

OPERATING CASE STUDY / OPERATIONAL EXCELLENCE

Standard Work & Escalation Design

A documented procedure does not establish why performance missed plan

An operating-project case on connecting defect categories to governing procedures, practical actions and accountable owners. It examines a policy-coverage diagnostic, corrects its causal interpretation and designs a controlled test of evidence-based routing.

STANDARD WORK & OPERATING ACCOUNTABILITY

BLUF / BOTTOM LINE UP FRONT

Verify applicability, execution and measurement before assigning a failure to noncompliance.

RECALCULATED FINDING

Analytical finding: identical plan shortfalls can arise from different combinations of rollout, adherence, efficacy and measurement. A policy lookup alone cannot distinguish them. Missing documentation can also coexist with active operating programs.

CONTRIBUTION IN THE SOURCE RECORD

Provisional contribution: research design, mapping and recommendations. Policy authority and program delivery were separately owned.

EVIDENCE STATUS

Historical originals were not recovered. Analytical design case; adoption, customer improvement and savings remain unverified.

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

STRUCTURED ABSTRACT / DECISION BRIEF

Assign the next action from evidence, not the existence of a document

[AI] A standard-work map becomes useful when it distinguishes what is documented, what is executable and what the operating evidence actually establishes.

Context

[HR] A delivery-network leader mapped a governed defect taxonomy to operating procedures, tracking tools, owner roles and review cadences. The diagnostic separated apparent compliance gaps from capability gaps and proposed routing work to the actor able to resolve it. AI-assisted research supported the document review.

Objective

[HR] The mapping sought to replace undifferentiated performance pressure with specific actions: apply an existing standard, develop missing guidance, transfer a capability, repair measurement or escalate a dependency. The record describes a completed reference artifact and a proposed automated lookup; downstream implementation is not established.

Approach

[AI] The proposed Evidence-to-Action Contract (EAC) tests identity, scope, execution authority, measurement and diagnostic evidence separately. It preserves uncertain classifications and distinguishes a permissible action from a causal finding about why performance failed. Exact examples expose the limits of coverage ratios and plan attainment.

Findings

[AI] The same plan shortfall can occur with full adherence or poor adherence. A missing policy search result can coexist with funded activity. Reclassification can improve a reported category without reducing total customer defects. Static lookup rules can outlive the evidence that justified their instructions.

Decision value

[PD] The operating sponsor should appoint a policy owner and measurement owner before expanding automated routing. Release a recommendation only when its scope, current authority, feasible action and evidence status are explicit. Route unresolved cases to a paid verification queue with an owner and deadline; compare the workflow with simpler rules.

Implications

[AI] The supported contribution is an analytical reference design, with provisional personal attribution. Its useful extension is a testable routing contract. Neither the original recommendations nor the hypothetical evaluation establish adopted decisions, changed behavior, mature customer benefits or savings.

[HR] Selected source register S1; original recovery and historical coordinates are private. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical causal impact is not established.

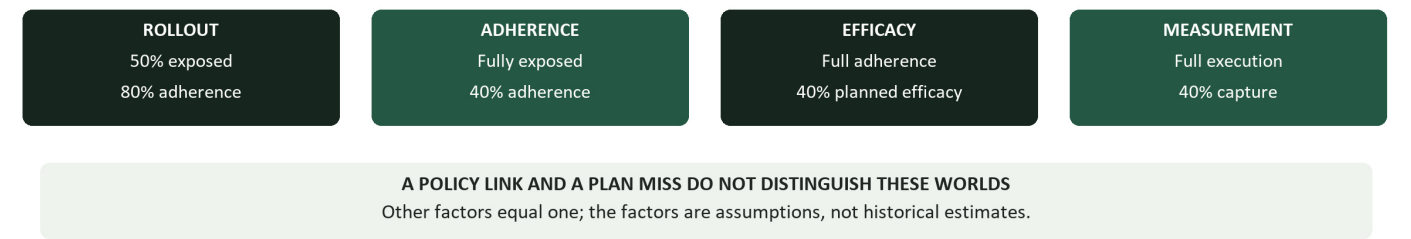
EXECUTIVE EXHIBIT 1 / CAUSAL DIAGNOSIS

A plan miss does not identify an adherence failure

[HR] The source used the existence of mature procedures and below-plan program contributions to support a compliance-gap diagnosis.

