

Endpoint-Calibrated Safety Margins for Lifecycle Control of Model-Informed Precision-Dosing Algorithms

Redouane Boudjemaa

`boudjemaa_redouane@univ-blida.dz`

University Saad Dahlab

Mohamed Boukhari

University Saad Dahlab

Research Article

Keywords: model-informed precision dosing, endpoint calibration, safety margin, competing risks, algorithm updating, exact testing

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10949627/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Endpoint-Calibrated Safety Margins for Lifecycle Control of Model-Informed Precision-Dosing Algorithms

Redouane Boudjemaa Mohamed Boukhari

LAMDA-RO, University Saad Dahlab Blida 1, Blida, Algeria

Corresponding author: Redouane Boudjemaa; boudjemaa_redouane@univ-blida.dz

Mohamed Boukhari: mohamed_boukhari@univ-blida.dz

Abstract

Model-informed precision-dosing algorithms are periodically refitted, whereas their safety margins are usually expressed in units of a version-specific uncertainty scale. The same numerical margin can therefore produce different doses and endpoint risks after an update. We develop an endpoint-calibrated margin whose meaning is preserved across accepted algorithm versions. First, we prove that reciprocal rescaling of the reported uncertainty and margin leaves every dose unchanged, establishing that the numerical margin is not identifiable from the policy alone. We then define common toxicity–progression risk envelopes and show that exact endpoint tests with lifecycle error spending give each accepted version–margin pair the same finite-sample risk-box interpretation. For initial selection, a fixed-sequence intersection–union procedure applied to randomized competing endpoints controls the probability of selecting a policy that violates either population-risk ceiling. A joint dominance condition transfers this certificate without new endpoints; all other refits require two exact tests at the deployed margin. In an impulsive pharmacokinetic–pharmacodynamic simulation, the matched generator at $z = 4$ had toxicity and progression risks of 7.49% and 20.82%, and 82.80% of 5,000 randomized blocks of size 750 certified both endpoints. A scale-collapsed refit increased toxicity risk to 10.11% and passed the stricter cross-version calibration gate in only 4.04% of repeated blocks. No realistic refit examined in the case study satisfied the outcome-free joint transfer condition, making single-anchor recertification the operative update path. These results establish a finite-sample route from version-specific model uncertainty to a reproducible endpoint meaning for the deployed margin.

Keywords: model-informed precision dosing; endpoint calibration; safety margin; competing risks; algorithm updating; exact testing.

Introduction

Model-informed precision dosing (MIPD) combines a pharmacokinetic– pharmacodynamic model, sequential patient data, and a constrained dose rule. Its clinical behavior is therefore a property of the complete deployed algorithm, not of the pharmacokinetic model in isolation. This distinction becomes critical when the population model is refitted, the estimator or uncertainty calculation changes, or the dose solver is replaced. Existing MIPD and model-informed drug-development guidance emphasizes validation relative to the question of interest, robustness, applicability, and prospective evidence [10, 11, 23]. A complementary statistical question remains: what endpoint claim attached to a dosing rule survives a change in the algorithm that generated it?

Consider a policy generator $\mathcal{G}$ that maps patient history x to an estimate $\hat{\theta}_{\mathcal{G}}(x)$, a reported scale $s_{\mathcal{G}}(x) > 0$, and a dose. At margin z , it acts through

$$q_{\mathcal{G}}(x; z) = \hat{\theta}(x) - zs(x), \quad \pi_{\mathcal{G},z}(x) = D_{\mathcal{G}}\{x, q_{\mathcal{G}}(x; z)\}. \quad (1)$$

The coordinate z counts units of the generator’s own scale. Without an external condition on s_G , that coordinate has no invariant meaning: shrinking s_G while increasing z reciprocally leaves the policy unchanged, and changing s_G alone changes the dose. This paper makes the numerical margin comparable by calibrating its consequence against observed endpoints. A common pair of toxicity and progression envelopes $(b_T(z), b_P(z))$ is fixed across versions, and a version–margin pair is admitted only when exact endpoint tests support both inequalities. The result is a common operational meaning for z , expressed as a bivariate risk box rather than as an unverifiable claim about the internal standard error.

The initial margin is selected from a prospective randomized ladder $z_1 < \dots < z_J$. Both toxicity and progression are required because a dose of zero can suppress treatment toxicity while producing unacceptable disease progression. We therefore formulate the certificate as an intersection–union test for two competing-event cumulative incidences. The lifecycle problem then separates into two cases. A release whose two risk curves are dominated by those of the certified generator inherits the endpoint claim. A refit that cannot establish this joint order is tested at the single deployed margin, avoiding repetition of the complete ladder.

The contributions are as follows.

1. We adapt standard intersection–union and fixed-sequence testing to an ordered *model-output* coordinate, obtaining an exact bivariate certificate from randomized competing endpoints without requiring correct model-based endpoint probabilities.
2. Motivated by the elementary rescaling ambiguity of an uncalibrated margin, we prove that common endpoint envelopes with lifecycle error spending give z a cross-version risk interpretation.
3. We characterize the dominance envelope under which that endpoint claim transfers without new observations and derive the exact endpoint cost of single-anchor recertification outside the envelope.
4. We evaluate the theory in a positive impulsive PK–PD system with matched, misspecified, and refitted policy generators, including repeated finite-sample calibration and surveillance experiments.

State of the art

MIPD reviews distinguish technical validation of a pharmacometric model from prospective evidence that model-guided dosing improves decisions and outcomes [10, 11]. ICH M15 similarly asks that model performance be evaluated in relation to the question of interest and that assumptions, robustness, code verification, and applicability be documented [23]. Safe-update methods identify model-parameter regions in which a formal specification remains verified [13], and change-control guidance emphasizes bounded modifications, prospective validation, impact assessment, and monitoring [39, 41]. The present work connects these strands at the policy level: the target is the observed endpoint distribution after the model, estimator, constraints, and dose map have acted, and the update region is defined by preservation of that distributional claim.

Joint efficacy–toxicity dose finding is an established design problem, from Bayesian trade-off methods to the randomized benefit–risk emphasis of Project Optimus [37, 40]. Likewise, intersection–union testing and fixed-sequence multiplicity control are standard ingredients. The bivariate certificate below does not claim a new intersection–union principle. Its increment is to make a policy generator’s margin—rather than administered dose itself—the ordered coordinate, to define its two risks as competing-event cumulative incidences, and to carry the resulting endpoint claim through algorithm versions.

Sequential approval of software modifications has been studied directly by Feng, Emerson, and Simon, who formulate an automatic algorithmic change protocol as online hypothesis testing

and control bad-approval measures that prevent predictive-accuracy bio-creep [14]. Their work establishes the need for error control across successive modifications. The present lifecycle rule uses a different currency: jointly bounded observed toxicity and progression risk in a fixed target population. It also admits a distinct outcome-free route when the updated policy is proved jointly dominated; otherwise the update receives exact bivariate recertification at the deployed anchor. Parameter-space safe-update regions provide a further contrast because they preserve a formal model specification rather than an observed endpoint claim [13].

Adaptive control and control-oriented identification distinguish prediction accuracy from downstream decision performance [3, 7, 16, 18]. Robust and stochastic predictive control translate uncertainty sets or probability models into constraint margins [29, 34]. These approaches motivate the generator in Eq. (1), but do not make the units of a learned uncertainty scale comparable after refitting. Our endpoint envelope is designed for precisely this gap.

Conformal prediction supplies finite-sample marginal guarantees through ranks and exchangeability [33, 43]. Risk-controlling prediction sets extend this principle to general losses [4]. Learn-Then-Test (LTT) reframes calibration as multiple hypothesis testing over a prespecified parameter family [1]. Decision-aware conformal methods choose uncertainty sets for downstream value or robust optimization [25, 26, 30, 45], and conformal control methods regulate learned dynamics or deviation from a reference policy [8, 32, 36, 47]. Here, each participant reveals one competing endpoint under one randomized margin policy. This is a current-status or quantal-bioassay observation scheme [15, 38]: the methodological advance is to use a model-generated margin as the ordered treatment coordinate and to determine when that coordinate remains meaningful across algorithm versions.

Impulsive treatment models, direct pulse optimization, and robust PK–PD feedback are well established [2, 5, 42]. Moving-horizon estimation has been coupled to chemotherapy MPC [9]; monotonicity has reduced set-based PK–PD design to extreme parameter values [27]; and chance-constrained nonlinear MPC has treated adverse-effect constraints in stochastic cancer models [17]. Memory in drug disposition has been represented with fractional dynamics, fading kernels, and augmented compartments [12, 31, 46]. The positive transit augmentation used below is a linear-chain realization [22, 28, 35]. It supplies a controlled setting in which matched and omitted-memory generators can be compared under the same sparse observation design. The resulting study links three established components—recursive PK estimation, exact ordered testing, and prospective endpoint observation—through the new questions of margin calibration and certificate transfer.

Theoretical

We first define the closed-loop generator and the endpoint estimands, then derive the initial and cross-version certificates.

Flux-memory pharmacokinetics

Let A_1 and A_2 be central and peripheral drug amounts. The nonnegative states L_{12} and L_{21} store drug in the central-to-peripheral and peripheral-to-central transfer paths, respectively, with mean transit time $\tau_m = 1/c$. Between doses,

$$\dot{A}_1 = -(k_{10} + k_{12})A_1 + cL_{21}, \quad \dot{L}_{12} = k_{12}A_1 - cL_{12}, \quad (2a)$$

$$\dot{A}_2 = -k_{21}A_2 + cL_{12}, \quad \dot{L}_{21} = k_{21}A_2 - cL_{21}. \quad (2b)$$

For zero initial transit mass, the transfer paths have the Volterra form

$$L_{12}(t) = \int_0^t k_{12}e^{-c(t-s)}A_1(s)ds, \quad L_{21}(t) = \int_0^t k_{21}e^{-c(t-s)}A_2(s)ds, \quad (3)$$

with the corresponding shifted history after each impulse. Thus [equation \(2\)](#) is the one-stage exponential instance of the classical linear-chain trick [22, 28, 35]. The augmentation supplies the finite-dimensional realization used to connect memory time directly to estimation and endpoint calibration. The parent-drug plasma concentration is $C_1 = A_1/(V_1/F)$. The effect-site concentration of total active drug follows

$$\dot{C}_e = k_{e0} \left(\frac{A_1}{(V_1/F)f_p} - C_e \right), \quad (4)$$

where f_p is the parent fraction of total active drug. An administered dose is represented as the jump

$$A_1(t_n^+) = A_1(t_n^-) + D_n, \quad (A_2, L_{12}, L_{21}, C_e)(t_n^+) = (A_2, L_{12}, L_{21}, C_e)(t_n^-). \quad (5)$$

This instantaneous input is a computational approximation to oral dosing, not a mechanistic absorption model.

The effect compartment is driven by A_1 so that k_{e0} retains its usual concentration-to-effect interpretation. Memory changes the effect indirectly through the altered central trajectory.

Writing $x = (A_1, A_2, L_{12}, L_{21}, C_e)^\top$, the between-dose flow is $x(t + \Delta t) = \exp(G\Delta t)x(t)$. Summing the four PK states gives

$$\frac{d}{dt}(A_1 + A_2 + L_{12} + L_{21}) = -k_{10}A_1, \quad (6)$$

so internal transfer cancels exactly. The transit variables are continuous at a dose, and the model reduces to the classical two-compartment flow as $\tau_m \downarrow 0$.

Patient variation is introduced through apparent clearance alone. The micro-rate constants are

$$k_{10} = \frac{CL/F}{V_1/F}, \quad k_{12} = \frac{Q/F}{V_1/F}, \quad k_{21} = \frac{Q/F}{V_2/F}, \quad (7)$$

with Q/F and both volumes fixed at their reference values. Consequently, k_{12} and k_{21} remain positive across the simulated clearance domain.

Proposition 1 (Mass balance, positivity, and classical limit). *For positive rate constants and $\tau_m > 0$, the four-state PK generator in [equation \(2\)](#) is Metzler. Hence nonnegative initial states and nonnegative dose impulses generate nonnegative trajectories. The total amount satisfies [equation \(6\)](#). As $\tau_m \downarrow 0$, the transfer outputs cL_{12} and cL_{21} converge to $k_{12}A_1$ and $k_{21}A_2$, yielding the classical two-compartment equations.*

Proof. With $x_{\text{PK}} = (A_1, A_2, L_{12}, L_{21})^\top$, the generator is

$$G_{\text{PK}} = \begin{pmatrix} -(k_{10} + k_{12}) & 0 & 0 & c \\ 0 & -k_{21} & c & 0 \\ k_{12} & 0 & -c & 0 \\ 0 & k_{21} & 0 & -c \end{pmatrix}.$$

All off-diagonal entries are nonnegative, hence $\exp(G_{\text{PK}}t)$ is nonnegative and both the flow and nonnegative dose jumps preserve the positive orthant. Summing the four equations cancels the two transfer fluxes and gives [equation \(6\)](#). Finally,

$$cL_{12}(t) = \int_0^t ce^{-c(t-s)}k_{12}A_1(s)ds, \quad cL_{21}(t) = \int_0^t ce^{-c(t-s)}k_{21}A_2(s)ds,$$

apart from exponentially vanishing inter-dose initial terms. Positivity and the mass balance bound the states on every finite horizon, while the kernels $ce^{-c(t-s)}$ form an approximate identity. Thus, uniformly on compact inter-dose intervals, $cL_{12} \rightarrow k_{12}A_1$ and $cL_{21} \rightarrow k_{21}A_2$ as $c \rightarrow \infty$. Substitution into [equation \(2\)](#) yields the classical two-compartment flow with the same dose jumps. $\square$

Effect and competing-risk outcomes

The normalised effect is

$$E(t) = \frac{C_e(t)}{EC_{50} + C_e(t)}. \quad (8)$$

For disease progression, let $s = t/24$ be time in days and define the exposure-dependent scale $\phi(C_e) = \phi_0 + b_\phi(24C_e)$. Here $24C_e$ is a daily-AUC-equivalent exposure (mg h/L) under a local steady-exposure approximation; consequently b_ϕ has units d L/(mg h). The cause-specific hazard in hours is

$$\lambda_1(t) = \frac{\log 2}{24} \frac{\gamma_w}{\phi(C_e)} \left(\frac{s}{\phi(C_e)} \right)^{\gamma_w - 1}, \quad t > 0. \quad (9)$$

For constant exposure, the factor $\log 2$ makes ϕ the median progression time. With time-varying C_e , [equation \(9\)](#) is a Weibull-type hazard with exposure-dependent scale, not a Weibull event-time law. The toxicity hazard is the investigator-calibrated link

$$\lambda_2(t) = \lambda_{02} \exp\{\beta_2 E(t)\}. \quad (10)$$

These links adapt published sunitinib exposure–response quantities [\[20\]](#); they are not a validated joint clinical model.

With $\Lambda = \lambda_1 + \lambda_2$, survival S and cumulative incidences F_1, F_2 are updated on the numerical grid by

$$S^{m+1} = S^m e^{-\Lambda^m \Delta t}, \quad F_j^{m+1} = F_j^m + \frac{\lambda_j^m}{\Lambda^m} (S^m - S^{m+1}), \quad (11)$$

with zero incidence increment when $\Lambda^m = 0$. For hazards held constant on each time cell, this update is exact and preserves $S + F_1 + F_2 = 1$ to machine precision.

Observation and decision rules

At observation times τ_n , the stochastic reference system returns

$$y_n = C_1(\tau_n) \exp(\sigma Z_n), \quad Z_n \sim \mathcal{N}(0, 1), \quad (12)$$

independently across assays. Hence $\log y_n - \log C_1(\tau_n)$ is zero-mean with variance σ^2 .

At day 7, the decision rule retains the already administered induction doses. Write the terminal toxicity functional for a candidate continuation dose d and log-clearance θ as

$$\mathcal{J}(d, \theta) := F_2(T; d, \theta). \quad (13)$$

The certainty-equivalent rule selects the largest d satisfying

$$\mathcal{J}(d, \hat{\theta}_7) \leq \theta_{\text{tox}}, \quad 0 \leq d \leq 75 \text{ mg}. \quad (14)$$

Monotonicity on the verified dose range permits bisection. This is a one-dimensional continuation rule rather than a multi-variable open-loop optimiser.

For a selected radius $z \geq 0$, let

$$\mathcal{U}_z(\hat{\theta}_7, P_7) = \left[\hat{\theta}_7 - z\sqrt{P_7}, \hat{\theta}_7 + z\sqrt{P_7} \right], \quad \theta_{7,z}^- = \hat{\theta}_7 - z\sqrt{P_7}. \quad (15)$$

The robust continuation problem is

$$d_z^* = \max \left\{ d \in [0, 75] : \sup_{\theta \in \mathcal{U}_z(\hat{\theta}_7, P_7)} \mathcal{J}(d, \theta) \leq \theta_{\text{tox}} \right\}. \quad (16)$$

Proposition 2 (Saturation identity, chance level, and robust endpoint). *Suppose that $\mathcal{J}$ is continuous, strictly increasing in d , and strictly decreasing in θ on the dose–parameter rectangle under consideration. Suppose also that the robust feasible set is nonempty and that the solution is interior with respect to both dose bounds; equivalently,*

$$\mathcal{J}(0, \theta_{7,z}^-) < \theta_{\text{tox}} < \mathcal{J}(75, \theta_{7,z}^-). \quad (17)$$

Then the endpoint rule satisfies $\mathcal{J}(d_z^*, \theta_{7,z}^-) = \theta_{\text{tox}}$ and, for a correctly specified deterministic reference system,

$$\{\mathcal{J}(d_z^*, \theta^*) > \theta_{\text{tox}}\} = \{\theta^* < \hat{\theta}_7 - z\sqrt{P_7}\}. \quad (18)$$

Moreover, monotonicity makes [equation \(16\)](#) exactly equivalent to

$$\mathcal{J}(d, \theta_{7,z}^-) \leq \theta_{\text{tox}}. \quad (19)$$

The probability statement has two distinct readings. In a Bayesian reading, if $\theta^* \mid \mathcal{F}_7 \sim \mathcal{N}(\hat{\theta}_7, P_7)$, the conditional violation probability is $\Phi(-z)$, so certainty equivalence is the 50% posterior chance constraint. In a frequentist reading over repeated realizations, writing $E_7 = \hat{\theta}_7 - \theta^*$, the rate is $\Pr\{E_7 > z\sqrt{P_7}\}$ under the joint distribution of (E_7, P_7) . At $z = 0$ this equals one half when E_7 is median-unbiased (with no atom at zero). For $z > 0$, the frequentist rate is governed by the joint empirical distribution of estimation error and filter scale.

