

OPERATING CASE STUDY / RISK, CONTROLS & GOVERNANCE

Scan Integrity & Outcome Verification

Make instrumentation earn its connection to service

An operating-project case on disputed return attribution in a contracted delivery network: auditing nested opportunity estimates, testing scan-data readiness and defining an authorized response that can be evaluated against physical customer outcomes.

EVIDENCE RECONCILIATION & OPERATING DESIGN

BLUF / BOTTOM LINE UP FRONT

Validate the signal before assigning cause. Test the response before scaling it.

RECALCULATED FINDING

Analytical finding: correcting a nested denominator does not resolve overlap between opportunity tiers. A missing scan can reflect an inapplicable step or failed capture; a cleaner recorded metric can therefore leave the physical service outcome unchanged.

CONTRIBUTION IN THE SOURCE RECORD

Recorded leadership: arithmetic challenge, data-readiness inquiry and response governance. Contribution evidence remains provisional.

EVIDENCE STATUS

Source-register reconstruction; originals unrecovered. No verified causal effect, realized saving or completed production rollout.

BLUF: analytical inference. Source contribution: historical record. Client and internal identities anonymized.

STRUCTURED ABSTRACT / EXECUTIVE DECISION

Separate a better record from a better delivery

[AI] The investment decision is whether validated event evidence can support a timely, authorized recovery action. More complete logging, a smaller attributed bucket and a stronger correlation do not answer that question.

Context

[HR] A delivery network was investigating packages marked missing and later returned. A team analysis ranked site-level associations, a leadership discussion assembled a purportedly controllable share, and a parallel workstream sought driver-device evidence. A default ownership rule and incomplete event coverage complicated interpretation.

Objective

[HR/AI] The recorded operating lead challenged the opportunity arithmetic, commissioned a data-readiness determination and reframed the response through the organization that owned the partner relationship. The decision was whether the evidence justified a specific mechanism, an analytical build and a supported accountability claim.

Approach

[RC/AI] Reconcile each population and denominator; distinguish a subset correlation from an independent cause; test whether scenario points belong to the stated line; separate applicable process steps from captured scans; preserve query defects, maturity and ownership boundaries. The technical layer specifies an independent physical reference and a prospective policy comparison.

Findings

[RC/AI] The corrected nested calculation still requires overlap reconciliation. The recorded scenario points cannot all follow one common straight line as stated. In a constructive measurement example, improved capture changes the apparent missing-scan rate without changing the underlying violation rate. A cycle-type breakdown therefore answers a composition question while leaving guardrail-specific causation open.

Decision value

[PD] The accountable operations sponsor should authorize a bounded validation stage. The data owner must reconcile eligibility, event grain and reference accuracy; the relationship owner must approve the response. Release live prompts only after these gates, and expand only when a prospective comparison clears economic break-even and protects customer outcomes.

Implications

[AI] This is a diagnostic and design case. The source record supports analytical challenge and operating judgment; it does not establish corrected-number adoption, production signal readiness, a completed outreach tool or a mature service improvement. The proposed framework and numerical trial are new analysis, not historical results.

Claim classes: [HR] historical record; [RC] recalculated; [AI] analytical inference; [PD] proposed design. Source basis S1, the supplied project register; original artifacts remain unrecovered.

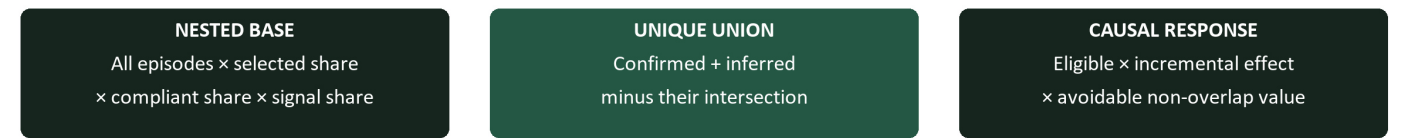
EXHIBIT 1 / OPPORTUNITY INTEGRITY

Fix the denominator, then reconcile the union

[RC/AI] A corrected component can remain an invalid total. Nested populations, overlapping signals and causal responsiveness are separate tests.

DENOMINATOR / OVERLAP / RESPONSE

Three separate reasons an opportunity estimate can fail



Correct multiplication repairs the base; it does not establish disjointness or preventability.
Adding a child category to its parent counts the child twice.

Conceptual identities. Each factor needs an explicit population, event grain, version and maturity window.

[HR] The recorded challenge concerned an inferred component applied to the full same-driver population even though the surrounding discussion described the compliant subset. A second challenge asked whether the inferred and scan-confirmed tiers were mutually exclusive. The record preserves the competing denominator interpretations; it does not confirm which estimate was retired.

Test	Required evidence	Decision consequence
Base	Identify the population on which each conditional share was measured.	Use the complete nested product; retain the alternative if scope remains unresolved.
Union	Join the signal memberships at the same episode grain and maturity.	Subtract the intersection. Independence does not justify adding shares.
Response	Establish that the authorized action changes a mature outcome.	A diagnostic candidate is not yet a controllable or avoidable-cost share.

[RC/AI] The proposed additional recovery tier shares a parent population with the scan-confirmed tier. Under that nesting, adding the child again counts the same eligible episodes twice. The category flags inside the inferred tier can also overlap; their individual shares cannot be summed into a union without a joint-membership table.

EXECUTIVE ACTION

[PD] Freeze a single population contract and episode-level membership table before pricing the intervention. Carry unresolved overlap as bounds, and keep causal preventability outside the arithmetic total.

S1: nested opportunity chain and recorded challenges. Recalculations, alternative-base checks and overlap bounds are retained in the private annex. No historical raw share or dollar anchor is disclosed.

EXHIBIT 2 / ANALYTICAL IDENTIFICATION

A strong subset correlation does not identify a lever

[AI] The high-ranking return category is already part of the total defect measure. Its statistical association can be strong even when it provides no independent explanation of the remaining defects.

PART-WHOLE CORRELATION / A CONSTRUCTIVE LIMIT

Total defect rate = selected component + residual component

DESCRIPTIVE ASSOCIATION

The selected component is already in the total.
Shared exposure can affect both rates.
A high correlation does not isolate a mechanism.

INTERVENTION QUESTION

Does the approved operating action reduce independently verified mature failures?
The source comparison does not identify this.

Correlation is scale-free; a regression slope requires the standard-deviation ratio.
A common fitted line must also pass the recorded scenario points.

New line-consistency analysis is retained privately; no confidential rate or correlation anchor is exposed.

[HR] The source describes a cross-sectional station ranking, outcome differences between selected groups and a modeled improvement from moving a return metric toward a benchmark. These are descriptive associations and a scenario. No controlled intervention or recovered station-level data establishes a closable gap.

Source proposition	Analytical test	Conclusion
Dominant correlation	Write total = selected component + residual; inspect shared exposure and covariance.	Part-whole structure and a common denominator can generate association.
One fitted line	Compare the slopes implied by the three recorded scenario points.	The points are incompatible with one common line as stated.
Worst-site opportunity	Compare untreated recurrence, exposure, mix and selection reliability.	Benchmark gaps require a causal test; ranked recovery can include reversion.

[RC/AI] The first two recorded scenario points imply a materially different third prediction from the stated low-end benchmark. The near match between the implied slope and the reported correlation is a diagnostic clue, not proof of the author’s algorithm: correlation becomes a regression slope only after the outcome-to-input standard-deviation ratio is supplied.