DIAGNOSTIC IDENTIFICATION / THE SAME GAP HAS DIFFERENT CAUSES

Independent hypothetical illustration: planned contribution 100; booked contribution 40 in every world



A targeted observation or experiment is needed to identify the mechanism.

[PD] Independent illustration: a plan of 100 produces a booked contribution of 40. One compatible world has full rollout and adherence of 0.40. Another has full adherence but efficacy of 0.40 relative to plan. A third has full execution and efficacy but booking captures only 0.40. Limited rollout can generate the same result. These are constructions, not estimates.

Observation	What it establishes
Procedure located	A candidate source exists. Confirm current version, governing authority, applicability and whether the action can be performed locally.
Booked contribution below plan	A gap in the selected planning measure. Separate exposure, execution, efficacy, measurement and the realism of the plan before diagnosing its cause.
Repeated booked zero	A reason to investigate launch, observability, efficacy and execution. Zero alone cannot distinguish these explanations or rank their probabilities.

[RC] One cited program appears closer to plan when measured as a rate contribution than when measured as contributed units. If definitions match, the discrepancy implies different exposure bases; otherwise the measures are incompatible. Neither version is a direct adherence measure.

OPERATING DECISION
[PD] Commission a mechanism audit before using plan attainment to assign corrective accountability. A verified, applicable procedure may justify offering a safe action; it does not by itself establish who caused the shortfall.

[HR] Selected source register S1; original recovery and historical coordinates are private. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical causal impact is not established.

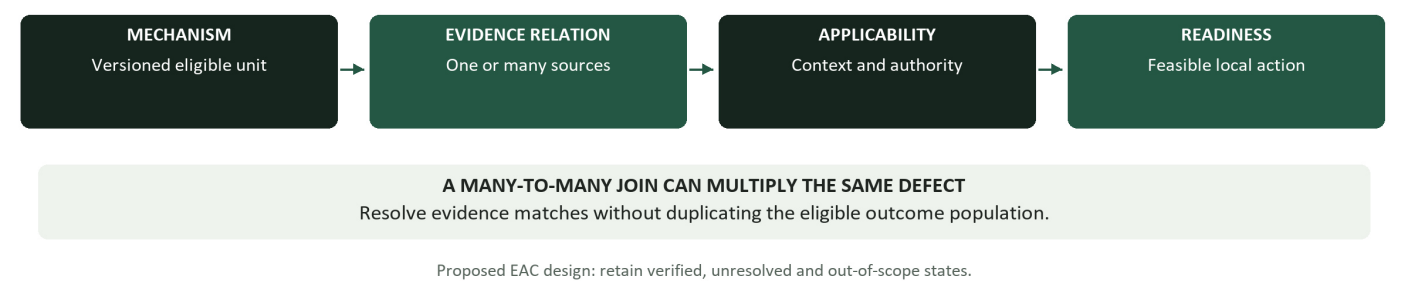
EXECUTIVE EXHIBIT 2 / COVERAGE VALIDITY

Measure eligible mechanisms rather than counting policy links

[HR] The source combined an analyst-defined mechanism census, a differently sized mapping table and multiple document references. Several coverage claims lacked directly verifiable identifiers.

COVERAGE CONTRACT / FIX THE UNIT BEFORE COMPUTING THE RATIO

Document links, mapped rows, unique mechanisms and eligible episodes are different populations



Coverage contract	Required evidence
Unit and taxonomy	One stable eligible mechanism, with a versioned definition and explicit cross-tags. Document count, mapping-row count and mechanism count are different denominators.
Matched policy evidence	The document identity, applicable clause, effective version, geography and operating mode. One policy may cover several mechanisms; several policies may govern one mechanism.
Operational coverage	A feasible action, authorized actor, required resources and observable outcome. Documentary coverage is one input to this assessment, not its conclusion.

[AI] Joining defect data to every matching policy can multiply rows and inflate apparent coverage or defect volume. Keep the mechanism population fixed, store the evidence relation separately and declare how conflicts and multiple matches resolve. A coverage percentage has no stable meaning until that contract is fixed.