Proof. Write $\theta^- = \hat{\theta}_7 - z\sqrt{P_7}$. Since $\mathcal{J}(d, \theta)$ is decreasing in θ ,

$$\sup_{\theta \in [\theta^-, \theta^+]} \mathcal{J}(d, \theta) = \mathcal{J}(d, \theta^-),$$

which proves the deterministic robust counterpart. Continuity, strict increase in d , and [equation \(17\)](#) give a unique interior root d_z^* satisfying $\mathcal{J}(d_z^*, \theta^-) = \theta_{\text{tox}}$. Strict decrease in θ then yields

$$\mathcal{J}(d_z^*, \theta^*) > \theta_{\text{tox}} \iff \theta^* < \theta^- \iff \hat{\theta}_7 - \theta^* > z\sqrt{P_7}.$$

Under the stated conditional Gaussian law, standardization gives probability $\Phi(-z)$. Under repeated realizations the last event has probability $\Pr\{E_7 > z\sqrt{P_7}\}$, which equals 1/2 at $z = 0$ under median unbiasedness and no atom at zero. $\square$

The identity is used only on the verified dose–parameter domain and only when the dose cap is inactive. If the 75 mg cap binds, the robust inequality may be slack and [equation \(18\)](#) need not hold. With process perturbations, the identity is a noise-free benchmark; the randomized endpoint procedure below calibrates the realised stochastic policy directly.

Observed-endpoint certification of a margin ladder

Let $0 \leq z_1 < \dots < z_J$ be a margin ladder fixed before endpoint outcomes are observed. The index z is dimensionless and labels a family of personalized policies; it is not a physical dose and is not a certified-ML perturbation radius. Block randomization assigns exactly N_j independent systems to arm j . System i assigned to $A_i = j$ receives $d_{z_j}(Z_i)$.

Endpoint and causal estimands

Let $C_i(z) \in \{0, P, T\}$ be the first terminal event through the fixed 84-day horizon under margin z : no event, progression, or toxicity. Define

$$Y_i^T(z) = \mathbf{1}\{C_i(z) = T\}, \quad Y_i^P(z) = \mathbf{1}\{C_i(z) = P\},$$

and the population cumulative-incidence risks

$$R_T(z) = \Pr\{C_i(z) = T\} = \mathbb{E}[F_{2,i}(T; d_z)], \quad R_P(z) = \Pr\{C_i(z) = P\} = \mathbb{E}[F_{1,i}(T; d_z)]. \quad (20)$$

These are the certification estimands. The planner instead uses the individual model ceiling $\theta_{\text{tox}} = 0.10$ inside its dose map. The distinct latent diagnostic

$$V_T(z) = \Pr\{F_{2,i}(T; d_z) > \theta_{\text{tox}}\} \quad (21)$$

is used only as a simulation diagnostic and is never supplied to the endpoint certifier. In the primary certificate we prespecify population-event ceilings $\alpha_T = 0.12$ and $\alpha_P = 0.25$. Thus θ_{tox} is a planner tuning quantity, whereas (α_T, α_P) define the empirical claim.

The causal interpretation of [equation \(20\)](#) requires consistency (the observed endpoint equals the potential endpoint under the assigned policy), no interference between systems, randomized assignment independent of all potential endpoints, positivity for every arm to be tested, and a fixed target population, horizon, treatment strategy, and endpoint definition. The primary analysis assumes complete 84-day ascertainment and classifies the first event by mutually exclusive adjudication rules. Death or another intercurrent event must be prespecified as progression, toxicity, or a third competing event; it cannot silently be coded as no event. With censoring or loss to follow-up, the simple binomial law below is unavailable. A protocol would then need a separately identified treatment-policy estimand and an appropriate competing-risk analysis, with its own uncertainty and missingness assumptions.

For toxicity, the only outcome supplied to the certifier is

$$(A_i, Y_i^T), \quad Y_i^T \mid (Z_i, A_i = j) \sim \text{Bernoulli}(F_{2,i}(T; d_{z_j}(Z_i))). \quad (22)$$

Thus one assigned policy yields one observed binary terminal event; neither the individual probability $F_{2,i}$ nor outcomes under unassigned policies are calibration data. Progression is observed simultaneously with conditional probability $F_{1,i}$. The indicators are dependent and mutually exclusive, but each has the required marginal Bernoulli law.

For the next proposition abbreviate $Y_i = Y_i^T$, $p_i = p_i^T = F_{2,i}$, $R_j = R_{T,j}$, and $K_j = K_{T,j}$. For analysis, attach an independent $U_i \sim \text{Uniform}(0, 1)$ to each system and use the same U_i across the policy family. Write $p_i(z) = F_{2,i}(T; d_z(Z_i))$ and define the event-critical index

$$\tilde{\zeta}_i = \inf\{z \geq 0 : U_i \geq p_i(z)\}. \quad (23)$$

This coupling is a representation of Bernoulli potential outcomes, not an additional observation requirement.

Proposition 3 (Observed current-status identity). *Suppose $p_i(z)$ is nonincreasing and right-continuous. Under randomized arm assignment independent of (Z_i, U_i) ,*

$$Y_i = \mathbf{1}\{U_i < p_i(z_{A_i})\} = \mathbf{1}\{\tilde{\zeta}_i > z_{A_i}\}, \quad R_j := \Pr(Y_i = 1 \mid A_i = j) = \Pr(\tilde{\zeta}_i > z_j) = \mathbb{E}[p_i(z_j)]. \quad (24)$$

Consequently, with fixed block size $N_j > 0$, $K_j = \sum_{i:A_i=j} Y_i \sim \text{Binomial}(N_j, R_j)$.

Proof. Fix a system i and define

$$\mathcal{A}_i = \{z \geq 0 : U_i \geq p_i(z)\}.$$

We establish the event identity, the arm risk, and the count distribution.

Step 1: The crossing set is an upper interval. Let $z \in \mathcal{A}_i$ and let $z' \geq z$. Since p_i is nonincreasing,

$$p_i(z') \leq p_i(z) \leq U_i.$$

Hence $z' \in \mathcal{A}_i$. Thus $\mathcal{A}_i$ is an upper interval, and by [equation \(23\)](#) its lower endpoint is $\tilde{\zeta}_i = \inf \mathcal{A}_i$, with the convention $\inf \emptyset = +\infty$.

Step 2: Current-status event identity. First suppose $U_i < p_i(z)$. For any $0 \leq r \leq z$, monotonicity gives $p_i(r) \geq p_i(z) > U_i$. Therefore no $r \leq z$ belongs to $\mathcal{A}_i$. To exclude an open crossing immediately to the right of z , set

$$\epsilon = \frac{p_i(z) - U_i}{2} > 0.$$

Right-continuity of p_i at z gives an $\eta > 0$ such that

$$|p_i(r) - p_i(z)| < \epsilon \quad \text{whenever } z \leq r < z + \eta.$$

For such r ,

$$p_i(r) > p_i(z) - \epsilon = \frac{p_i(z) + U_i}{2} > U_i,$$

so these points also do not belong to $\mathcal{A}_i$. Hence $\mathcal{A}_i \cap [0, z + \eta) = \emptyset$, and

$$\tilde{\zeta}_i = \inf \mathcal{A}_i \geq z + \eta > z,$$

with the same conclusion when $\mathcal{A}_i = \emptyset$ under the convention $\inf \emptyset = +\infty$. Conversely, if $U_i \geq p_i(z)$, then $z \in \mathcal{A}_i$, so $\tilde{\zeta}_i \leq z$. Taking contrapositives and combining both directions gives the exact equivalence

$$U_i < p_i(z) \iff \tilde{\zeta}_i > z.$$

Substituting the assigned margin $z = z_{A_i}$ yields

$$Y_i = \mathbf{1}\{U_i < p_i(z_{A_i})\} = \mathbf{1}\{\tilde{\zeta}_i > z_{A_i}\}.$$

Step 3: Conditional and marginal arm risk. On the event $A_i = j$, condition on Z_i . The quantity $p_i(z_j)$ is then fixed and U_i is independent uniform, so

$$\begin{aligned} \Pr(Y_i = 1 \mid Z_i, A_i = j) &= \Pr\{U_i < p_i(z_j) \mid Z_i, A_i = j\} \\ &= p_i(z_j). \end{aligned}$$

Random assignment is independent of (Z_i, U_i) , hence the distribution of Z_i conditional on $A_i = j$ is its population distribution. The tower property therefore gives

$$\begin{aligned} R_j &= \Pr(Y_i = 1 \mid A_i = j) \\ &= \mathbb{E}[\Pr(Y_i = 1 \mid Z_i, A_i = j) \mid A_i = j] \\ &= \mathbb{E}[p_i(z_j)]. \end{aligned}$$

Step 2 also gives

$$R_j = \Pr(\tilde{\zeta}_i > z_j \mid A_i = j) = \Pr(\tilde{\zeta}_i > z_j),$$

where the final equality again uses assignment independence.

Step 4: Binomial arm count. Within a fixed arm, the N_j randomized systems are independent draws from the same population. Steps 2–3 show that their outcomes are independent Bernoulli variables with common success probability R_j . Consequently, for every $k = 0, 1, \dots, N_j$,

$$\Pr(K_j = k) = \binom{N_j}{k} R_j^k (1 - R_j)^{N_j - k},$$

which is precisely $K_j \sim \text{Binomial}(N_j, R_j)$.

□

When the margin ordering is well defined

The latent tolerance representation is available only when increasing z makes the policy no more aggressive and the *actual* endpoint risk is nondecreasing in that ordered action. The first property is a software property of the policy generator. The second is a scientific property of the interventional dose–endpoint relation. A numerical surface of the PK–PD model can check internal implementation consistency, but it cannot establish the second property for patients. Randomized endpoints at two or more ordered actions can test population-level violations and should be included when the scientific monotonicity premise is uncertain; they cannot prove monotonicity at every unobserved history without additional structure.

Competing risks give a concrete failure mechanism. With constant cause-specific hazards $\lambda_1(a)$ and $\lambda_2(a)$ over horizon T , the toxicity cumulative incidence is

$$F_2(T; a) = \frac{\lambda_2(a)}{\lambda_1(a) + \lambda_2(a)} \left[1 - e^{-\{\lambda_1(a) + \lambda_2(a)\}T} \right]. \quad (25)$$

Even if λ_2 is nondecreasing, F_2 need not be. For example, take $T = 10$, $\lambda_2(a) = 0.035(1 - e^{-1.3a})$, and $\lambda_1(a) = 0.003 + 0.0015a^3$. Then F_2 at $a = 1, 2, 4, 6$ is respectively 0.220, 0.258, 0.192, and 0.094. For a 0.20 ceiling the safe action set has two components: low actions are safe, intermediate actions are unsafe, and sufficiently high actions appear safe because the competing event occurs first. Thus $\max\{a : F_2(T; a) \leq 0.20\}$ can select an action that dominates unsafe ones.

When such folding is scientifically plausible, the scalar tolerance index in [proposition 3](#), the certified upper ray, and the free transfer theorem below do not apply. The candidate policies must instead be treated as an unordered or partially ordered family with multiplicity control, or analyzed using dose-finding designs that relax a simple toxicity order [\[24, 44\]](#). This limitation concerns the factual endpoint ordering, not merely the smoothness of a simulator curve.

The event-critical index $\tilde{\zeta}_i$ is a representation of the monotone arm-risk curve. The deployable certificate itself uses only the randomized binary endpoints and does not require individual critical indices to be observed or estimated.

For a toxicity population-event ceiling α_T and error level δ , the single-component construction tests (temporarily writing $\alpha = \alpha_T$)

$$H_j : R_j \geq \alpha \quad \text{against} \quad R_j < \alpha \quad (26)$$

using the exact lower-tail value

$$q_j = \Pr\{\text{Binomial}(N_j, \alpha) \leq K_j\}. \quad (27)$$

with $q_j = 1$ when $N_j = 0$. The experiments use fixed positive N_j by block randomization. Starting at $j = J$, reject H_j when $q_j \leq \delta$ and continue to $j - 1$; stop at the first non-rejection. If at least one hypothesis is rejected, return the smallest-index policy in this uninterrupted rejected sequence. Otherwise return no certified policy. This is the fixed-sequence specialization of Learn-Then-Test [\[1\]](#); its role here is to certify the observed endpoint attached to the margin-indexed controller.

Proposition 4 (Finite-sample observed-endpoint certificate). *Suppose the ladder is fixed independently of the endpoint outcomes, arm sizes are fixed and positive, and systems are independent within arms. Let $\hat{J}$ be the policy index returned by the ordered rule above. Then, for every collection of arm risks $(R_1, \dots, R_J)$,*

$$\Pr\{\hat{J} \text{ exists and } R_{\hat{J}} \geq \alpha\} \leq \delta. \quad (28)$$

If R_j is nonincreasing in j , the rule is the least conservative policy reached before the first policy that cannot be certified. The bound remains valid without this ordering property.

Proof. For each arm define the exact critical count

$$c_j = \max \{k \in \{-1, 0, \dots, N_j\} : \Pr\{\text{Binomial}(N_j, \alpha) \leq k\} \leq \delta\}.$$

The convention at $k = -1$ makes this definition valid even when no nonnegative count has lower-tail probability at most δ .

Step 1: Equivalent rejection region. The map $k \mapsto \Pr\{\text{Binomial}(N_j, \alpha) \leq k\}$ is nondecreasing. Therefore

$$\begin{aligned} q_j \leq \delta &\iff \Pr\{\text{Binomial}(N_j, \alpha) \leq K_j\} \leq \delta \\ &\iff K_j \leq c_j. \end{aligned}$$

Step 2: Size of each exact lower-tail test. Let H_j be true, so $R_j \geq \alpha$. A binomial variable with success probability R_j is stochastically no smaller than one with success probability α . Equivalently, for every integer c ,

$$\Pr\{\text{Binomial}(N_j, R_j) \leq c\} \leq \Pr\{\text{Binomial}(N_j, \alpha) \leq c\}.$$

Applying this inequality at $c = c_j$ and using Step 1 gives

$$\begin{aligned} \Pr_{R_j}(q_j \leq \delta) &= \Pr_{R_j}(K_j \leq c_j) \\ &\leq \Pr_{\alpha}\{\text{Binomial}(N_j, \alpha) \leq c_j\} \\ &\leq \delta. \end{aligned}$$

Step 3: The first true null in the testing order. The prespecified order is $J, J-1, \dots, 1$. If every H_j is false, no selected arm is unsafe and the desired event is empty. Otherwise let j_0 denote the first true null encountered in this order; equivalently, j_0 is the largest index satisfying $R_{j_0} \geq \alpha$. To reject any true null, the procedure must reach and reject H_{j_0} . Hence

$$\{\text{at least one true null is rejected}\} \subseteq \{q_{j_0} \leq \delta\}.$$

The event of reaching j_0 can only reduce this probability, because

$$\{\text{reach } j_0 \text{ and } q_{j_0} \leq \delta\} \subseteq \{q_{j_0} \leq \delta\}.$$

By Step 2,

$$\Pr\{\text{at least one true null is rejected}\} \leq \Pr(q_{j_0} \leq \delta) \leq \delta.$$

Step 4: Unsafe selection implies rejection of a true null. If $\hat{J}$ exists and $R_{\hat{J}} \geq \alpha$, then $H_{\hat{J}}$ is true. The rule returns an arm only after rejecting its null, so

$$\{\hat{J} \text{ exists and } R_{\hat{J}} \geq \alpha\} \subseteq \{\text{at least one true null is rejected}\}.$$

Combining this containment with Step 3 proves [equation \(28\)](#). Notice that no cross-arm independence was invoked: only the marginal size bound for the first true null is needed.

Step 5: Interpretation under ordered risks. If R_j is nonincreasing in j , increasing the margin cannot increase the population risk. The descent $J, J-1, \dots$ therefore moves from more conservative toward less conservative policies, and the returned arm is the smallest index in the uninterrupted certified suffix. This ordering explains the policy interpretation but is not used in Steps 1–4.

□

Theorem 5 (Bivariate efficacy–toxicity certificate). *For each arm let $K_{T,j} = \sum_{i:A_i=j} Y_i^T$ and $K_{P,j} = \sum_{i:A_i=j} Y_i^P$, and define*

$$q_{T,j} = \Pr\{\text{Binomial}(N_j, \alpha_T) \leq K_{T,j}\}, \quad (29)$$

$$q_{P,j} = \Pr\{\text{Binomial}(N_j, \alpha_P) \leq K_{P,j}\}. \quad (30)$$

Test the composite unacceptable-policy null

$$H_j^{\text{joint}} : R_T(z_j) \geq \alpha_T \quad \text{or} \quad R_P(z_j) \geq \alpha_P \quad (31)$$

by rejecting only when $\max(q_{T,j}, q_{P,j}) \leq \delta$. Apply the same fixed-sequence rule in a prespecified arm order. Then

$$\Pr\left[\hat{J} \text{ exists and } \{R_T(z_{\hat{J}}) \geq \alpha_T \text{ or } R_P(z_{\hat{J}}) \geq \alpha_P\}\right] \leq \delta. \quad (32)$$

No independence between the two endpoint counts is required.

Proof. Step 1: Separate the two component nulls. Write $H_{T,j} : R_T(z_j) \geq \alpha_T$ and $H_{P,j} : R_P(z_j) \geq \alpha_P$. Equation (31) is the union $H_j^{\text{joint}} = H_{T,j} \cup H_{P,j}$.

Step 2: Establish each marginal exact-test bound. Under $H_{T,j}$, the same stochastic-order calculation as in Step 2 of proposition 4 gives

$$\Pr(q_{T,j} \leq \delta) \leq \delta.$$

Under $H_{P,j}$ the identical calculation, with $(K_{T,j}, \alpha_T)$ replaced by $(K_{P,j}, \alpha_P)$, gives

$$\Pr(q_{P,j} \leq \delta) \leq \delta.$$

Step 3: Control the intersection–union test for one arm. If H_j^{joint} is true, at least one component null is true. If $H_{T,j}$ is true, then

$$\begin{aligned} \Pr\{\max(q_{T,j}, q_{P,j}) \leq \delta\} &= \Pr(q_{T,j} \leq \delta, q_{P,j} \leq \delta) \\ &\leq \Pr(q_{T,j} \leq \delta) \leq \delta. \end{aligned}$$

If instead $H_{P,j}$ is true, the same containment gives the bound through $q_{P,j}$. This is an intersection–union test, so no multiplicity split and no independence between toxicity and progression is needed.

