When can a negative recovery assay exclude a recoverable subpopulation?

Sharp coverage, stage mismatch and allocation bounds for paired observations

17 September 2026

Abstract

A finite recovery experiment that records no qualifying event is compatible both with the absence of recoverable units and with a failure to observe their recovery. We give a sharp conditional separation of the two explanations for a gated two-stage source observed under two alternative conditions, with activation and subsequent progeny probabilities (x, y) and (a, b). If the two conditions cover each stage, $x + a \geq c$ and $y + b \geq c$, and the within-condition stage mismatch obeys $|x - y| \leq d$, the equally weighted complete-path response is at least $R_{\text{full}}(c, d) = \left(c^2 - t^2\right)/4$ with $t = \min\{d, \min(c, 2 - c)\}$, and this is attained. Because c bounds a sum of two probabilities, its range is $[0, 2]$; on the restricted range $c \leq 1$ the bound reduces to $\max\{0, c^2 - d^2\}/4$, and for $c > 1$ a positive floor of at least $c - 1$ holds with no mismatch information at all. We compute the exact worst-case response of every fixed allocation weight w, which has three regimes, is maximized at $w = 1/2$, and is strictly suboptimal away from $w = 1/2$ exactly when the mismatch budget binds. The uniform mismatch bound can be weakened to a mean-square budget $\mathbb{E}[(x - y)^2] \leq D^2$ without changing the floor. With conditional recording probability at least κ on a covered fraction $1 - \eta$ of recoverable types, the usable floor is $g = (1 - \eta)\kappa R_{\text{full}}(c, D)$. For an even number n of independent eligible units split equally between the two conditions, and recoverable fraction at least θ, the worst-case all-negative probability is exactly $(1 - \theta g)^n$; a conditional calibration-failure allowance δ gives the sharp total bound $\delta + (1 - \delta)(1 - \theta g)^n$. A finite balanced design achieves total error α if and only if $\theta g > 0$ and $\delta < \alpha$, or $\theta g = 1$ and $\delta \leq \alpha$; no history-dependent reallocation of the same budget improves the guarantee. The coverage, allocation, mean-square, population and decision inequalities are verified in Lean 4. A synthetic example needs 2,300 units to exclude a recoverable fraction of one per cent at five per cent error under its declared premises, and 750 units when the coverage assumptions are strengthened beyond the restricted range. The result identifies a joint calibration requirement; it does not validate any particular assay, and non-detection is not evidence of irreversible death.

1 The question answered by a negative record

Failure to observe progeny after an exposure is consistent with several explanations: there were no recoverable units; recovery was delayed past the observation deadline; activation occurred but reproduction did not qualify; units were lost in handling; or the instrument failed to record a genuine event. These explanations support very different decisions. Statistical regrowth assays and the minimum-duration-for-killing framework are established tools for quantifying tolerance [4, 5], individual-level recovery has been followed since the phenotypic-switch experiments of Balaban et al. [2], and the vocabulary of resistance, tolerance and persistence has agreed definitions [3]. We do not propose new biological definitions. Our question is narrower and quantitative:

Under which explicit, checkable recovery assumptions does an all-negative finite experiment exclude a recoverable fraction of at least θ, and how many units does that take?

Without a restriction connecting recoverability to a finite observation, no uniform answer exists. If a recoverable fraction p has conditional detection-time distribution F, the observable cumulative event probability through the deadline T is $r(t) = pF(t)$; for every $p' \geq r(T)$ one may set $F' = r/p'$ on $[0, T]$ and place the remaining mass after T, reproducing the record exactly. This is the elementary form of a familiar obstruction in censored-event and cure-fraction models [1, 10], and is a partial-identification statement in the sense of Manski [11]. It is also why the classical zero-numerator devices [6, 9, 14] answer a different question: the rule of three bounds a *detectable* event rate, not a latent recoverable fraction, and the gap between the two is exactly what must be assumed or calibrated. The viable-but-nonculturable literature [12] is a standing reminder that the gap can be large.

The restriction studied here concerns two serial stages and two alternative observation conditions. It is not enough for the two conditions to cover each stage somewhere: the source state $(x, y, a, b) = (1, 0, 0, 1)$ has full first-stage coverage in condition 1 and full second-stage coverage in condition 2, yet neither condition completes a path, so no number of units observed under these conditions can detect it. We quantify the additional connection required. The cross-condition coverage level c must exceed the within-condition stage mismatch d, and the resulting floor composes with recording losses, hidden types, a fixed balanced design and a calibration allowance into an explicit finite sample size.

1.1 Contributions

1. **A sharp coverage floor on the full parameter range** (Theorem 4.1). Since c bounds the sum of two probabilities, the natural range is $0 \leq c \leq 2$. The sharp equally weighted complete-path floor is $R_{\text{full}}(c, d) = (c^2 - t^2)/4$ with $t = \min\{d, h(c)\}$ and $h(c) = \min(c, 2 - c)$. On $c \leq 1$ this is the positive-part formula $\max\{0, c^2 - d^2\}/4$; for $c > 1$ it is at least $c - 1$ *with no mismatch hypothesis* (Corollary 4.3). The frequently quoted ceiling of $1/4$ for such floors is an artefact of the restricted range, not a limit on paired recovery sensitivity.

2. **The exact allocation profile** (Theorem 5.1). For every allocation weight $w \in [0, 1]$ we compute the exact worst-case response, which has three regimes in $A = |2w - 1|$. Equal allocation is maximin; it is the *unique* maximin allocation exactly when the mismatch budget binds ($t < h(c)$), and otherwise the optimum is a plateau of width $h(c)/c$ in A (Corollaries 5.2 and 5.3). This replaces a numerical grid search, which cannot distinguish a plateau from a flat numerical optimum.

3. **Mismatch in mean square** (Theorem 6.1). The uniform per-type bound $|x_z - y_z| \leq d$ may be replaced by the population budget $\mathbb{E}[(x - y)^2] \leq D^2$ without changing the floor, which is again $R_{\text{full}}(c, D)$ and again attained. The uniform bound is the special case $D = d$. We also show which pooled summaries are *not* sufficient: equal stage means permit a zero floor, and the pooled second-stage conditional probability is $\mathbb{E}[xy]/\mathbb{E}[x]$, not $\mathbb{E}[y]$ (Examples 6.2 and 6.3).

4. **An exact finite decision rule** (Theorem 7.3 and Section 8). The all-negative certificate has worst-case error exactly $\delta + (1 - \delta)(1 - \theta g)^n$ with $g = (1 - \eta)\kappa R_{\text{full}}(c, D)$, the likelihood being derived from the source paths rather than postulated. A finite balanced design suffices if and only if $\theta g > 0$ and $\delta < \alpha$, or $\theta g = 1$ and $\delta \leq \alpha$ (Theorem 8.1); the second regime exists only because the coverage range was extended.

5. **No gain from adaptive reallocation** (Theorem 9.1). On the least detectable source both conditions respond identically, so every history-dependent allocation policy with the same budget and the same qualifying event faces the same bound.

6. **Machine-checked core.** The coverage bound and its witnesses, the allocation profile and its witnesses, the mean-square population floor, the balanced sampling identity and inequality, the calibration composition, and the feasibility characterization are verified in Lean 4 [7] against a pinned Mathlib [13]. Section 12 states exactly which results are compiled and which are conventional.

The mathematical ingredients are elementary: polynomial inequalities on a rectangle, finite probability sums, and a monotone one-dimensional optimisation. The contribution is the sharp, explicit and composable route from a source assumption to a finite decision, together with a precise account of what such a decision cannot do. In particular, nothing here licenses the inference that a unit which fails to recover by T is dead.

2 Target, source events and sampling units

Fix an exposure duration τ and a separate recovery-observation deadline T; these are different experimental clocks. Fix an intermediate activation time $t_A < T$ and a prespecified *reference task* that defines recoverability independently of whether the test conditions detect a unit by T. Let $p \in [0, 1]$ denote the fraction of reference-recoverable units in the population from which eligible units are sampled independently. A unit may be a cell, a cluster or a lineage; its meaning, sampling scheme and denominator must be fixed before p is interpreted.

For a recoverable latent type z, let A_j be activation by t_A and D_j the specified qualifying progeny event by T, in condition $j \in \{1, 2\}$, and write

$$x_z = \mathbb{P}(A_1 \mid z), \quad y_z = \mathbb{P}(D_1 \mid A_1, z), \quad a_z = \mathbb{P}(A_2 \mid z), \quad b_z = \mathbb{P}(D_2 \mid A_2, z). \tag{1}$$

By the chain rule the complete early-activation paths have probabilities $x_z y_z$ and $a_z b_z$; no independence between the stages is assumed. If the instrument also records qualifying paths with later activation, these two paths are sufficient subevents of the positive event, and the lower bounds below remain valid. When an activation probability is zero, the corresponding second-stage parameter is an off-path source parameter: it cannot be estimated by conditioning on observations that never occur, and any constraint involving it is structural or counterfactual rather than an observed conditional frequency.