EXECUTIVE ACTION

[PD] Obtain the original station extract, model specification and weighting before using the projection. Re-estimate on an independent period and test the operating action, rather than pricing the cross-sectional association.

S1: station correlation matrix, benchmark groups and modeled scenario. The line-consistency calculation is new analysis. Association, projection and causal treatment effect remain distinct.

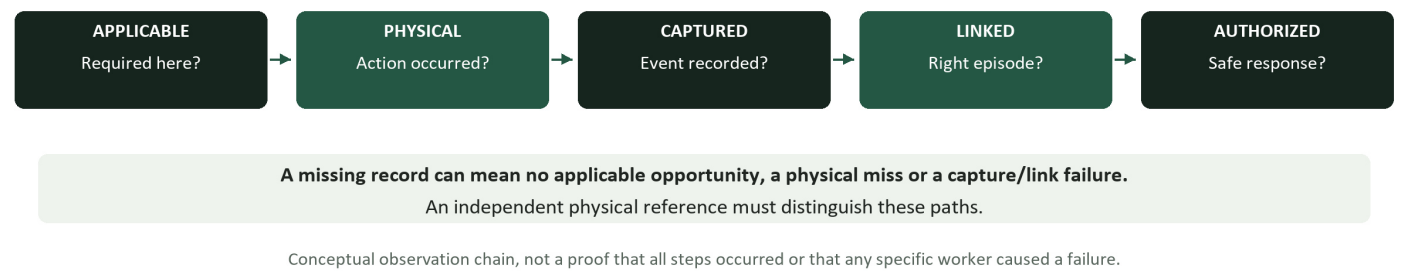
EXHIBIT 3 / SCAN MEASUREMENT

Establish an applicable opportunity before calling a missing scan a violation

[AI] A scan record is the result of several processes: an applicable requirement, physical activity, capture, transmission, linkage and classification. Failure at any stage changes what the dashboard can observe.

SCAN-EVIDENCE AND ELIGIBILITY CONTRACT / PROPOSED

Distinct stages must remain visible in the data and in the operating decision



[HR] The compliance description invoked a full operational scan sequence, while the documented analytical schema exposed only part of that sequence. The readiness inquiry identified the missing fields and unresolved guardrail-specific evidence. No row-level query ran in that inquiry; schema inspection did not validate episode-level classification.

Observed condition	Interpretation retained	Response
Required step, validated absence	Candidate physical process exception.	Investigate the mechanism against the independent reference.
Step not applicable	No opportunity to satisfy that scan requirement.	Exclude from that compliance denominator, retain in original service cohort.
Capture or link unknown	Observation is incomplete.	Preserve unknown; do not impute worker fault or physical absence.

[AI] The new constructive example holds physical compliance fixed while increasing capture. The recorded missing-scan measure improves, but the true eligible violation rate does not. The converse is also possible: better detection can increase recorded violations while physical failures decline. Neither direction establishes success without an independent outcome.

EXECUTIVE ACTION

[PD] Require a versioned eligibility rule, a traceable capture chain and a blinded physical reference. Publish separate coverage, eligible compliance and mature service measures; do not combine them into one compensating score.

S1: schema and instrumentation readiness record. The scan-observation decomposition and constructive examples are new analysis; their hypothetical counts are not source operational data.

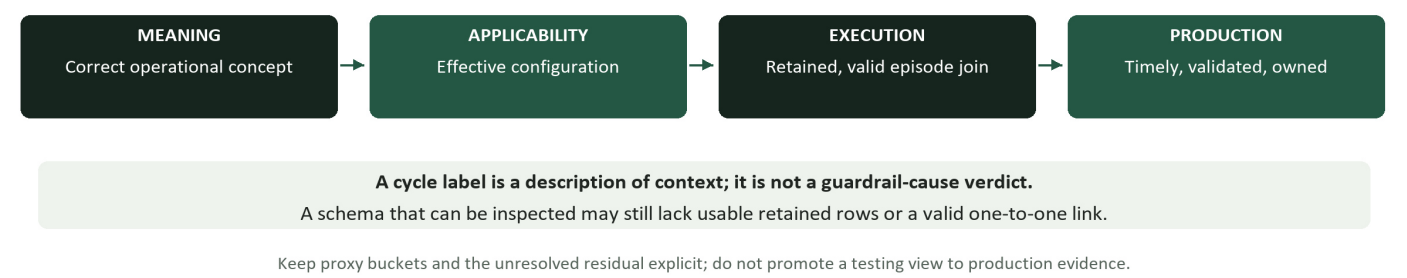
EXHIBIT 4 / READINESS AND SCOPE

Do not turn cycle composition into cause

[AI] Sorting exceptions by operating cycle identifies where cases occur. It does not establish which guardrail applied, whether it was enabled, what action occurred or why a physical failure followed.

DATA READINESS HAS SEPARATE ACCEPTANCE TESTS

Passing an earlier test does not establish the next one



[HR] The metadata inquiry found partial feasibility: access and schema leads, missing configuration history, ambiguous override evidence and short retention. A separately reported cycle-type breakdown answered a composition question; it did not establish guardrail-specific causes.

Readiness layer	Acceptance requirement
Meaning	Confirm the field represents the relevant package guardrail, not a labor cap or another workflow with a similar name.
Applicability	Recover configuration effective at the station, cycle and event time; unknown enablement remains unknown.
Execution	Probe retained episodes, cardinality and timestamp ordering; metadata access is not a successful analytical join.
Production	Validate independent accuracy, latency, safe fallback and an accountable owner before live operational use.

[AI] A station-actor touch may represent a legitimate intervention or a different exception. Subtracting proxy buckets does not create a clean residual without verified coverage, mutual exclusivity and grain. A separately reported descriptive pull does not validate those assumptions.

EXECUTIVE ACTION

[PD] Commission the smallest discriminating probes first. Retain a cycle composition view for planning, while holding guardrail-specific attribution until effective configuration and episode evidence can be joined.

S1: partial data-readiness determination, named false matches, retention boundary and separately reported cycle pull. Original queries and row-level validation have not been recovered.

EXHIBIT 5 / CUSTOMER AND FINANCIAL EXPOSURE

Reconcile exposure before pricing value

[AI] Outcome concentration can guide investigation. A cost map with inconsistent scope, immature outcomes or repeated join rows cannot establish preventability, causal ownership or financial benefit.

COST ATTRIBUTION REQUIRES MORE THAN A STATUS LABEL

Independent reconciliation gates for a decision-grade cost map

COMMON COHORT
One eligible universe

UNIQUE EPISODE
Controlled join cardinality

MATURE OUTCOME
Common follow-up clock

INCREMENTAL VALUE
Effect, overlap and cost

Concession count share is not automatically dollar-cost share.
A virtual on-road status does not exclude an upstream physical failure.

Historical query defects remain unresolved; incurred exposure is not an identified savings opportunity.

[HR] The record separates later-recovered packages from unresolved and lost outcomes, and challenges a claim that every unrecovered package generates a loss concession. It also documents a package-universe mismatch, a many-to-many join problem and unusable owner-ranking queries. Those defects constrain both the cost narrative and the attribution claim.

Failure mode	What must be repaired
Population mismatch	Apply one network and eligibility definition to the package cohort, outcome records and concession ledger.
Join multiplication	Freeze an episode key, order repeat events and reconcile every row expansion before aggregating units or money.
Incomplete maturity	Follow the original cohort to a common horizon, preserve unresolved outcomes and distinguish missed promise from eventual recovery.
Virtual status as cause	Validate physical custody independently; an itinerary status cannot exclude upstream failure or identify personal fault.

