widehat{\delta}_{\max,t} - \delta\}. \quad (6)$$

A large positive value means the current best pairwise estimate exceeds δ . Due to winner’s curse, it is not safe to monitor G_t directly because the maximum operation creates positive selection drift.

Equal-mean calibration: exact endpoint laws and a useful heterogeneous envelope



The equal-mean law is a calibration reference, not a claim that dose effects are equal. The foundation paper proves its envelope property; this panel illustrates the finite-sample behavior.

Figure 1: Panel A uses $K = 2$, state $u = (d, 0)$, and iid $N(0, 1)$ fresh dose noises. Panel B uses $r \in \{1, 2, 4, 8\}$ tied leaders with iid $\text{Bernoulli}(\theta)$ outcomes; in the binary trial reference, $\theta = p_0 + \delta$. Panel C uses a heterogeneous near-tie Gaussian example with $K = 4$, independent $N(0, 1)$ errors, and means $(.25, .24, .23, .22)$. The red curve is the exact expected fixed-look winner optimism $\mathbb{E}(\max_k \bar{X}_{t,k} - \mu_*)$. The green curve is the mean state-adaptive equal-mean selection charge $\mathbb{E}(A_t/t)$ from 3,500 simulated paths; the shaded band is a Monte Carlo 95% interval. Both curves decay with information.

Let $a_t(\delta)$ denote the selection charge computed from information available immediately before update t ; equivalently, $a_t(\delta) = \text{sp}_{\mathcal{L}_t(\delta)}(\text{current state})$ for the appropriate null increment law and current score state. Define

$$A_t(\delta) := \sum_{s=1}^t a_s(\delta), \quad R_t(\delta) := G_t(\delta) - A_t(\delta), \quad V_t := \sum_{s=1}^t v_s, \quad (7)$$

where both $a_s(\delta)$ and v_s are $\mathcal{F}_{s-1}$ -measurable, hence predictable.

The next theorem is the bridge from the selection-premium identity to an anytime-valid decision rule. At the equal-mean reference, the running maximum has predictable conditional increment

$$\mathbb{E}\{\Delta M_t \mid \mathcal{F}_{t-1}\} = \text{sp}_{\mathcal{L}_t}(\text{current state}) \geq 0.$$

Equivalently, the maximum is a submartingale and the evaluated selection premium is its predictable upward drift. Once this charge is subtracted, the residual has nonpositive conditional drift and can be turned into an e-process.

Theorem 3.1 (selection premium to e-process). *Fix a candidate δ . Suppose that under every distribution in $H_0(\delta)$,*

$$\mathbb{E}\{\Delta G_t(\delta) \mid \mathcal{F}_{t-1}\} \leq a_t(\delta), \quad (8)$$

$$\mathbb{E}\left[\exp\left\{\lambda(\Delta G_t(\delta) - \mathbb{E}[\Delta G_t(\delta) \mid \mathcal{F}_{t-1}])\right\} \mid \mathcal{F}_{t-1}\right] \leq \exp(\lambda^2 v_t/2) \quad (9)$$

for every $\lambda \geq 0$. Then, for each fixed $\lambda \geq 0$,

$$L_t(\lambda; \delta) := \exp\{\lambda R_t(\delta) - \lambda^2 V_t/2\} \quad (10)$$

is a nonnegative supermartingale with $L_0 = 1$.

The second step uses the standard method of mixtures for nonnegative supermartingales [Howard et al., 2020]. For any probability distribution on $\lambda \geq 0$, the mixture of the fixed- λ supermartingales remains a supermartingale.

If $\rho > 0$ is fixed before monitoring and L_t is mixed against the half-normal density

$$\pi_\rho(\lambda) = \sqrt{2\rho/\pi} \exp(-\rho\lambda^2/2) \mathbf{1}\{\lambda \geq 0\},$$

then

$$\mathcal{E}_t(\delta) = \sqrt{\frac{\rho}{V_t + \rho}} \left[1 + \operatorname{erf} \left\{ \frac{R_t(\delta)}{\sqrt{2(V_t + \rho)}} \right\} \right] \exp \left\{ \frac{R_t(\delta)^2}{2(V_t + \rho)} \right\} \quad (11)$$

is an e -process. Consequently,

$$\sup_{P \in H_0(\delta)} \mathbb{P}_P \left\{ \sup_{t \geq 1} \mathcal{E}_t(\delta) \geq \frac{1}{\alpha} \right\} \leq \alpha. \quad (12)$$

Proof. Write

$$b_t := a_t(\delta) - \mathbb{E}\{\Delta G_t(\delta) \mid \mathcal{F}_{t-1}\} \geq 0$$

and $Z_t := \Delta G_t(\delta) - \mathbb{E}\{\Delta G_t(\delta) \mid \mathcal{F}_{t-1}\}$. Then $\Delta R_t = Z_t - b_t$. For $\lambda \geq 0$,

$$\mathbb{E}(e^{\lambda \Delta R_t} \mid \mathcal{F}_{t-1}) = e^{-\lambda b_t} \mathbb{E}(e^{\lambda Z_t} \mid \mathcal{F}_{t-1}) \leq e^{-\lambda b_t} e^{\lambda^2 v_t/2} \leq e^{\lambda^2 v_t/2}.$$

Multiplying by $L_{t-1}(\lambda; \delta) e^{-\lambda^2 v_t/2}$ and conditioning proves $\mathbb{E}\{L_t(\lambda; \delta) \mid \mathcal{F}_{t-1}\} \leq L_{t-1}(\lambda; \delta)$.