Step 4: Locate the first true joint null. If all joint nulls are false, every selectable arm satisfies both ceilings. Otherwise let j_0 be the first true H_j^{joint} encountered in the fixed testing order. Rejection of any true joint null requires the procedure to reach and reject $H_{j_0}^{\text{joint}}$. Therefore

$$\Pr\{\text{at least one true joint null is rejected}\} \leq \Pr\{\max(q_{T,j_0}, q_{P,j_0}) \leq \delta\} \leq \delta.$$

Step 5: Translate rejection into selection. If the returned arm violates either ceiling, its joint null is true and was rejected. Hence the event in equation (32) is contained in the event bounded in Step 4, which proves the claim. $\square$

The two risks move in opposite directions in the case study: increasing z reduces toxicity but increases progression. Consequently, the bivariate claim does not certify an infinite upper ray. It certifies only the tested arm, or a prespecified finite set of arms separately reached by the fixed sequence. This qualification prevents the zero-continuation policy from satisfying the clinical certificate merely because its toxicity risk is small.

Corollary 6 (Certified continuous upper ray). *If $R(z) = \mathbb{E}[p_i(z)]$ is nonincreasing, then returning arm $\hat{J}$ simultaneously certifies every policy in the continuous upper ray $[z_{\hat{J}}, \infty)$: the probability that this reported ray contains an index z with $R(z) \geq \alpha$ is at most δ .*

Proof. Step 1: Define the reported set. On the event that the rule returns $\hat{J}$, write

$$\mathcal{R}_{\hat{J}} = [z_{\hat{J}}, \infty).$$

Therefore every $z \in \mathcal{R}_{\hat{J}}$ satisfies $z - z_{\hat{J}} \geq 0$.

Step 2: Propagate risk ordering from the anchor. Since R is nonincreasing, $z \geq z_{\hat{J}}$ implies

$$R(z) \leq R(z_{\hat{J}}).$$

If some point in the reported ray is unsafe, then there is a $z \geq z_{\hat{J}}$ for which $R(z) \geq \alpha$. Combining both inequalities yields the fully ordered chain

$$R(z_{\hat{J}}) \geq R(z) \geq \alpha.$$

Hence the anchor is unsafe whenever any member of its upper ray is unsafe.

Step 3: Write the event containment explicitly. Define

$$\begin{aligned} A &= \{\hat{J} \text{ exists and } \exists z \geq z_{\hat{J}} : R(z) \geq \alpha\}, \\ B &= \{\hat{J} \text{ exists and } R(z_{\hat{J}}) \geq \alpha\}. \end{aligned}$$

Step 2 proves $A \subseteq B$.

Step 4: Apply the anchor certificate. Because $R(z_{\hat{J}}) = R_{\hat{J}}$,

$$\begin{aligned} \Pr(A) &\leq \Pr(B) \\ &= \Pr\{\hat{J} \text{ exists and } R_{\hat{J}} \geq \alpha\} \\ &\leq \delta, \end{aligned}$$

where the last inequality is [proposition 4](#). □

Proposition 7 (Exact selection and power law). *Condition on fixed positive arm sizes $N_1, \dots, N_J$ and independent arms. Define*

$$\begin{aligned} c_j &= \max\{k \in \{-1, 0, \dots, N_j\} : \Pr\{\text{Binomial}(N_j, \alpha) \leq k\} \leq \delta\}, \\ \beta_j &= \Pr\{\text{Binomial}(N_j, R_j) \leq c_j\}, \end{aligned} \tag{33}$$

where the probability at $k = -1$ is zero, and put $\beta_0 = 0$. Then

$$\Pr\{\hat{J} \text{ does not exist}\} = 1 - \beta_J, \tag{34}$$

$$\Pr\{\hat{J} = j\} = (1 - \beta_{j-1}) \prod_{\ell=j}^J \beta_{\ell}, \quad j = 1, \dots, J. \tag{35}$$

Consequently, for any deterministic policy summary g_j ,

$$\mathbb{E}[g_{\hat{J}} \mid \hat{J} \text{ exists}] = \frac{\sum_{j=1}^J g_j (1 - \beta_{j-1}) \prod_{\ell=j}^J \beta_{\ell}}{\beta_J}. \tag{36}$$

Proof. Step 1: Compute each arm's certification probability. For each arm j , write

$$\mathcal{C}_j = \{K_j \leq c_j\}$$

for its certification event. The definition of c_j in [equation \(33\)](#) makes the exact-binomial test reject its null hypothesis precisely on $\mathcal{C}_j$. Since $K_j \sim \text{Binomial}(N_j, R_j)$ under the fixed arm risk R_j ,

$$\begin{aligned} \Pr(\mathcal{C}_j) &= \Pr\{K_j \leq c_j\} \\ &= \Pr\{\text{Binomial}(N_j, R_j) \leq c_j\} \\ &= \beta_j. \end{aligned}$$

Step 2: Derive the no-selection probability. The ordered procedure inspects arms from J downward. Therefore it makes at least one selection if and only if the first inspected arm, arm J , is certified. Consequently,

$$\begin{aligned} \{\widehat{J} \text{ exists}\} &= \mathcal{C}_J, \\ \Pr\{\widehat{J} \text{ exists}\} &= \Pr(\mathcal{C}_J) = \beta_J, \\ \Pr\{\widehat{J} \text{ does not exist}\} &= 1 - \beta_J, \end{aligned}$$

which proves [equation \(34\)](#).

Step 3: Derive the probability of selecting a particular arm. Fix $j \in \{1, \dots, J\}$. The selected index equals j exactly when every arm from J through j is certified and the next arm, $j - 1$, is not certified. To include $j = 1$ in the same notation, define $\mathcal{C}_0 = \emptyset$, so that $\Pr(\mathcal{C}_0) = \beta_0 = 0$ and $\Pr(\mathcal{C}_0^c) = 1 - \beta_0 = 1$. Then

$$\{\widehat{J} = j\} = \mathcal{C}_{j-1}^c \cap \bigcap_{\ell=j}^J \mathcal{C}_\ell.$$

No condition is imposed on arms $1, \dots, j - 2$, because the procedure stops as soon as arm $j - 1$ fails certification. The arm samples are independent, so the events $\mathcal{C}_1, \dots, \mathcal{C}_J$ are independent. Hence

$$\begin{aligned} \Pr\{\widehat{J} = j\} &= \Pr(\mathcal{C}_{j-1}^c) \prod_{\ell=j}^J \Pr(\mathcal{C}_\ell) \\ &= (1 - \beta_{j-1}) \prod_{\ell=j}^J \beta_\ell, \end{aligned}$$

which is [equation \(35\)](#).

Step 4: Expand the conditional policy summary. The events $\{\widehat{J} = j\}$, $j = 1, \dots, J$, are disjoint and their union is $\{\widehat{J} \text{ exists}\}$. Therefore

$$\begin{aligned} \mathbb{E}\left[g_{\widehat{J}} \mathbf{1}_{\{\widehat{J} \text{ exists}\}}\right] &= \sum_{j=1}^J g_j \Pr\{\widehat{J} = j\} \\ &= \sum_{j=1}^J g_j (1 - \beta_{j-1}) \prod_{\ell=j}^J \beta_\ell. \end{aligned}$$

Dividing this identity by $\Pr\{\hat{J} \text{ exists}\} = \beta_J$ gives

$$\begin{aligned}\mathbb{E}[g_{\hat{J}} \mid \hat{J} \text{ exists}] &= \frac{\mathbb{E}[g_{\hat{J}} \mathbf{1}_{\{\hat{J} \text{ exists}\}}]}{\Pr\{\hat{J} \text{ exists}\}} \\ &= \frac{\sum_{j=1}^J g_j (1 - \beta_{j-1}) \prod_{\ell=j}^J \beta_{\ell}}{\beta_J},\end{aligned}$$

which proves [equation \(36\)](#). □

The integer cutoff c_j exposes a second source of discreteness beyond the margin ladder. When all blocks have size n , the largest observable event rate that can pass is $c(n)/n$. The normal approximation gives

$$\frac{c(n)}{n} = \alpha - \Phi^{-1}(1 - \delta) \sqrt{\frac{\alpha(1 - \alpha)}{n}} + O(n^{-1}), \quad (37)$$

so increasing n improves statistical resolution only at the $n^{-1/2}$ rate; once this improvement is smaller than the change in risk between adjacent ladder arms, policy granularity becomes the binding source of conservatism. Changes in $c(n)$ by one event also explain small nonmonotone steps in unsafe-selection probability across nearby sample sizes.

For comparison, a simultaneous exact-binomial certifier tests every arm at level δ/J and returns the smallest index among the rejected hypotheses. The union bound gives the same unsafe-selection guarantee without using the order. Under independent arms its exact selection law is

$$\Pr\{\hat{J}_B = j\} = \beta_j^B \prod_{\ell < j} (1 - \beta_{\ell}^B), \quad \Pr\{\hat{J}_B \text{ absent}\} = \prod_{\ell=1}^J (1 - \beta_{\ell}^B), \quad (38)$$

where β_j^B uses the exact cutoff at level δ/J . This comparator isolates the efficiency gained by the prespecified order.

The results separate validity from power. A conservative fallback can make non-selection unlikely, while sample size and $\alpha - R_j$ determine how often the rule advances to a less conservative policy. Every random variable used by [proposition 4](#) is observed under the assigned policy.

Lifecycle certification across model versions

Assumptions and composition of claims

The finite-sample results require the stable unit treatment value assumption (SUTVA), namely consistency and no interference between participants, randomized allocation with positive probability for every registered arm, independent participants, complete ascertainment of the competing endpoint at the fixed horizon, and a target population, treatment strategy, endpoint definition, ladder, and testing order fixed before outcomes are examined. Correct model-based event probabilities are not required. The current-status representation additionally requires the toxicity risk to be nonincreasing in z ; the bivariate certificate itself does not require this order or independence between toxicity and progression. Structural transfer requires an unchanged target population and endpoint mechanism together with the stated risk-dominance premises.

The transfer and calibration results compose as follows. Fresh randomized blocks are required for a version–margin pair admitted through the empirical calibration route. If instead a release is jointly dominated by an already calibrated version, then for $e \in \{T, P\}$, $R_{e, \mathcal{G}'}(z) \leq R_{e, \mathcal{G}}(z) < b_e(z)$; the release therefore inherits both the endpoint certificate and the calibrated meaning of z without a new block. Toxicity-only action dominance transfers neither the progression component nor this bivariate calibration statement. Any update for which joint dominance cannot be established must undergo bivariate single-anchor recertification, including the calibration-envelope tests if the numerical value of z is to be compared across versions.

Prospective registration and independent verification

Sponsor-controlled choices are locked before certification endpoints are observed. Amendments create a new protocol version and cannot retroactively preserve a certificate. Table 1 distinguishes claims that can be recomputed from released evidence from those that require source access or scientific attestation; this distinction prevents a signed release record from being presented as independent empirical verification.

Table 1: Compact adversarial-control and verification register.

Object	Prospective control	Verification status
Policy family	Lock ladder, testing order, current-practice anchor, both risk ceilings, fallback and stopping rules before allocation or outcome access.	The exact p -values and selection decision are independently reproducible from arm sizes and endpoint counts.
Scale and generator	Lock θ_{tox} , definition and provenance of s , training data, endpoint envelopes, solver and executable; prohibit silent rescaling or relabeling of z .	Dataset, code, randomization, adjudication and release hashes are independently auditable when source access is provided.
Transfer domain	Register eligibility-linked $\mathcal{X}$, tail and boundary cases, missing-data states and the signed conformance test. Shrinking $\mathcal{X}$ voids transfer.	Completeness of $\mathcal{X}$ and biological action-risk monotonicity remain sponsor attestations unless an independent assessor receives the generator, test harness and supporting data.
Endpoints and monitor	Lock estimands, horizon, competing-event priority, censoring, adherence, adjudication, analysis population, mixture alternatives, weights, alarm spending and attribution plan.	Endpoint arithmetic is reproducible; ascertainment and adjudication are auditable; transportability and unmeasured biological premises remain scientific attestations.
Governance	Record protocol version, approval by the investigator and independent statistician, database lock, evidence identifiers and signatories.	Each Label field is marked independently reproduced, independently audited, or sponsor attestation.

The registered objects also instantiate the three elements of a predetermined change-control plan [39]: the description of modifications is the finite update class, frozen ladder, domain $\mathcal{X}$, and explicit exclusions; the modification protocol comprises signed joint-dominance tests, the exact bivariate single-anchor design, calibration spending, monitoring and triage; and the impact assessment records action differences, verification status, component endpoints, transport checks, and residual risk. At release, these records lead to one of four decisions: transfer under joint dominance, collect new endpoints, select a jointly certified fallback, or declare the certificate void.

The certified object and its semantics

Let $\mathcal{X}$ be the space of admissible patient histories available at the decision time. A software version, or *policy generator*, is a measurable map $\mathcal{G}$ that turns $x \in \mathcal{X}$ and a margin z into an action $\pi_{\mathcal{G},z}(x)$. For a fixed target population and endpoint definition, let $Y_{\mathcal{G}}(z) \in \{0, 1\}$ denote the potential adverse endpoint under that action and define

$$R_{\mathcal{G}}(z) = \Pr\{Y_{\mathcal{G}}(z) = 1\}. \quad (39)$$

In this subsection $Y_{\mathcal{G}} = Y_{\mathcal{G}}^T$ and $R_{\mathcal{G}} = R_{T,\mathcal{G}}$ abbreviate the toxicity component. The clinically operative certificate is the bivariate claim in theorem 5, which additionally requires $R_{P,\mathcal{G}}(z) < \alpha_P$. This notation includes every internal component of the deployed version: the PK model, estimator, uncertainty calculation, constraints, numerical solver, and dose map. It excludes changes to the target population or endpoint ascertainment; those are distributional changes addressed by the surveillance procedure below.

Proposition 8 (Non-identifiability of the numerical margin). *Suppose a generator uses $q(x; z) = \hat{\theta}(x) - zs(x)$ and action $D\{x, q(x; z)\}$. For any constant $c > 0$, define $\tilde{s}(x) = cs(x)$ and $\tilde{z} = z/c$. Then the two representations generate identical actions for every history. In particular, choosing $c < 1$ makes the reported numerical margin $\tilde{z}$ larger without making the policy more conservative.*

Consequently, z has no cross-version safety meaning without an externally fixed scale convention or calibration condition.

Proof. For every x ,

$$\tilde{q}(x; \tilde{z}) = \hat{\theta}(x) - \frac{\tilde{z}}{c} \{c, s(x)\} = \hat{\theta}(x) - z s(x) = q(x; z).$$

Applying the same deterministic dose map gives identical actions and hence identical endpoint distributions. If $0 < c < 1$, however, $\tilde{z} = z/c > z$. Thus the numerical magnitude of z depends on its representation even though the deployed policy does not. $\square$

Definition 9 (Endpoint-calibrated margin scale). Fix before observing outcomes a common calibration grid $\mathcal{Z}_C$, a target population and endpoint definition, and upper-envelope functions $b_T(z)$ and $b_P(z)$. A generator's reported scale s_g is *endpoint calibrated on $\mathcal{Z}_C$* when its induced deployed policies satisfy

$$R_{T,g}(z) < b_T(z), \quad R_{P,g}(z) < b_P(z), \quad z \in \mathcal{Z}_C. \quad (40)$$

The definition calibrates the operational consequence of the scale, not a latent standard error. A numerically understated scale is therefore prohibited exactly when it makes the same registered z violate the common endpoint envelope. Rescaling that leaves all actions unchanged is immaterial.

Theorem 10 (Cross-version meaning from endpoint calibration). *Consider any finite collection $\mathcal{I}$ of registered version-margin pairs (v, j) , with $z_j \in \mathcal{Z}_C$. Pair (v, j) receives a fresh randomized block of $N_{v,j}$ completely observed competing endpoints. Define*

$$q_{v,j}^T = \Pr\{\text{Binomial}(N_{v,j}, b_T(z_j)) \leq K_{v,j}^T\}, \quad (41)$$

$$q_{v,j}^P = \Pr\{\text{Binomial}(N_{v,j}, b_P(z_j)) \leq K_{v,j}^P\}. \quad (42)$$

Accept endpoint calibration of that pair only when $\max(q_{v,j}^T, q_{v,j}^P) \leq \delta_{v,j}$, where the lifecycle spending rule satisfies $\sum_{(v,j) \in \mathcal{I}} \delta_{v,j} \leq \delta_C$. Then

$$\Pr[\exists (v, j) \in \mathcal{I} : (v, j) \text{ is accepted and } \{R_{T,v}(z_j) \geq b_T(z_j) \text{ or } R_{P,v}(z_j) \geq b_P(z_j)\}] \leq \delta_C. \quad (43)$$

Consequently, apart from a lifecycle error event of probability at most δ_C , the same numerical z_j has the common cross-version meaning $(R_T, R_P) \in [0, b_T(z_j)) \times [0, b_P(z_j))$ for every accepted release.

Proof. *Step 1:* Write the unacceptable envelope null. For pair (v, j) define

$$H_{v,j}^C = \{R_{T,v}(z_j) \geq b_T(z_j)\} \cup \{R_{P,v}(z_j) \geq b_P(z_j)\}.$$

If this union is true, at least one component null is true.

Step 2: *Control one registered pair.* Suppose first that the toxicity component is true. Exact binomial stochastic ordering gives

$$\Pr(q_{v,j}^T \leq \delta_{v,j}) \leq \delta_{v,j}.$$

Since acceptance requires both component p -values to be small,

$$\begin{aligned} \Pr\{\max(q_{v,j}^T, q_{v,j}^P) \leq \delta_{v,j}\} &\leq \Pr(q_{v,j}^T \leq \delta_{v,j}) \\ &\leq \delta_{v,j}. \end{aligned}$$

If the progression component is true, the identical calculation with P in place of T yields the same bound. Dependence between the two endpoints is irrelevant because acceptance is an intersection.

Step 3: Form the lifecycle bad-acceptance events. Let $B_{v,j}$ be the event that pair (v,j) is accepted while $H_{v,j}^C$ is true. Step 2 gives $\Pr(B_{v,j}) \leq \delta_{v,j}$ for every registered pair.

Step 4: Expand the union bound across releases and margins. The event in [equation \(43\)](#) is $\bigcup_{(v,j) \in \mathcal{I}} B_{v,j}$. Therefore

$$\begin{aligned} \Pr\left(\bigcup_{(v,j) \in \mathcal{I}} B_{v,j}\right) &\leq \sum_{(v,j) \in \mathcal{I}} \Pr(B_{v,j}) \\ &\leq \sum_{(v,j) \in \mathcal{I}} \delta_{v,j} \\ &\leq \delta_C. \end{aligned}$$

Step 5: State the common coordinate meaning. Outside that union event, every accepted pair satisfies both strict inequalities in [equation \(40\)](#). Because the functions (b_T, b_P) and the numerical grid are common and registered, the risk box attached to z_j is identical across versions. This is the asserted cross-version meaning. $\square$

This theorem is deliberately endpoint-verifiable. It does not claim that s_G estimates a statistical standard deviation. It rules out policy-relevant understatement of s by requiring the actions produced at a common z to pass the same observed-endpoint envelope after every material version change. Structural joint dominance may avoid new endpoints under [proposition 13](#); otherwise the error-spent calibration block is mandatory.