Figure 1 shows the terminal source events; Table 1 fixes the symbols and their domains.

Type populations are finite, with weights $\mu_z \geq 0$ and $\sum_z \mu_z = 1$ *conditional on reference recoverability.* The same population supplies both groups, and assignment is independent of latent type; randomly assigning independent units to two fixed-size groups is one realization. Sorted subpopulations, repeated images of one unit, related descendants of one founder and sampling without replacement from a fixed specimen are *not* this model. Non-target units do not generate a positive in the exact source law; background events that only add positives preserve the one-sided negative-record bound, although reading a positive as proof of the target would require a separate specificity argument.

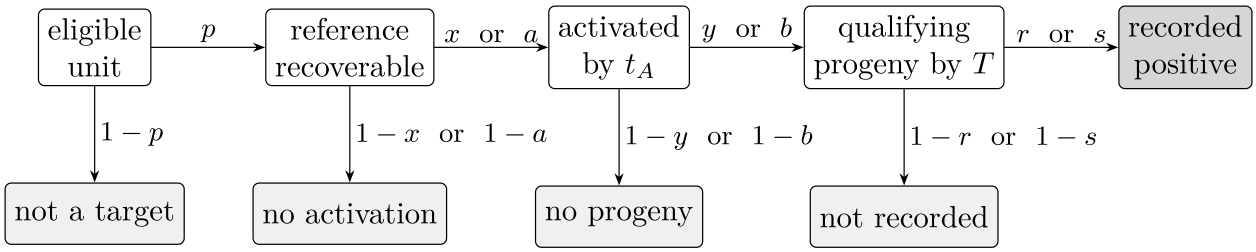

Figure 1: Terminal source events for one unit. Arrow labels are conditional probabilities for a fixed recoverable type: the first entry of each pair applies in condition 1 and the second in condition 2, so the two conditions differ only in which pair of stage probabilities governs the two gates. The four shaded states on the bottom row all produce a negative record, as does an unrecorded completed path. Failure at a gate contributes to the negative record but does not imply irreversible incapacity. The recording probabilities r and s are conditional on the completed path and the type.

Symbol	Domain	Meaning
p, θ	$p \in [0,1]$, $0 < \theta \le 1$	Recoverable fraction and exclusion threshold
x, y, a, b	$[0,1]$	Conditional stage probabilities of a recoverable type
c	$[0,2]$	Cross-condition stage-coverage lower bound
d, D	$[0,1]$	Uniform and mean-square within-condition mismatch bounds
$h(c), t$	$h(c) = \min(c, 2-c)$, $t = \min(d, h(c))$	Attainable stage width and the binding mismatch
κ, η	$[0,1]$	Conditional recording floor and exceptional recoverable mass
δ, α	$0 \le \delta < \alpha < 1$	Calibration-failure allowance and total error budget
$n = 2m$	$m \ge 1$ integer	Independent eligible units, allocated equally
g	$[0,1]$	Average recorded detection floor among recoverable units

Table 1: Symbols and domains. The floor g is neither an instrumental accuracy score nor a lower bound on p. The coverage level c bounds a *sum* of two probabilities, which is why its range is $[0,2]$.

3 Why a finite negative record alone decides nothing

We record the three obstructions that motivate the source restriction. They are elementary and are stated to fix the scope of the later results, not claimed as new.

Proposition 3.1 (Finite-follow-up identified set)**.** *Suppose the conditional detection-time distribution among recoverable units is F and the observable cumulative event probability is $r(t) = pF(t)$ for $0 \le t \le T$. Then the set of recoverable fractions compatible with exact knowledge of r on $[0,T]$, with no restriction on the distribution after T, is $[r(T), 1]$.*

Proof. Any compatible p' satisfies $p' \ge p'F(T) = r(T)$. Conversely, for $p' \in (0,1]$ with $p' \ge r(T)$ put $F'(t) = r(t)/p'$ on $[0,T]$, which is nondecreasing with $F'(T) = r(T)/p' \le 1$, and place the remaining mass $1 - F'(T)$ after T. Then $p'F' = r$ on $[0,T]$. If $r(T) = 0$ the explanation $p' = 0$ also fits, with any F'. □

An all-negative finite sample carries sampling uncertainty *in addition* to this structural ambiguity; more precise measurement of the same finite event law does not shrink the identified set.

Proposition 3.2 (Successful-only timing is uninformative for the denominator)**.** *Among events detected by T, the timing distribution is $F(t)/F(T)$. For any $\lambda \in (0,1]$, replacing F by λF on*

$[0,T]$ *and moving the removed mass after* T *leaves* $F(t)/F(T)$ *unchanged while multiplying the total detection probability by* λ*.*

Selected resuscitation analyses normalize by observed recovery events and study their timing and trajectories [8]. Such data inform recovery dynamics and motivate separating the stages, but by Proposition 3.2 they do not by themselves calibrate the unseen recoverable mass or the paired coverage used below.

Example 3.3 (The serial-path obstruction)**.** Let $(x, y, a, b) = (1, 0, 0, 1)$. Then $x + a = y + b = 1$, so each stage is covered by one of the two conditions, yet $xy = ab = 0$: neither condition completes a path. Every unit of this type is negative in both conditions with probability one, so no number of units observed under these conditions can detect it. Stage-wise coverage and complete-path coverage are different quantities, and the mismatch $|x - y| = 1$ is what separates them.

4 The sharp coverage floor

Throughout this section, fix $0 \le c \le 2$ and $d \ge 0$ and consider the source set

$$\mathcal{A}(c, d) = \big\{(x, y, a, b) \in [0, 1]^4 : \ x + a \ge c,\ y + b \ge c,\ |x - y| \le d\big\}. \tag{2}$$

The description is asymmetric: coverage is required of both stages across the two conditions, while mismatch is controlled within condition 1. Relabelling the conditions is permitted only with a correspondingly justified mismatch bound. Exact complementarity, $a = 1 - x$ and $b = 1 - y$, is *not* assumed; it is the special case treated in Corollary 4.5. Write

$$h(c) = \min(c, 2 - c), \qquad t = \min\{d, h(c)\}, \qquad R_{\text{full}}(c, d) = \frac{c^2 - t^2}{4}. \tag{3}$$

Theorem 4.1 (Sharp coverage floor)**.** *For* $0 \le c \le 2$ *and* $d \ge 0$*,*

$$\min_{(x,y,a,b)\in\mathcal{A}(c,d)} \frac{xy + ab}{2} = R_{\text{full}}(c, d),$$

and the minimum is attained at

$$(x, y, a, b) = \left(\tfrac{c+t}{2},\ \tfrac{c-t}{2},\ \tfrac{c-t}{2},\ \tfrac{c+t}{2}\right), \tag{4}$$

where both complete-path products equal $R_{\text{full}}(c, d)$*. Consequently, for every allocation weight* $w \in [0, 1]$ *the response* $w\,xy + (1 - w)\,ab$ *of the source* (4) *equals* $R_{\text{full}}(c, d)$*, so equal allocation is maximin and no allocation can guarantee more.*

Proof. Lower bound. Put $\ell = \max(0, c - 1)$ and $u = \min(1, c)$, so that $u - \ell = h(c)$ and $\ell + u = c$. From $x + a \ge c$ and $a \le 1$ we get $x \ge c - 1$, and $x \ge 0$, hence $x \ge \ell$; likewise $y \ge \ell$. Set $X = \min(x, u)$ and $Y = \min(y, u)$, so $X, Y \in [\ell, u]$ and, by nonnegativity, $xy \ge XY$.

We claim $a \ge c - X$. If $x \le u$ then $X = x$ and $a \ge c - x$. Otherwise $u < x \le 1$, which forces $u < 1$ and therefore $u = \min(1, c) = c$, so $c - X = c - u = 0 \le a$. The same argument gives $b \ge c - Y$. Since $X, Y \le u \le c$, both $c - X$ and $c - Y$ are nonnegative, so $ab \ge (c - X)(c - Y)$.