ANALYST ACTION

[PD] Publish verified, unresolved and out-of-scope counts on the same eligible frame. Do not substitute document richness for operational readiness, or add cross-tagged categories into a network total.

[HR] Selected source register S1; original recovery and historical coordinates are private. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical causal impact is not established.

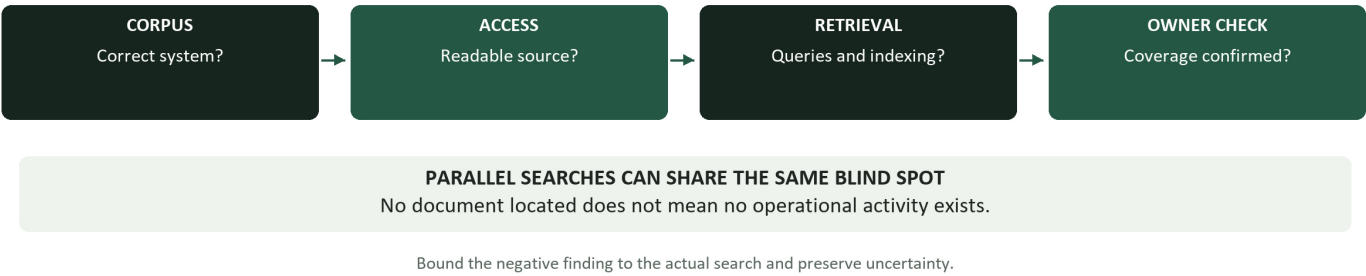
EXECUTIVE EXHIBIT 3 / SEARCH EVIDENCE

A missing search result is a verification task, not proof of absence

[HR] The diagnostic treated unsuccessful searches as evidence of policy gaps. The wider record also describes access failures, incorrect source labels and searches conducted in the wrong repository.

SEARCH VALIDITY / CHECK WHAT THE SEARCH COULD HAVE FOUND

An inaccessible or incorrectly chosen repository does not test policy nonexistence



Search state	Permitted interpretation
Correct system; verified document	A source can be inspected. Its title, version, scope and authority still need validation before supporting an action.
Access denied or wrong system	The search does not test the intended absence claim. Restore authorized access or use the correct repository; preserve the unresolved state.
Document not located in scoped search	A bounded finding about retrieval under the stated corpus and method. Ask the accountable policy owner to confirm coverage before asserting nonexistence.

[AI] Parallel searches can share the same inaccessible corpus, vocabulary and indexing blind spot. Repetition does not create independent evidence. Active programs can also exist without a dedicated policy document, so a documentation gap and an operational capability gap must remain separate classifications.

RESEARCH OWNER ACTION

[PD] Record the corpus, query families, access status, searched versions and owner response. A confirmed gap should specify whether the missing element is guidance, authority, equipment, information or measurement.

[HR] Selected source register S1; original recovery and historical coordinates are private. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical causal impact is not established.

EXECUTIVE EXHIBIT 4 / MATURE OUTCOMES

Correct classification without claiming a prevented customer defect

[HR] One cited intervention moved incorrectly coded events between defect categories. Other mapping rows paired a funded outcome with a tracking view that measured a different construct.

MEASUREMENT BRIDGE / RELABELLING IS NOT DEFECT PREVENTION

Independent hypothetical construction: the same 100 customer events are recoded



[PD] Independent illustration: category A falls from 40 to 25 while category B rises from 60 to 75 because the same events are recoded. The total remains 100. The classification becomes more accurate, but the arithmetic does not establish that any customer failure was prevented.

Measurement issue	Decision consequence
Category attribution	Audit the code change and preserve the full outcome bridge. A cleaner category can improve diagnosis without producing net service benefit.
Proxy versus funded outcome	Specify the causal or accounting link. A general tracking view cannot certify a narrower outcome merely because both relate to the same program.
Permitted observation grain	Keep site-level evidence at the site level. A drill-down ambition does not authorize individual-level inference or attribution.

MEASUREMENT OWNER ACTION

[PD] Freeze the eligible cohort and mature outcome definition before comparing performance. Validate taxonomy reliability, denominator stability and cross-group meaning; retain unresolved and missing cases in the evaluation.

[HR] Selected source register S1; original recovery and historical coordinates are private. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical causal impact is not established.