Theorem 11 (Certificate semantics). *Under the assumptions of [propositions 3](#) and [4](#), the ordered endpoint procedure certifies the pair $(\mathcal{G}, z_{\hat{\mathcal{J}}})$ in the sense that*

$$\Pr\left\{\hat{\mathcal{J}} \text{ exists and } R_{\mathcal{G}}(z_{\hat{\mathcal{J}}}) \geq \alpha\right\} \leq \delta. \quad (44)$$

The probability is over the randomized calibration experiment. No assertion that the internal PK, observation, or outcome model is correctly specified is needed for this statement.

Proof. Step 1: Identify the observed arm mean. For a system assigned to arm j , [proposition 3](#) gives

$$\begin{aligned} \Pr(Y_i = 1 \mid A_i = j) &= \mathbb{E}[\Pr\{Y_i = 1 \mid X_i, A_i = j\} \mid A_i = j] \\ &= \mathbb{E}[r\{X_i, \pi_{\mathcal{G}, z_j}(X_i)\}] \\ &= R_{\mathcal{G}}(z_j). \end{aligned}$$

Thus the arm parameter R_j in the exact-binomial construction is precisely $R_{\mathcal{G}}(z_j)$.

Step 2: Substitute the policy risk into the unsafe-selection event. The event in [equation \(28\)](#) becomes

$$\{\hat{\mathcal{J}} \text{ exists and } R_{\hat{\mathcal{J}}} \geq \alpha\} = \{\hat{\mathcal{J}} \text{ exists and } R_{\mathcal{G}}(z_{\hat{\mathcal{J}}}) \geq \alpha\}.$$

Step 3: Apply the finite-sample bound. Taking probabilities and invoking [proposition 4](#) yields

$$\Pr\left\{\hat{\mathcal{J}} \text{ exists and } R_{\mathcal{G}}(z_{\hat{\mathcal{J}}}) \geq \alpha\right\} \leq \delta.$$

The internal model is used before observation to generate the fixed actions; no equality between model-predicted and actual endpoint probabilities occurs in these steps. $\square$

Thus the statement is neither “the model is valid” nor “every patient is safe.” Applied through [theorem 5](#), it is a repeated-experiment guarantee against selecting a deployed policy whose toxicity or progression cumulative-incidence risk reaches its prespecified ceiling.

Safety order and free transfer

Fix an anchor z_0 . For two versions evaluated in the same target population with the same endpoint, define

$$\mathcal{G}' \preceq_{z_0} \mathcal{G} \iff R_{\mathcal{G}'}(z) \leq R_{\mathcal{G}}(z) \text{ for every } z \geq z_0. \quad (45)$$

Proposition 12 (Version preorder and transfer). *The relation $\preceq_{z_0}$ is reflexive and transitive. It becomes a partial order after versions with identical risk curves on $[z_0, \infty)$ are treated as equivalent. If the base procedure returns $z_{\hat{J}} \geq z_0$ and an accepted update satisfies $\mathcal{G}' \preceq_{z_0} \mathcal{G}$ pointwise for every possible returned anchor, then*

$$\Pr\left\{\hat{J} \text{ exists and } R_{\mathcal{G}'}(z_{\hat{J}}) \geq \alpha\right\} \leq \delta. \quad (46)$$

No new endpoint is required for this transfer.

Proof. Step 1: Verify reflexivity. For every version $\mathcal{G}$ and every $z \geq z_0$,

$$R_{\mathcal{G}}(z) \leq R_{\mathcal{G}}(z).$$

Therefore $\mathcal{G} \preceq_{z_0} \mathcal{G}$.

Step 2: Verify transitivity. Suppose $\mathcal{G}'' \preceq_{z_0} \mathcal{G}'$ and $\mathcal{G}' \preceq_{z_0} \mathcal{G}$. For each $z \geq z_0$,

$$R_{\mathcal{G}''}(z) \leq R_{\mathcal{G}'}(z) \leq R_{\mathcal{G}}(z).$$

Hence $\mathcal{G}'' \preceq_{z_0} \mathcal{G}$.

Step 3: Identify the quotient partial order. Define $\mathcal{G} \sim_{z_0} \mathcal{G}'$ when $R_{\mathcal{G}}(z) = R_{\mathcal{G}'}(z)$ for all $z \geq z_0$. Two distinct implementations may satisfy this equality, so antisymmetry need not hold on implementations. If both equivalence classes satisfy $[\mathcal{G}'] \preceq [\mathcal{G}]$ and $[\mathcal{G}] \preceq [\mathcal{G}']$, then for every $z \geq z_0$,

$$R_{\mathcal{G}'}(z) \leq R_{\mathcal{G}}(z) \text{ and } R_{\mathcal{G}}(z) \leq R_{\mathcal{G}'}(z),$$

so the curves are equal and $[\mathcal{G}'] = [\mathcal{G}]$. This is antisymmetry on the quotient.

Step 4: Transfer the finite-sample event. On every outcome for which $\hat{J}$ exists, $z_{\hat{J}} \geq z_0$ and

$$R_{\mathcal{G}'}(z_{\hat{J}}) \leq R_{\mathcal{G}}(z_{\hat{J}}).$$

Consequently,

$$\begin{aligned} &\{\hat{J} \text{ exists and } R_{\mathcal{G}'}(z_{\hat{J}}) \geq \alpha\} \\ &\subseteq \{\hat{J} \text{ exists and } R_{\mathcal{G}}(z_{\hat{J}}) \geq \alpha\}. \end{aligned}$$

Taking probabilities and applying [theorem 11](#) gives [equation \(46\)](#). □

The preorder in [equation \(45\)](#) concerns toxicity alone. For the bivariate certificate define

$$\mathcal{G}' \preceq_{z_0}^{\text{joint}} \mathcal{G} \iff \begin{cases} R_{T,\mathcal{G}'}(z) \leq R_{T,\mathcal{G}}(z), \\ R_{P,\mathcal{G}'}(z) \leq R_{P,\mathcal{G}}(z), \end{cases} \text{ for every declared } z \geq z_0. \quad (47)$$

Proposition 13 (Joint certificate transfer). *If the base bivariate procedure returns $z_{\hat{J}} \geq z_0$ and an update satisfies [equation \(47\)](#) for every possible returned anchor, then the updated version retains the bound in [equation \(32\)](#) without new endpoints.*

Proof. Step 1: On every outcome for which $\hat{J}$ exists, joint dominance gives $R_{T,\mathcal{G}'}(z_{\hat{J}}) \leq R_{T,\mathcal{G}}(z_{\hat{J}})$ and $R_{P,\mathcal{G}'}(z_{\hat{J}}) \leq R_{P,\mathcal{G}}(z_{\hat{J}})$.

Step 2: If the updated arm violates either ceiling, the corresponding base-arm risk is at least as large and violates that same ceiling. Thus the updated unacceptable-selection event is contained in the base unacceptable-selection event.

Step 3: Taking probabilities and applying [theorem 5](#) gives the required bound δ . □

[Equation \(45\)](#) is a semantic definition, not by itself an outcome-free verification method. Estimating $R_{\mathcal{G}'}$ from simulations generated by $\mathcal{G}'$ would make transfer depend on the model being updated. The following sufficient condition avoids that circularity.

Let the action space be ordered so that $a' \leq a$ means that a' is no more aggressive than a ; in the case study this is the usual order on continuation dose. Let $r(x, a) = \Pr(Y = 1 \mid X = x, \text{do}(A = a))$ be the actual conditional endpoint risk.

Theorem 14 (Structural dominance envelope). *Assume that $r(x, a)$ is nondecreasing in a for every admissible history x and that the population and endpoint mechanism remain unchanged. If an update satisfies*

$$\pi_{\mathcal{G}',z}(x) \leq \pi_{\mathcal{G},z}(x) \quad \text{for every } x \in \mathcal{X} \text{ and } z \geq z_0, \quad (48)$$

then $\mathcal{G}' \preceq_{z_0} \mathcal{G}$.

A checkable sufficient condition for [equation \(48\)](#) is

$$q_{\mathcal{G}'}(x; z) \leq q_{\mathcal{G}}(x; z), \quad (49)$$

$$D_{\mathcal{G}'}(x, p') \leq D_{\mathcal{G}}(x, p) \quad \text{whenever } p' \leq p, \quad (50)$$

throughout the prespecified software test domain.

Proof. Step 1: Compare actions at a fixed history and margin. Fix arbitrary $x \in \mathcal{X}$ and $z \geq z_0$. By [equation \(48\)](#),

$$a' = \pi_{\mathcal{G}',z}(x) \leq a = \pi_{\mathcal{G},z}(x).$$

Step 2: Propagate the action order to actual endpoint risk. Because $a \mapsto r(x, a)$ is nondecreasing,

$$a' \leq a \implies r(x, a') \leq r(x, a).$$

Substituting the definitions of a' and a gives

$$r\{x, \pi_{\mathcal{G}',z}(x)\} \leq r\{x, \pi_{\mathcal{G},z}(x)\}.$$

Step 3: Integrate over the unchanged population. Pointwise order is preserved by expectation, so

$$\begin{aligned} R_{\mathcal{G}'}(z) &= \int_{\mathcal{X}} r\{x, \pi_{\mathcal{G}',z}(x)\} dP_X(x) \\ &\leq \int_{\mathcal{X}} r\{x, \pi_{\mathcal{G},z}(x)\} dP_X(x) \\ &= R_{\mathcal{G}}(z). \end{aligned}$$

Since $z \geq z_0$ was arbitrary, this is $\mathcal{G}' \preceq_{z_0} \mathcal{G}$.

Step 4: Verify the coordinate-level sufficient condition. Under [equation \(49\)](#), put $p' = q_{\mathcal{G}'}(x; z)$ and $p = q_{\mathcal{G}}(x; z)$, so $p' \leq p$. [Equation \(50\)](#) then gives

$$\begin{aligned}\pi_{\mathcal{G}', z}(x) &= D_{\mathcal{G}'}(x, p') \\ &\leq D_{\mathcal{G}}(x, p) \\ &= \pi_{\mathcal{G}, z}(x),\end{aligned}$$

which recovers [equation \(48\)](#). □

This structural theorem proves only toxicity-component dominance. In a setting where lower dose can increase progression, pointwise “no more aggressive” action does not imply [equation \(47\)](#). Free transfer of the bivariate claim therefore requires either action equivalence or independent scientific evidence that both endpoint risks are dominated; otherwise both components must be recertified at the anchor.

The premise about the actual endpoint response is a scientific causal assumption. The numerical surface check reported in the Results verifies monotonicity of the model functional $\mathcal{J}$, not monotonicity of $r(x, a)$, and therefore provides no evidence for this premise. Software conformance to [equations \(49\)](#) and [\(50\)](#) can be checked by formal analysis or exhaustive testing on a documented finite deployment domain. Evidence for the action–risk premise requires external pharmacology or randomized endpoints at two or more ordered actions; developing a finite-sample monotonicity gate for that experiment is a natural extension.

Corollary 15 (A nontrivial uncertainty-expansion class). *Suppose the versions share a nondecreasing dose map and use $q_{\mathcal{G}}(x; z) = \hat{\theta}(x) - zs(x)$. If*

$$s'(x) \geq s(x), \quad \hat{\theta}'(x) - \hat{\theta}(x) \leq z_0\{s'(x) - s(x)\} \tag{51}$$

for every admissible x , then $\mathcal{G}' \preceq_{z_0} \mathcal{G}$. In particular, increasing the uncertainty scale while leaving the estimate and dose map fixed transfers the certificate for every $z_0 \geq 0$.

Proof. Step 1: Expand the coordinate difference. For fixed x and $z \geq z_0$,

$$\begin{aligned}q_{\mathcal{G}'}(x; z) - q_{\mathcal{G}}(x; z) &= \{\hat{\theta}'(x) - zs'(x)\} - \{\hat{\theta}(x) - zs(x)\} \\ &= \hat{\theta}'(x) - \hat{\theta}(x) - zs'(x) + zs(x) \\ &= \hat{\theta}'(x) - \hat{\theta}(x) - z\{s'(x) - s(x)\}.\end{aligned}$$

Step 2: Use the margin and scale orders. Since $z \geq z_0$ and $s'(x) - s(x) \geq 0$,

$$-z\{s'(x) - s(x)\} \leq -z_0\{s'(x) - s(x)\}.$$

Therefore

$$\begin{aligned}q_{\mathcal{G}'}(x; z) - q_{\mathcal{G}}(x; z) &\leq \hat{\theta}'(x) - \hat{\theta}(x) - z_0\{s'(x) - s(x)\} \\ &\leq 0,\end{aligned}$$

where the final inequality is [equation \(51\)](#).

Step 3: Complete the transfer argument. Thus $q_{\mathcal{G}'}(x; z) \leq q_{\mathcal{G}}(x; z)$ for every admissible x and $z \geq z_0$. The shared nondecreasing dose map preserves this order, and [theorem 14](#) gives $\mathcal{G}' \preceq_{z_0} \mathcal{G}$. If $\hat{\theta}' = \hat{\theta}$, the second condition in [equation \(51\)](#) reduces to $0 \leq z_0(s' - s)$, which holds for every $z_0 \geq 0$. □

The condition is restrictive for statistical refits because a history-dependent movement of $\hat{\theta}$ can violate the uniform inequality even when predictive error improves. In the case study, the one-unit effective-margin update satisfies the condition, but it is a deliberately conservative release patch rather than a model refit. The only realistic refit examined—the held-out shrinkage refit—reduces the reported scale by the factor $\sqrt{0.001}$ and therefore violates the required condition $s'(x) \geq s(x)$ for every member of the cohort. Thus no realistic refit in this case study belongs to the outcome-free transfer class; in practice, every refit proceeds to the single-anchor endpoint procedure below. This is the principal operational role of the recertification theorem.

Bounded single-anchor recertification

Suppose an update fails the structural screen. Keep the previously selected anchor z^* fixed and collect n new, independent endpoints under π_{G', z^*} . Write K_n for their event count and

$$c_n(\alpha, \delta_r) = \max \{k \in \{-1, 0, \dots, n\} : \Pr\{\text{Binomial}(n, \alpha) \leq k\} \leq \delta_r\}. \quad (52)$$

Theorem 16 (Exact recertification cost). *The rule that confirms the update when $K_n \leq c_n(\alpha, \delta_r)$ has*

$$\sup_{R_{G'}(z^*) \geq \alpha} \Pr\{K_n \leq c_n(\alpha, \delta_r)\} \leq \delta_r. \quad (53)$$

For a desired power $1 - \beta$ at risk $\alpha - \gamma$, the exact required increment is

$$n_{\text{exact}}(\alpha, \delta_r, \beta, \gamma) = \min \left\{ n \geq 1 : \Pr_{\alpha - \gamma} \{K_n \leq c_n(\alpha, \delta_r)\} \geq 1 - \beta \right\}. \quad (54)$$

It tests one arm regardless of the original ladder size J . A conservative bound of order

$$n = O[\gamma^{-2} \max \{\log(1/\delta_r), \log(1/\beta)\}] \quad (55)$$

follows from Hoeffding's inequality.

Proof. Step 1: Write the exact rejection region. Let $r = R_{G'}(z^*)$. Under independent sampling at the fixed anchor,

$$K_n \sim \text{Binomial}(n, r).$$

By [equation \(52\)](#), confirmation occurs exactly on the event $\mathcal{C}_n = \{K_n \leq c_n(\alpha, \delta_r)\}$.

Step 2: Bound confirmation under every unsafe risk. For each integer c , the lower tail of a binomial variable is nonincreasing in its success probability. Hence, for $r \geq \alpha$,

$$\begin{aligned} \Pr_r(\mathcal{C}_n) &= \Pr\{\text{Binomial}(n, r) \leq c_n(\alpha, \delta_r)\} \\ &\leq \Pr\{\text{Binomial}(n, \alpha) \leq c_n(\alpha, \delta_r)\} \\ &\leq \delta_r. \end{aligned}$$

Taking the supremum over $r \geq \alpha$ proves [equation \(53\)](#).

Step 3: Compute exact power and invert it. At the design alternative $r = \alpha - \gamma$, the confirmation probability is

$$\text{Power}(n) = \sum_{k=0}^{c_n(\alpha, \delta_r)} \binom{n}{k} (\alpha - \gamma)^k (1 - \alpha + \gamma)^{n-k}.$$

Requiring this sum to be at least $1 - \beta$ and choosing the smallest integer n is exactly [equation \(54\)](#).

Step 4: Derive a conservative order bound. Define

$$t_n = \sqrt{\frac{\log(1/\delta_r)}{2n}}, \quad h_n = \lfloor n(\alpha - t_n) \rfloor.$$

Under $r = \alpha$, Hoeffding's inequality gives

$$\begin{aligned} \Pr_{\alpha}(K_n \leq h_n) &\leq \Pr_{\alpha}\left(\frac{K_n}{n} - \alpha \leq -t_n\right) \\ &\leq \exp(-2nt_n^2) \\ &= \exp\{-\log(1/\delta_r)\} = \delta_r. \end{aligned}$$

Thus h_n is an admissible lower-tail cutoff and, by maximality in [equation \(52\)](#), $h_n \leq c_n(\alpha, \delta_r)$. Under $r \leq \alpha - \gamma$ and $t_n \leq \gamma/2$,

$$\begin{aligned} \Pr_r(K_n > h_n) &\leq \Pr_r\left\{\frac{K_n}{n} > \alpha - t_n\right\} \\ &\leq \exp\{-2n(\gamma - t_n)^2\} \\ &\leq \exp(-n\gamma^2/2). \end{aligned}$$

Therefore power is at least $1 - \beta$ whenever

$$n \geq \frac{2}{\gamma^2} \max\{\log(1/\delta_r), \log(1/\beta)\},$$

apart from the harmless integer rounding in h_n . This yields [equation \(55\)](#). □

Corollary 17 (Bivariate single-anchor recertification). *At a fixed updated anchor, compute the toxicity and progression lower-tail tests with ceilings (α_T, α_P) and confirm only when both reject at level δ_r . Then the probability of confirming an anchor that violates either ceiling is at most δ_r . If marginal design powers $1 - \beta_T$ and $1 - \beta_P$ are required, choosing $n \geq \max(n_T, n_P)$, with each n_e obtained from [equation \(54\)](#), meets both marginal power requirements; the joint power must be evaluated under the specified competing-event law.*

Proof. Step 1: Confirmation is the intersection of the two component rejection events. Under the union null, at least one component null is true.

Step 2: The intersection is contained in the rejection event for that true component, whose probability is at most δ_r by [theorem 16](#). This proves the size statement without assuming endpoint independence.