Clipping at the common endpoint u is nonexpansive, so $|X - Y| \le |x - y| \le d$; and $X, Y \in [\ell, u]$ gives $|X - Y| \le u - \ell = h(c)$. Hence $|X - Y| \le t$ and

$$\frac{xy + ab}{2} \ \ge\ \frac{XY + (c - X)(c - Y)}{2} = \frac{(X + Y - c)^2 + c^2 - (X - Y)^2}{4} \ \ge\ \frac{c^2 - t^2}{4}. \tag{5}$$

Attainment. Since $0 \le t \le h(c)$ we have $t \le c$ and $t \le 2 - c$, so all four entries of (4) lie in $[0, 1]$; moreover $x + a = y + b = c$ and $|x - y| = t \le d$, so (4) lies in $\mathcal{A}(c, d)$, and $xy = ab = (c^2 - t^2)/4$.

Maximin. Equal allocation guarantees $R_{\text{full}}(c, d)$ by the lower bound. On the source (4) every allocation weight yields the same response $w\,xy + (1-w)\,ab = R_{\text{full}}(c, d)$, so no allocation guarantees more. □

Corollary 4.2 (Restricted range). *If $0 \le c \le 1$ then $h(c) = c$ and $R_{\text{full}}(c, d) = \max\{0, c^2 - d^2\}/4$, which is positive exactly when $d < c$.*

Proof. For $c \le 1$ we have $c \le 2-c$, so $h(c) = c$ and $t = \min(d, c)$. If $d \le c$ then $t = d$ and $c^2 - d^2 \ge 0$; if $d \ge c$ then $t = c$ and $c^2 - d^2 \le 0$, and the formula returns 0 in both cases. □

Corollary 4.3 (High coverage without mismatch information). *If $1 \le c \le 2$ then $R_{\text{full}}(c, d) \ge c - 1 > 0$ for every $d \ge 0$, including $d = 1$.*

Proof. Here $t \le h(c) = 2 - c$, so $c^2 - t^2 \ge c^2 - (2 - c)^2 = 4c - 4$. □

Remark 4.4 (The 1/4 ceiling is an artefact). On the restricted range the floor never exceeds $c^2/4 \le 1/4$. This has been read as a ceiling on what paired recovery designs can guarantee. It is not: it is the value of the restriction $c \le 1$. Figure 2 shows R_{full} on the full range, where the floor rises to 1 at $c = 2$, and the mismatch-free lower envelope $c - 1$ of Corollary 4.3. Section 10 gives a planning row that the restricted theorem cannot license.

Corollary 4.5 (Approximate complementarity, sharply). *Let $0 \le \varepsilon \le 1$ and suppose $|a-(1-x)| \le \varepsilon$, $|b - (1 - y)| \le \varepsilon$, together with $x, y, a, b \in [0, 1]$ and $|x - y| \le d$. Then*

$$\frac{xy + ab}{2} \;\ge\; R_{\text{full}}(1 - \varepsilon, d) = \frac{\max\{0, (1 - \varepsilon)^2 - d^2\}}{4},$$

which is positive exactly when $d + \varepsilon < 1$. The bound is attained within this smaller class: the source (4) with $c = 1 - \varepsilon$ has both complementarity deviations exactly ε.

Proof. The deviation bounds give $a \ge 1 - x - \varepsilon$ and $b \ge 1 - y - \varepsilon$, that is, membership in $\mathcal{A}(1 - \varepsilon, d)$; apply Theorem 4.1 and Corollary 4.2. For the witness, write $c = 1 - \varepsilon$ and $t = \min(d, c)$; then $a - (1 - x) = \frac{c-t}{2} - 1 + \frac{c+t}{2} = c - 1 = -\varepsilon$ and likewise $b - (1 - y) = c - 1$. Taking $\varepsilon = 0$ recovers the exactly complementary floor $(1 - d^2)/4$. □

The witnesses explain the obstruction. At the zero-floor boundary the two conditions cover different gates without completing either serial path; in the positive region the least detectable source balances the two complete-path probabilities, which is why allocation alone cannot improve on it.

5 The exact allocation profile

Theorem 4.1 shows that equal allocation is maximin, but not by how much other allocations lose, nor whether the optimum is unique. Both matter in practice: a design that must depart from a balanced split needs the exact price, and a numerical search over a grid cannot distinguish a unique optimum from a plateau. We compute the profile exactly.

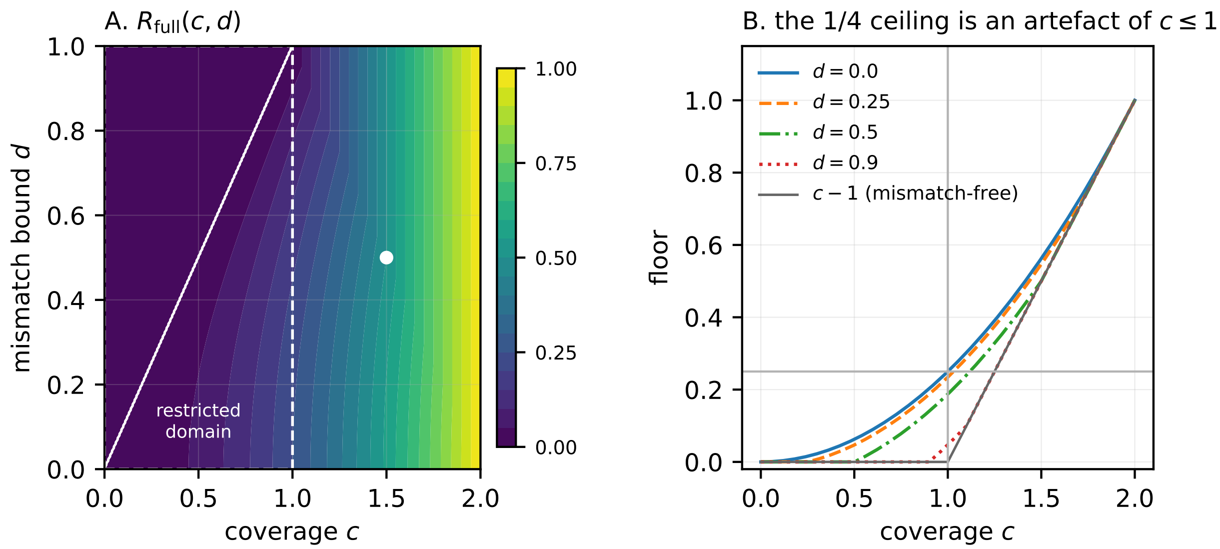

Figure 2: The sharp coverage floor on the full range. **A.** $R_{\mathrm{full}}(c,d)$ for $0 \le c \le 2$; the dashed box is the domain of the restricted theorem and the white curve is the zero-floor boundary $d = c$, which exists only for $c \le 1$. The marked point is the planning row $(c,d) = (1.5, 0.5)$ of Table 2. **B.** Sections at fixed d, with the mismatch-free envelope $c - 1$ of Corollary 4.3; the horizontal line is the value 1/4 discussed in Remark 4.4. These are exact theory curves, not fitted data.

For $w \in [0,1]$ write $A = |2w-1|$ and let

$$M(w;c,d) = \begin{cases} \dfrac{c^2 - t^2 - A^2c^2}{4}, & Ac \le h(c) - t, \\ \dfrac{c^2 + h(c)^2 - 2h(c)t - 2Ac(h(c)-t)}{4}, & h(c) - t \le Ac \le h(c), \\ \dfrac{c^2 + h(c)^2 - 2Ac\,h(c)}{4}, & Ac \ge h(c). \end{cases} \tag{6}$$

The three expressions agree on the overlaps, so M is continuous, and it is nonincreasing in A.

Theorem 5.1 (Allocation profile)**.** *For $0 \le c \le 2$, $d \ge 0$ and $w \in [0,1]$,*

$$\min_{(x,y,a,b)\in\mathcal{A}(c,d)} \Big\{ w\,xy + (1-w)\,ab \Big\} = M(w;c,d),$$

and the minimum is attained. In particular $M(1/2;c,d) = R_{\mathrm{full}}(c,d)$.