Nonnegative mixing preserves the supermartingale property by Tonelli's theorem. Note that A different pre-specified mixing law would give a different valid boundary. To obtain the closed form, substitute (10) and complete the square:

$$\begin{aligned} \mathcal{E}_t(\delta) &= \sqrt{\frac{2\rho}{\pi}} \int_0^\infty \exp \left\{ \lambda R_t - \frac{V_t + \rho}{2} \lambda^2 \right\} d\lambda \\ &= \sqrt{\frac{2\rho}{\pi}} \exp \left\{ \frac{R_t^2}{2(V_t + \rho)} \right\} \int_0^\infty \exp \left[-\frac{V_t + \rho}{2} \left\{ \lambda - \frac{R_t}{V_t + \rho} \right\}^2 \right] d\lambda. \end{aligned}$$

With $y = \sqrt{(V_t + \rho)/2} \{\lambda - R_t/(V_t + \rho)\}$, the lower integration limit is $-R_t/\sqrt{2(V_t + \rho)}$, and the Gaussian tail integral is $\sqrt{\pi/[2(V_t + \rho)]} [1 + \operatorname{erf}\{R_t/\sqrt{2(V_t + \rho)}\}]$. This yields (11). Ville's inequality gives (12). $\square$

3.2 The GO rule

For a variance budget $v \geq 0$, define $q_\alpha(v; \rho)$ as the nonnegative solution z of

$$\sqrt{\frac{\rho}{v + \rho}} \left[1 + \operatorname{erf} \left\{ \frac{z}{\sqrt{2(v + \rho)}} \right\} \right] \exp \left\{ \frac{z^2}{2(v + \rho)} \right\} = \frac{1}{\alpha}. \quad (13)$$

The solution exists and is unique. At $z = 0$, the left side equals $\sqrt{\rho/(v + \rho)} \leq 1 < 1/\alpha$, while it diverges as $z \rightarrow \infty$. Moreover, for $z \geq 0$ its log derivative is

$$\frac{z}{v + \rho} + \frac{\sqrt{2/\pi} \exp\{-z^2/[2(v + \rho)]\}}{\sqrt{v + \rho} [1 + \operatorname{erf}\{z/\sqrt{2(v + \rho)}\}]} > 0,$$

so the left side is strictly increasing.

At the protocol margin δ_0 ,

$$R_t(\delta_0) = n_t\{\widehat{\delta}_{\max,t} - \delta_0\} - A_t(\delta_0).$$

Because (11) is increasing in R_t on the relevant nonnegative side, $\mathcal{E}_t(\delta_0) \geq 1/\alpha$ is equivalent to $R_t(\delta_0) \geq q_\alpha(V_t; \rho)$. Dividing by n_t yields the operational rule

$$\boxed{\widehat{\delta}_{\max,t} \geq \delta_0 + \frac{A_t(\delta_0)}{n_t} + \frac{q_\alpha(V_t; \rho)}{n_t}.} \quad (14)$$

Every term has a separate role: $\widehat{\delta}_{\max,t}$ is the raw best effect, $A_t(\delta_0)/n_t$ is the selection-premium charge, δ_0 is the clinical requirement, and $q_\alpha(V_t; \rho)/n_t$ is the extra margin needed for safe repeated monitoring.

3.3 Choosing the half-normal tuning constant ρ in practice

The parameter ρ is the precision of the half-normal mixing distribution over the betting parameter λ . The choice of ρ does not change Type-I validity but can affect the exact Type-I error.

For $v > 0$, write $c = \rho/v$ and $z = \sqrt{v}x$ in (13). Then

$$q_\alpha(v; \rho) = \sqrt{v} Q_\alpha(\rho/v), \quad (15)$$

where $Q_\alpha(c) := q_\alpha(1; c)$. Hence only the ratio ρ/v matters. If the trial team has a plausible variance budget V_* at which a go decision would be most valuable, a natural targeted choice is

$$\rho = c_\alpha^* V_*, \quad c_\alpha^* := \arg \min_{c>0} Q_\alpha(c). \quad (16)$$

4 Gaussian outcomes

4.1 Model and computable premium

Gaussian outcomes cover continuous Phase II endpoints such as change from baseline in blood pressure, symptom score, or a biomarker. We use balanced unit blocks for clarity, so $n_t = t$ in the displayed formulas.

Assumption 4.1 (Gaussian shared-control blocks). *Across blocks, $(X_{t,1}, \dots, X_{t,K}, X_{t,0})$ are independent. Within a block,*

$$X_{t,1:K} \sim N_K(\mu_{1:K}, \Sigma_D), \quad X_{t,0} \sim N(\mu_0, \sigma_0^2),$$

and the dose vector is independent of control. Before monitoring begins, the analysis specifies deterministic bounds $\bar{\Sigma}_D \succeq \Sigma_D$ and $\bar{\sigma}_0^2 \geq \sigma_0^2$.

$\bar{\Sigma}_D$ and $\bar{\sigma}_0^2$ can be based on external information or an independent variance-estimation stage.

For a centered Gaussian dose-noise vector $\varepsilon^G \sim N_K(0, \bar{\Sigma}_D)$, define the Gaussian selection premium

$$\text{sp}_G(u; \bar{\Sigma}_D) := \mathbb{E} \left[\max_k (u_k + \varepsilon_k^G) \right] - \max_k u_k. \quad (17)$$

Translation invariance makes it computable from the observed cumulative dose vector $U_{t-1} = (U_{t-1,1}, \dots, U_{t-1,K})$ without knowing μ_0 or δ . Set

$$A_t^G := \sum_{s=1}^t \text{sp}_G(U_{s-1}; \bar{\Sigma}_D), \quad V_t^G := t\bar{v}_G, \quad \bar{v}_G := \max_k (\bar{\Sigma}_D)_{kk} + \bar{\sigma}_0^2. \quad (18)$$

The shared control is absent from sp_G because a common subtraction cannot change which dose is largest.

For two doses, if the current score gap is $d = |u_1 - u_2|$ and $\tau^2 = \text{Var}(\varepsilon_1^G - \varepsilon_2^G)$, then

$$\text{sp}_G(u) = \tau\varphi(d/\tau) - d\{1 - \Phi(d/\tau)\}, \quad (19)$$

where φ and Φ are the standard normal density and distribution function. For independent doses with standard deviations σ_k , general K can be computed by the one-dimensional density of a maximum,

$$\mathbb{E} \max_k (u_k + Z_k) = \int_{-\infty}^{\infty} x \sum_{k=1}^K \frac{1}{\sigma_k} \varphi\left(\frac{x - u_k}{\sigma_k}\right) \prod_{j \neq k} \Phi\left(\frac{x - u_j}{\sigma_j}\right) dx. \quad (20)$$

4.2 Gaussian Decision rule

The following proposition is the bias-accounting part of the Gaussian application. It tells a trial team that at the equal-mean reference, A_t^G/n_t is the exact expected upward bias of the reported maximum and acts as an upper bound for the expected selection bias for arbitrary true effect configuration.