Step 3: For endpoint e , the exact power is nondecreasing once a design size satisfying its inversion is selected. Thus $n \geq n_e$ meets its marginal target. Taking the maximum satisfies both marginal requirements, while the probability that both reject depends on their joint distribution. □

The initial study pays for exploration of a family; recertification pays only for confirmation of the already deployed coordinate. This is an $O(1)$ versus $O(J)$ comparison in the number of arms, not a claim that the endpoint count is uniformly small. Near the risk ceiling, the one-arm sample can still be a large fraction of the discovery trial. The Hoeffding calculation establishes only the γ^{-2} order and is far too conservative for study design; exact binomial inversion should be used. If the single-anchor test fails, no safety conclusion is drawn. The sponsor may then collect more data under a new, prespecified design or return to a more conservative margin whose status is separately justified.

Anytime surveillance

The likelihood-ratio process used here is a standard nonnegative supermartingale construction for time-uniform inference [21]. It is included as the monitoring component of the lifecycle protocol, not as a new sequential-testing result. After certification, let Y_t be the successive observed adverse endpoints under the deployed version and margin, and let $\mathcal{F}_{t-1}$ contain the past. For a fixed alternative $q \in (\alpha, 1)$ define

$$E_t(q) = \prod_{s=1}^t \left(\frac{q}{\alpha}\right)^{Y_s} \left(\frac{1-q}{1-\alpha}\right)^{1-Y_s}, \quad E_0(q) = 1. \quad (56)$$

A mixture $E_t = \int E_t(q) dW(q)$ over a prespecified probability measure W on $(\alpha, 1)$ avoids committing to one deterioration magnitude. For a Bernoulli risk r , the expected log-evidence increment at alternative q is

$$g(r; q) = r \log \frac{q}{\alpha} + (1-r) \log \frac{1-q}{1-\alpha}. \quad (57)$$

This quantity determines the first-order detection-delay benchmark used in the simulation.

Theorem 18 (Anytime-valid deterioration signal). *Under the sequential safety null*

$$\mathbb{E}(Y_t \mid \mathcal{F}_{t-1}) \leq \alpha \quad \text{for every } t, \quad (58)$$

$\{E_t(q)\}$ and every fixed mixture $\{E_t\}$ are nonnegative supermartingales. Consequently, for any $\eta \in (0, 1)$,

$$\Pr \left\{ \sup_{t \geq 1} E_t \geq 1/\eta \right\} \leq \eta. \quad (59)$$

The crossing time is valid under continuous inspection and optional stopping.

Proof. Step 1: Separate the one-step likelihood factor. Define

$$L_t(q) = \left(\frac{q}{\alpha}\right)^{Y_t} \left(\frac{1-q}{1-\alpha}\right)^{1-Y_t},$$

so that $E_t(q) = E_{t-1}(q)L_t(q)$ and $L_t(q) > 0$.

Step 2: Expand its conditional expectation. Put $\mu_t = \mathbb{E}(Y_t \mid \mathcal{F}_{t-1})$. Since Y_t is binary,

$$\begin{aligned} \mathbb{E}\{L_t(q) \mid \mathcal{F}_{t-1}\} &= \Pr(Y_t = 1 \mid \mathcal{F}_{t-1}) \frac{q}{\alpha} \\ &\quad + \Pr(Y_t = 0 \mid \mathcal{F}_{t-1}) \frac{1-q}{1-\alpha} \\ &= \mu_t \frac{q}{\alpha} + (1-\mu_t) \frac{1-q}{1-\alpha} \\ &= \frac{1-q}{1-\alpha} + \mu_t \left\{ \frac{q}{\alpha} - \frac{1-q}{1-\alpha} \right\}. \end{aligned}$$

The coefficient of μ_t is

$$\begin{aligned} \frac{q}{\alpha} - \frac{1-q}{1-\alpha} &= \frac{q(1-\alpha) - \alpha(1-q)}{\alpha(1-\alpha)} \\ &= \frac{q-\alpha}{\alpha(1-\alpha)} > 0. \end{aligned}$$

Step 3: Use the sequential safe null. Because the conditional expectation is increasing in μ_t and $\mu_t \leq \alpha$,

$$\begin{aligned}\mathbb{E}\{L_t(q) \mid \mathcal{F}_{t-1}\} &\leq \alpha \frac{q}{\alpha} + (1 - \alpha) \frac{1 - q}{1 - \alpha} \\ &= q + (1 - q) = 1.\end{aligned}$$

Therefore

$$\begin{aligned}\mathbb{E}\{E_t(q) \mid \mathcal{F}_{t-1}\} &= E_{t-1}(q) \mathbb{E}\{L_t(q) \mid \mathcal{F}_{t-1}\} \\ &\leq E_{t-1}(q),\end{aligned}$$

proving the nonnegative-supermartingale property.

Step 4: Pass to a mixture and apply Ville's inequality. For a fixed probability measure W , conditional Tonelli's theorem gives

$$\begin{aligned}\mathbb{E}(E_t \mid \mathcal{F}_{t-1}) &= \int \mathbb{E}\{E_t(q) \mid \mathcal{F}_{t-1}\} dW(q) \\ &\leq \int E_{t-1}(q) dW(q) = E_{t-1}.\end{aligned}$$

Also $E_0 = 1$. Ville's inequality for a nonnegative supermartingale now yields

$$\Pr \left\{ \sup_{t \geq 1} E_t \geq \frac{1}{\eta} \right\} \leq \eta E_0 = \eta,$$

which is [equation \(59\)](#). □

The monitor tests the safe null $\mu_t \leq \alpha$ against deterioration. This orientation is essential: an e-process for the opposite null would accumulate evidence of safety and cannot serve as a harm alarm. For the bivariate claim, run one process with (Y_t^T, α_T) and one with (Y_t^P, α_P) . If their alarm budgets are η_T and η_P , the union bound gives familywise false-alarm probability at most $\eta_T + \eta_P$, regardless of endpoint dependence.

At a toxicity alarm, increasing the margin is a plausible mitigation; at a progression alarm, the same change may be harmful. Therefore no automatic one-direction escalation is authorized. The response is to pause or constrain deployment, review endpoint ascertainment and population drift, and choose only among prespecified arms that already possess a valid joint certificate; any other action requires renewed bivariate confirmation. When processes are restarted after interventions, stagewise component budgets must satisfy $\sum_s (\eta_{T,s} + \eta_{P,s}) \leq \eta_{\text{total}}$.

An e-process alarm identifies deterioration of the registered endpoint process but does not attribute it to the algorithm. The registered triage therefore verifies event coding and follow-up, compares the current covariate and site distributions with the certified population, audits the deployed executable and input pipeline, and, when ethically and operationally feasible, examines a prespecified concurrently randomized reference-policy stream. A toxicity alarm does not authorize automatic increase of z , because the same action may worsen progression. Deployment is paused or constrained to a jointly certified fallback until attribution is resolved; population or measurement change voids transport of the old certificate, whereas an algorithm-attributed change requires renewed bivariate confirmation.

Dosing Algorithm Safety Label

Every released generator is paired with a versioned Dosing Algorithm Safety Label. The Label prevents the numerical margin from being separated from the population, endpoint, evidence, and change-control conditions that give it meaning. The companion `dosing_algorithm_safety_label.json` stores the same fields in a machine-readable form.

Table 2: Minimum fields of the Dosing Algorithm Safety Label. The entries shown are the worked simulation instance, not a clinically authorized claim.

Field	Required versioned content
Certified object	Generator version, software and data-pipeline identifiers, deployed margin z^* , admissible-history domain $\mathcal{X}$, and release date. The worked generator uses $z^* = 4$.
Claim and scope	Target population, treatment strategy, horizon, competing toxicity and progression endpoints, and ceilings. Here $(\alpha_T, \alpha_P) = (0.12, 0.25)$; the separate planner ceiling is $\theta_{\text{tox}} = 0.10$.
Endpoint calibration	Common functions (b_T, b_P) , registered grid, fresh version-margin evidence where required, and lifecycle allocation $\sum_{v,j} \delta_{v,j} \leq \delta_C$. This calibrates the policy consequence of s , not its interpretation as a standard error.
Evidence identity	Protocol, dataset, code and environment hashes; trial/site identifier; enrollment and analysis dates; arm sizes and endpoint counts; and responsible signatories. The present evidence is synthetic.
Transfer and recertification	Joint-risk dominance permits transfer. Any other refit requires bivariate exact testing at the deployed anchor. A lower action alone transfers at most the toxicity component.
Monitoring and response	Registered toxicity and progression e-process mixtures, stagewise error budgets, attribution triage, jointly certified fallback, pause criteria, and renewed-confirmation rule.
Void conditions	Unregistered change in population, eligibility, endpoint, horizon, competing-event handling, adherence, ascertainment, generator, software, or data pipeline; unresolved monitoring alarm; or failure of a transfer premise.
Non-claims	No individual-patient guarantee, no implication that prediction accuracy alone establishes safety, and no clinical or regulatory authorization from the simulation.

Methods

Case-study generator and online estimator

Reference system and synchronized predictor

The *reference system* is the hidden stochastic data-generating process with parameter $\theta^* = \log(\text{CL}/F)^*$, process perturbations, dose jumps, and measurement channel [equation \(12\)](#). It is intentionally generated from nearly the same deterministic structure as the predictor. This choice creates a controlled numerical experiment; it does not mix parameter uncertainty with uncontrolled between-subject differences. A separate omitted-memory regime supplies the structural-mismatch test.

The *synchronized predictor* consists of the deterministic augmented flow, the parameter belief $(\hat{\theta}_n, P_n)$, and the model state regenerated from the known impulse history. At each observation it predicts the log concentration, applies the scalar update below, and recomputes future trajectories. The static comparator fixes $\hat{\theta}_n = \log(51.8)$ and ignores observations. In the continuation experiment the MAP and filter point estimates use [equation \(14\)](#), whereas the endpoint rule uses [equation \(19\)](#). This synchronized estimate-predict-decide loop provides the common architecture for every planner comparison.

Recursive clearance estimation

Scalar extended Kalman parameter filter.

Only apparent clearance is estimated. Set $\theta = \log(\text{CL}/F)$ and let

$$h_n(\theta) = \log C_1(\tau_n; \theta, D_{0:n}) \quad (60)$$

be the deterministic log-trough predicted from the dosing history available at τ_n . The filter treats

θ as a slowly varying scalar with variance inflation Q . Given $(\hat{\theta}_{n-1}, P_{n-1})$, it computes

$$P_n^- = P_{n-1} + Q, \quad (61)$$

$$J_n = \left. \frac{\partial h_n}{\partial \theta} \right|_{\hat{\theta}_{n-1}}, \quad K_n = \frac{P_n^- J_n}{J_n^2 P_n^- + R_y}, \quad (62)$$

$$\hat{\theta}_n = \hat{\theta}_{n-1} + K_n \{\log y_n - h_n(\hat{\theta}_{n-1})\}, \quad (63)$$

$$P_n = (1 - K_n J_n)^2 P_n^- + K_n^2 R_y. \quad (64)$$

The final line is the scalar Joseph covariance update. The sensitivity J_n is evaluated by a centred difference with step 0.02 in log clearance. The scalar filter uses no clearance projection, and P_n is treated as an algorithmic local scale whose calibration is evaluated empirically.

The model state is regenerated from the known dose history under the current clearance estimate, keeping the adaptive layer focused on the parameter that drives the terminal decision. A maximum-a-posteriori comparator minimises the same log-domain residual sum of squares plus the Gaussian population-prior penalty.

Algorithm 1 One observation update of the scalar clearance filter.

- 1: **Input:** $(\hat{\theta}_{n-1}, P_{n-1})$, dose history, y_n at τ_n
 - 2: Predict $h_n(\hat{\theta}_{n-1})$ with the exact propagator $\exp(G(\hat{\theta}_{n-1})\Delta t)$
 - 3: Evaluate J_n by two additional propagations at $\hat{\theta}_{n-1} \pm 0.02$
 - 4: Apply [equations \(63\) and \(64\)](#)
 - 5: **Output:** $(\hat{\theta}_n, P_n)$ and forecasts under $\hat{\theta}_n$
-

Choice of estimated parameter.

Apparent clearance is the reported population-PK quantity and directly controls the derived elimination rate $k_{10} = (\text{CL}/F)/(V_1/F)$. The study therefore estimates this one scalar while fixing volume and transfer parameters. Daily troughs have non-zero local sensitivity to clearance throughout the simulated schedule; the observed innovation sequence provides the empirical check reported in the Results. Joint clearance–volume estimation is reserved for richer sampling designs with an explicit sensitivity-rank analysis.

Computational experiment

Simulation design

The horizon is two sunitinib Schedule-4/2 cycles: 28 daily dosing days followed by 14 off-treatment days, repeated once, for $T = 84$ days and 56 dose events. The exact matrix exponential advances the PK state on a $\Delta t = 1$ h grid. Unless otherwise specified, tracking uses 37.5 mg on each scheduled dosing day. Observations are taken one hour before dosing on days 1–27 of the first cycle.

Stochastic reference ensemble and observation noise

The hidden apparent clearance is generated as

$$(\text{CL}/F)^* = 51.8 \exp(\eta), \quad \eta \sim \mathcal{N}(0, \omega^2), \quad \omega = \sqrt{\log(1 + 0.38^2)} = 0.3673. \quad (65)$$

The 38 % coefficient of variation is the sunitinib interindividual clearance variability reported by [19]. With master seed 2024, the 2000 draws span 13.01–163.01 L/h and have an empirical clearance coefficient of variation of 38.5 %. The first 150 draws used for the trajectory-level tracking diagnostic span 18.13–121.46 L/h. For each draw, $Q/F = 107.32$ L/h, $V_1/F = 2030$

L and $V_2/F = 1453.43$ L are held fixed, while k_{10} and the decay modes change with clearance. Here “terminal half-life” means $\log(2)/|\alpha|$, where α is the spectral abscissa of the augmented 4×4 PK-memory block (the effect-site mode is excluded). Under this definition the expanded cohort spans 23.5–221.3 h; classical two-compartment macro-constants give a different range.

Each reference trajectory applies a mean-preserving multiplicative perturbation to the central amount at each step,

$$A_1 \leftarrow A_1 \exp \left\{ \sigma_{\text{proc}} \sqrt{\Delta t} Z - \frac{1}{2} \sigma_{\text{proc}}^2 \Delta t \right\}, \quad \sigma_{\text{proc}} = 0.03 \text{ h}^{-1/2}, \quad (66)$$

with independent standard-normal Z . Assay observations follow [equation \(12\)](#) with $\sigma = 0.10$. The process perturbation is a simulation device and is not a fitted physiological stochastic differential equation.

For a fixed realization of the multiplicative factors in [equation \(66\)](#), every hourly update is linear in the state: it is a matrix propagation followed by a diagonal multiplication. Induction and continuation dose responses therefore superpose exactly along the shared noise path. Induction and unit-continuation basis trajectories are computed once, and every candidate continuation is their linear combination. This identity preserves the paired design while making scalar calibration inexpensive.

Filter settings

The filter begins at $\hat{\theta}_0 = \log(51.8)$ with $P_0 = \omega^2$, uses $Q = 10^{-5}$ and $R_y = 0.10^2$, and applies [algorithm 1](#) once per observation. The maximum-a-posteriori comparator uses the same Gaussian prior for θ . Its scalar numerical search interval extends six log-standard deviations on either side of the population value; this interval does not alter or truncate the virtual cohort.

For each of 150 paired stochastic realizations, we record three realization-level errors:

1. the final absolute relative error in CL/F after the day-27 update;
2. the root-mean-square error of the filtered trough prediction over the 27 observation days; and
3. the absolute error in the terminal toxicity incidence $F_2(T)$ forecast made with the day-27 clearance estimate.

The static comparator uses 51.8 L/h for every realization. Reported percentage reductions are $100(1 - \bar{e}_{\text{adaptive}}/\bar{e}_{\text{static}})$. Paired 95% intervals are percentile intervals from 4000 realization-level bootstrap resamples, with fixed resampling seeds.

Paired day-7 continuation experiment

Every realization first receives 25 mg on each scheduled dosing day before day 7. One-hour pre-dose observations on days 1–7 are used to form the day-7 belief. The full-cohort diagnostic compares four planners that choose the common dose for every remaining scheduled dosing day:

Static population clearance 51.8 L/h;

MAP the posterior mode of log clearance;

Certainty-equivalent
 the scalar filter estimate $\hat{\theta}_7$;

Nominal radius, $z = 1$
 the lower log-clearance endpoint $\hat{\theta}_7 - \sqrt{P_7}$, or equivalently the clearance $\exp(\hat{\theta}_7 - \sqrt{P_7})$.

The primary decision experiment uses all 2000 systems. The dose curve is computed by 24 bisection iterations on an 801-point log-clearance grid and linearly interpolated for each planner; values beyond the grid are assigned the corresponding dose bound. The chosen continuation is then applied to the hidden reference system. All four arms use the same parameter realization, induction history, observations, and process-noise seed; their trajectories are identical up to the decision time. Thus, only the chosen post-decision continuation differs between planners.

The primary certification experiment implements the exact randomized endpoint procedure above. Before any randomized endpoint is generated, we fix the broad ordered margin ladder

$$\mathcal{Z} = (0, 0.2, 0.5, 1, 2, 4, 8.5, 10, 16) \quad (67)$$

to cover the observed risk transitions in both regimes and include a fallback. Within each fixed randomization block, a system is sampled from the population and Y_i is generated from [equation \(22\)](#). We evaluate $N_j \in \{250, 500, 1000, 2000\}$, use $(\alpha_T, \delta) = (0.12, 0.05)$, and apply the ordered exact tests in [equation \(27\)](#). The primary operating characteristics use the exact law in [proposition 7](#); 1000 complete block-randomized trials at $N_j = 1000$ check the full sampling pipeline.

That nine-arm calculation is a methodological stress test, not a proposed trial. The clinically interpretable design is centered at the worked current-practice anchor $z = 4$, which reduces mean continuation dose by 8.11 mg relative to $z = 0$. Its ethical fallback set is restricted to $z \in \{4, 5, 6\}$; patients are not randomized to the near-unadjusted $z = 0$ or extreme $z = 16$ policies. A primary $n = 750$ single-anchor confirmation records the mutually exclusive toxicity and progression indicators and applies [theorem 5](#) with $(\alpha_T, \alpha_P, \delta) = (0.12, 0.25, 0.05)$. Failure yields no deployment claim. Any later fallback is evaluated with fresh prespecified error spending, not selected opportunistically from the same endpoints.

Ladder sensitivity compares the primary ladder with a five-arm coarse ladder and with transition-enriched ladders. A simultaneous exact-binomial Bonferroni rule provides a multiplicity-controlled comparator. The primary method is also evaluated over $\alpha \in \{0.08, 0.10, 0.12\}$ and over $\delta \in \{0.01, 0.05, 0.10\}$. Finally, multiplying the terminal hazard slope β_2 by 0.9 or 1.1 evaluates outcome-layer perturbation while leaving the controller’s planning map fixed. Arm risks and their Monte Carlo standard errors are computed from the 2000-system ensemble; those standard errors describe numerical integration over the ensemble, not external-population uncertainty.