[RC/AI] A concession-count share does not prove a dollar-cost share without group-specific amounts. A query that duplicates package rows can also duplicate money unless its financial aggregation is separately protected. The private audit retains these unresolved issues; the public case does not price them as corrected historical savings.

EXECUTIVE ACTION
[PD] Ask Finance and the data owner for an episode-level reconciliation before approving an opportunity estimate. Then apply the incremental mature effect, avoidable share, non-overlap and full cost to the same eligible population.

S1: outcome/concession query summaries and documented defects. Cost is an incurred exposure, not a benefit. The proposed economic model is separately labeled and has no historical calibration.

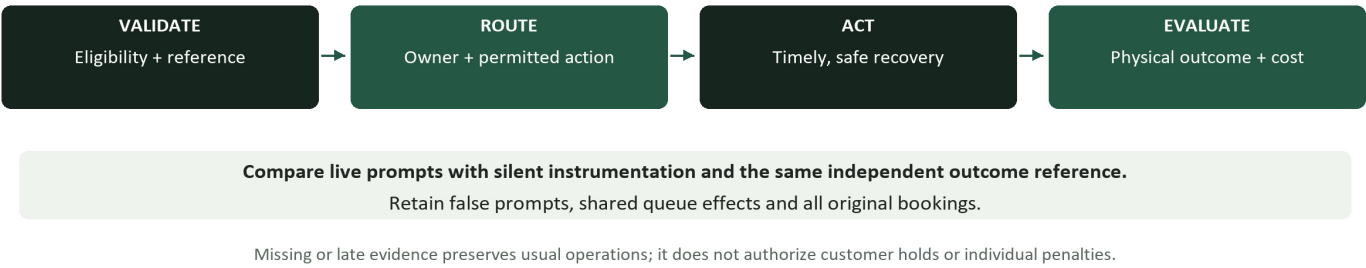
EXHIBIT 6 / OPERATING RESPONSE

Give every signal an authorized action and a closure test

[AI] A ranked list, email reply or new reporting split cannot produce an outcome without an executable response. Authority and closure evidence are part of that response.

FROM SIGNAL TO VERIFIED SERVICE / PROPOSED

Authority, safe fallback and original-cohort measurement are part of the mechanism



[HR] The outreach design was paused after a challenge to its tracking mechanism. A replacement tool was proposed but completion is unverified. A reporting split acquired named owners; its impact remained unsized. The partner-facing route was reframed through the relationship owner, while an intended outreach draft remained unsent.

Role	Proposed responsibility	Closure evidence
Operations sponsor	Set the bounded service decision, resource ceiling and escalation rule.	Original-cohort physical result and complete resource record.
Data and signal owner	Validate eligibility, reference accuracy, event grain and timeliness.	Versioned replay, independent audit and visible unknowns.
Relationship owner	Approve the partner-facing response within delegated authority.	Documented acceptance, action and unresolved conditions.

[PD] The proposed trial compares live, approved evidence prompts with the same underlying instrumentation operating silently. Both arms retain usual recovery and an independent physical outcome reference. Late or missing signals do not authorize a penalty, customer hold or presumption of fault. False prompts and shared capacity costs remain in the evaluation.

EXECUTIVE ACTION

[PD] Fund shadow validation first. Require a feasible, powered comparison before field expansion; a cleaner dashboard or a favorable point estimate alone does not satisfy the investment gate.

S1: outreach pivot, relationship-governance boundary, unsent draft and reporting action. The live-versus-silent trial is a proposed design, not a recorded deployment.

EXHIBIT 7 / BOUNDED CONTRIBUTION

Credit the contribution by workstream

[HR/AI] The record describes several lanes with different owners. Managerial review, original analysis, signal development, approval and implementation are separate contributions.

ONE PROJECT / DISTINCT CONTRIBUTION LANES

The source register supports bounded roles, not universal ownership

LEAD Challenge + readiness	ANALYST Station study + outreach	SIGNAL OWNER Telemetry + cycle pull	RELATIONSHIP OWNER Partner-facing decision
--	--	---	--

An accepted invitation is evidence of participation, not convening.
Current proofs and the prospective evaluation are new analysis, not historical achievements.

Contribution grades are provisional while original artifacts are unrecovered; no realized impact is claimed.

Contribution claim	Process-tracing disposition
Arithmetic and overlap challenge	Provisional hoop at register level: the record attributes the calculation and prepared challenge to the operating lead. Original notes must authenticate it.
Readiness inquiry and governance framing	Provisional hoop: commissioning, the partial verdict and a bounded response draft are recorded. A draft does not establish external delivery or adoption.
Review and shared workstream ownership	Provisional hoop for the specified review and named role. The underlying station analysis and signal build belong to other contributors.
Meeting, production and results	Acceptance of an invitation supports participation only. Convening, production readiness and mature impact are not established for the lead.

[AI] No personal claim passes a smoking-gun or doubly-decisive test solely because the register labels its underlying source primary. The original artifact must be recovered and tested against timing, authorship and alternatives. Adjacent-project results remain outside this case’s achievement claims.

[AI] The line-consistency test, scan-observation proof, overlap analysis and proposed evaluation are current analytical contributions. They are not retroactive descriptions of the original project team’s methods or a claim that those methods already changed network performance.

PORTFOLIO VALUE

[AI] The supported capability is disciplined operating judgment: challenge a load-bearing estimate, expose data limits, preserve ownership boundaries and turn an ambiguous signal into a testable response.

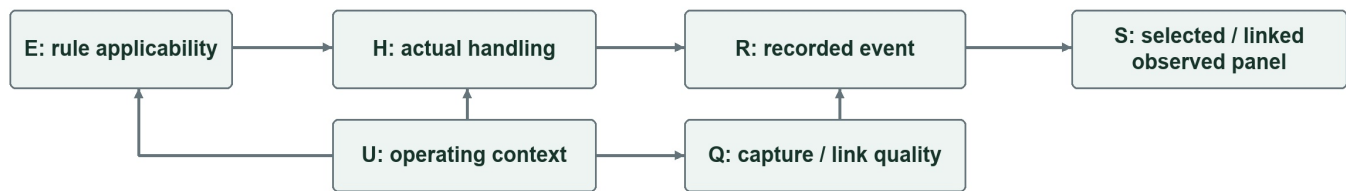
S1: lane-specific ownership, review thread, personal-note attribution, signal-owner records, co-owned reporting action and invitation acceptance. Historical outcomes and downstream adoption remain unverified.