Proposition 4.2 (Gaussian equal-margin bias identity). *Suppose Assumption 4.1 holds with exact covariance bounds and $\mu_1 - \mu_0 = \dots = \mu_K - \mu_0 = \delta$. Then for every fixed look t ,*

$$\mathbb{E}\{\hat{\delta}_{\max,t} - \delta\} = \frac{1}{n_t} \mathbb{E}(A_t^G). \quad (21)$$

For every bounded stopping time τ ,

$$\mathbb{E}\{\max_k U_{\tau,k} - C_\tau - n_\tau \delta\} = \mathbb{E}(A_\tau^G). \quad (22)$$

Proof. Under the equal-margin reference, $U_{t,k} - n_t(\mu_0 + \delta)$ are centered Gaussian dose walks. Translation invariance identifies their one-step premium with (17) evaluated at U_{t-1} . Apply the fixed- and stopped-horizon identities from [de la Pena et al., 2026], then subtract the centered control sum. $\square$

The next theorem is the Type-I error control. The equal-margin premium becomes a safe predictable drift charge over the entire composite null, including heterogeneous configurations in which some doses are below the margin. In practice this means the rule does not require knowing which dose is best, or how far the other doses are below it, before monitoring begins.

Theorem 4.3 (Gaussian anytime-valid global test). *Under Assumption 4.1, for every candidate δ and every parameter vector satisfying $\delta_k \leq \delta$ for all k ,*

$$\mathbb{E}\{\Delta G_t(\delta) \mid \mathcal{F}_{t-1}\} \leq \text{sp}_G(U_{t-1}; \bar{\Sigma}_D). \quad (23)$$

Moreover, the centered increment of $G_t(\delta)$ is conditionally sub-Gaussian with variance proxy $\bar{v}_G$. Therefore Theorem 3.1, with $A_t = A_t^G$ and $V_t = V_t^G$, yields an e-process $\mathcal{E}_t^G(\delta)$ and

$$\sup_{\delta_* \leq \delta} \mathbb{P} \left\{ \sup_t \mathcal{E}_t^G(\delta) \geq 1/\alpha \right\} \leq \alpha.$$

Equivalently, the first look satisfying

$$\widehat{\delta}_{\max, t} \geq \delta_0 + \frac{A_t^G}{n_t} + \frac{q_\alpha(t\bar{v}_G; \rho)}{n_t} \quad (24)$$

is a level- α global test of $H_0(\delta_0)$ at arbitrary data-dependent looks.

We omit the proof because the above theorem is a special case of Theorem 3.1.

4.3 Finite selection cost under a unique best Gaussian dose

Proposition 4.4 (Gaussian charge is summable after separation). *Under Assumption 4.1, suppose a unique best dose k^* has $\mu_{k^*} > \max_{j \neq k^*} \mu_j$. Then*

$$A_\infty^G := \sum_{t=1}^{\infty} \text{sp}_G(U_{t-1}; \bar{\Sigma}_D) < \infty \quad \text{almost surely.} \quad (25)$$

Consequently, $A_t^G/n_t \rightarrow 0$ almost surely.

Proof. By the strong law, $U_{t,k^*} - U_{t,j}$ grows linearly and is eventually positive for every $j \neq k^*$. Once k^* is the current leader, let $d_{t,j} = U_{t,k^*} - U_{t,j}$. For $\varepsilon^G \sim N_K(0, \bar{\Sigma}_D)$,

$$\text{sp}_G(U_t) \leq \sum_{j \neq k^*} \mathbb{E}[(\varepsilon_j^G - \varepsilon_{k^*}^G - d_{t,j})_+].$$

If $\tau_j^2 = \text{Var}(Z_j - Z_{k^*})$, then $\mathbb{E}[(N(0, \tau_j^2) - d)_+] \leq \tau_j \varphi(d/\tau_j)$ for $d > 0$. Linear growth of $d_{t,j}$ therefore makes the displayed upper bound summable. $\square$

5 Binary outcomes

5.1 Model and closed-form tie premium

Binary outcomes cover response, remission, and other success–failure endpoints. Here the unit information update contains one new binary outcome from each dose and one from control, so the per-arm sample size is $n_t = t$.

Assumption 5.1 (Binary shared-control blocks). *At update t , $X_{t,k} \in \{0, 1\}$ is observed from dose k and $X_{t,0} \in \{0, 1\}$ from control. Updates are independent. Within an update the $K + 1$ outcomes are independent, with $X_{t,k} \sim \text{Bernoulli}(p_k)$ and $X_{t,0} \sim \text{Bernoulli}(p_0)$.*

At look t , define the cumulative dose and control counts by

$$U_{t,k} := \sum_{s \leq t} X_{s,k}, \quad C_t := \sum_{s \leq t} X_{s,0}.$$

Immediately before update t , let

$$L_t := \{k : U_{t-1,k} = \max_j U_{t-1,j}\}, \quad r_t := |L_t|, \quad (26)$$

where r_t is the number of tied leaders before update t .

Only a current leader can increase the dose maximum in one binary update. Define

$$B_t := \mathbf{1} \left\{ \sum_{k \in L_t} X_{t,k} \geq 1 \right\},$$

the indicator that at least one current leader responds.

For candidate margin $\delta \in [-1, 1]$, the equal-margin reference uses dose response probability θ and control probability $\theta - \delta$. Feasibility requires

$$\theta \in \Theta(\delta) := [\max(0, \delta), \min(1, 1 + \delta)]. \quad (27)$$

The exact one-step selection premium is

$$\text{sp}_r^B(\theta) := \mathbb{P}(B_t = 1) - \theta = 1 - (1 - \theta)^r - \theta. \quad (28)$$

This is the same object as the generic $\text{sp}_\mathcal{L}$ in (4), specialized to a binary state with r tied leaders. When $r = 1$, $\text{sp}_1^B(\theta) = 0$: a unique leader creates no immediate option value. When several leaders are tied, the maximum rises if any of them responds, and the option to switch leaders has positive value.