Proof. Reduction. Let ℓ, u, X, Y be as in the proof of Theorem 4.1. As shown there, $xy \ge XY$ and $ab \ge (c-X)(c-Y)$ with all factors nonnegative, and $w, 1-w \ge 0$, so

$$w\,xy + (1-w)\,ab \;\ge\; w\,XY + (1-w)(c-X)(c-Y),$$

with $X, Y \in [\ell, u]$ and $|X - Y| \le t$. Conversely, for any $X, Y \in [\ell, u]$ with $|X - Y| \le d$ the quadruple $(X, Y, c-X, c-Y)$ lies in $\mathcal{A}(c,d)$, because $c - X, c - Y \in [c-u, c-\ell] = [\ell, u] \subseteq [0,1]$. Hence the minimum equals

$$\min\Big\{ w\,XY + (1-w)(c-X)(c-Y) : \; X, Y \in [\ell, u], \; |X - Y| \le d \Big\}.$$

Coordinates. Put $\sigma = X + Y - c$ and $\tau = X - Y$. Since $\ell + u = c$ and $u - \ell = h(c)$, the constraint $X, Y \in [\ell, u]$ is exactly $|\sigma| + |\tau| \le h(c)$, and the mismatch constraint is $|\tau| \le d$; together, $|\tau| \le \min\{d, h(c) - |\sigma|\}$. A direct computation gives

$$w\,XY + (1-w)(c-X)(c-Y) = \frac{\sigma^2 + c^2 - \tau^2 + 2(2w-1)c\,\sigma}{4}. \tag{7}$$

Minimising over the sign of σ replaces $2(2w-1)c\sigma$ by $-2Acv$ with $v = |\sigma| \in [0, h(c)]$, and minimising over τ takes $|\tau| = \min\{t, h(c) - v\}$. So the problem reduces to minimising

$$f(v) = v^2 + c^2 - \min\{t, h(c) - v\}^2 - 2Acv, \qquad v \in [0, h(c)],$$

and the minimum of (7) is $\min_v f(v)/4$.

The one-dimensional minimisation. On $[0, h(c) - t]$ we have $f(v) = v^2 + c^2 - t^2 - 2Acv$, a convex parabola minimised at $v = Ac$. On $[h(c) - t, h(c)]$ we have $f(v) = c^2 - h(c)^2 + 2v(h(c) - Ac)$, which is affine in v and whose value at $v = h(c) - t$ agrees with the first branch. Three cases: if $Ac \le h(c) - t$, the minimum is at $v = Ac$, with value $c^2 - t^2 - A^2c^2$; if $h(c) - t \le Ac \le h(c)$, the first branch decreases on $[0, h(c) - t]$ and the second is nondecreasing, so the minimum is at $v = h(c) - t$, with value $c^2 + h(c)^2 - 2h(c)t - 2Ac(h(c) - t)$; if $Ac \ge h(c)$, the second branch is nonincreasing, so the minimum is at $v = h(c)$, where $\tau = 0$, with value $c^2 + h(c)^2 - 2Ach(c)$. These are the three lines of (6).

Attainment. Each case exhibits an explicit minimiser $(v, |\tau|)$, namely (Ac, t), $(h(c) - t, t)$ and $(h(c), 0)$, all satisfying $v + |\tau| \le h(c)$. Choosing the sign of σ opposite to that of $2w - 1$ and setting $X = (c + \sigma + \tau)/2$, $Y = (c + \sigma - \tau)/2$, $a = c - X$, $b = c - Y$ gives an admissible source attaining the value. Finally, at $w = 1/2$ we have $A = 0 \le h(c) - t$, and the first branch gives $(c^2 - t^2)/4 = R_{\text{full}}(c, d)$. □

Corollary 5.2 (Strict loss away from balance)**.** *Let $0 < c \le 2$ and suppose the mismatch budget binds, $t < h(c)$. Then $M(w; c, d) < R_{\text{full}}(c, d)$ for every $w \ne 1/2$, so equal allocation is the unique maximin allocation.*

Proof. Write $\Delta = R_{\text{full}}(c, d) - M(w; c, d)$ and $A = |2w - 1| > 0$, so $Ac > 0$. In the first branch $4\Delta = A^2c^2 > 0$. In the second branch $4\Delta = (h(c) - t)\big(2Ac - (h(c) - t)\big) > 0$, because $h(c) - t > 0$ and $2Ac > h(c) - t$ by the branch condition. In the third branch $4\Delta = 2Ac\,h(c) - h(c)^2 - t^2 \ge 2h(c)^2 - h(c)^2 - t^2 = h(c)^2 - t^2 > 0$, again by the branch condition. □

Corollary 5.3 (Plateau when the budget does not bind)**.** *If $h(c) \le d$, so that $t = h(c)$, then $M(w; c, d) = R_{\text{full}}(c, d)$ for every w with $|2w - 1|\,c \le h(c)$, and M is strictly decreasing in A beyond that point. For $c \le 1$ this plateau is the whole interval $w \in [0, 1]$ at the value 0; for $c > 1$ it is an interval of half-width $(2 - c)/(2c)$ around $w = 1/2$ at the positive value $(c^2 - (2 - c)^2)/4 = c - 1$.*

Proof. With $t = h(c)$ the first branch applies only at $A = 0$ and the second branch gives $(c^2 + h(c)^2 - 2h(c)^2 - 0)/4 = (c^2 - h(c)^2)/4 = R_{\text{full}}(c, d)$ on the whole range $Ac \le h(c)$. Beyond it the third branch is strictly decreasing in A when $h(c) > 0$. □

Corollary 5.3 is the honest reading of an earlier numerical observation that a grid search over w failed to select $w = 1/2$: on a non-binding mismatch budget there is nothing to select, because a whole interval of allocations is optimal. Figure 3 plots both regimes together with brute-force minima computed independently of (6).

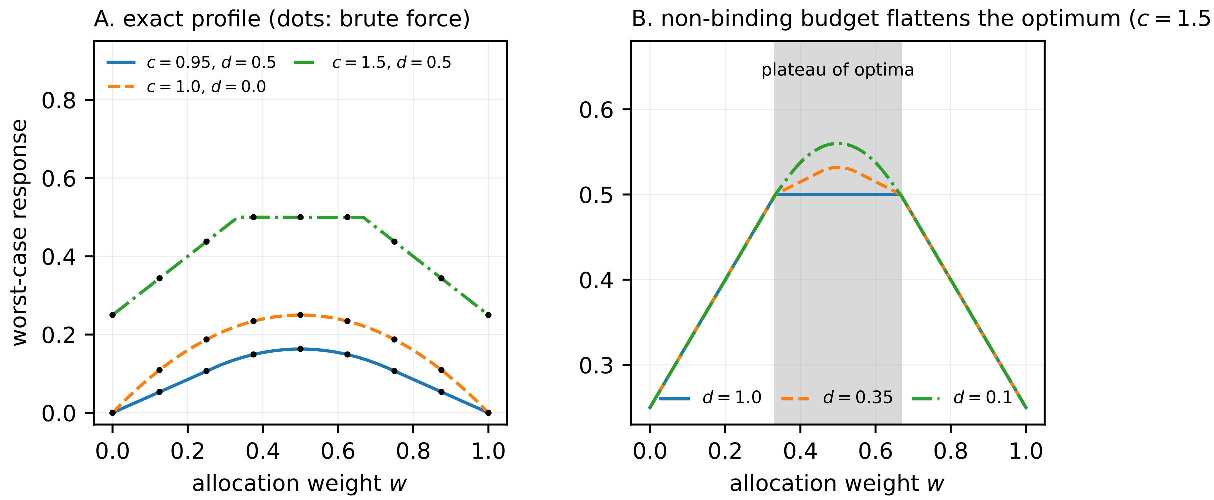

Figure 3: Worst-case response as a function of the allocation weight. **A.** Exact profile (6) (curves) against an independent brute-force minimisation over the admissible source set (dots); the kinks are the branch boundaries of Theorem 5.1. **B.** A non-binding mismatch budget produces a plateau of optima (shaded), of half-width $(2-c)/(2c)$ in w, as in Corollary 5.3. All curves are exact theory.

6 Mismatch in mean square

A uniform per-type bound $|x_z - y_z| \le d$ is a strong calibration requirement: it must hold for every recoverable type, including types that are rare. The next result shows that only a population average is needed, with the same floor.

Theorem 6.1 (Mean-square mismatch). *Let μ be a probability weight on recoverable types, let every type satisfy the coverage constraints $x_z + a_z \ge c$ and $y_z + b_z \ge c$ with $x_z, y_z, a_z, b_z \in [0,1]$ and $0 \le c \le 2$, and suppose*

$$\mathbb{E}_\mu\left[(x-y)^2\right] = \sum_z \mu_z (x_z - y_z)^2 \le D^2, \qquad D \ge 0.$$

Then

$$\mathbb{E}_\mu\Big[\frac{xy+ab}{2}\Big] \;\ge\; R_{\text{full}}(c, D),$$

and the bound is attained by the constant type (4) *with $t = \min\{D, h(c)\}$. The uniform-mismatch case of Theorem 4.1 is the special case $D = d$.*

Proof. Apply the lower bound (5) to each type with its own mismatch, which gives, with $t_z = \min\{|x_z - y_z|, h(c)\}$,

$$\frac{x_z y_z + a_z b_z}{2} \;\ge\; \frac{c^2 - t_z^2}{4}.$$

Averaging, $\mathbb{E}_\mu[(xy+ab)/2] \ge (c^2 - \mathbb{E}_\mu[t^2])/4$. Now $t_z^2 \le (x_z - y_z)^2$ and $t_z^2 \le h(c)^2$ pointwise, so $\mathbb{E}_\mu[t^2] \le \min\{D^2, h(c)^2\} = \min\{D, h(c)\}^2$, which is the claim. The constant type (4) with $t = \min\{D, h(c)\}$ has mean-square mismatch exactly $t^2 \le D^2$ and attains the value. A uniform bound $|x_z - y_z| \le d$ implies $\mathbb{E}_\mu[(x-y)^2] \le d^2$, giving the stated special case. $\square$

Theorem 6.1 is genuinely weaker as a hypothesis: a small population of badly mismatched types is tolerated provided the mean square stays within budget. What cannot be weakened is the requirement that the average be taken over the *recoverable* population with its own weights. Two pooled summaries that look similar are not sufficient.