TECHNICAL ANALYSIS

Separate required work, physical action and recorded evidence

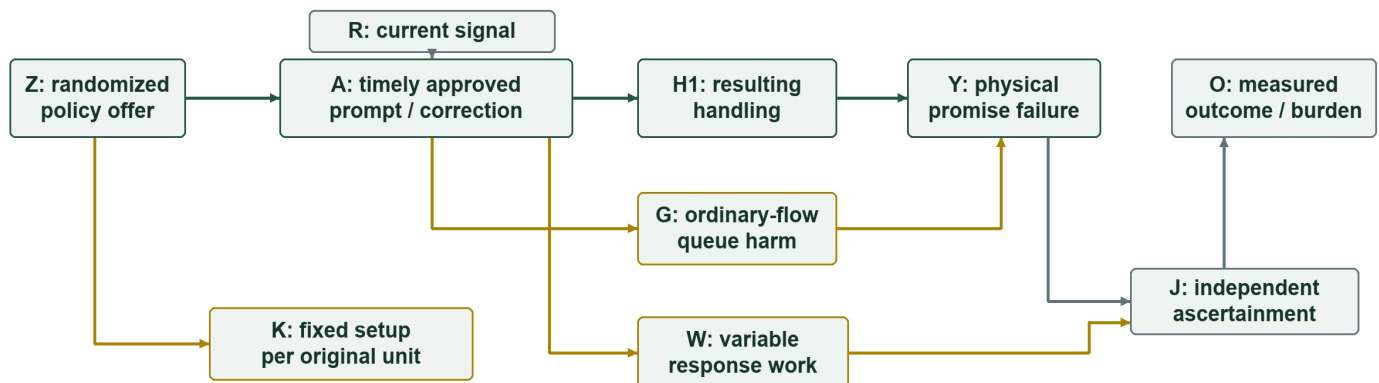
[AI/PD] The proposed Scan-Evidence and Eligibility Contract (SEEC) separates whether a scan was required, what physically happened, what was recorded and whether a timely signal supports an authorized action. Missing telemetry alone cannot close this chain.

MEASUREMENT: ABSENCE IS NOT A PHYSICAL-OMISSION LABEL



Context also affects outcomes and selection. Conditioning on S can change the mix; E is fixed before action.

POLICY TRIAL: LIVE PROMPTS VERSUS THE SAME SILENT LOGGING



Y and W form variable burden B. J measures Y and W; K is charged once outside B.

Baseline risk affects R and Y. Feed outages remain in the original cohort; no remedy or blame rule changes.

[AI] Applicability E and actual handling H affect recorded event R; capture and linkage quality Q also affect R. Context U affects E, H, Q and physical outcome Y. Selection on returned, conceded or successfully linked episodes changes the observed mix. A virtual status or default identity label is not an independent physical reference.

[PD] Randomized offer Z enables live prompts A using timely signal R. Prompts can improve subsequent handling H1, but also consume work W and displace ordinary-flow capacity G, changing Y. Independent ascertainment J measures actual Y and W to produce O. Identical silent logging in both arms separates the operating offer from merely increasing recorded capture.

[PD] Primary ITT follows the original manifest, including outages and absent signals. Do not adjust it for alerts, completed corrections or successful links. Fixed setup K is outside variable burden B. Shared feed and capacity require domain-level randomization. These paths are proposed mechanisms, not identified historical effects.

Basis: source measurement and feasibility distinctions [HR/AI]; proposed SEEC trial [PD]; Hernan and Cole (2009). Reference independence concerns the policy decision, not statistical independence from the outcome it measures.

TECHNICAL ANALYSIS

Better capture can change the metric without changing handling

[AI] These independent hypothetical examples are not indexed historical data. They show why the applicability denominator and reference measurement must be fixed before interpreting a scan-based performance change.

Original cohort of 100	Units	Recorded scans
Scan not applicable	40	0
Required action omitted	6	0
Required action completed	54	27

[AI] Missing records are 73: 40 inapplicable episodes, six true omissions and 27 unrecorded completed actions. Perfect capture reduces missing records to 46, while the omission rate among eligible opportunities stays $6/60 = 10\%$. Treating all missing scans as noncompliance would punish inapplicability and measurement failure.

[AI] Let q be the eligible fraction, v the true omission rate among eligible episodes and s the probability of recording a completed applicable action. Under this stated model with no false scans, missing records/all episodes = $1 - q(1 - v)s$. Even at $s = 1$, absence need not vanish. This equation is a conditional measurement model, not an empirical correction fitted to the source.

[AI] Detection can also reverse apparent direction. Suppose true violations fall from 20% to 10%, sensitivity rises from 25% to 90%, and specificity stays 100%. Recorded positives rise from 5% to 9%. Thus a rising recorded problem need not mean physical deterioration, and a falling missing-scan count need not mean physical improvement.

[PD] With p the true violation fraction and r the recorded-positive fraction, $r = \text{sensitivity} \times p + (1 - \text{specificity}) \times (1 - p)$. Recovering p requires independently identified, transportable accuracy parameters and a nonzero sensitivity + specificity - 1. Report uncertainty and unsupported strata rather than treating a calibrated point estimate as truth.

[AI] Rogan and Gladen motivate separating screening frequency from prevalence. Stone and colleagues show the value of independent recording-time evidence, while their study also cautions against treating an instrumented record as the underlying state or importing its empirical compliance rates here.

Basis: independently enumerated synthetic examples and measurement identity [AI]; Rogan and Gladen (1978), Stone et al. (2002). No historical violation prevalence or misconduct is estimated.

TECHNICAL ANALYSIS

A decomposition needs set membership, not just multiplied shares

[HR/AI] The source contains alternative denominators, potentially overlapping signal tiers and a branch nested within another branch. These are separate questions from whether any identified episode is preventable by a particular control.

Check	What it establishes; what remains open
Eligibility base	Which original episodes could require this action, under the rule effective at that time. A cycle name alone is not the enabled-rule export.
Nested branch	If B is a subset of A on the same entities and predicates, A union B equals A. Adding both counts B twice.
Separate tiers	For validated sets of size a and b in universe N, their union lies from max(a,b) to min(N,a+b) without the joint table.
Control effect	A signal can identify a candidate for review without identifying fault, avoidability or the benefit of a specific remedy.

[HR/AI] The lower-confidence percentage in the register has alternative all-defect and compliant-subset bases. The reported rounded-count chain also differs from multiplying rounded percentages. Preserve both until a joint episode table establishes the intended population. Renormalizing within a selected residual does not convert that composition into an all-population effect.

[AI] A strong part-whole correlation is similarly insufficient. If $Y = S + R$, then $Cov(S,Y) = Var(S) + Cov(S,R)$. With independent components having variances four and one, the correlation squared is 0.80 even though no causal intervention has been identified. Shared exposure can add another source of association.

[RC/AI] A private algebraic audit finds that the source scenarios cannot all lie on the one affine relationship they invoke. Two stated pairs determine a line that misses the third scenario substantially. Correlation is not the regression slope without the standard-deviation ratio; resemblance of two coefficients does not establish a coding error.

[PD] Reproduce the fitted specification, units, weights, intercept, uncertainty and out-of-sample population before using a projection for funding. Reconcile the episode-level sets before summing opportunities; do not force a union of proxies into an exclusive causal Pareto.

Basis: source denominator and projection forks [HR/RC]; independently derived set bounds and finite part-whole example [AI]. Historical raw values and alternative calculations remain private.

TECHNICAL ANALYSIS

A discoverable schema is not an executed causal trace

[HR/AI] The register describes a bounded metadata and access investigation, a separately reported partner cycle breakdown and a test-stage telemetry view. These support different evidence claims and cannot be combined into a recovered production query.

Layer	Required evidence before release
Applicability	Rule version and effective station/cycle/time configuration; explicitly distinguish required, optional, disabled and unknown.
Episode trace	Original shipment episode, event and posting timestamps, actor role, authorized transfer, correction history and linkage confidence.
Customer outcome	Original promise and independent receipt/nonreceipt disposition at a common horizon, retaining unresolved episodes.
Resource ledger	Unique charge or transaction identity, its eligible scope and episode allocation; audit one-to-many joins before summing cost.