Because θ is unknown in routine use, define the nuisance-robust premium

$$\overline{\text{sp}}_r^B(\delta) := \sup_{\theta \in \Theta(\delta)} \text{sp}_r^B(\theta), \quad \bar{A}_t^B(\delta) := \sum_{s=1}^t \overline{\text{sp}}_{r_s}^B(\delta), \quad V_t^B := t/2. \quad (29)$$

The binary case is attractive operationally because the one-step charge is closed form. The only state variable needed for the selection premium is r_t , the number of current response-count leaders before the next update. The lemma below removes the remaining nuisance parameter by maximizing over the feasible boundary response probability.

Lemma 5.2 (Closed-form robust binary charge). *For $r = 1$, $\overline{\text{sp}}_1^B(\delta) = 0$. For $r \geq 2$, sp_r^B is strictly concave and has unconstrained maximizer*

$$\theta_r^* = 1 - r^{-1/(r-1)}.$$

Let $\ell(\delta) = \max(0, \delta)$ and $u(\delta) = \min(1, 1 + \delta)$. Then

$$\overline{\text{sp}}_r^B(\delta) = \text{sp}_r^B(\min\{u(\delta), \max\{\ell(\delta), \theta_r^*\}\}). \quad (30)$$

Moreover $h_r(\delta) := \delta + \overline{\text{sp}}_r^B(\delta)$ is nondecreasing on $[-1, 1]$.

Proof. Differentiate: $\{\text{sp}_r^B\}'(\theta) = r(1 - \theta)^{r-1} - 1$ and $\{\text{sp}_r^B\}''(\theta) = -r(r-1)(1 - \theta)^{r-2} < 0$. Therefore the constrained maximizer is the projection of θ_r^* onto $\Theta(\delta)$, giving (30). For monotonicity, first note that $h_1(\delta) = \delta$. For $r \geq 2$, consider the feasible interval separately. If $\delta \geq 0$, then $\Theta(\delta) = [\delta, 1]$. When $\delta \leq \theta_r^*$, the optimizer is interior and $h_r'(\delta) = 1$; when $\delta > \theta_r^*$, the optimizer is $\theta = \delta$ and

$$h_r(\delta) = 1 - (1 - \delta)^r,$$

whose derivative is $r(1 - \delta)^{r-1} \geq 0$. If $\delta < 0$, then $\Theta(\delta) = [0, 1 + \delta]$. When $1 + \delta \geq \theta_r^*$, again $h_r'(\delta) = 1$; otherwise the optimizer is $\theta = 1 + \delta$ and

$$h_r(\delta) = -(-\delta)^r,$$

whose derivative is $r(-\delta)^{r-1} \geq 0$. Continuity at the case boundaries completes the proof. $\square$

5.2 Exact reference bias and anytime-valid test

The binary bias identity has the same interpretation as the Gaussian one, but it is even more transparent: ties among current response-count leaders create the winner's curse, and a unique leader creates no one-step binary selection charge.

Proposition 5.3 (Binary equal-margin bias identity). *Suppose $p_1 = \dots = p_K = \theta$ and $p_0 = \theta - \delta$. Define $A_t^B(\theta) := \sum_{s=1}^t \text{sp}_{r_s}^B(\theta)$. Then*

$$\mathbb{E}\{\widehat{\delta}_{\max,t} - \delta\} = \frac{1}{n_t} \mathbb{E}\{A_t^B(\theta)\} \quad (31)$$

for every fixed look t , and

$$\mathbb{E}\{\max_k U_{\tau,k} - C_\tau - n_\tau \delta\} = \mathbb{E}\{A_\tau^B(\theta)\} \quad (32)$$

for bounded stopping times τ ; in the unit-block model $n_\tau = \tau$.

Proof. See de la Pena et al. [2026]. □

Similarly, we have the following decision rule.

Theorem 5.4 (Binary anytime-valid global test). *Under Assumption 5.1, for every candidate $\delta \in [-1, 1]$ and every configuration with $p_k - p_0 \leq \delta$ for all k ,*

$$\mathbb{E}\{\Delta G_t(\delta) \mid \mathcal{F}_{t-1}\} \leq \overline{\text{sp}}_{r_t}^B(\delta). \quad (33)$$

The centered increment is conditionally sub-Gaussian with variance proxy $1/2$. Hence Theorem 3.1, with $A_t = \bar{A}_t^B(\delta)$ and $V_t = t/2$, gives an e -process $\mathcal{E}_t^B(\delta)$ and an anytime-valid level- α global test. At the clinical margin,

$$\boxed{\widehat{\delta}_{\max,t} - \geq \delta_0 + \frac{\bar{A}_t^B(\delta_0)}{n_t} + \frac{q_\alpha(t/2; \rho)}{n_t}, \quad n_t = t.} \quad (34)$$

Proof. Given $\mathcal{F}_{t-1}$, the score increment is $\Delta G_t(\delta) = B_t - X_{t,0} - \delta$. If $p_0 + \delta \leq 1$, set $\theta = p_0 + \delta \in \Theta(\delta)$. Since $p_k \leq p_0 + \delta$ under the null,

$$\begin{aligned} \mathbb{E}\{\Delta G_t(\delta) \mid \mathcal{F}_{t-1}\} &= 1 - \prod_{k \in L_t} (1 - p_k) - p_0 - \delta \\ &\leq 1 - (1 - p_0 - \delta)^{r_t} - (p_0 + \delta) \\ &= \text{sp}_{r_t}^B(\theta) \leq \overline{\text{sp}}_{r_t}^B(\delta). \end{aligned}$$

If $p_0 + \delta > 1$, the conditional mean is at most $1 - p_0 - \delta < 0 \leq \overline{\text{sp}}_{r_t}^B(\delta)$. The centered variables $B_t - \mathbb{E}(B_t \mid \mathcal{F}_{t-1})$ and $X_{t,0} - p_0$ are independent and each is $1/4$ -sub-Gaussian by Hoeffding's lemma. Their difference has proxy $1/2$. Apply Theorem 3.1. □

Proposition 5.5 (Finite binary selection charge under a unique best dose). *If one binary dose has strictly larger response probability than every other dose, then for every fixed candidate δ ,*

$$\bar{A}_\infty^B(\delta) < \infty \quad \text{almost surely.}$$

Proof. The strong law implies that the cumulative response count of the unique best dose eventually exceeds every other dose count and remains ahead. Thereafter $r_t = 1$ for t large enough, and $\overline{\text{sp}}_1^B(\delta) = 0$. □

5.3 Exact correction at the equal-effect reference

Propositions 4.2 and 5.3 say that, at a specified equal-margin reference, the accumulated premium divided by n_t is the exact expected optimism of the raw largest pairwise estimate. This is useful as a design-stage calculation of expected winner optimism when dose effects are close and also as a conservative correction for the true best program effect.