Example 6.2 (Equal stage means do not suffice)**.** Let two types have weight 1/2 each, with $(x, y) = (1, 0)$ and $(x, y) = (0, 1)$, and exactly complementary partners. Then $\mathbb{E}[x] = \mathbb{E}[y] = 1/2$, yet $\mathbb{E}[xy] = 0$ and the complementary complete paths also vanish, so the average complete-path response is 0. The mean-square mismatch is 1, and Theorem 6.1 correctly returns $R_{\text{full}}(1, 1) = 0$. Equality of the two stage means carries no information about their joint behaviour within a type.

Example 6.3 (The pooled conditional is not the mean second stage)**.** For the mixture of Example 6.2, the pooled second-stage conditional probability is

$$\mathbb{P}(D_1 \mid A_1) = \frac{\mathbb{E}[xy]}{\mathbb{E}[x]} = 0 \neq \frac{1}{2} = \mathbb{E}[y],$$

whenever $\mathbb{E}[x] > 0$. Averaging conditional probabilities across latent types without activation weights can manufacture an apparent alignment that the source does not possess. A calibration that reports $\mathbb{E}[x]$ and $\mathbb{E}[y]$ separately does not bound d or D.

7 From source paths to a finite negative certificate

We now compose the source floor with recording losses, exceptional types, an explicit sampling law and a calibration allowance.

7.1 Recorded response of a heterogeneous population

Let a covered subset of recoverable types carry mass at least $1 - \eta$. Only covered types must satisfy the coverage constraints, with conditional mismatch budget D in the sense of Theorem 6.1; their completed paths are recorded with conditional probabilities $r_z, s_z \geq \kappa$. All remaining contributions may be zero. Define the recorded per-condition responses

$$u = \sum_z \mu_z x_z y_z r_z, \qquad v = \sum_z \mu_z a_z b_z s_z. \tag{8}$$

Proposition 7.1 (Recording and exceptional mass)**.** *Under these assumptions,*

$$\frac{u + v}{2} \;\geq\; g := (1 - \eta)\, \kappa\, R_{\text{full}}(c, D), \tag{9}$$

and the bound is sharp in the class that admits invisible exceptional types and recording probability exactly κ on covered paths.

Proof. On each covered type, nonnegativity of the path probabilities and $r_z, s_z \geq \kappa$ give $(x_z y_z r_z + a_z b_z s_z)/2 \geq \kappa(x_z y_z + a_z b_z)/2$. Summing over covered types and applying Theorem 6.1 to the covered submixture, whose total mass $m \geq 1 - \eta$, yields $\sum_{\text{covered}} \mu_z (x_z y_z r_z + a_z b_z s_z)/2 \geq \kappa\, m\, R_{\text{full}}(c, D) \geq \kappa(1 - \eta) R_{\text{full}}(c, D)$, using $R_{\text{full}} \geq 0$. Discarding the nonnegative contributions of exceptional types gives (9). For sharpness, place mass $1 - \eta$ on the attaining type (4), set both recording probabilities to κ, and place mass η on a type with $x = y = a = b = 0$; then $u = v = g$. □

No independence of recording and phenotype is used: the recording bound is stated conditionally on the completed path and the type. A product of a retention bound and an instrumental detection bound is legitimate only if each applies conditionally at its own stage; two marginal success rates do not automatically justify their product. The interpretation of η is equally strict: it bounds uncovered mass among true reference-recoverable units, not the share of unusual types among observed positives.

7.2 The observation law is derived, not postulated

Lemma 7.2 (Terminal masses and the all-negative event)**.** *Fix a condition and a type with stage probabilities (x, y) and recording probability r. The five terminal source events of Figure 1 carry masses*

$$1-p, \qquad p(1-x), \qquad px(1-y), \qquad pxy(1-r), \qquad pxyr, \tag{10}$$

which are nonnegative and sum to one; the first four, which are exactly the negative records, sum to $1-pxyr$. Mixing over types, a unit in condition 1 is negative with probability $1-pu$ and a unit in condition 2 with probability $1-pv$. For $n = 2m$ independent units with m assigned to each condition independently of type,

$$\mathbb{P}(\textit{all negative}) = (1-pu)^m(1-pv)^m. \tag{11}$$

Proof. Nonnegativity and normalisation of (10) are immediate, and the first four entries sum to $1-pxyr$. Averaging over the type weights gives $1-pu$ and $1-pv$ by (8). Independent units have product masses over histories; summing those products over all histories in which every unit is negative factors into (11). □

Equation (11) is a consequence of the source events, not an extra modelling assumption about wells. This matters, because it is precisely the step at which an unjustified independence or an unnoticed change of denominator would enter.

7.3 Balanced design and conditional calibration

Theorem 7.3 (Balanced sampling with a calibration allowance)**.** *Take $n = 2m$ independent eligible units from the common source, m per condition. Suppose that on every* good *calibration environment the source and recording contract of Proposition 7.1 holds with fixed c, D, κ, η, and that the probability of a bad environment is at most $\delta \in [0, 1]$. Then for $0 < \theta \le p \le 1$ the rule that certifies $p < \theta$ exactly on an all-confirmed-negative record has false-certification probability at most*

$$\delta + (1-\delta)\,(1-\theta g)^n, \qquad g = (1-\eta)\kappa R_{\mathrm{full}}(c, D). \tag{12}$$

The bound is attained in the stated class, by the mixture of Proposition 7.1 with $p = \theta$, bad mass exactly δ, and bad environments that certify with probability one.

Proof. Condition on a good environment and use Lemma 7.2. All bases are nonnegative, since $pu, pv \le 1$. The identity

$$\Big(1-p\frac{u+v}{2}\Big)^2 - (1-pu)(1-pv) = \frac{p^2(u-v)^2}{4} \;\ge\; 0 \tag{13}$$

gives $(1-pu)^m(1-pv)^m \le \big(1-p(u+v)/2\big)^{2m}$, and $p \ge \theta$ with (9) gives $\big(1-p(u+v)/2\big)^n \le (1-\theta g)^n =: B$. If the actual bad mass is $\beta \le \delta$, the total false-certificate probability is at most $(1-\beta)B + \beta \le (1-\delta)B + \delta$, because $B \le 1$. Attainment: the mixture of Proposition 7.1 has $u = v = g$, so with $p = \theta$ the conditional bound holds with equality, and equality in the composition requires exactly the stated bad-environment behaviour. □

The composition does not require independence between calibration and the future experiment. It does require the sampling guarantee to hold *conditionally on every good environment.* A marginal confidence statement about a calibration estimate does not by itself supply that contract when data are reused, when several bounds are selected after inspection, or when batch effects are shared; in those cases the environment allowance δ must be constructed to cover the selection.

Remark 7.4 (Independent random allocation, and unbalanced designs). If each unit independently receives condition 1 or 2 with probability 1/2, its negative mass is $1 - p(u+v)/2$ and the all-negative law is $\big(1 - p(u+v)/2\big)^n$ for any integer n, with the same worst-case bound. By (13), for a fixed response pair the balanced fixed design is no worse than random assignment at even n. An unbalanced fixed design instead has $(1-pu)^{n_1}(1-pv)^{n_2}$; substituting an observed allocation fraction into the random-assignment formula does not reproduce that law.