For each planner we report the toxicity-ceiling violation rate with a Wilson 95% interval, mean realised $F_1(T)$ and $F_2(T)$, and the mean and standard deviation of the maintenance dose. Violation intervals are Wilson 95% intervals. The $N = 1000$ evaluation half gives an approximately 3.7 percentage-point 95% interval width near a 10% rate; the full $N = 2000$ paired experiment can resolve differences of about two percentage points with approximately 80% power. The $N = 150$ tracking experiment supports the trajectory-level bootstrap summaries; safety-rate estimation uses the $N = 2000$ ensemble.

Sensitivity experiments repeat the paired comparison for $\tau_m \in \{0, 0.5, 2.6667, 8, 24\}$ h and $\theta_{\text{tox}} \in \{0.075, 0.10, 0.125, 0.15\}$. A structural-mismatch case uses $\tau_m = 8$ h in the reference system and a classical ($\tau_m = 0$) predictor and planner. The same hidden clearances, assay draws, and process-noise paths are reused in every paired comparison.

Lifecycle experiments

Seven prespecified studies illustrate the lifecycle results without claiming clinical validation. First, the matched and omitted-memory risk curves are treated as two versions in the same synthetic population; this is the negative transfer example. Second, a positive transfer update replaces every requested margin $z \geq 4$ by the base generator’s action at $z + 1$. This release patch is pointwise no more aggressive by construction. We verify the action inequality on all 2,000 simulated histories and compare both population-risk curves on $z \in \{4, 5, 6, 8.5, 10, 12, 16\}$.

Third, for $(\alpha_T, \delta_r, 1 - \beta) = (0.12, 0.05, 0.80)$ we evaluate [equation \(54\)](#) over updated risks from 0.06 to 0.11. For the $n = 218$ design point at risk 0.07 we additionally run 5,000 end-to-end confirmation trials. Each trial resamples patients from the $N = 2,000$ closed-loop population—including filter error and process noise—and draws one Bernoulli endpoint per patient. The heterogeneous probabilities are scaled to marginal risks 0.12 and 0.07 to study size and designed power.

Fourth, the accuracy–safety counterexample is moved into the PK–PD simulator. On one half of the cohort, a scalar shrinkage refit $\hat{\theta}' = \mu + \lambda(\hat{\theta} - \mu)$ selects λ from a fixed grid by log-trough RMSE; performance is evaluated on the other half. The updated software reports 0.001 times the old posterior variance. Both versions then act at $z = 4$ in the same noisy reference system. This construction tests the claim at the estimator–policy interface rather than in an external Gaussian toy.

Fifth, the endpoint-scale calibration study registers at $z = 4$ the common envelope $(b_T, b_P) = (0.10, 0.25)$. It applies the bivariate exact gate to 5,000 fresh blocks of $n = 1000$ under both the matched generator and the collapsed-scale refit. This directly tests whether the two implementations can attach the same endpoint meaning to the same numerical coordinate.

Sixth, monitoring operating characteristics use 3,000 repeated toxicity streams per condition over 4,000 endpoints. Boundary-null streams have risk 0.12; drifted streams change after endpoint 200 from 0.075 to 0.1262, 0.14, or 0.16. The primary equal-weight mixture $\{0.13, 0.14, 0.16, 0.18, 0.22\}$ spans small through severe operationally relevant deterioration. Sensitivity mixtures are a local grid $\{0.125, 0.13, 0.14, 0.15\}$ and a broad grid $\{0.13, 0.17, 0.22, 0.30\}$. We report boundary false alarms, detection probability, and the median and 10th–90th percentile detection delay. A fixed seed-2029 path is retained only as an auditable example of escalation.

Seventh, [equation \(25\)](#) is evaluated on a dense action grid to show the disconnected safe set that invalidates the scalar margin order. The primary ladder $(0, 0.2, 0.5, 1, 2, 4, 8.5, 10, 16)$ is intentionally irregular: it combines a broad stress-test range with finer resolution around the empirical risk transition. It was fixed before endpoint draws and is not claimed to be optimal.

Numerical verification

A 31×801 dose–log-clearance surface verifies the two monotonicity conditions used by [proposition 2](#): strictly positive dose differences and strictly negative log-clearance differences at every sampled adjacent pair. Zero-continuation risk is evaluated over the same grid to check feasibility of every reported toxicity ceiling.

The verification suite additionally checks the algebraic mass-balance row, the Metzler property, generator eigenvalues, the competing-risk simplex identity, the $\tau_m \downarrow 0$ limit, and dose-reachable impulse responses at 0.1 h resolution over the full horizon and clearance grid. Terminal outcomes are recomputed with 1, 0.5, 0.25, and 0.125 h hazard grids.

Reproducibility

The accompanying Python package is supplied as Supplementary Information. It includes the analysis implementation, result files, and figure-generation scripts. Random streams use NumPy’s `default_rng`(2024) with the PCG64 bit generator; package versions and all secondary seeds are recorded in the environment file.

Results

Model checks

The implemented generator is Metzler and satisfies the augmented amount balance algebraically. Its mass row has coefficient $-k_{10}$ on A_1 and zero on every other state, with maximum absolute

error 6.94×10^{-18} in double precision. The full generator spectral abscissa is $-3.610 \times 10^{-3} \text{ h}^{-1}$, exactly the effect-site mode $-k_{e0}$; the four-state PK-memory block has spectral abscissa $-1.160 \times 10^{-2} \text{ h}^{-1}$ at reference parameters.

Proposition 1 gives positivity for the continuous model. The implementation also returned nonnegative impulse responses on a 0.1 h grid over 2016 h, the full clearance range, and every swept memory time. The exact competing-risk balance $S(T) + F_1(T) + F_2(T) = 1$ was preserved to 2.44×10^{-15} , and the memory model approached the classical flow with relative central-concentration error 1.83×10^{-6} at $\tau_m = 10^{-4} \text{ h}$. These checks support the numerical decision study on its reported dose-reachable domain.

Table 3: Algebraic and numerical checks for the dynamical model.

Check	Result
mass-row maximum absolute error	6.94×10^{-18}
full-generator spectral abscissa	$-3.610 \times 10^{-3} \text{ h}^{-1}$
PK-memory-block spectral abscissa	$-1.160 \times 10^{-2} \text{ h}^{-1}$
PK block Metzler	yes
dose-reachable values, all swept τ_m	nonnegative
direct-recursive terminal discrepancy	7.33×10^{-21}
maximum $ S + F_1 + F_2 - 1 $	2.44×10^{-15}
relative C_1 error at $\tau_m = 10^{-4} \text{ h}$ versus classical flow	1.83×10^{-6}

Tracking and terminal-outcome forecasting

The adaptive filter reduces every prespecified tracking endpoint (table 4). The final mean relative clearance error is 4.66 %, compared with 28.63 % for the static population value. The filtered-trough RMSE is 0.00254 mg/L, compared with 0.00746 mg/L, and the mean absolute terminal toxicity-incidence forecast error is 0.00513, compared with 0.03530. The paired bootstrap intervals exclude zero reduction for all three endpoints.

Table 4: Static and adaptive tracking errors in 150 stochastic realizations. Intervals are paired percentile-bootstrap intervals for the percentage reduction.

Endpoint	Static	Adaptive	Reduction (95% interval)
final relative CL/F error	0.2863	0.0466	83.7% (80.3–86.4%)
filtered-trough RMSE (mg/L)	0.007462	0.002544	65.9% (62.0–69.3%)
terminal F_2 forecast error	0.03530	0.005131	85.5% (82.9–87.9%)

Figure 1 shows a pre-specified lower-quartile realization, chosen by clearance and not by estimator performance. The physical trajectory contains process perturbations that neither deterministic predictor reproduces, but the recursive clearance estimate moves away from the population prior and toward the hidden value. The cohort curves in figure 2 show that parameter error falls over the observation period while concentration error remains below the static comparator. The larger relative improvement in terminal F_2 forecasting is consistent with its use of the full estimated exposure trajectory, whereas a trough reflects one time point. These results motivate calibrating the remaining decision error empirically.

The 4050 log-domain innovations have mean 0.00240 and standard deviation 0.167. In the expanded $N = 2000$ day-7 experiment, the innovation standard deviation is 0.183 (variance ratio 3.36 relative to $R_y = 0.01$), mean posterior standard deviation is 0.0521, and the log-clearance error standard deviation is 0.1208. The radius is therefore interpreted as a filter-scale multiple and calibrated empirically. Inflating R_y to 0.0336 raises nominal one-radius coverage from 35.9%

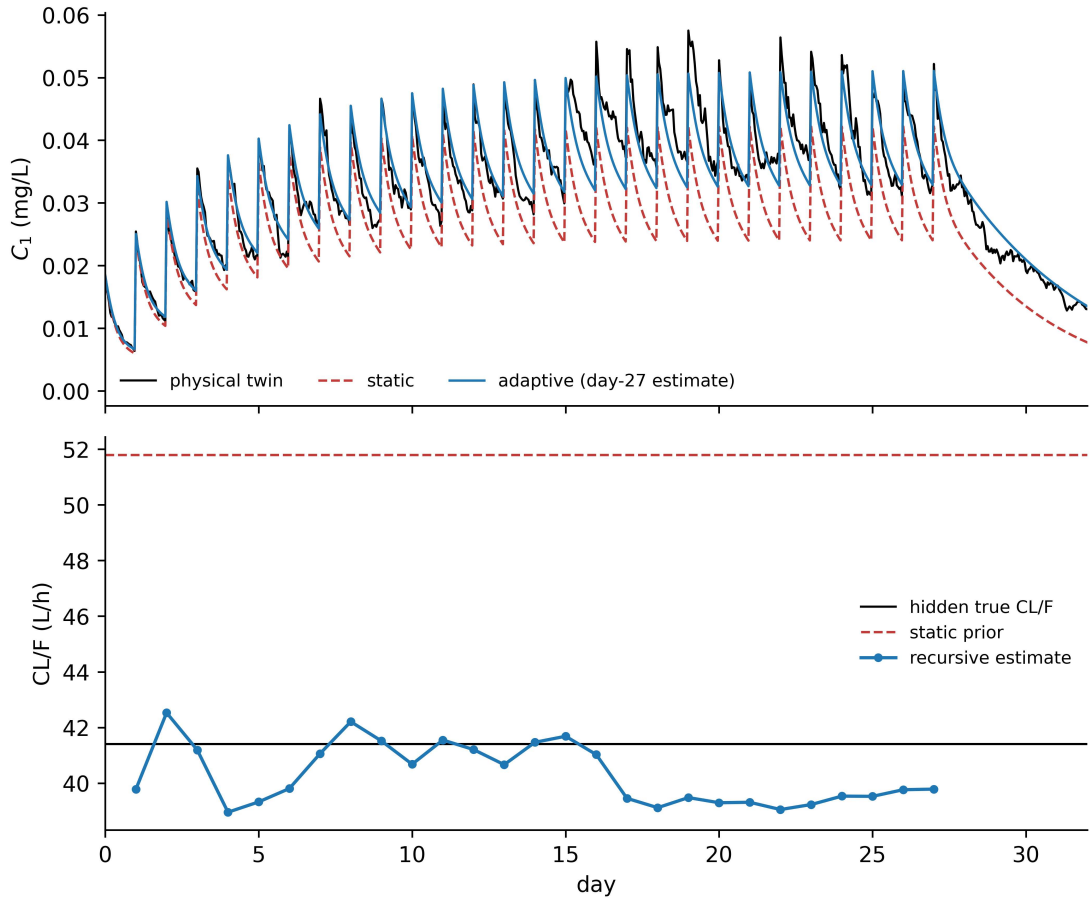


Fig. 1 Pre-specified lower-quartile stochastic realization. Top: stochastic physical concentration and deterministic predictions. Bottom: hidden apparent clearance, static prior and recursive estimate

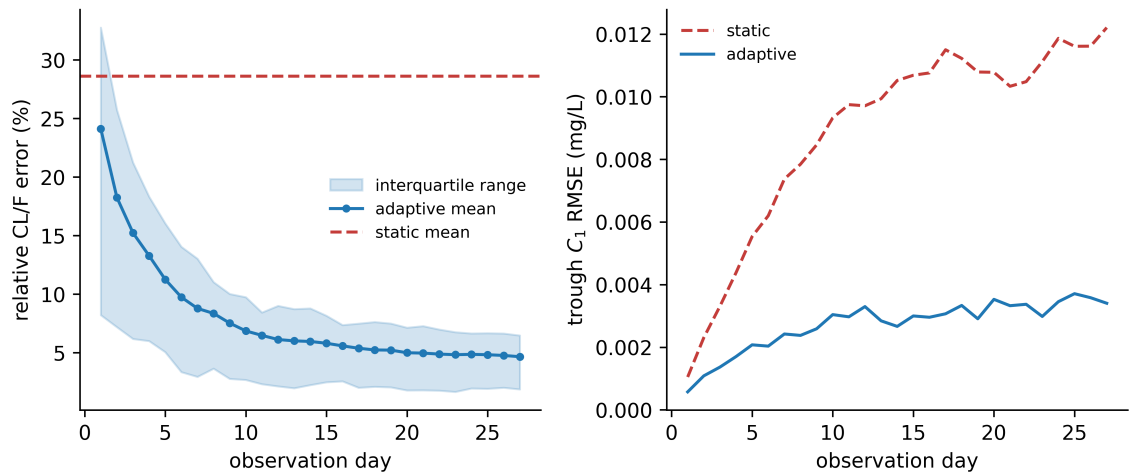


Fig. 2 Population tracking across the 27 observations. Left: adaptive relative clearance error with interquartile range and the static mean. Right: trough-concentration RMSE across realizations

to 60.9%. The omitted-memory run also remains stable with the positive transit realization: the maximum absolute log-clearance error is 0.955, and none of the 2000 systems has error exceeding one.

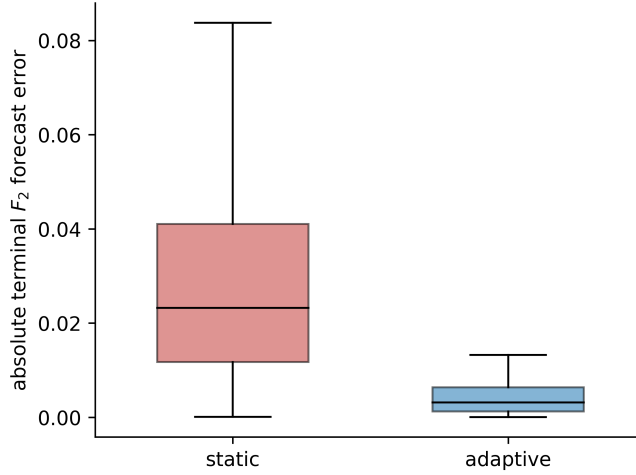


Fig. 3 Absolute terminal toxicity-incidence forecast error after the day-27 update, static and adaptive models ($N = 150$)

Day-7 continuation and realised safety

The expanded paired experiment shows the value of personalization before calibration (table 5). Certainty-equivalent adaptation reduces mean $F_2(T)$ from 0.12176 to 0.10145 (16.7%), while the nominal-radius rule reduces it further to 0.09387. The corresponding change in $F_1(T)$ and dose makes the efficacy–toxicity trade-off explicit. The near-50% zero-radius rate agrees with the saturation identity at the population-median design and motivates selecting the radius from held-out decision outcomes.

Table 5: Realised outcomes and post-day-7 doses in the paired continuation experiment. “Rate” is the latent oracle diagnostic $V_T = \Pr\{F_2(T; d) > \theta_{\text{tox}}\}$ estimated from $N = 2000$ paired systems, with Wilson 95 % intervals; F_1 and F_2 are ensemble means, not binary-event confidence limits.

Planner	Violations /2000	Rate (95% CI)	$F_1(T)$	$F_2(T)$	Dose, mg mean (SD)
Static	1016	50.8% [48.6, 53.0]	0.18120	0.12176	48.18 (0) ^a
MAP	951	47.6% [45.4, 49.7]	0.17643	0.10044	49.80 (16.12)
Certainty-equivalent	986	49.3% [47.1, 51.5]	0.17544	0.10145	50.04 (16.01)
Nominal radius, $z = 1$	686	34.3% [32.3, 36.4]	0.18296	0.09387	47.92 (16.24)

^aThe static dose has zero standard deviation by construction.

Observed-endpoint certification

The primary ladder in equation (67) spans the observable event-risk transitions in both dynamical regimes (table 6). In the matched system every tested toxicity-event risk is below 12%; in the omitted-memory system the risk crosses 12% between $z = 4$ and $z = 8.5$. The ladder was fixed without using randomized endpoint realizations. Its low, intermediate, and high anchors deliberately provide resolution near both transitions and a strong fallback policy. Across ladder arms, the ensemble Monte Carlo standard errors range from 0.019 to 0.051 percentage points in

the matched regime and from 0.029 to 0.115 percentage points under omitted memory. Because risks are averages of individual event probabilities, no zero-count substitution is involved.

Table 6: Population operating characteristics of the primary margin ladder. Risk is $R_T(z) = \mathbb{E}[F_2(T; d_z)]$, the toxicity-event cumulative incidence in the $N = 2000$ Monte Carlo system ensemble; no sampling interval is asserted for a clinical population. Dose is the ensemble mean.

Margin	Matched		Omitted memory	
	Risk	Dose, mg	Risk	Dose, mg
0.00	10.14%	50.04	15.64%	63.02
0.20	9.99%	49.61	15.48%	62.70
0.50	9.76%	48.98	15.24%	62.23
1.00	9.39%	47.92	14.84%	61.44
2.00	8.69%	45.86	14.07%	59.83
4.00	7.49%	41.94	12.62%	56.57
8.50	5.47%	34.09	9.83%	49.33
10.00	4.96%	31.76	9.05%	47.02
16.00	3.47%	23.73	6.58%	38.56

The exact law in [proposition 7](#) gives the toxicity-component operating characteristics in [table 7](#). At $N_j = 1000$ and $(\alpha_T, \delta) = (0.12, 0.05)$, the selected policies average 9.50% and 9.31% toxicity-event risk, with mean doses 48.19 and 47.68 mg in the matched and omitted-memory regimes. The probabilities of selecting a toxicity-unsafe arm are 0 and 0.634%, respectively; non-selection is negligible in both regimes. These unsafe-selection probabilities lie well below $\delta = 0.05$ because δ is a uniform finite-sample upper bound rather than a target operating probability. The separation quantifies the additional safety margin produced by the exact discrete cutoff and the risk spacing of the prespecified ladder, while selection probability and retained dose measure operating efficiency.