[HR/AI] The inspected schema did not contain every field needed to reproduce the stated complete scan path, and no row-level SELECT ran in that feasibility probe. This does not prove those events never existed elsewhere. The separately reported partner cycle mix answers a composition question; it does not supply the absent rule-enablement history or validate a causal noncompliance split.

[AI] Preserve event episodes rather than collapsing all recurring identifiers to one lifetime package row. A latest-status or MAX flag answers a particular existence question, not custody at an earlier mark. A virtual on-road status may describe an itinerary; it does not establish physical presence, misconduct or the absence of an upstream failure.

[AI] Join fanout can inflate financial sums as well as counts. Deduplicate true duplicate charge records at their own grain before allocating to episodes; do not remove distinct legitimate charges with blanket DISTINCT. A selected never-scanned cell cannot rule out a station mechanism when capture is imperfect.

[AI] Eventual delivery at extraction does not establish timely delivery, ultimate recovery or low customer burden. Concession records depend on eligibility, policy, posting and maturity as well as physical outcomes. The Process Mining Manifesto supports explicit event semantics and completeness scope; it does not certify this historical log.

Basis: source feasibility, join and clock limitations [HR/AI]; van der Aalst et al. (2012). Data availability, successful execution, production use and causal validity are distinct stages.

TECHNICAL ANALYSIS

Audit eligibility, capture and action validity separately

[PD] Before a live test, use a probability sample from the original manifest, stratified by rule applicability, cycle, signal presence, feed timeliness and linkage status. Include unknown and signal-negative episodes so verification does not depend on the detector.

Reference question	Validation output
Was the action required?	Independent effective-rule audit; applicable, inapplicable or unresolved. Do not infer eligibility from a subsequent alert.
Was it physically completed?	Authenticated trace or prospective observation, separate from the candidate logging system.
Was it captured and linked in time?	Sensitivity/specificity by applicable stratum; latency, false links, missed links and correction rates.
Was the response safe and useful?	Approved action record, customer acceptance when relevant, work incurred and independently mature promise outcome.

[PD] Freeze a codebook and have two blinded reviewers code independently before adjudication. Report the full confusion matrix, class prevalences, raw agreement and kappa with uncertainty; use weighted kappa only for a justified ordered scale. High agreement on an invalid proxy remains invalid measurement.

[PD] As a preliminary precision screen, 200 independent reference episodes give a worst-case approximate 95% proportion interval of plus or minus 6.9 percentage points. This is not an adequate guarantee for rare harmful errors, subgroup accuracy or correlated domains. Plan stratum sample sizes and cluster-adjusted intervals from the actual error frequencies.

[AI] Hui and Walter show how error rates may be identified without a perfect standard under restrictive multi-population and conditional-independence assumptions. Two feeds derived from the same device, default logic or upstream table may violate those assumptions. More agreeing telemetry does not automatically create independent truth.

[PD] Keep unresolved reference states; bound missing binary outcomes on the original denominator rather than discarding failed links. Set false-alert, safety and latency thresholds before holdout evaluation. The operating rule uses decision-time information only. Later independent reference evidence stays hidden during tuning and policy use; blinded adjudication may use it to reconstruct the actual outcome.

Basis: proposed blinded validation [PD]; Cohen (1960), Hui and Walter (1980), Hernan and Cole (2009). The precision example is independent and hypothetical; no historical audit sample is claimed.

TECHNICAL ANALYSIS

Do not mistake selected extremes or changing labels for a mechanism

[HR/AI] The source combines cross-sectional ranks, selected residual compositions, gapped performance quotes and measurement versions. None is a recovered stable intervention phase, and a later increase does not itself establish either implementation failure or an instrumentation artifact.

[RC] The frozen inventory contains 10 source/sensitivity screens, nine separate source-version audits, five mechanical plan audits and one cross-sectional profile. All 10 source/sensitivity screens lie inside fitted tentative X/MR limits; none establishes a stable phase. No eligible Laney panel exists. The X multiplier is 2.66 and MR upper multiplier 3.267; numerical coordinates remain private.

[AI] Compare fitted and frozen-reference limits only on commensurate observations. A crossing after a changed definition requires investigation; refitting the limits can absorb a step. No stable-phase claim follows from all points lying inside wide limits. Large exposures alone do not establish a valid binary numerator or justify a Laney chart.

[AI] Worst-site selection on the total also selects its component. Under an explicitly assumed linear conditional-mean model, untreated follow-up is $m + \rho * (x - m)$, where m is the population mean, x the selected value and ρ assumed persistence. Values 0.4, 0.7 and 0.9 imply 60%, 30% and 10% reversion of selected excess. No empirical reversion fraction is identified here.

[PD] Freeze the focus cohort before follow-up and compare repeated pretreatment levels, mix and mature outcomes with a credible comparison. Do not select successful corrections or later returns. Prespecify moderators: rule applicability, device coverage, feed delay, cycle configuration, baseline risk, staffing capacity and cross-domain routing. A lower signal rate can reflect poorer detection rather than better operations.

[HR/AI] The partner-authored residual cycle mix is not the causal share reachable by a cycle-specific control. Likewise, many affected people having one observed event does not mean most events are isolated or that exposure-adjusted coaching would be effective or ineffective. The sibling identity panel uses a different population and role comparison; it cannot settle this case by transporting its labels.

[PD] Define independent capacity domains around shared feeds, staff and routing. If spillovers persist, estimate exposure mappings or redesign assignment. Ordinary-flow harm must be measured beyond the selected alert population.

Basis: source selection/version limits [HR/AI]; proposed SPC and RTM sensitivity [PD/AI]. Final inventory is independently reproduced privately; no observed treatment effect or ranking reliability is claimed.

TECHNICAL ANALYSIS

Credit the review and keep deployment and outcomes unclaimed

[HR/AI] Contribution is a provisional registry-level hoop: the derivative record supports specific review, design and coordination claims, but original artifacts have not been recovered. Present analytical corrections are distinguished from the historical participant's recorded work.

Recorded role	Supported boundary
Review and delivery design	Reviewed a team and AI-assisted report and challenged whether outreach would lead to accountable action; not sole report authorship.
Denominator and feasibility work	Recorded caveats about branch bases and overlap, and commissioned a bounded data-availability investigation; no claim of completed row-level reproduction.
Governance and framing	Prepared a hypothesis-based partner handoff and revised the narrative; the draft remained unsent and another forum carried the discussion.
Measurement development	Named co-owner of a reporting split and participant in another owner's telemetry work; no claim of building the dashboard or convening that session.

[AI] Competing explanations include team authorship, later retrospective synthesis and common source lineage. Repetition across a memory note, call notes and a derivative register is not independent corroboration. Attendance acceptance is not convening, and a planned date passing does not prove a capability was never deployed elsewhere.

[HR/AI] The record describes a test-stage telemetry view and open implementation questions. It does not establish a mature customer-outcome panel, executed production safeguards, approved individual action or realized savings. A subsequent performance quote is a different observation, not an untreated counterfactual.

[AI] Consolidated limits: applicability and physical-reference accuracy are unresolved; overlap and join grain limit opportunity sizing; concession scope and maturity limit value estimates; correlated measurement limits triangulation; selected sites and cycles limit transport; shared capacity limits independence. Funding and authority remain separate from a technically available signal.

[PD] Release requires an accountable data owner, approved operating response, independently calibrated reference and a bounded trial decision. Preserve a reversible no-new-action state when mandatory evidence is unresolved. Route partner-facing decisions through the authorized relationship owner; do not convert an uncertain signal into individual blame.