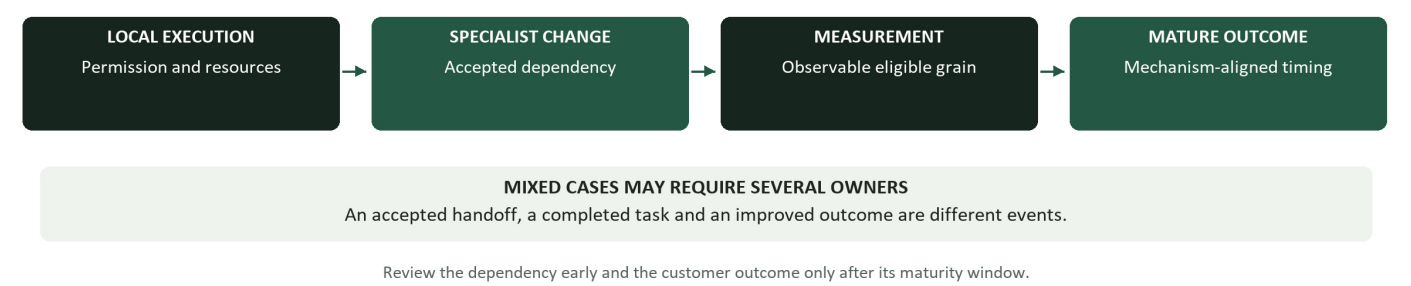
EXECUTIVE EXHIBIT 5 / OPERATIONAL READINESS

Match the action and review cadence to the actor who can deliver

[HR] The mapping distinguished actions a local team could take promptly from changes requiring infrastructure, specialist intervention, policy authorship or a measurement capability.

EXECUTION CONTRACT / ROUTE TO THE ACTOR WHO CONTROLS THE CONSTRAINT

A feasible action requires more than a documented procedure



Readiness question	Accountable response
Can the local team execute now?	The operating owner verifies resources, permissions and prerequisites. A documented procedure without these conditions is an incomplete intervention.
Does another owner control the constraint?	Route a bounded request to that owner with acceptance criteria and a response deadline. Keep interim local responsibilities explicit.
Can the outcome be measured at this grain?	The measurement owner establishes observability before a target is assigned. A missing dashboard is an evidence gap, not proof that the underlying outcome cannot be observed.

[AI] Compliance and capability can coexist within a defect category. Some episodes may need better execution while others require a system or resource change. Review cadence should follow the mechanism and its lead time, with early dependency checks and later mature-outcome checks.

SPONSOR ACTION

[PD] Require an accepted owner, feasible action and evidence of completion for every route. An escalation named in a document is not an accepted handoff; an accepted handoff is not a demonstrated customer outcome.

[HR] Selected source register S1; original recovery and historical coordinates are private. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical causal impact is not established.

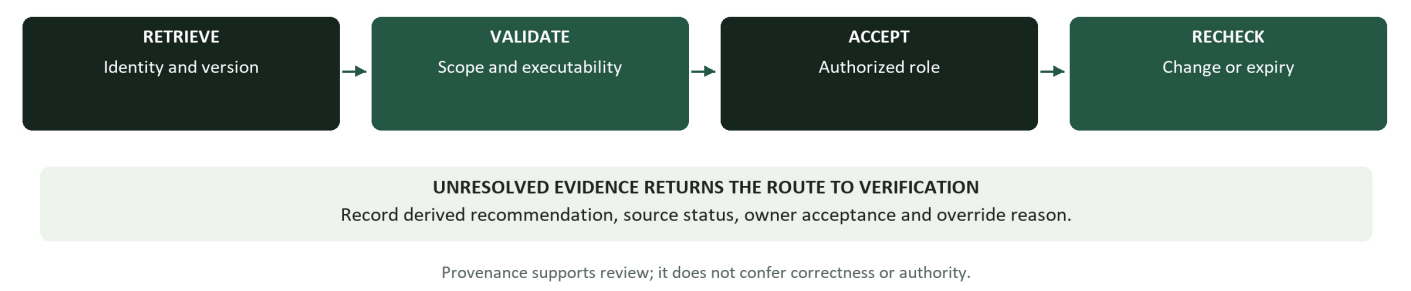
EXECUTIVE EXHIBIT 6 / AUTOMATION CONTRACT

Let evidence expire before an automated instruction becomes stale

[HR] The source proposed a structured lookup that would generate actions from defect categories. The record does not establish that it was built, and it identifies no systematic refresh mechanism.