8 Sample size, feasibility and reporting

Assume $0 \le \delta < \alpha < 1$, $0 < \theta \le 1$ and $g > 0$, and set

$$B_* = \frac{\alpha - \delta}{1 - \delta}, \qquad z = \theta g \in (0, 1]. \tag{14}$$

The bound (12) is at most α exactly when $(1-z)^n \le B_*$. For $z < 1$, taking logarithms gives the random-assignment minimum

$$n_{\min} = \left\lceil \frac{\log B_*}{\log(1-z)} \right\rceil, \tag{15}$$

and for a fixed balanced design one takes the smallest even integer at least $n_{\min}$. These are minima within the stated design class and for the stated all-negative rule; they are not claims about every possible observation technology, and a rule that never certifies trivially has zero false-certificate probability while performing no task. Writing $H = \log(1/B_*)$ and integrating $1/(1-s)$ on $[0, z]$ gives $z \le -\log(1-z) \le z/(1-z)$, hence the unrounded requirement n_* obeys

$$\frac{H(1-z)}{z} \;\le\; n_* \;\le\; \frac{H}{z}. \tag{16}$$

The simple bound $\lceil H/z \rceil$, rounded up to an even integer, is always sufficient.

On the restricted range $g \le 1/4$, so $z < 1$ automatically and (15) always applies. On the full range g can reach 1 — at $c = 2$, $\kappa = 1$, $\eta = 0$ — and then $z = \theta g$ can equal 1, in which case a single unit already has zero good-environment negative probability. The feasibility boundary must therefore be stated with that case separated.

Theorem 8.1 (Exact finite feasibility)**.** *Let $0 \le z \le 1$, $0 \le \delta \le 1$ and $0 < \alpha < 1$. There exists an integer $m \ge 1$ with $\delta + (1-\delta)(1-z)^{2m} \le \alpha$ if and only if*

$$(z > 0 \text{ and } \delta < \alpha) \quad \text{or} \quad (z = 1 \text{ and } \delta \le \alpha).$$

Proof. ($\Rightarrow$) Suppose $m \ge 1$ satisfies the bound. Since $\delta \le 1$ and $(1-z)^{2m} \ge 0$, the added term is nonnegative, so $\delta \le \alpha < 1$ and hence $1 - \delta > 0$. If $z = 0$ the left-hand side is $\delta + (1-\delta) = 1 > \alpha$, so $z > 0$. If moreover $\delta \ge \alpha$, then $(1-\delta)(1-z)^{2m} \le \alpha - \delta \le 0$, and $1 - \delta > 0$ with $(1-z)^{2m} \ge 0$ forces $(1-z)^{2m} = 0$, hence $z = 1$; the displayed bound then reads $\delta \le \alpha$.

($\Leftarrow$) If $z = 1$ take $m = 1$: the error is $\delta \le \alpha$. If $z > 0$ and $\delta < \alpha$, then $0 \le 1 - z < 1$, so there is k with $(1-z)^k < \alpha - \delta$. Take $m = k + 1$, so $2m \ge k$ and hence $(1-z)^{2m} \le (1-z)^k$; since $0 \le 1 - \delta \le 1$, $\delta + (1-\delta)(1-z)^{2m} \le \delta + (1-z)^k < \alpha$. □

Corollary 8.2 (Feasibility in terms of the source parameters)**.** *With $g = (1-\eta)\kappa R_{\text{full}}(c, D)$ and $0 < \alpha < 1$, a finite balanced sample attains uniform total error at most α over the full sharpness class if and only if $\theta > 0$, $\kappa > 0$, $\eta < 1$, $R_{\text{full}}(c, D) > 0$ and $\delta < \alpha$, or else $\theta g = 1$ and $\delta \le \alpha$. Moreover $R_{\text{full}}(c, D) > 0$ holds exactly when $c > 0$ and $\min\{D, h(c)\} < c$; on the restricted range this is $c > D$, while for $c > 1$ it is automatic.*

Proof. Combine Theorem 8.1 with $z = \theta g$ and the positivity statements: $g > 0$ iff each factor is positive, and $R_{\text{full}}(c, D) = (c^2 - t^2)/4 > 0$ iff $t < c$, where $t = \min\{D, h(c)\} \le h(c)$. For $c > 1$, $h(c) = 2 - c < c$, so $t < c$ always; for $c \le 1$, $h(c) = c$ and $t < c$ iff $D < c$. □

This is a statement about a *uniform guarantee over a model class.* A real assay may perform well when the available bounds are too weak to certify it; failure of the sufficient information is not evidence of biological failure.

Proposition 8.3 (All-negative confidence endpoint)**.** *For $n > 0$ and $g > 0$ define, on an evaluable all-negative record,*

$$p_U = \min\Big\{1,\ \frac{1 - B_*^{1/n}}{g}\Big\}, \tag{17}$$

and return the endpoint 1 on every other record. Then the interval $[0, p_U]$ has coverage at least $1 - \alpha$ for the true recoverable fraction.

Proof. Coverage can fail only on an all-negative record with $p > p_U$, which requires $p_U < 1$ and hence $pg > 1 - B_*^{1/n}$, that is $(1 - pg)^n < B_*$. By Theorem 7.3 applied with $\theta = p$, the probability of that record is at most $\delta + (1-\delta)(1-pg)^n < \delta + (1-\delta)B_* = \alpha$. □

This is a deliberately limited procedure: positive or unevaluable records return an uninformative endpoint rather than an unsupported point estimate, and the statement concerns repeated sampling, not the posterior probability that a particular specimen is sterile. For computation use $p_U = -\text{expm1}(\log(B_*)/n)/g$, capped at one, which avoids cancellation when the root is close to one; if $g = 0$, $n = 0$ or $\delta \ge \alpha$, report the certificate as unavailable rather than evaluating (17).

Proposition 8.4 (Cost of a closing margin)**.** *As $z \to 0$, $n_* = H\big(1/z - 1/2 + O(z)\big)$. With $\theta, \kappa, 1-\eta$ fixed and positive and $0 < c \le 1$, the requirement diverges as the coverage margin closes:*

$$n_* \ \sim\ \frac{2H}{\theta(1-\eta)\kappa\, c\,(c - D)} \qquad (D \uparrow c),$$

and in the rare-detection regime $\partial \log n_*/\partial c \approx -2c/(c^2 - D^2)$ and $\partial \log n_*/\partial D \approx 2D/(c^2 - D^2)$. Integer rounding changes n_* by a bounded additive amount. As $\delta \uparrow \alpha$, $H \to \infty$: no increase in the number of units crosses the calibration-failure boundary.

Proof. Expand $-\log(1-z) = z + z^2/2 + O(z^3)$ in $n_* = H/(-\log(1-z))$. With $R_{\text{full}}(c, D) = (c-D)(c+D)/4$ on the restricted range and $D \uparrow c$, $z = \theta(1-\eta)\kappa(c-D)(c+D)/4 \sim \theta(1-\eta)\kappa c(c-D)/2$, and $n_* \sim H/z$. The derivatives follow from $n_* \approx 4H/[\theta(1-\eta)\kappa(c^2 - D^2)]$, and $H = \log\big((1-\delta)/(\alpha-\delta)\big) \to \infty$ as $\delta \uparrow \alpha$. □

9 Adaptive reallocation does not help

Theorem 4.1 assumes the allocation is fixed in advance. A natural question is whether choosing each unit's condition in the light of previous outcomes improves the guarantee. It does not.

Theorem 9.1 (Adaptive policies face the same bound)**.** *Consider a budget of n fresh independent units. A policy may choose the condition of each unit as a (possibly randomized) function of the outcomes of the previously observed units, but cannot change a unit's condition after observing its own stage trajectory; the qualifying event is the same in both conditions and the certificate is issued exactly on an all-negative record. Then for every such policy the worst-case false-certificate probability over the class of Theorem 7.3 is at least $\delta + (1-\delta)(1-\theta g)^n$, which the fixed balanced design attains.*

Proof. Evaluate the policy on the attaining population of Proposition 7.1, which has recorded response $u = v = g$ in *both* conditions, with $p = \theta$. Let N_i be the event that the first i units are negative. Conditional on N_{i-1} and on the policy's choice for unit i — which is measurable with respect to the first $i-1$ outcomes and independent randomisation — unit i is recorded positive with probability $\theta u = \theta v = \theta g$, since the unit is fresh and independent of the past. Hence $\mathbb{P}(N_i \mid N_{i-1}) = 1 - \theta g$ whatever the choice, and $\mathbb{P}(N_n) = (1-\theta g)^n$ by the chain rule. Adding bad environments of mass δ that certify with probability one gives the stated value, and Theorem 7.3 shows the balanced design attains it. □

The scope of Theorem 9.1 should be read carefully. Adaptivity may well improve performance on populations other than the least detectable one; and the result says nothing about designs that switch conditions *within* a unit, introduce a new readout, change the reference task, or decline to certify some all-negative records. Switching within a unit is a genuinely different causal question, discussed in Section 11.

10 A reproducible synthetic planning example

Table 2 fixes $\theta = 0.01$, $\alpha = 0.05$, the exposure and the observation deadline. Every parameter is synthetic; no row is an estimated operating characteristic of any assay. The rows successively relax ideal coverage, allow recording loss and an exceptional recoverable fraction, allocate error to calibration failure, and finally exploit the full coverage range.