For the actual trial with heterogeneous dose-effect configuration, define the endpoint-specific robust charge $A_t(\delta)$ as in Sections 4 or 5. At the unknown true best effect δ_* , every dose satisfies the null inequality $\delta_k \leq \delta_*$. The drift calculations therefore imply the following.

Proposition 5.6 (Heterogeneous best-effect bias envelope). *For either endpoint model and every fixed look t ,*

$$\mathbb{E} \left[n_t \{ \hat{\delta}_{\max, t} - \delta_* \} \right] \leq \mathbb{E} \{ A_t(\delta_*) \}. \quad (35)$$

The same inequality holds at bounded stopping times on the cumulative score scale.

See [de la Pena et al., 2026] for proof.

6 Anytime lower confidence sequence for the best effect

For each candidate δ , compute the corresponding e-process $\mathcal{E}_t(\delta)$. The rejected candidate values are those for which the data provide anytime-valid evidence that $\delta_* > \delta$.

Testing one candidate margin is useful for a go/no-go decision. Inverting all candidate margins gives a more interpretable object: an anytime lower floor for the best effect available in the tested dose range. The theorem below formalizes that inversion.

Theorem 6.1 (Anytime lower confidence sequence). *Assume either the Gaussian or binary model. Define*

$$\underline{\delta}_t := \sup \{ \delta \in \mathcal{D} : \mathcal{E}_t(\delta) \geq 1/\alpha \}, \quad (36)$$

where $\mathcal{D} = \mathbb{R}$ for Gaussian outcomes and $\mathcal{D} = [-1, 1]$ for the displayed binary risk-difference analysis, with the supremum of the empty set taken as the lower endpoint of $\mathcal{D}$. Then

$$\mathbb{P} \{ \underline{\delta}_t \leq \delta_* \text{ for every } t \geq 1 \} \geq 1 - \alpha. \quad (37)$$

For Gaussian outcomes, $\mathcal{E}_t(\delta)$ is nonincreasing in δ . For binary outcomes, the same is true because $\delta + \overline{\text{sp}}_r^B(\delta)$ is nondecreasing. Thus (36) is obtained by one-dimensional bisection.

Proof. If $\underline{\delta}_t > \delta_*$ for some t , then the level- α test has rejected the true null $H_0(\delta_*)$ at some time. The probability of this event is at most α by Theorems 4.3 or 5.4. Monotonicity follows directly for Gaussian outcomes because the charge is independent of δ . For binary outcomes, $n_t \delta + \bar{A}_t^B(\delta) = \sum_{s=1}^t \{ \delta + \overline{\text{sp}}_{r_s}^B(\delta) \}$ is nondecreasing by Lemma 5.2. $\square$

The global go rule at the clinical margin is simply

$$\underline{\delta}_t > \delta_0.$$

The finite-sample rule is conservative early because it pays for both selection and repeated monitoring. The asymptotic result below explains why this cost is transient when one dose is truly best: the selection ledger per subject vanishes, the monitoring boundary per subject vanishes, and the lower floor approaches the true best effect.

Corollary 6.2 (Consistency and evidence growth). *Suppose a unique best dose exists and $\eta := \delta_* - \delta_0 > 0$. Under either endpoint model,*

$$\underline{\delta}_t \rightarrow \delta_* \quad \text{almost surely,}$$

and the global test crosses in finite time almost surely. In the Gaussian model,

$$\frac{1}{n_t} \log \mathcal{E}_t^G(\delta_0) \rightarrow \frac{\eta^2}{2\bar{v}_G};$$

in the binary model,

$$\frac{1}{n_t} \log \mathcal{E}_t^B(\delta_0) \rightarrow \eta^2.$$

Proof. The strong law gives $\hat{\delta}_{\max,t} \rightarrow \delta_*$. Propositions 4.4 and 5.5 give $A_t/n_t \rightarrow 0$. The half-normal mixture boundary satisfies $q_\alpha(V_t; \rho) = O\{\sqrt{V_t \log V_t}\}$, hence $q_\alpha(V_t; \rho)/n_t \rightarrow 0$. The lower sequence therefore converges to δ_* . Substituting $R_t(\delta_0) = n_t\eta + o(n_t)$ into (11) gives the stated log-growth limits. $\square$

7 A binary example

Suppose the control response probability is 0.30 and the six dose response probabilities are

$$(.32, .38, .44, .50, .47, .42),$$

so the true dose-control effects are (.02, .08, .14, .20, .17, .12). The clinical margin is $\delta_0 = .05$, the nominal level is $\alpha = .05$, and $\rho = 13.41$. One outcome is observed from each dose and the control at every update, so in this balanced illustration $t = n_t$ is also the number of subjects per arm.

Before observing update s , the software counts the number r_s of doses tied for the cumulative response-count lead and adds the robust binary premium $\bar{\text{sp}}_{r_s}^B(\delta_0)$ to the running ledger

$$\bar{A}_t^B(\delta_0) = \sum_{s=1}^t \bar{\text{sp}}_{r_s}^B(\delta_0).$$

After the outcomes at look t arrive, it computes the raw best risk difference, the selection charge $\bar{A}_t^B/t$, the monitoring margin $q_{.05}(t/2; 13.41)/t$, and their difference. See table 1 for the quantities a data-monitoring report needs.

Table 1: One representative path. The first global go occurs at $t = 216$ subjects per arm. Percentages are risk-difference points.