Basis: lane-specific registry attribution [HR/AI]; Collier process-tracing principles and Kellogg, Orlikowski and Yates (2006). Neither an original implementation nor a causal operating benefit is established.

TECHNICAL ANALYSIS

Test live prompts against identical silent logging

[PD] Randomize independent operating-capacity domains before dispatch. Both arms retain the same logging and independent outcome reference; the offer activates approved timely reconciliation prompts, while the comparator preserves usual response. This tests an operating policy, not whether adding a field changes a dashboard.

Estimand	Population, clock and contrast
Physical ITT benefit	Usual minus offer original-promise failure risk across all originally booked units, adjudicated at original promise plus 28 days. Later recovery does not erase a missed promise.
Resource ITT	Usual minus offer mature failure burden and variable alert/correction work, per original unit; compare with fixed incremental setup cost once.
Ordinary-flow safety	Offer minus usual failure risk among the ordinary stratum fixed before assignment; a prespecified noninferiority margin requires authorized acceptance.

[PD] Scan-applicable opportunities and ordinary bookings are frozen upstream. The policy sees the current signal, not the latent risk used in a planning model. Outages and late signals retain usual operations and remain in the original denominator. Charge every alert and every successful correction, including corrections of units that would not otherwise fail. No customer remedy, promise or disciplinary policy changes.

[PD] The separate planning model derives finite conditional outcome moments with common feed dependence, then uses explicitly approximate Gaussian decision bounds and a conservative joint-power floor. Validate candidate designs using direct finite-process simulation at physical-null, economic-boundary and ordinary-safety-boundary laws, with Monte Carlo uncertainty. These checks do not establish real-world calibration.

[PD] Require detectable economic break-even and joint physical, economic and safety gates. Pilot measurement, variance, outage and recruitment feasibility first; no global sample-size optimum is needed. Discount only favorable avoided harm when testing realization assumptions, charging adverse ordinary-flow effects and work fully. Information value is a separate decision model, not an extra operating benefit.

[PD] If later adoption is staggered rather than randomized, estimate group-time ATT against not-yet-treated domains. Check anticipation, common support, pretreatment event-study and equivalence diagnostics, mix and interference; mature outcome counts and a new power analysis are required. The randomized ITT above and this unestimated field extension are different estimands.

Basis: proposed SEEC trial and wholly hypothetical planning assumptions [PD]; Callaway and Sant'Anna (2021). No historical effect, approved harm margin, practical recruitment or customer benefit is asserted.

TECHNICAL ANALYSIS

Four propositions make the measurement-to-action claim falsifiable

[AI/PD] SEEC is an untested framework. It contributes explicit conditions under which richer telemetry can change a decision without being mistaken for proof of physical improvement or causal responsibility.

Proposition and mechanism	Observable test and failure condition
P1: Eligibility separates missingness	An independently versioned applicability audit should reduce false noncompliance classification at a prespecified sensitivity floor. Fails if error does not decline on a held-out original cohort.
P2: Capture improves knowledge first	Better logging should improve reference-verified completeness and latency. Physical benefit is a separate claim; fails if calibrated capture does not improve or if increased capture is mislabeled as reduced harm.
P3: Action needs value and safety	Live prompts should improve mature promise outcomes and net burden while meeting ordinary-flow safety. The funding rule fails if any confidence-bound gate fails; that alone does not prove zero effect.
P4: Ownership requires acceptance	An authorized handoff with an explicit evidence gap and next step should improve verified response completion. Fails if accepted actions or learning do not improve at the prespecified effort budget.

[AI] Deming cautions against tampering with unstable signals; here, selected-site pressure should not become a control before the measurement phase is stable. Campbell and Goodhart explain why targeting scan completion can alter the proxy. Ittner-Larcker-Randall motivates testing the signal-to-action-to-customer chain. Tucker-Edmondson warns that a local workaround can complete work without feeding learning back to its owner. Composite-indicator guidance rules out treating nested, separately based signals as additive causal value.

[AI] The fresh measurement literature adds a sharper distinction: an event log needs defined case, activity, time and completeness scope; calibrated detectors still need an independent reference; cross-boundary visibility does not itself transfer decision rights. The case extends these ideas into an eligibility-aware release contract with an explicit ordinary-flow harm gate.

[PD] Reproducibility materials retain the frozen source ledger, alternative arithmetic, exact constructive examples, model code, independent solver and page reviews privately. Publish protected conclusions and independent hypotheticals. The next useful result may be a validated no-action decision rather than a larger apparent opportunity.

Basis: proposed SEEC propositions [AI/PD]; fresh measurement/process/coordination sources and continuing control, proxy-incentive, causal-measurement and learning literature. Falsifiers and thresholds must be registered before observing results.

Test what live scan evidence changes beyond logging

[PD] Randomize independent station/cycle capacity domains to live reconciliation prompts or silent instrumentation with existing response. Both arms retain the same permitted logging, customer remedies and independent physical reference.

[PD] Freeze 400 externally booked units per domain: 80 scan-applicable opportunities and 320 ordinary bookings, over four scheduled cohorts. Routing/configuration evidence fixes applicability before assignment. Policies observe signals, not hidden omissions or future outcomes. All bookings remain; missing telemetry preserves usual operations.

Wholly hypothetical input	Base	Stress
Timely feed / common domain state	80%	65%
Signal sensitivity / actual omission	80%	80%
False signal / no omission	10%	14%
Safe accepted correction succeeds	65%	65%
Ordinary active-feed harm / absolute	0.30 pp	0.40 pp

[PD] Within applicable opportunities, actual omission prevalence is 30%; usual promise-failure risks are 45% with an omission and 3% otherwise. A true alert permits verified correction; a false alert can delay an otherwise on-time unit with 4% probability. Ordinary risk starts at 2%; shared queue harm affects that entire stratum.

OBSERVATION IS NOT PREVENTION

[PD] With no false positives, doubling capture of unchanged physical failures doubles recorded failures while service stays unchanged. Reconstruct original-promise receipt by promise + 28 days; later recovery never erases a missed deadline. No worker blame, sanctions or altered refunds follow from a signal.

[PD] All inputs and the reference horizon are synthetic. Independent custody/eligibility validation precedes live use; unresolved truth stays in the original cohort.

Require value, service improvement and ordinary-flow protection

[PD] Exact finite-process moments include the shared feed state. Gaussian planning uses conservative variance envelopes; its MDE, confidence limits and rejection probabilities are approximations, not exact finite-tail guarantees.

Design	Per arm	MDE (index)	BE detection	Joint floor
Initial	96	9.01	70.1%	58.7%
Base candidate	192	6.37	95.6%	90.6%
Same size / stress	192	6.51	94.7%	20.6%
Conditional redesign	1024	2.82	>99.9%	87.5%

[PD] Break-even is 8.11 index points. Require 80% detection at break-even and an 80% three-gate joint floor: resource lower bound above fixed cost, physical-failure reduction lower bound above zero, and ordinary-harm upper bound below 0.80 percentage points. This margin is a hypothetical policy choice requiring approval.

[PD] Each directional gate uses a 97.5% normal bound; the zero test is two-sided 5%. The joint floor is $\max(0, P(\text{value}) + P(\text{service}) + P(\text{safety}) - 2)$, without assuming independent endpoints. Variance envelopes span correction success 0-100% and active-feed ordinary harm 0-2 percentage points.