The exact toxicity cutoff is $K_{T,j} \leq 102$ at $N_j = 1000$. The certifier receives only event counts and fixed block sizes. These single-component results illustrate the ladder mechanics; clinical acceptance additionally requires the progression gate in [theorem 5](#).

Table 7: Exact toxicity-component ordered certification at $(\alpha_T, \delta) = (0.12, 0.05)$. Selected risk is $R_T = \mathbb{E}[F_2(T; d_z)]$ conditional on selection; entries are exact population operating characteristics, not confidence intervals. N_j is per arm.

Regime	N_j	Unsafe select	No select	Selected risk	Dose, mg
Matched	250	0	0.0072%	7.46%	41.44
Matched	500	0	< 0.001%	8.59%	45.35
Matched	1000	0	< 0.001%	9.50%	48.19
Matched	2000	0	< 0.001%	10.00%	49.65
Omitted memory	250	0.225%	10.26%	7.69%	42.32
Omitted memory	500	0.395%	0.618%	8.39%	44.64
Omitted memory	1000	0.634%	0.00064%	9.31%	47.68
Omitted memory	2000	0.516%	< 0.001%	9.77%	49.14

Bivariate certificate

[Table 8](#) closes the zero-dose loophole and separates the stress test from a plausible confirmation study. At the worked current-practice anchor $z = 4$, mean continuation dose is 41.94 mg, 8.11 mg below the near-unadjusted $z = 0$ policy. Its toxicity and progression risks are 7.49% and 20.82%. Among 5,000 independent synthetic $n = 750$ trials, 82.80% reject both component nulls (Wilson

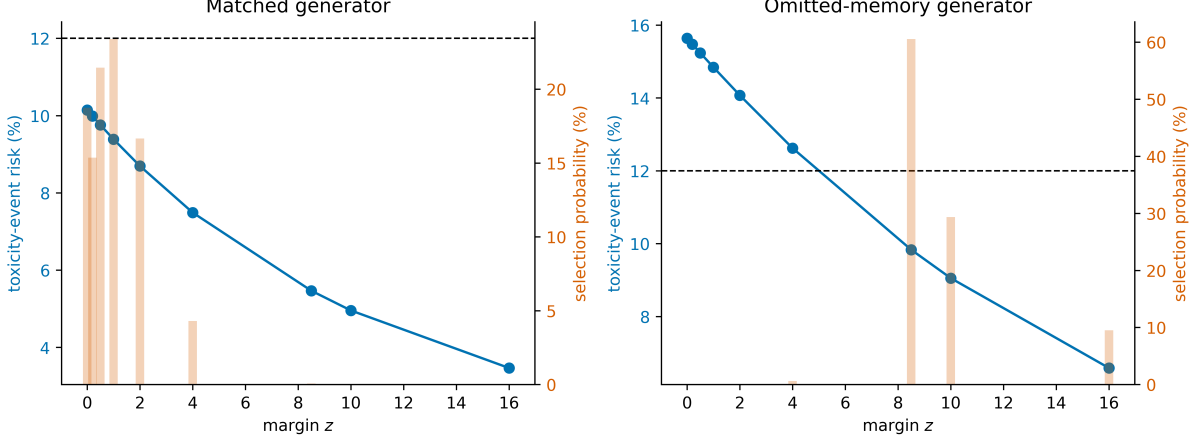


Fig. 4 Toxicity-component observed-endpoint certification on the primary margin ladder. Curves show $R_T = \mathbb{E}[F_2]$ in the $N = 2000$ Monte Carlo population and bars show exact fixed-sequence selection probabilities for $N_j = 1000$, $\alpha_T = 0.12$, and $\delta = 0.05$; no interval is plotted

95% interval 81.73%–83.82%). By contrast, zero continuation has toxicity risk only 1.01% but progression risk 87.68%; none of the 5,000 trials jointly certifies it (interval 0–0.077%).

Table 8: Bivariate feasibility operating points. $R_T = \mathbb{E}[F_2]$ and $R_P = \mathbb{E}[F_1]$ are $N = 2000$ Monte Carlo population cumulative incidences; dose is the ensemble mean. The $z = 4$ and zero-dose certification probabilities use 5,000 independent trials of $n = 750$ with exact cutoffs $(K_T, K_P) \leq (75, 167)$; parentheses are Wilson 95% intervals across trials.

Policy	Dose, mg	R_T	R_P	Joint certify
Current anchor $z = 4$	41.94	7.49%	20.82%	82.80% (81.73–83.82)
Ethical fallback $z = 5$	40.08	6.97%	21.77%	not simulated
Ethical fallback $z = 6$	38.29	6.49%	22.76%	not simulated
Zero continuation	0.00	1.01%	87.68%	0% (0–0.077)

The proposed evidence plan enrolls at the anchor rather than allocating 1,000 systems to each of nine widely separated margins. The guarded $\{4, 5, 6\}$ set is centered on current practice, uses no zero-dose or extreme arm, and supplies only prespecified fallbacks. This reduces the immediate endpoint burden from 9,000 to 750 but gives up discovery of a continuous upper ray and comparisons across the broad ladder. A data-monitoring plan, early stopping for excess toxicity or progression, multicenter transport checks, and an independent confirmatory cohort would still be required before clinical use. Response-adaptive allocation is possible only with a separately prespecified adaptive design and valid conditional or martingale inference; the fixed-binomial proof cannot simply be reused after outcome-driven reallocation.

Cross-version interpretation of the margin

We instantiate [definition 9](#) at the shared coordinate $z = 4$ with the stricter common envelope $(b_T(4), b_P(4)) = (0.10, 0.25)$. This is distinct from the deployment toxicity ceiling $\alpha_T = 0.12$: it is the price of claiming that “four scale units” has the same operational meaning after an update. With $n = 1000$, the exact component cutoffs are $(K_T, K_P) \leq (84, 227)$. The matched generator lies inside the envelope and is accepted in 81.96% of 5,000 repeated calibration blocks (Wilson 95% interval 80.87%–83.00%).

The shrinkage refit improves held-out log-trough RMSE but multiplies the reported variance by 0.001. At the same numerical $z = 4$, its mean dose rises from 41.94 to 49.73 mg, toxicity

risk rises from 7.49% to 10.11%, and progression risk falls from 20.82% to 17.56%. Because the toxicity component now lies outside $b_T(4) = 0.10$, accepting the common coordinate meaning is a type-I error; it occurs in only 4.04% of repeated blocks (Wilson interval 3.53%–4.62%), consistent with the exact 5% bound. Thus the collapsed scale is not merely an anecdotal failure of prediction-based safety: the endpoint-calibration gate detects its policy-relevant understatement while leaving harmless reparameterizations alone.

Table 9: Endpoint calibration of the common coordinate $z = 4$. Risks are $R_T = \mathbb{E}[F_2]$ and $R_P = \mathbb{E}[F_1]$ in the $N = 2000$ Monte Carlo population. Acceptance probabilities use 5,000 fresh synthetic blocks of $n = 1000$ and exact envelope cutoffs (84, 227); parentheses are Wilson 95% intervals across repeated blocks.

Version	Dose, mg	R_T	R_P	Envelope accepted
Matched scale	41.94	7.49%	20.82%	81.96% (80.87–83.00)
Collapsed-scale refit	49.73	10.11%	17.56%	4.04% (3.53–4.62)

The fixed-sequence order improves efficiency relative to the simultaneous Bonferroni certifier. At $N_j = 1000$, its mean selected dose is 42.43 mg versus 39.93 mg in the matched regime and 40.77 mg versus 39.19 mg under omitted memory. A coarse five-arm ladder gives 41.48 and 40.73 mg; adding arms near the matched and omitted-memory transitions raises these values to 43.63 and 42.68 mg in their respective regimes. These comparisons leave the finite-sample guarantee unchanged and quantify only design efficiency.

Across the primary (α, δ) sweep at $N_j = 1000$, the selected mean risk ranges from 5.31% to 9.77% in the matched regime and from 6.58% to 9.61% under omitted memory. Increasing the event-hazard slope by 10% moves the two $z = 0$ arm risks to 12.83% and 20.44%, while the $z = 16$ fallback remains at 3.97% and 8.00%. Thus observed-endpoint certification retains a certifiable fallback even when the outcome layer, not only the kinetic predictor, is perturbed.

Lifecycle stress-test results

Figure 5 reports the negative transfer, endpoint-cost, simulator-refit, and auditable monitoring-path calculations. At the worked anchor $z = 4$, the matched population risk is 7.49%, but the same coordinate under the omitted-memory version has risk 12.62%. Thus the base matched certificate cannot be transferred merely because both versions use the same dimensionless margin. The updated implementation must either pass the structural envelope in [theorem 14](#) or obtain new endpoints. Escalating the omitted-memory version to $z = 8.5$ lowers its ensemble risk to 9.83% (Monte Carlo 95% interval 9.73%–9.93%), but this numerical fact is simulation evidence, not an outcome-free transfer proof and not evidence that escalation restores safety in a changed clinical population.

Table 10 states the practical endpoint comparison. The saving is dramatic for a comfortable margin but falls as the updated risk approaches 12%. The Hoeffding calculation remains an order result, not a study-design recommendation.

Table 10: Exact one-anchor recertification burden relative to the initial nine-arm, 9,000-endpoint ladder; 80% power, $(\alpha_T, \delta_r) = (0.12, 0.05)$. Risk is the toxicity-event estimand $R_T = \mathbb{E}[F_2]$; n_{exact} is an exact binomial design quantity, not an interval estimate.

Updated risk (%)	6.0	7.0	8.0	9.0	10.0	10.5	11.0
n_{exact}	149	218	353	659	1536	2771	6332
Initial-ladder endpoints (%)	1.7	2.4	3.9	7.3	17.1	30.8	70.4

The 5,000 repeated $n = 218$ toxicity-component confirmation trials give 5.12% boundary

confirmations at $R_T = 0.12$ (Wilson 95% interval 4.54%–5.77%) around the exact 4.99% size, and 81.14% confirmations at $R_T = 0.07$ (interval 80.03%–82.20%) around the exact 80.81% power. Thus the simulation reproduces both the finite-sample boundary calculation and the designed alternative in the heterogeneous closed-loop population.

The positive-transfer safety patch increases the effective margin from z to $z + 1$; it is an engineered conservatism rule, not a model refit. Across all 2,000 histories and seven tested margins on the certified ray, the largest updated-minus-base dose is 0 mg. The updated risk at $z = 4$ is 6.97% versus 7.49% for the base, and the updated curve remains below the base at every tested point. This exercises only the toxicity-component implication of [theorem 14](#); because progression increases as dose falls, it is not a free transfer of the bivariate certificate.

The held-out shrinkage refit selects $\lambda = 0.925$ and improves log-trough RMSE from 0.08154 to 0.08010. After its reported variance is multiplied by 0.001, however, its $z = 4$ mean dose rises from 41.94 to 49.73 mg and policy risk rises from 7.49% to 10.11%. A predictive improvement can therefore worsen toxicity risk and cross the planner’s 10% target when the uncertainty channel is changed, although this particular refit remains below the distinct 12% certification ceiling. Because this refit has $s'(x) = \sqrt{0.001} s(x) < s(x)$ throughout the cohort, it fails [equation \(51\)](#). It is the only realistic refit examined; accordingly, no realistic refit in this study receives outcome-free transfer, and every refit requires single-anchor bivariate recertification.

For the primary toxicity e-process mixture, 3.70% of 3,000 boundary-null streams alarm by endpoint 4,000 (Wilson interval 3.08%–4.44%). At drift risks 0.1262, 0.14, and 0.16, detection probabilities are 3.87%, 68.47%, and 99.97%; median delays among detected streams are 2740, 2327, and 1031 endpoints. At risk 0.14, local, primary, and broad grids have median delays 2179, 2327, and 2834. The single seed-2029 path crosses at endpoint 3080 and is illustrative only. For the primary $M = 5$ mixture and $\eta = 0.05$, the registered growth benchmark in [equation \(57\)](#) is approximately 2544 endpoints at $r_{\min} = 0.14$. At the much smaller drift 0.1262 it is about 39,255, well beyond the simulated horizon; this makes the low 3.87% detection probability transparent rather than presenting the mixture as uniformly fast.

Table 11: Repeated-run e-process sensitivity over 3,000 streams per cell. False alarm is evaluated for Y^T at boundary risk $R_T = 0.12$ through endpoint 4,000; entries are median post-change delays among detected streams, with 3,000 repetitions rather than confidence intervals for a clinical population.

Mixture grid	False alarm (%)	Drift 0.1262	Drift 0.14	Drift 0.16
Local (.125, .13, .14, .15)	3.00	2467	2179	944
Primary (.13, .14, .16, .18, .22)	3.70	2740	2327	1031
Broad (.13, .17, .22, .30)	3.57	3130	2834	1256

Figure 6 adds the positive theory, repeated recertification, and the competing-risk scope failure. In the fold example the 20% safe set is disconnected: incidence rises from 0 at action 0 to 25.8% at action 2, then falls to 9.4% at action 6 because the competing event increasingly pre-empts toxicity. The scalar upper-ray certificate is therefore inapplicable to that mechanism.

Decision-map verification

The complete dose–parameter surface is strictly decreasing in log clearance and strictly increasing in dose at all sampled adjacent pairs. The least-negative clearance difference is -3.56×10^{-6} and the least-positive dose difference is 2.89×10^{-4} . At zero continuation dose the planner toxicity functional is small, but the population experiment gives $(R_T, R_P) = (0.0101, 0.8768)$. Thus zero dose makes the univariate toxicity constraint feasible precisely by abandoning efficacy. The lower bound is not a clinically acceptable fallback; bisection verifies only the planner’s toxicity geometry, whereas deployment requires the bivariate endpoint gate.

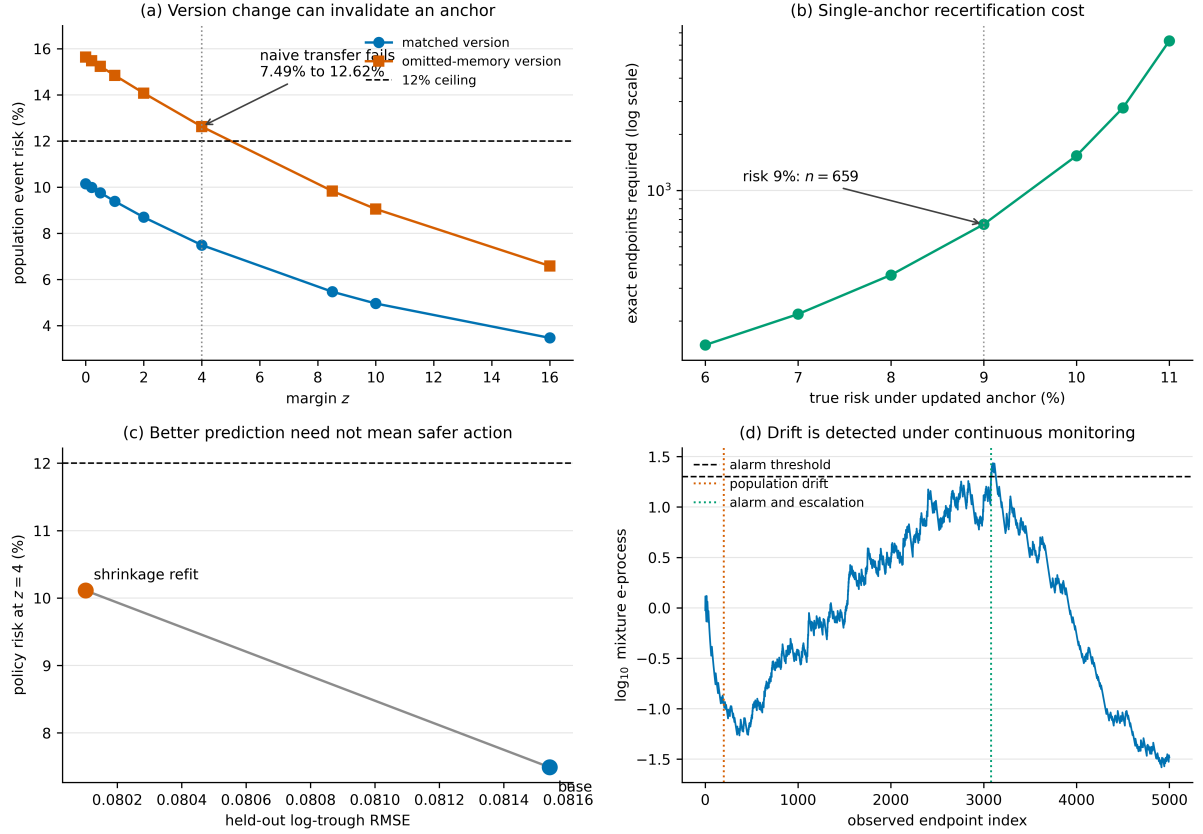


Fig. 5 Lifecycle stress tests. (a) The matched $z = 4$ anchor becomes unsafe under the omitted-memory policy generator, so version identity of the margin does not imply transfer. (b) Exact single-anchor sample sizes for 80% power at $(\alpha_T, \delta_r) = (0.12, 0.05)$. (c) A held-out shrinkage refit with lower trough RMSE but understated uncertainty has higher deployed risk. (d) A fixed mixture e-process detects the imposed risk increase under continuous monitoring; the vertical green line marks alarm and escalation