$t = n_t$	Current leader count versus control	Raw best	Selection charge	Monitoring margin	Anytime lower floor	Decision
50	Dose 4: 29/50 vs 18/50	22.0%	5.1%	29.5%	−12.6%	Continue
100	Dose 4: 54/100 vs 33/100	21.0%	6.3%	19.8%	−5.1%	Continue
150	Doses 3 and 4: 77/150 vs 48/150	19.3%	6.5%	16.0%	−3.2%	Continue
200	Dose 4: 109/200 vs 64/200	22.5%	5.0%	13.8%	3.7%	Continue
216	Dose 4: 116/216 vs 66/216	23.15%	4.64%	13.32%	5.19%	GO

At the crossing look,

$$\underbrace{0.2315}_{\text{raw best risk difference}} - \underbrace{0.0464}_{\text{selection charge}} - \underbrace{0.1332}_{\text{monitoring margin}} = \underbrace{0.0519}_{\text{anytime lower floor}} > \underbrace{0.0500}_{\text{clinical margin}}. \quad (38)$$

The trial therefore stops for a global proof-of-concept conclusion. Up to the crossing, 179 pre-update states had one leader, 34 had a two-way tie, two had a four-way tie, and the initial state had a six-way tie. These states produce the cumulative charge $\bar{A}_{216}^B = 10.027$. The practical message is simpler than the arithmetic: early dose competition is charged, while a persistent unique leader adds no further binary tie charge.

8 Simulation

At update t , one new binary outcome is generated in each of the six dose groups and in the common control group:

$$C_t \sim \text{Bernoulli}(p_0), \quad X_{t,k} \sim \text{Bernoulli}(p_k), \quad k = 1, \dots, 6, \quad (39)$$

independently across groups and updates, with $p_0 = .30$. The clinical margin is $\delta_0 = .05$, the anytime level is $\alpha = .05$, and the pre-specified mixture tuning parameter is $\rho = 13.41$.

Table 2 shows three the models used in the simulation.

Table 2: Three models with different levels of effect signal. The control probability is $p_0 = .30$ in every row. The dose effects are $\delta_k = p_k - p_0$, $\delta_* = \max_k \delta_k$, and $\eta = \delta_* - \delta_0$ is the amount by which the best true effect exceeds the clinical margin. Each row uses 12,000 independent trials through 500 subjects per arm.

Scenario	Dose response probabilities ($p_1, \dots, p_6$)	True risk differences ($\delta_1, \dots, \delta_6$)	δ_*	η
Near threshold	(.31, .35, .38, .40, .39, .36)	(.01, .05, .08, .10, .09, .06)	.10	.05
Moderate	(.32, .37, .41, .45, .43, .39)	(.02, .07, .11, .15, .13, .09)	.15	.10
Strong	(.32, .38, .44, .50, .47, .42)	(.02, .08, .14, .20, .17, .12)	.20	.15

Immediately before update t , the rule counts the number r_t of dose groups tied for the current cumulative response-count lead and adds the nuisance-robust binary premium $\overline{\text{sp}}_{r_t}^B(\delta_0)$ to the ledger. After the new outcomes are observed, it declares the first global GO when

$$\widehat{\delta}_{\max, t} - \frac{\bar{A}_t^B(\delta_0)}{t} - \frac{q_{.05}(t/2; 13.41)}{t} \geq .05. \quad (40)$$

The cap of 500 per arm is a simulation horizon, not a proposed sample size for Phase II trial.

8.1 Finite-horizon power of the anytime rule

For an anytime rule, “power” must specify a horizon. We report

$$\text{Power}_T(p) := \mathbb{P}_p\{\tau \leq T\},$$

where τ is the first global GO time and T is the maximum subjects per arm. This is the probability that the trial reaches a valid global GO by the operational cap. Table 3 uses the same six-dose shared-control binary design, continuous monitoring, $\alpha = .05$, $\rho = 13.41$, and $T = 500$.

Table 3: Finite-horizon power of the proposed anytime selection-premium rule. Each alternative uses 20,000 trials and the same models as Table 2.

Scenario	δ_*	$\delta_* - .05$	GO by 100	GO by 200	GO by 500	50% GO time
Near threshold	.10	.05	1.7%	4.3%	11.0%	Not reached
Moderate	.15	.10	6.4%	19.0%	56.5%	440 per arm
Strong	.20	.15	18.7%	50.6%	94.6%	198 per arm

As the best effect moves farther above the margin, both the GO probability and the time to GO improve. This is the desired expected behavior: larger true gaps create persistent leaders, which reduce future selection charges and make the corrected evidence grow faster.

8.2 Comparison to other procedures

We used the same six-dose shared-control binary model as the main experiment. Under the all-boundary null,

$$C_t \sim \text{Bernoulli}(.30), \quad X_{t,k} \sim \text{Bernoulli}(.35), \quad k = 1, \dots, 6,$$

so every true dose-control risk difference equals the clinical margin .05. See Table 4 for the Type-1 error behavior for different methods. The first two rows are anytime-valid procedures and may stop at any synchronized update. The next two rows show planned Dunnett-style analyses used only at the looks for which their thresholds were calibrated. These rows are valid for their stated schedules, but they are not anytime-valid. The red rows then show what happens when the same planned thresholds are reused at additional, unplanned looks. Under the all-boundary null, this reuse changes the statistic being monitored from a maximum over doses at planned times into a maximum over doses and many extra times, and the false-GO probability rises accordingly. The final row shows the necessity of the selection charge.

Table 4: Simulation demonstration that planned Dunnett-style thresholds are not anytime-valid. All rows use the all-boundary null above, with 60,000 independent trials through 500 subjects per arm. Rows labelled “misuse” intentionally apply those planned thresholds at unplanned looks.