[PD] Direct finite simulations change the actual success or harm law at each boundary. In 30,000 trials, redesigned joint success was 87.3% (95% Monte Carlo interval 86.9-87.7%). Simulated boundary error estimates and intervals are retained privately; they do not validate unknown field inputs.

CALIBRATE BEFORE RECRUITING

[PD] The conditional redesign needs 2,048 independent domains and 819,200 bookings; availability is unestablished. Begin with bounded applicability, reference, feed and queue calibration. Re-estimate variance and use studentized field inference before committing a trial.

[PD] MDE and cost use index points; usual modeled burden = 100. Direct candidate tests, no minimum claim. Severe delay, damage, authority and workforce safeguards remain separate.

Charge false alerts and queue harm before claiming value

[PD] A promise failure costs 12 hypothetical resource units, each alert 0.10 and each successful physical correction 0.25. Correction work is paid even when that booking would otherwise have met its promise.

[PD] Common reserved capacity is 0.05 per original booking in both arms. Incremental K = 0.05 covers configuration 0.02, monitoring/owner support 0.02 and renewed linkage validation 0.01. Prompt work and correction effort are outside K. Research-only outcome adjudication has a separate study budget.

[PD] $Net_raw = 12[a h A - F - Q] - W - K$. A is avoided target failure, F added false-alert failure, Q added ordinary failure and W incremental response work, all per original booking. Avoidability a and non-overlap h discount only favorable prevention. Added harms and all new costs remain fully charged.

Avoidability a	h = 0.50	h = 0.75	h = 1.00
0.50	-9.05	-6.31	-3.58
0.75	-6.31	-2.21	1.89
1.00	-3.58	1.89	7.35

[PD] Table values are net index points, $100 \times Net_raw / 0.6164$. Full-realization base net is 7.35; stress net is 3.11. At K = 0.03 / 0.05 / 0.07, base net is 10.60 / 7.35 / 4.11. Changed prices require new power.

KEEP THE ORIGINAL DENOMINATOR

[PD] With verified resource records, epsilon unknown or wrong outcome labels in each arm can shift the resource contrast by 24 epsilon and a risk contrast by 2 epsilon; ordinary risk uses its own denominator. Day 28 must be validated. Field value requires randomized same-cohort effects, not changed scan counts or withheld payments.

[PD] Costs and realization shares are assumptions, not historical dollars. Delayed receipt, missing truth and inapplicable scans cannot be removed to improve the result.

PROPOSED INFORMATION VALUE / 4 OF 4

Buy evidence that separates logging value from live-action value

[PD] A separate three-state decision model asks whether to fund a bounded live-response test, validate logging only, or reconcile eligibility before expansion. It does not presume that more instrumentation improves service.

Action / state	Live justified	Logging only	Unresolved
Bounded live test	14	-6	-12
Logging validation	4	7	-3
Reconcile first	2	3	4

[PD] The mutually exclusive states apply an evidence/authority gate first: unresolved mandatory requirements take precedence. Otherwise classify whether live action passes its value and service criteria or only logging is justified. Priors are 0.35 / 0.40 / 0.25; the true state is reported with probability q and each wrong state with (1 - q)/2.

Signal q	Gross EVSI	Acquisition	Net EVSI
0.45	0.1550	0.40	-0.2450
0.70	2.2125	1.20	1.0125
0.90	4.2375	2.60	1.6375

[PD] All 27 signal-contingent policies are enumerated. Logging validation is the prior optimum, worth 3.4500; EVPI is 5.2500. Payoffs already include subsequent test expense; acquisition cost is separate. Weak information costs more than the decisions it improves in this example.

A NEW FIELD IS NOT AN ACTION MANDATE

[PD] Apply source validity, safe response and decision-authority requirements before using a signal. EVSI is a separate decision-utility model; never add it to the operating benefit or use it to authorize individual consequences.

[PD] Priors, likelihoods and utilities are hypothetical. Neither a testing tab nor improved observation establishes live adoption, preventability or mature customer benefit.

Reproduce the scope, the observation and the response

[AI] SEEC separates applicability, physical activity, recorded evidence and authorized action. A reproducible result must preserve those distinctions through the historical audit and the prospective evaluation.

Claim class	Replication boundary
[HR] Historical record	The register describes the arithmetic challenge, partial readiness inquiry, mechanism pivot and team instrumentation work. It preserves missing originals, source conflicts and bounded ownership.
[RC] Recalculated	Nested products, conditional union bounds, line consistency, source-version checks, query reconciliation and tentative process limits. Historical magnitudes remain private.
[AI] Analytical inference	Part-whole correlation is not an independent mechanism. Missing scans combine distinct paths. Cycle composition does not identify guardrail cause, and exposure cost is not preventable benefit.
[PD] Proposed design	Eligibility and reference contracts, approved response policy, power and economic scenarios. The numerical law is hypothetical; simulation validates computations under that law, not real-world assumptions.

Replication sequence

[PD] Authenticate the original analysis, call notes, schemas, query outputs and correspondence. Recover the actual model and event versions before accepting the historical forecast or attributing intellectual ownership. Do not infer a delivered action from an unsent draft or a completed tool from a commitment.

[PD] Rebuild population memberships and joins at a frozen episode grain. Test the alternative denominator readings, within-tier and between-tier intersections, retained date coverage and effective guardrail configuration. Reconcile package counts and concession amounts separately; retain unknown reference states.

[PD] Reproduce the constructive examples and proposed policy law from primitive inputs. Compare mature physical outcomes, full resources and effects on ordinary flow. Distinguish analytical moments, approximate inference and Monte Carlo validation. Preserve the random seed, simulation error and calibration requirements.

THE INTERPRETATION THIS CASE CHANGES

[AI] Instrumentation is an investment in evidence. Its operating value depends on whether that evidence supports an authorized response that improves physical service after false prompts, shared capacity effects and full costs are included.

The separate private package contains source identities, exact locators, raw values, alternative calculations and review records. The public case contains protected conclusions and disclosed methods.

RESEARCH REFERENCES / MEASUREMENT AND EVENT EVIDENCE

Validate what the instrument can observe

Primary research supports the methods and interpretation. It does not authenticate the unrecovered project artifacts, establish historical fault or supply a transferable effect size.

Rogan, W. J., & Gladen, B. (1978). Estimating Prevalence from the Results of a Screening Test. *American Journal of Epidemiology*, 107(1), 71–76. Original abstract. Recorded positives differ from underlying prevalence. A correction requires identified, transportable accuracy parameters; this case's examples derive the observation law explicitly rather than estimating historical prevalence.

Hernán, M. A., & Cole, S. R. (2009). Invited Commentary: Causal Diagrams and Measurement Bias. *American Journal of Epidemiology*, 170(8), 959–962. Indexed original abstract and selected text. Represent true action, recorded evidence and physical outcome separately; common measurement causes can bias their relationships. Missing records are not a causal category.

Hui, S. L., & Walter, S. D. (1980). Estimating the Error Rates of Diagnostic Tests. *Biometrics*, 36(1), 167–171. Selected original pages. Identification without a perfect reference requires restrictive assumptions, including distinct prevalences, conditionally independent errors and invariant accuracy. Two views of the same event feed do not automatically meet them.

van der Aalst, W., et al. (2012). Process Mining Manifesto. *Business Process Management Workshops, LNBIP 99*, 169–194. Author manuscript, selected sections. Case identity, event meaning, sequence and completeness need explicit definitions. Schema availability does not establish a valid, retained episode history or a reproducible compliance classifier.