Planning assumptions	g	$n_{\min}$	Balanced n	Total error
$c = 1$, $d = .5$	0.187500000	1597	1598	0.0498338
$c = .95$, $d = .5$	0.163125000	1835	1836	0.0499157
Also $\kappa = .90$, $\eta = .05$	0.139471875	2147	2148	0.0498894
Also $\delta = .01$	0.139471875	2300	2300	0.0499494
Full range, $c = 1.5$, $d = .5$	0.427500000	749	750	0.0498285

Table 2: Exact integer planning checks; displayed probabilities are rounded. Unspecified factors are ideal and $\delta = 0$ except where stated. The column $n_{\min}$ uses independent random assignment and the next column is the smallest even sample for the fixed balanced theorem. The last row uses $c = 1.5$ with $\kappa = .90$, $\eta = .05$, $\delta = .01$, and is licensed only by Theorem 4.1 on the full range.

In the fourth row, $R_{\text{full}} = 261/1600$, $g = 44631/320000$ and $B_* = 4/99$, and the integer threshold is confirmed by exact rational arithmetic:

$$\left(1 - \frac{44631}{32000000}\right)^{2300} \le \frac{4}{99} < \left(1 - \frac{44631}{32000000}\right)^{2299}, \tag{18}$$

the corresponding total errors being 0.0499494 and 0.0500052; so 2299 units narrowly fail the budget and 2300 succeed. The all-negative endpoint (17) at $n = 2300$ is $p_U \approx 0.0099961$. In the last row,

$R_{\text{full}}(1.5, 0.5) = 1/2$ and $g = 171/400 = 0.4275$, giving 749 and 750 units at total error 0.0498285: a factor of about three fewer units than the fourth row, obtained entirely by strengthening the coverage premise beyond $c = 1$. On the restricted range no premise whatever can beat $g \leq 1/4$.

The earlier compiled regression of this model uses $\theta = 1/5$, $c = 1$, $d = 1/2$ and $n = 100$, with $(1-\theta R)^{100} = (77/80)^{100} \approx 0.0218813 < 0.05$; its exact minima are 79 units under random assignment and 80 under a balanced design, so the 100-unit example is valid but deliberately conservative.

Figure 4 shows why the margin, rather than the number of units, dominates feasibility. Its curves use the unrounded requirement; the table gives integer designs. A permissible report would read:

> "All 2,300 eligible tracked units had no qualifying recorded event. Under the stated recovery-coverage, exceptional-mass, recording and calibration assumptions, this procedure excludes a reference-recoverable fraction of 1% or more with false-certificate probability at most 5%."

The eligible denominator must be stated explicitly. If eligibility selects intact cells found after exposure, the fraction refers to that selected population; it does not directly quantify killing among the originally exposed cells.

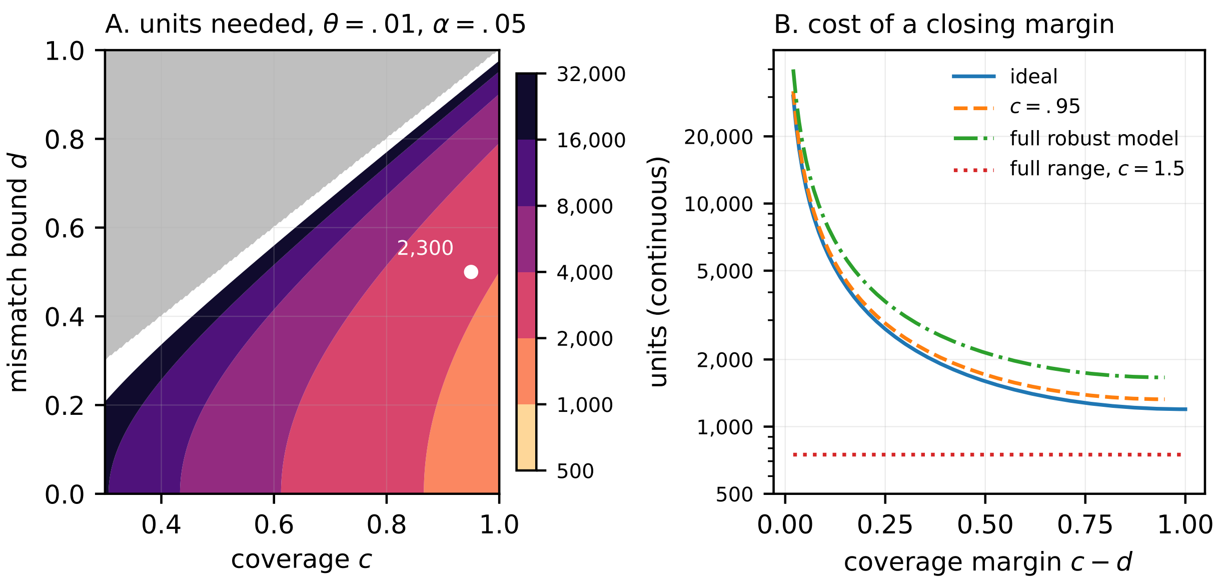

Figure 4: Synthetic design consequences. **A.** Units required on the restricted range for $\theta = .01$, $\alpha = .05$, $\kappa = .90$, $\eta = .05$, $\delta = .01$; the grey region has no positive uniform floor and the marked point is the fourth row of Table 2. **B.** Cost of a closing coverage margin $c - D$ on a logarithmic scale, including the full-range row with $c = 1.5$. These are exact theory curves, not fitted data or Monte Carlo estimates.

11 What must be established experimentally

The decisive unresolved question is empirical: whether an exposed target population and a paired recovery design support the joint bounds of Table 3. Neither random assignment alone nor pooled stage means identify within-type potential responses in two conditions, and a single unit generally cannot undergo both potential recovery histories. Mechanistic transport, replicate lineages or calibrated strata can supply the missing information only with their own explicit assumptions.

Premise	Required information	Insufficient substitute
Reference task	Defined recoverability endpoint and target population	Calling every finite non-crosser dead
c	Stage coverage across conditions, transported to exposed types	High response in an unrelated healthy control
d or D	Joint type-level stage information, or a structural bound	Equal marginal stage means (Examples 6.2, 6.3)
η	Uncovered mass among reference-recoverable types	Unusual-type fraction among observed positives
κ	Recording bound conditional on completed paths and phenotype	Pooled instrumental accuracy
δ	Simultaneous calibration and a conditional future-experiment contract	Unadjusted marginal intervals or adaptive data reuse
Independent units	Sampling, assignment, batch and lineage provenance	Counting images or descendants as new replicates

Table 3: The mathematical assumptions define an empirical validation task. They are not established by the theorems or by the synthetic example, and the numbers c, d and D are properties of a declared event model and population, not universal physiological constants.

Shared failures are not independent recording losses. If an entire run fails with probability ρ, the negative probability can be $\rho + (1-\rho)(1-pg)^n$. The floor ρ does not shrink by adding units to the same failing run; independent batches may help under a separate, justified batch model. The calibration allowance δ and a physical shared failure ρ should not be equated or added without specifying their relationship.

Losses must not be silently dropped. One safe rule certifies only when every required record is evaluable and negative, which shrinks the certification event and preserves the bound. Alternatively, losses can enter a justified conditional recording model. Discarding lost units and recomputing an apparently clean denominator requires a new sampling argument.

A recovery certificate is not a killing curve. For two exposures with comparable original denominators, write the observable event fractions as $r_0 = p_0 g_0$ and $r_\tau = p_\tau g_\tau$. Then

$$\frac{p_\tau}{p_0} = \frac{r_\tau}{r_0} \cdot \frac{g_0}{g_\tau}, \tag{19}$$

so a tenfold decrease in recorded events is equally consistent with a tenfold decrease in recovery sensitivity: $p_\tau = p_0$ with $g_0 = 0.8$ and $g_\tau = 0.08$ produces it. If separately justified bounds $g_j \in [L_j, U_j]$ are available, then $\frac{r_\tau}{r_0}\frac{L_0}{U_\tau} \le \frac{p_\tau}{p_0} \le \frac{r_\tau}{r_0}\frac{U_0}{L_\tau}$, and finite-sample uncertainty in r_0, r_τ must be propagated in addition. Equation (19) is algebra, not a new estimator; it explains why a recovery certificate and a quantitative killing curve [5] require related but distinct calibration tasks.