Analysis rule	When may it stop?	Anytime-valid?	False GO by 500
Proposed selection-premium rule	Any synchronized update	Yes	1.35% (809/60,000)
Armwise union e-process	Any synchronized update	Yes	0.80% (479/60,000)
Fixed-final many-to-one max-z test	Only final look $t = 500$	No	4.78% (2,870/60,000)
Planned five-look many-to-one max-z test	Planned looks only	No	4.53% (2,720/60,000)
Misuse: five-look boundary checked every 10 subjects from $t = 100$	Extra unplanned looks	No	7.77% (4,664/60,000)
Misuse: five-look boundary checked after every update from $t = 100$	Continuous extra looks	No	9.39% (5,636/60,000)
Misuse: Fixed-final max-z cutoff checked every 10 subjects from $t = 100$	Extra unplanned looks	No	20.91% (12,546/60,000)
Misuse: Fixed-final max-z cutoff checked after every update from $t = 100$	Continuous extra looks	No	24.27% (14,562/60,000)
Misuse: Same anytime boundary but no selection charge	Any update	No	7.26% (4,358/60,000)

Table 5 reports performance for the valid uses of the same comparators. The fixed-final method cannot stop before 500 per arm, and the five-look method can stop only at the planned stages.

Table 5: Stopping behavior of valid uses of the methods. Each cell reports “GO by 500; mean stop-or-cap enrollment per arm.”

Analysis rule	Near threshold $\delta_* = .10$	Moderate $\delta_* = .15$	Strong $\delta_* = .20$
Proposed selection-premium rule	11.5%; 472.7	57.1%; 369.5	94.8%; 223.2
Armwise union e-process	10.0%; 478.9	60.3%; 369.9	97.0%; 217.7
Fixed-final many-to-one max-z test	38.5%; 500.0	91.0%; 500.0	99.9%; 500.0
Five-look many-to-one max-z test	29.3%; 444.2	84.4%; 308.0	99.6%; 190.3

The previous tables describe finite-horizon operating characteristics through 500 subjects per arm. Figure 2 instead checks the long-run mechanism behind Corollary 6.2. Under the strong profile, one dose has the unique largest response probability, so eventually the empirical leader separates from the other doses. Once this happens, the binary tie premium is usually zero, because $r_t = 1$. The figure tracks whether the three asymptotic claims become visible numerically: the selection charge per subject should vanish, the anytime lower floor should approach the true best effect, and the normalized log e-value should approach its theoretical growth rate.

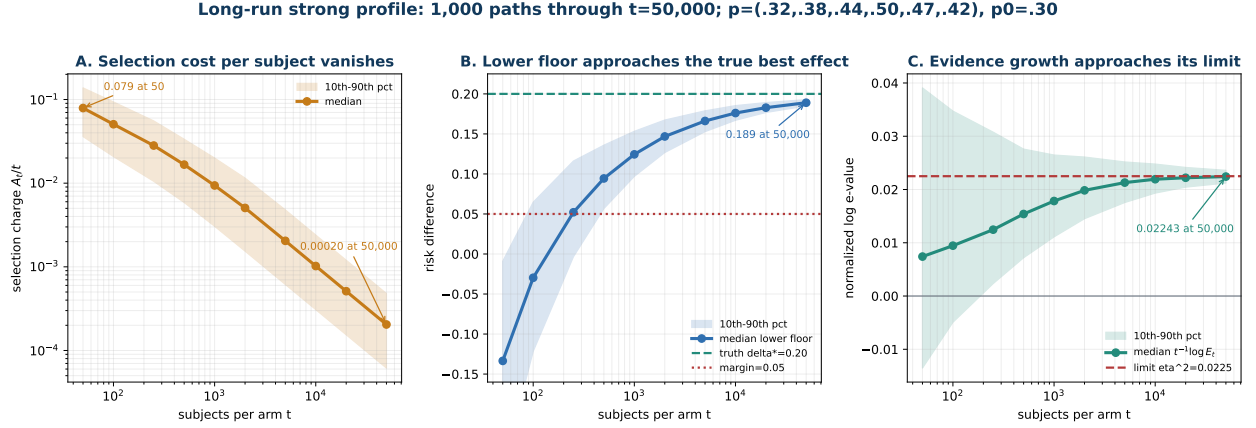


Figure 2: The exact strong-profile probabilities are $(.32, .38, .44, .50, .47, .42)$ with control probability $.30$. Over 1,000 paths through 50,000 subjects per arm, Panel A shows the selection charge per subject becoming negligible, Panel B shows the median anytime lower confidence floor rising toward the true best effect $.20$, and Panel C shows the median normalized log e-value approaching the theoretical evidence-growth limit $(.20 - .05)^2 = .0225$. Shaded bands are the 10th–90th percentiles.

9 Limitations and extensions

Balanced information blocks. The present theorems use synchronized blocks with one observation per active dose and control. Unequal randomization, delayed outcomes, covariate adjustment, arm addition, and arm dropping require an information-time or weighted-score extension.

Gaussian scale specification. The Gaussian theorem requires a pre-specified covariance or a deterministic upper bound. Re-estimating variance from the same unblinded efficacy stream and inserting it without additional protection is not justified by the current proof. Independent or blinded variance adaptation is a natural extension.

Binary nuisance robustness. Maximizing over the unknown boundary rate is safe but can be conservative, especially near extreme response probabilities.

Global rather than named-dose inference. The lower confidence sequence concerns δ_* . It does not establish that the empirical leader is the true best dose or quantify its selected-arm effect. A separate localization procedure remains necessary.

No universal dominance. The method can be more or less efficient than existing procedures depending on K , signal strength, score separation, etc. The benefit of the proposed procedure is interpretable, state-adaptive selection accounting under flexible monitoring.

10 Conclusion

This paper shows how the selection-premium identity of [de la Pena et al. \[2026\]](#) can be turned into a practical anytime-valid procedure for dose-ranging trials. The key observation is that the maximum over dose-control contrasts has a predictable upward drift: before the next information block is observed, retaining the option to re-select the empirical leader has positive expected value. This one-step value is exactly the selection premium. At an equal-effect reference it gives the exact expected winner optimism, and under the composite null it provides a conservative drift charge for the selected maximum.

Subtracting this predictable charge separates the two sources of distortion in a dose-ranging analysis. The cumulative premium A_t/n_t accounts for selection optimism from reporting the current empirical winner, while the mixture boundary $q_\alpha(V_t; \rho)/n_t$ accounts for repeated monitoring over time. The resulting decision rule has the effect-scale form

$$\hat{\delta}_{\max,t} \geq \delta_0 + \frac{A_t(\delta_0)}{n_t} + \frac{q_\alpha(V_t; \rho)}{n_t}.$$

Thus a global go decision is made only when the raw best observed effect clears the clinical margin after paying both the selection charge and the anytime monitoring margin.