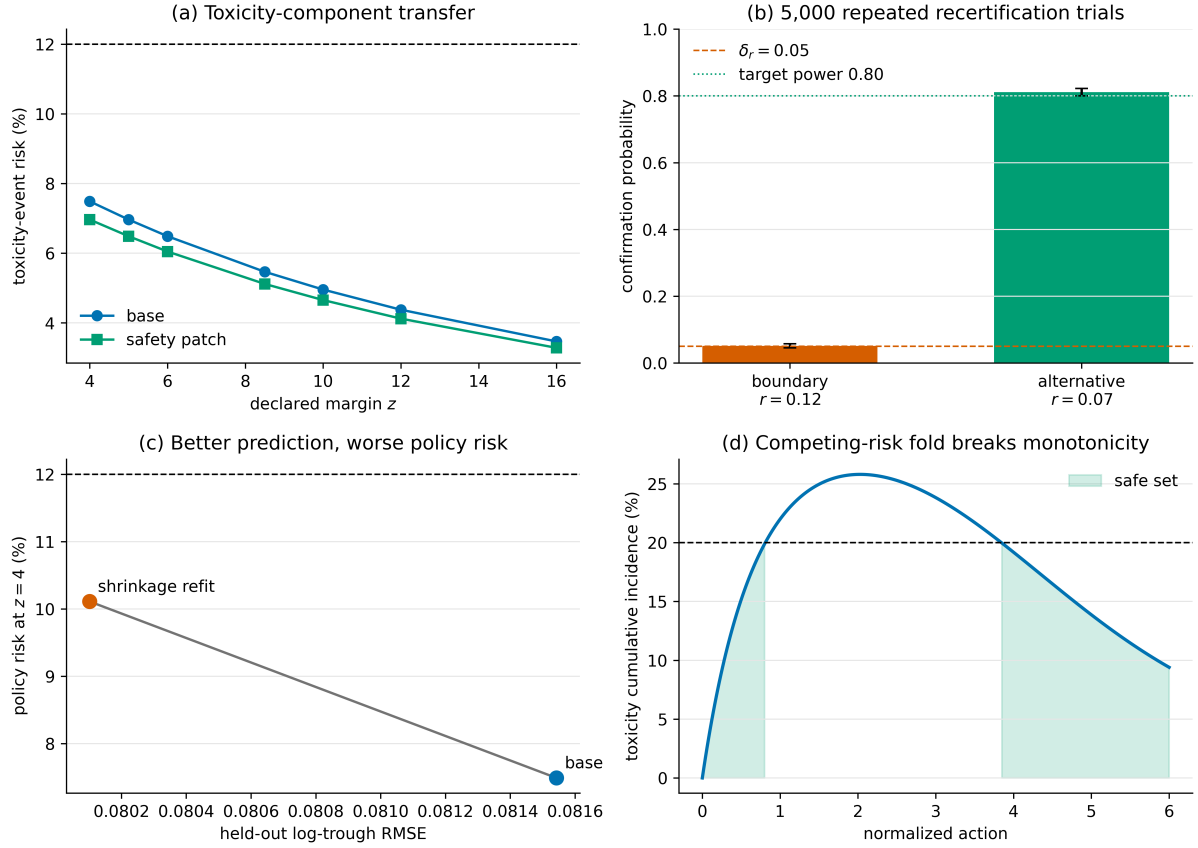


Fig. 6 Lifecycle evidence. (a) A one-unit effective-margin safety patch satisfies pointwise action dominance and has no larger simulated toxicity risk $R_T = \mathbb{E}[F_2]$ on the declared ray; this is not joint transfer. (b) Repeated exact one-anchor toxicity trials reproduce boundary size and designed power; bars show Wilson 95% intervals. (c) A held-out shrinkage refit improves trough prediction but, after severe variance understatement, worsens policy risk. (d) A numerical competing-risk mechanism folds, producing a disconnected safe-action set

Discussion

The central result is that a model-generated margin becomes comparable across algorithm versions only after its endpoint consequences are calibrated on a common scale. The algebraic rescaling result shows that the pair (s, z) has no unique numerical representation at the policy level. The common functions (b_T, b_P) remove this ambiguity by assigning each registered value of z a toxicity–progression risk box. Exact component tests and lifecycle error spending then make that interpretation simultaneous across accepted releases. In the case study, this distinction was consequential: the matched generator at $z = 4$ lay inside the common envelope, whereas collapsing its reported scale increased toxicity risk from 7.49% to 10.11% and reduced the empirical acceptance probability from 81.96% to 4.04%.

The bivariate formulation is essential. Toxicity decreases as the continuation dose becomes more conservative, but progression moves in the opposite direction. Zero continuation consequently achieved a toxicity risk of only 1.01% while producing an 87.68% progression risk. Requiring both exact component tests eliminates this degeneracy without assuming independence between the competing outcomes. It also explains why a continuous upper-ray claim is appropriate for the toxicity component but not for the joint clinical endpoint.

The two update routes certify different evidence and compose without conflict. A jointly dominated release inherits the base endpoint certificate and, if the base pair was endpoint calibrated, also inherits the same risk-box meaning of z . No fresh endpoint block is needed because $R_{e,G'}(z) \leq R_{e,G}(z) < b_e(z)$ for both components. Toxicity-only action dominance is insufficient because it says nothing about progression. When joint dominance cannot be established, the new version is a new empirical object and must be confirmed at the deployed margin.

This latter route is the practically relevant one for refitting. The one-unit effective-margin construction demonstrates the structural theorem, but it is an engineered conservative patch. The held-out shrinkage update is a genuine refit: it slightly improves log-trough prediction while reducing the reported scale by $\sqrt{0.001}$, so it fails the transfer condition throughout the cohort and increases the deployed dose. No realistic refit studied here therefore belongs to the free-transfer class. This negative finding gives the single-anchor theorem its main operational significance: an update can be evaluated without repeating the complete margin-selection experiment.

The endpoint burden depends primarily on the gap between the updated risk and its ceiling. At toxicity risk 0.07, $n = 218$ endpoints give approximately 80% power at a one-sided 0.05 level; the repeated simulation recovered 81.14% confirmation and 5.12% boundary size. The required size rises rapidly near the 0.12 ceiling, so single-anchor recertification should be viewed as removal of the ladder multiplicity rather than as a uniformly small study. After deployment, the endpoint-specific e-processes preserve false-alarm control under continuous inspection. Their simulated delays also show that small deteriorations can require many endpoints, reinforcing the separation between prospective recertification and surveillance.

The conclusions rely on a fixed target population and horizon, valid competing endpoint ascertainment, randomized positive allocation, and a prespecified policy family. Structural transfer additionally requires scientific support for the relevant action–risk order; monotonicity of the internal simulator is not sufficient. The PK–PD system updates apparent clearance alone and uses a literature-informed competing-risk layer, so the numerical values illustrate method behavior rather than clinical performance of sunitinib dosing. Extension to censored endpoints, multiple adaptive decisions, rare toxicities, subgroup transport, and jointly estimated PK parameters requires additional design and inference.

Relative to model-space verification and decision-aware conformal calibration, the proposed framework places the invariant at the endpoint-policy interface. The internal model generates an ordered action family; randomized endpoints determine which member is acceptable; the common risk envelope gives its margin coordinate a version-independent interpretation; and

structural order or exact single-anchor tests determine what survives refitting. This division preserves the value of pharmacometric modeling while locating the statistical claim in observable clinical outcomes.

Conclusions

An uncertainty multiple is not a transferable safety margin unless its endpoint consequence is calibrated. Common toxicity–progression envelopes, exact randomized endpoint tests, and lifecycle error spending give the same registered z a common risk-box meaning across accepted algorithm versions. The bivariate selection rule controls erroneous certification in finite samples, including dependent competing endpoints.

Jointly dominated releases inherit both the certificate and the calibrated meaning of z without new endpoints. The realistic refit studied here does not satisfy that restrictive envelope and therefore requires bivariate single-anchor recertification, as do all realistic refits examined in the case study. The scale-collapse experiment and the zero-dose counterexample show why both requirements matter: uncertainty rescaling can invalidate cross-version interpretation, and toxicity control alone can select a clinically ineffective policy. Endpoint-calibrated margins therefore provide a coherent statistical coordinate for selecting, updating, and monitoring model-informed dosing algorithms.

Declarations

Funding. No funding was received for conducting this study.

Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval. This study did not involve human participants or animals; all data are simulated.

Data and code availability. The Python source code, fixed random seeds, numerical result files, and machine-readable Dosing Algorithm Safety Label needed to reproduce the results and figures are archived in Zenodo [6]. The archive includes a one-command reproduction driver and a checksum manifest.

Author contributions. Redouane Boudjemaa and Mohamed Boukhari contributed equally to conceptualization, methodology, software, formal analysis, investigation, visualization, and manuscript preparation.

Use of generative AI. Generative AI tools contributed to language, editing and code review. The authors independently reviewed all outputs and accept responsibility for the manuscript.

References

- [1] Angelopoulos AN, Bates S, Candès EJ, et al (2025) Learn then test: Calibrating predictive algorithms to achieve risk control. *The Annals of Applied Statistics* 19(2):1641–1662. <https://doi.org/10.1214/24-AOAS1998>

- [2] Bachmann F, Koch G, Pfister M, et al (2021) OptiDose: Computing the individualized optimal drug dosing regimen using optimal control. *Journal of Optimization Theory and Applications* 189(1):46–65. <https://doi.org/10.1007/s10957-021-01819-w>
- [3] Bar-Shalom Y, Tse E (1974) Dual effect, certainty equivalence, and separation in stochastic control. *IEEE Transactions on Automatic Control* 19(5):494–500. <https://doi.org/10.1109/TAC.1974.1100635>
- [4] Bates S, Angelopoulos AN, Lei L, et al (2021) Distribution-free, risk-controlling prediction sets. *Journal of the ACM* 68(6):43:1–43:34. <https://doi.org/10.1145/3478535>
- [5] Belfo JaP, Lemos JaM (2021) Optimal Impulsive Control for Cancer Therapy. *SpringerBriefs in Control, Automation and Robotics*, Springer, Cham, <https://doi.org/10.1007/978-3-030-50488-5>
- [6] Boudjemaa R, Boukhari M (2026) Policy guarantees for lifecycle change control. <https://doi.org/10.5281/zenodo.22541482>, URL <https://doi.org/10.5281/zenodo.22541482>
- [7] Campi MC, Kumar PR (1998) Adaptive linear quadratic Gaussian control: The cost-biased approach revisited. *SIAM Journal on Control and Optimization* 36(6):1890–1907. <https://doi.org/10.1137/S0363012997317499>
- [8] Chee KY, Silva TC, Hsieh MA, et al (2024) Uncertainty quantification and robustification of model-based controllers using conformal prediction. In: *Proceedings of the 6th Annual Learning for Dynamics and Control Conference*, pp 528–540, URL <https://proceedings.mlr.press/v242/chee24a.html>
- [9] Chen T, Kirkby NF, Jena R (2012) Optimal dosing of cancer chemotherapy using model predictive control and moving horizon state/parameter estimation. *Computer Methods and Programs in Biomedicine* 108(3):973–983. <https://doi.org/10.1016/j.cmpb.2012.05.011>
- [10] Darwich AS, Polasek TM, Aronson JK, et al (2021) Model-informed precision dosing: Background, requirements, validation, implementation, and forward trajectory of individualizing drug therapy. *Annual Review of Pharmacology and Toxicology* 61:225–245. <https://doi.org/10.1146/annurev-pharmtox-033020-113257>
- [11] Dibbets AC, Koldewei C, Osinga EP, et al (2025) Barriers and facilitators for bringing model-informed precision dosing to the patient’s bedside: A systematic review. *Clinical Pharmacology & Therapeutics* 117(3):633–645. <https://doi.org/10.1002/cpt.3510>
- [12] Dokoumetzidis A, Macheras P (2009) Fractional kinetics in drug absorption and disposition processes. *Journal of Pharmacokinetics and Pharmacodynamics* 36(2):165–178. <https://doi.org/10.1007/s10928-009-9116-x>
- [13] Elmecker-Plakolm L, Fasterling P, Sosnin P, et al (2025) Provably safe model updates. URL <https://arxiv.org/abs/2512.01899>, arXiv:2512.01899
- [14] Feng J, Emerson S, Simon N (2021) Approval policies for modifications to machine learning-based software as a medical device: A study of bio-creep. *Biometrics* 77(1):31–44. <https://doi.org/10.1111/biom.13379>
- [15] Finney DJ (1947) *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve*. Cambridge University Press, Cambridge
- [16] Formentin S, Chiuso A (2021) Control-oriented regularization for linear system identification. *Automatica* 127:109539. <https://doi.org/10.1016/j.automatica.2021.109539>

- [17] Hernández-Rivera A, Velarde P, Zafra-Cabeza A, et al (2025) Drug dosing for cancer therapy: A stochastic model predictive control perspective. *Journal of Theoretical Biology* 615:112255. <https://doi.org/10.1016/j.jtbi.2025.112255>
- [18] Hjalmarsson Hk (2005) From experiment design to closed-loop control. *Automatica* 41(3):393–438. <https://doi.org/10.1016/j.automatica.2004.11.021>
- [19] Houk BE, Bello CL, Kang D, et al (2009) A population pharmacokinetic meta-analysis of sunitinib malate (SU11248) and its primary metabolite (SU12662) in healthy volunteers and oncology patients. *Clinical Cancer Research* 15(7):2497–2506. <https://doi.org/10.1158/1078-0432.CCR-08-1893>
- [20] Houk BE, Bello CL, Poland B, et al (2010) Relationship between exposure to sunitinib and efficacy and tolerability endpoints in patients with cancer: Results of a pharmacokinetic/pharmacodynamic meta-analysis. *Cancer Chemotherapy and Pharmacology* 66(2):357–371. <https://doi.org/10.1007/s00280-009-1170-y>
- [21] Howard SR, Ramdas A, McAuliffe J, et al (2020) Time-uniform chernoff bounds via nonnegative supermartingales. *Probability Surveys* 17:257–317. <https://doi.org/10.1214/18-PS321>
- [22] Hurtado PJ, Kirosingh AS (2019) Generalizations of the linear chain trick: Incorporating more flexible dwell time distributions into mean field ODE models. *Journal of Mathematical Biology* 79(5):1831–1883. <https://doi.org/10.1007/s00285-019-01412-w>
- [23] International Council for Harmonisation (2026) ICH M15: General principles for model-informed drug development. Step 5 guideline, URL <https://www.fda.gov/media/184747/download>, accessed 5 September 2026
- [24] Jaki T, Clive S, Weir CJ (2021) A comparison of phase i dose-finding designs in clinical trials with monotonicity assumption violation. *Clinical Trials* 18(5):580–592. <https://doi.org/10.1177/17407745211010212>
- [25] Johnstone C, Cox B (2021) Conformal uncertainty sets for robust optimization. In: *Proceedings of the Tenth Symposium on Conformal and Probabilistic Prediction and Applications*, pp 72–90, URL <https://proceedings.mlr.press/v152/johnstone21a.html>
- [26] Kiyani S, Pappas GJ, Roth A, et al (2025) Decision theoretic foundations for conformal prediction: Optimal uncertainty quantification for risk-averse agents. In: *Proceedings of the 42nd International Conference on Machine Learning*, pp 30943–30965, URL <https://proceedings.mlr.press/v267/kiyani25a.html>
- [27] Lucia S, Schliemann-Bullinger M, Findeisen R, et al (2016) A set-based optimal control approach for pharmacokinetic/pharmacodynamic drug dosage design. *IFAC-PapersOnLine* 49(7):797–802. <https://doi.org/10.1016/j.ifacol.2016.07.286>
- [28] MacDonald N (1978) Time Lags in Biological Models, *Lecture Notes in Biomathematics*, vol 27. Springer-Verlag, Berlin, <https://doi.org/10.1007/978-3-642-93107-9>
- [29] Mesbah A (2018) Stochastic model predictive control with active uncertainty learning: A survey on dual control. *Annual Reviews in Control* 45:107–117. <https://doi.org/10.1016/j.arcontrol.2017.11.001>
- [30] Patel YP, Rayan S, Tewari A (2024) Conformal contextual robust optimization. In: *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, pp 2485–2493, URL <https://proceedings.mlr.press/v238/patel24a.html>

- [31] Popović JK, Atanacković MT, Pilipović AS, et al (2010) A new approach to the compartmental analysis in pharmacokinetics: fractional time evolution of diclofenac. *Journal of Pharmacokinetics and Pharmacodynamics* 37(2):119–134. <https://doi.org/10.1007/s10928-009-9147-3>
- [32] Prinster D, Fannjiang C, Park JW, et al (2026) Conformal policy control. <https://doi.org/10.48550/arXiv.2603.02196>, arXiv:2603.02196
- [33] Romano Y, Patterson E, Candès EJ (2019) Conformalized quantile regression. In: *Advances in Neural Information Processing Systems*, URL https://proceedings.neurips.cc/paper_files/paper/2019/hash/5103c3584b063c431bd1268e9b5e76fb-Abstract.html
- [34] Sasfi A, Zeilinger MN, Köhler J (2022) Robust adaptive MPC using control contraction metrics. URL <https://arxiv.org/abs/2209.11713>, arXiv:2209.11713
- [35] Smith HL (2011) *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, Texts in Applied Mathematics, vol 57. Springer, New York, <https://doi.org/10.1007/978-1-4419-7646-8>
- [36] Srinivasan A, Leeman A, Chou G (2026) Safety beyond the training data: Robust out-of-distribution MPC via conformalized system level synthesis. In: *Proceedings of the 8th Annual Learning for Dynamics and Control Conference*, pp 412–439, URL <https://proceedings.mlr.press/v331/srinivasan26a.html>
- [37] Thall PF, Cook JD (2004) Dose-finding based on efficacy–toxicity trade-offs. *Biometrics* 60(3):684–693. <https://doi.org/10.1111/j.0006-341X.2004.00218.x>
- [38] Turnbull BW (1976) The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society: Series B* 38(3):290–295
- [39] U.S. Food and Drug Administration (2025) Marketing submission recommendations for a predetermined change control plan for artificial intelligence-enabled device software functions: Guidance for industry and food and drug administration staff. Final guidance, URL <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence>, accessed 5 September 2026
- [40] U.S. Food and Drug Administration (2026) Project optimus: Reforming the dose optimization and dose selection paradigm in oncology. Oncology Center of Excellence web resource, URL <https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus>, accessed 6 September 2026
- [41] U.S. Food and Drug Administration, Health Canada, and Medicines and Healthcare products Regulatory Agency (2023) Predetermined change control plans for machine learning-enabled medical devices: Guiding principles. International guiding principles, URL <https://www.fda.gov/medical-devices/software-medical-device-samd/predetermined-change-control-plans-machine-learning-enabled-medical-devices-guiding-principles>, accessed 5 September 2026
- [42] Vitková Z, Dodek M, Miklovičová E, et al (2023) Robust control of repeated drug administration with variable doses based on uncertain mathematical model. *Bioengineering* 10(8):921. <https://doi.org/10.3390/bioengineering10080921>
- [43] Vovk V, Gammerman A, Shafer G (2005) *Algorithmic Learning in a Random World*. Springer, New York, <https://doi.org/10.1007/b106715>

- [44] Wages NA, Conaway MR, O’Quigley J (2011) Continual reassessment method for partial ordering. *Biometrics* 67(4):1555–1563. <https://doi.org/10.1111/j.1541-0420.2011.01560.x>
- [45] Yeh C, Christianson N, Wu A, et al (2024) End-to-end conformal calibration for optimization under uncertainty. URL <https://arxiv.org/abs/2409.20534>, [arXiv:2409.20534](https://arxiv.org/abs/2409.20534)
- [46] Ynineb AR, Yumuk E, Copot D, et al (2025) Modeling drug retention as memory effects in obese patients using fractional and augmented models. *Scientific Reports* 15:26234. <https://doi.org/10.1038/s41598-025-10172-1>
- [47] Zhou H, Zhang Y, Luo W (2024) Safety-critical control with uncertainty quantification using adaptive conformal prediction. In: 2024 American Control Conference, pp 3956–3962, <https://doi.org/10.23919/ACC60939.2024.10644391>