Stage switching is a separate causal question. If the condition could be randomized independently at each stage and the stage probabilities transported unchanged across switches, then in the exactly complementary model the four schedules average to $(xy+x(1-y)+(1-x)y+(1-x)(1-y))/4 =$

1/4, bypassing the mismatch obstruction entirely. But the unswitched observations identify only two of the four paths; cross-condition stage behaviour may differ, and with unconstrained cross responses switching need not help. This is a precisely stated future intervention hypothesis, not a laboratory recommendation: its missing premise is the behaviour of the activated state under a changed condition.

A shortcut when direct calibration is available. If independently ascertained reference-positive units provide a representative direct lower bound on the average recorded recovery g, the sampling results of Section 7 apply to that bound directly and the two-stage model is a diagnostic explanation rather than a mandatory calibration ritual. The ascertainment must not exclude the slow or difficult units the claim intends to cover; reference-positive units selected because the test already detected them make the argument circular.

12 Formal verification and reproducibility

The source is represented by normalized nonnegative finite path weights and independence by products over finite histories; the decision chain is therefore finite and elementary, and is machine-checkable end to end. Table 4 maps each result of this paper to its status.

Result	Status	Lean declaration (namespace `FunctionalViability.Robust`)
Theorem 4.1, bound	compiled	`extended_coverage_bound`
Theorem 4.1, witness and maximin	compiled	`extended_sharp_witness`, `extended_minimax`
Corollary 4.2	compiled	`extended_eq_recoveryFloor`
Corollary 4.3	compiled	`extended_high_coverage`
Corollary 4.5	compiled	`approximate_complementarity`, `approximate_complementarity_sharp`
Theorem 5.1	compiled	`allocation_profile` (`allocation_lower`, `allocation_attained`)
Corollaries 5.2, 5.3	compiled	`allocation_strict`, `allocation_plateau`, `allocation_half`
Theorem 6.1 and Proposition 7.1, bound	compiled	`moment_population_floor`
Lemma 7.2	compiled	`stage_normalized`, `stage_negative`, `balanced_identity`
Theorem 7.3, bound	compiled	`moment_source_certificate`, `moment_calibrated_certificate`
Theorem 8.1	compiled	`finite_feasibility`
Propositions 3.1, 3.2	conventional	identified set and timing rescaling
Sharpness constructions of Proposition 7.1, Theorem 7.3	conventional	exceptional mixture and bad environments
Equations (15)–(16), Corollary 8.2	conventional	logarithmic planning
Proposition 8.3, Proposition 8.4	conventional	confidence inversion, asymptotics
Theorem 9.1	conventional	adaptive-policy minimax
Section 11	conventional	interpretation and algebra

Table 4: Scope of the formal verification. "Compiled" means the stated inequality or identity is proved in Lean; the biological interpretation of any statement is never part of what a proof assistant checks.

The environment is Lean 4.30.0 with the repository's pinned Mathlib dependency manifest (SHA-256 `a8df0b60...6e9fac`). Compilation is strict, with warnings as errors; no proof uses `sorry` and every exported declaration depends only on the three standard Mathlib axioms `propext`, `Classical.choice` and `Quot.sound`. The new modules are `ExtendedCoverage.lean` (SHA-256 `3c7aef3b...3c9f8a10`) and `DesignLimits.lean` (SHA-256 `86d0340d...3ef26f61`), which build on the earlier chain `RecoverySource`, `Sampling`, `CoverageBounds`, `BalancedSampling` and `RobustCertificate` (SHA-256 `c29773c7...07d9306c6`). Verification receipts record the compiler command, the source hash, the manifest hash, the elaborated statement of each exported declaration and its axiom list.

A complementary script recomputes every printed constant: the exact rational values of R_{full} and g in Table 2, the integer budgets with their neighbouring powers as in (18), the endpoint (17), the asymptotic ratio of Proposition 8.4, and an independent brute-force minimisation of the source problem that reproduces R_{full} and the three-branch profile (6) to 10^{-11} across the full range. The figures are regenerated by the same script, so no number in the paper is transcribed by hand.

13 Discussion

The result is conditional but operational. A justified positive coverage margin, a recording floor, a bound on exceptional recoverable mass and a calibration allowance yield an explicit finite error guarantee and an explicit number of units; an admitted invisible type, a zero margin or a shared run failure prevents it, and no amount of additional sampling repairs any of these. Three points deserve emphasis.

First, the arithmetic of the floor is not an artefact of the restricted parameterisation. Extending c to its natural range $[0, 2]$ changes the sharp value, removes the apparent $1/4$ ceiling, produces a positive floor with no mismatch information at all once $c > 1$, and — as the last row of Table 2 shows — changes a planning answer by a factor of three. Readers who encounter such a ceiling in a related model should check whether it is a theorem or a domain convention.

Second, sharpness statements are claims about a class. Every optimum here is attained by an explicit source, so the bounds cannot be improved without adding a hypothesis that excludes those sources. That is a strength for design — one knows exactly what is being protected against — and a limitation for interpretation: a real population is usually not the least detectable one, so the guarantee is conservative by construction.

Third, the mathematics cannot supply the premises. Establishing c, D, κ, η and δ for an exposed population, with a reference task that is not defined circularly as "whatever this test detects", remains the central empirical task. Until that is done, an all-negative record is a statement about what was observed, not about what is alive; the viable-but-nonculturable literature [12] and the persistence literature [2, 3] both suggest that the gap between the two is where the biology lives.

Data and code availability. No new experimental data were generated or analysed. The Lean sources, verification receipts, the numerical script that reproduces every constant and figure, and the manuscript sources accompany this article.

References

[1] Maïlis Amico and Ingrid Van Keilegom. Cure models in survival analysis. *Annual Review of Statistics and Its Application*, 5:311–342, 2018. doi: 10.1146/annurev-statistics-031017-100101.

[2] Nathalie Q. Balaban, Jack Merrin, Remy Chait, Lukasz Kowalik, and Stanislas Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, 2004. doi: 10.1126/science.1099390.

[3] Nathalie Q. Balaban, Sophie Helaine, Kim Lewis, Martin Ackermann, Bree Aldridge, Dan I. Andersson, Mark P. Brynildsen, Dirk Bumann, Andrew Camilli, James J. Collins, Christoph Dehio, Sarah Fortune, Jean-Marc Ghigo, Wolf-Dietrich Hardt, Alexander Harms, Matthias Heinemann, Deborah T. Hung, Urs Jenal, Bruce R. Levin, Jan Michiels, Gisela Storz, Man-Wah Tan, Tanel Tenson, Laurence Van Melderen, and Annelies Zinkernagel. Definitions and guidelines for research on antibiotic persistence. *Nature Reviews Microbiology*, 17:441–448, 2019. doi: 10.1038/s41579-019-0196-3.

[4] Asher Brauner, Ofer Fridman, Orit Gefen, and Nathalie Q. Balaban. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nature Reviews Microbiology*, 14 (5):320–330, 2016. doi: 10.1038/nrmicro.2016.34.

[5] Asher Brauner, Noam Shoresh, Ofer Fridman, and Nathalie Q. Balaban. An experimental framework for quantifying bacterial tolerance. *Biophysical Journal*, 112(12):2664–2671, 2017. doi: 10.1016/j.bpj.2017.05.014.

[6] C. J. Clopper and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413, 1934. doi: 10.1093/biomet/26.4.404.

[7] Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28*, volume 12699 of *Lecture Notes in Computer Science*, pages 625–635. Springer, 2021. doi: 10.1007/978-3-030-79876-5_37.

[8] Xin Fang and Kyle R. Allison. Resuscitation dynamics reveal persister partitioning after antibiotic treatment. *Molecular Systems Biology*, 19:e11320, 2023. doi: 10.15252/msb.202211320.

[9] James A. Hanley and Abby Lippman-Hand. If nothing goes wrong, is everything all right? interpreting zero numerators. *JAMA*, 249(13):1743–1745, 1983.

[10] R. A. Maller and X. Zhou. Estimating the proportion of immunes in a censored sample. *Biometrika*, 79(4):731–739, 1992. doi: 10.1093/biomet/79.4.731.

[11] Charles F. Manski. *Partial Identification of Probability Distributions*. Springer, New York, 2003.

[12] James D. Oliver. Recent findings on the viable but nonculturable state in pathogenic bacteria. *FEMS Microbiology Reviews*, 34(4):415–425, 2010. doi: 10.1111/j.1574-6976.2009.00200.x.

[13] The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020)*, pages 367–381. ACM, 2020. doi: 10.1145/3372885.3373824.

[14] Frank Tuyl, Richard Gerlach, and Kerrie Mengersen. The rule of three, its variants and extensions. *International Statistical Review*, 77(2):266–275, 2009. doi: 10.1111/j.1751-5823.2009.00078.x.