For Gaussian shared-control outcomes, the selection charge is computed from the observed cumulative dose sums and a pre-specified covariance bound. For binary outcomes, the charge has a closed form depending on the number of current tied leaders; a unique leader contributes no new binary selection premium, while early ties produce a cumulative ledger that is later diluted on the effect scale. In both cases, the corrected residual yields an e-process and therefore controls the probability of ever making a false global go decision. Inverting the candidate-margin tests gives an anytime lower confidence sequence for δ_* , the best true effect among the tested doses.

The simulations support the theoretical message. Removing the selection charge while keeping an anytime boundary can inflate false global go rates, showing that optional-stopping protection alone is not enough for a repeatedly selected maximum. Planned Dunnett or MAMS-style thresholds behave correctly when used according to their planned schedules, but can lose their calibration when reused at unplanned additional looks. The proposed procedure is designed for the setting in which a trial team wants flexible monitoring together with an auditable decomposition of the current evidence into raw signal, selection opportunity, monitoring uncertainty, and clinical margin.

The scope is deliberately limited. The method gives a program-level proof-of-concept statement: it tests whether at least one studied dose exceeds the clinical margin and provides a lower confidence sequence for the best effect in the tested range. It does not identify the true best dose, give selected-arm inference for the empirical leader, or replace dose-response modelling, safety assessment,

pharmacokinetic analysis, or benefit-risk evaluation. Its value is therefore not universal power dominance over existing procedures, but a mathematically explicit and operationally transparent way to make winner-bias-aware, anytime-valid global decisions in exploratory dose-ranging trials.

References

- Bauer, P. and Köhne, K. (1994). Evaluation of experiments with adaptive interim analyses. *Biometrics*, 50(4), 1029–1041.
- Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press.
- U.S. Food and Drug Administration (2019). *Adaptive Designs for Clinical Trials of Drugs and Biologics: Guidance for Industry*.
- U.S. Food and Drug Administration (2022). *Multiple Endpoints in Clinical Trials: Guidance for Industry*.
- U.S. Food and Drug Administration (2024). *Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases: Guidance for Industry*.
- Jaki, T., Pallmann, P., and Magirr, D. (2019). The R package MAMS for designing multi-arm multi-stage clinical trials. *Journal of Statistical Software*, 88(4), 1–25.
- Friede, T. and Stallard, N. (2008). A comparison of methods for adaptive treatment selection. *Biometrical Journal*, 50(5), 767–781.
- Koenig, F., Brannath, W., Bretz, F., and Posch, M. (2008). Adaptive Dunnett tests for treatment selection. *Statistics in Medicine*, 27(10), 1612–1625.
- Magirr, D., Jaki, T., and Whitehead, J. (2012). A generalized Dunnett test for multi-arm multi-stage clinical studies with treatment selection. *Biometrika*, 99(2), 494–501.
- Ramdas, A., Grünwald, P., Vovk, V., and Shafer, G. (2023). Game-theoretic statistics and safe anytime-valid inference. *Statistical Science*, 38(4), 576–601.
- Stallard, N. and Todd, S. (2003). Sequential designs for Phase III clinical trials incorporating treatment selection. *Statistics in Medicine*, 22(5), 689–703.
- Wason, J. M. S. and Jaki, T. (2012). Optimal design of multi-arm multi-stage trials. *Statistics in Medicine*, 31(30), 4269–4279.
- Sandoval, R. J., Waudby-Smith, I., and Jordan, M. I. (2026). Multi-armed sequential hypothesis testing by betting. *arXiv:2603.17925*.

References

- Julia E Abery and Susan Todd. Comparing the mams framework with the combination method in multi-arm adaptive trials with binary outcomes. *Statistical Methods in Medical Research*, 28(6): 1716–1730, 2019.

- Isaiah Andrews, Toru Kitagawa, and Adam McCloskey. Inference on winners. *The Quarterly Journal of Economics*, 139(1):305–358, 2024.
- Peter Bauer and Karl Köhne. Evaluation of experiments with adaptive interim analyses. *Biometrics*, pages 1029–1041, 1994.
- Frank Bretz, José C Pinheiro, and Michael Branson. Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*, 61(3):738–748, 2005.
- Victor H de la Pena, Fangyuan Lin, and Victor K de la Pena. A selection premium decomposition for the expected maximum of random walks. *arXiv preprint arXiv:2602.19481*, 2026.
- Charles W Dunnett. A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association*, 50(272):1096–1121, 1955.
- Tim Friede and Nigel Stallard. A comparison of methods for adaptive treatment selection. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 50(5):767–781, 2008.
- Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform chernoff bounds via nonnegative supermartingales. 2020.
- Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, non-parametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021.
- Franz Koenig, Werner Brannath, Frank Bretz, and Martin Posch. Adaptive dunnett tests for treatment selection. *Statistics in medicine*, 27(10):1612–1625, 2008.
- Dominic Magirr, Thomas Jaki, and John Whitehead. A generalized dunnett test for multi-arm multi-stage clinical studies with treatment selection. *Biometrika*, 99(2):494–501, 2012.
- Aaditya Ramdas, Peter Grünwald, Vladimir Vovk, and Glenn Shafer. Game-theoretic statistics and safe anytime-valid inference. *Statistical Science*, 38(4):576–601, 2023.
- Ricardo J Sandoval, Ian Waudby-Smith, and Michael I Jordan. Multi-armed sequential hypothesis testing by betting. *arXiv preprint arXiv:2603.17925*, 2026.
- Nigel Stallard and Susan Todd. Sequential designs for phase iii clinical trials incorporating treatment selection. *Statistics in medicine*, 22(5):689–703, 2003.
- James Wason, Nigel Stallard, Jack Bowden, and Christopher Jennison. A multi-stage drop-the-losers design for multi-arm clinical trials. *Statistical methods in medical research*, 26(1):508–524, 2017.
- James MS Wason and Thomas Jaki. Optimal design of multi-arm multi-stage trials. *Statistics in medicine*, 31(30):4269–4279, 2012.