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2002). Patient Non-compliance with Paper Diaries. *BMJ*, 324, 1193–1194. Original article. Instrumented timing can disagree with reported completion. Recording time remains a proxy for the underlying activity; no clinical compliance estimate or causal effect is imported into this case.

Kellogg, K. C., Orlikowski, W. J., & Yates, J. (2006). Life in the Trading Zone: Structuring Coordination Across Boundaries in Postbureaucratic Organizations. *Organization Science*, 17(1), 22–44. Publisher abstract. Visible, revisable work can support coordination across boundaries. A signal still needs acceptance, jurisdiction and an executable handoff.

Prentice, R. L. (1989). Surrogate Endpoints in Clinical Trials: Definition and Operational Criteria. *Statistics in Medicine*, 8(4), 431–440. Original publisher abstract. Association with an outcome does not make an intermediate measure a valid replacement for a treatment comparison. Here scan coverage remains separate from the prospective physical-service endpoint.

Access scope is retained privately. Applications are methodological or conceptual; published empirical results are not attributed to this project.

RESEARCH REFERENCES / EVALUATION AND UNCERTAINTY

Evaluate the response on the original cohort

Primary research supports the methods and interpretation. It does not authenticate the unrecovered project artifacts, establish historical fault or supply a transferable effect size.

NIST/SEMATECH. e-Handbook of Statistical Methods, section 6.3.2.2: Individuals Control Charts. Official method. Calculate tentative limits for the sparse ordered observations, but separate monthly and weekly grains, incomplete periods and definition changes. Within-limit movement does not establish either improvement or a stable operating phase.

Laney, D. B. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531–537. Primary publisher abstract. Extra-binomial adjustment requires actual counts, denominators and rational subgroups. The historical register lacks an eligible repeated count panel; rounded rates are not converted into invented exposure counts.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37–46. Primary publisher record. Assess agreement and kappa in a blinded, stratified taxonomy audit. Report prevalence, confusion and uncertainty; high agreement between coders cannot authenticate a shared defective reference.

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the Mean: What It Is and How to Deal with It. *International Journal of Epidemiology*, 34, 215–220. Primary abstract. A selected worst-site cohort can recover without an intervention. The case uses explicitly assumed reliability sensitivity and requires independent recurrence and comparison data.

Hudgens, M. G., & Halloran, M. E. (2008). Toward Causal Inference with Interference. *Journal of the American Statistical Association*, 103, 832–842. Original manuscript. Prompt-driven work can affect other deliveries through shared queues. The proposed trial retains ordinary flow and groups shared resources within independent capacity domains; between-domain spillovers remain a design threat.

Callaway, B., & Sant’Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, 225, 200–230. Author manuscript. A nonrandom staggered field extension needs group-time effects, eligible controls and defensible parallel trends. It is distinct from the randomized policy ITT and requires its own power and sensitivity analysis.

Strong, M., Oakley, J. E., Brennan, A., & Breeze, P. (2015). Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample. *Medical Decision Making*, 35, 570–583. Author repository. Information has value when it changes a feasible decision under a specified prior, payoff and likelihood model. Evidence cost is subtracted; EVSI is not added to the trial’s operating benefit.

Access scope is retained privately. Applications are methodological or conceptual; published empirical results are not attributed to this project.

RESEARCH REFERENCES / CONTRIBUTION AND GOVERNANCE

Make the measure serve the operating decision

Primary research supports the methods and interpretation. It does not authenticate the unrecovered project artifacts, establish historical fault or supply a transferable effect size.

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44, 823–830. Primary publication. Grade each specific role and action against competing explanations. A derivative register supports a provisional contribution claim; a meeting acceptance, unsent draft or subordinate's artifact does not establish convening, delivery or sole authorship.

Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures. *California Management Review*, 45(2), 55–72. Author manuscript. Local workarounds can complete immediate tasks while preventing learning about recurrent system conditions. SEEC routes unresolved eligibility and capture errors to an owner alongside the recovery response; no healthcare effect is transferred.

The W. Edwards Deming Institute. The Funnel Experiment. Official exposition. Repeated adjustment to routine variation can worsen performance. The ranked-site mechanism therefore requires stable measurement, recurrence evidence and a prospective comparison; sparse dashboard fluctuations do not justify repeated changes to frontline instructions.

Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Original institutional record. Accountability measures can change both behavior and reporting. Track eligibility, capture, unknowns and independent service outcomes so that completing the data record cannot substitute for completing the required physical action.

Goodhart, C. A. E. Problems of Monetary Management: The U.K. Experience. 1975 conference; revised volume listed as 1976 in the RBA archive. A targeted proxy can lose its previous relationship with the intended outcome. Changing default ownership or scan rules can move a category without improving service.

Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance Implications of Strategic Performance Measurement in Financial Services Firms. *Accounting, Organizations and Society*, 28, 715–741. Author manuscript. Measurement needs an explicit connection to strategy and operating mechanisms. A ranked association or new dashboard tab does not verify the telemetry-to-action-to-customer chain.

OECD / European Commission JRC (2008). Handbook on Constructing Composite Indicators. Official handbook. Definitions, denominators, weights and compensation assumptions govern aggregation. Nested signal tiers cannot be added as independent opportunity; the proposed economic and customer gates remain distinct rather than permitting one favorable score to offset a failed requirement.

Access scope is retained privately. Applications are methodological or conceptual; published empirical results are not attributed to this project.

FINAL REFEREE REPORT / INTERNAL AGENT-ASSISTED REVIEW

Accept the measurement discipline; test the operating value

[AI] This author-led, agent-assisted review is not external peer review or client approval. It assesses an evidence-reconciliation case, constructive measurement analysis and a proposed operating-policy evaluation.

Major issue	Disposition in the revised case
A corrected component treated as a valid total	Preserved alternative bases and rounding; required joint membership and nested-set checks. Candidate exposure is separated from causal controllability and avoidable cost.
Correlation treated as an operating effect	Derived the part-whole limitation and audited the incompatible scenario points. Benchmark gaps and selected-cohort recovery remain descriptive until tested.
Absent scan treated as failed behavior	Separated eligibility, physical action, capture and linkage. Added constructive examples and a blinded independent-reference protocol.
Cycle composition treated as causal attribution	Kept effective configuration, override intent and episode validation unresolved. Metadata access and a testing tab do not establish production readiness.
A better record treated as customer benefit	Specified live prompts versus silent logging with original-manifest outcomes, false-alert work, ordinary-flow risk and economic break-even. Inference and simulation limits are explicit.

Partner critique: can a COO act on page one?

[AI] Yes: authorize bounded signal and reference validation, assign the relationship owner to approve a feasible response, and hold broad expansion until the evaluation can settle service and economic requirements. The cover identifies the operating problem, analytical finding, bounded contribution and evidence status.

Remaining conditions

[AI] Recover originals; resolve population, configuration and query defects; validate the reference and intervention authority; calibrate signal accuracy, feed availability, queue effects, costs and feasible independent-domain recruitment. No historical causal effect or realized saving is established.

REVIEW DISPOSITION / PASS WITHIN STATED SCOPE

[AI] The diagnostic case and proposed SEEC evaluation pass the stated internal gates. Acceptance is limited to the documented claims and model assumptions; it does not authorize deployment or certify a historical benefit.

Internal evidence, numerical, methods and executive review. The verdict applies to the released artifact and stated boundaries; no outside endorsement or ranking is implied.