AUTOMATION CONTRACT / PRESERVE THE EVIDENCE BEHIND EACH ROUTE

A static lookup must not turn a changed source into a current instruction



Contract element	Release behavior
Identity and applicability	Bind the recommendation to a specific document version, clause, operating context and access boundary. Resolve conflicting matches before choosing an action.
Verification and expiry	Record who checked the evidence and when it must be reconsidered. A changed source, withdrawn permission or failed prerequisite returns the route to review.
Human acceptance and audit	Separate a suggested action from authorization to execute. Preserve the evidence state, recipient, acceptance, override reason and later outcome.

[AI] Provenance records can show how a recommendation was derived and who contributed to it (Moreau and Missier, 2013). They do not make the recommendation substantively correct or confer operational authority. Evidence about evidence also needs a current owner and a verifiable source.

PRODUCT OWNER ACTION

[PD] Test the routing logic in shadow with explicit unresolved states and a strong simple comparator. Do not let a successful data join silently convert an unverified classification into an instruction.

[HR] Selected source register S1; original recovery and historical coordinates are private. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical causal impact is not established.

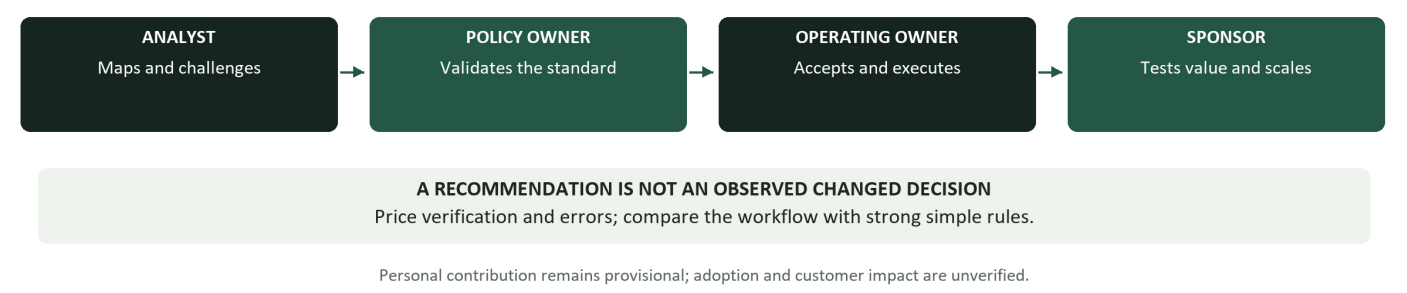
EXECUTIVE EXHIBIT 7 / CONTRIBUTION AND CONVERSION

Turn the diagnostic into an owned decision before claiming value

[HR] The register attributes the research design, mapping artifact and recommendations to the executive. It records no established sponsor, adoption forum or delivered implementation of the proposed lookup.

VALUE CONTRACT / SEPARATE CONTRIBUTION FROM IMPLEMENTED IMPACT

The analytical artifact, policy authority and operating intervention have different owners



Contribution or value claim	Evidence grade and boundary
Analytical design and authorship	Provisional registry-level hoop support: the record attributes the artifact and describes creator metadata. The original remains unauthenticated; exclusive intellectual authorship is not established.
Policy or program ownership	Unestablished for this project. Cited procedures, specialist programs and their reported results belong to separate owners and populations.
Adoption, benefit or avoided harm	Unestablished. Recommendations and an artifact do not demonstrate a changed decision, executed intervention or mature causal effect.

[PD] Begin with a sponsor-owned decision: which evidence routes are useful, what errors matter and what verification capacity is affordable? Evaluate all eligible requests, including unresolved cases and unaccepted handoffs. Compare the paid workflow against simple verification and routing rules before funding expansion.

INVESTMENT THRESHOLD

[PD] Adopt complexity only when its incremental decision value exceeds verification and operating cost, and the evaluation can detect that break-even. Preserve a simpler rule when the fuller workflow does not justify its additional burden.

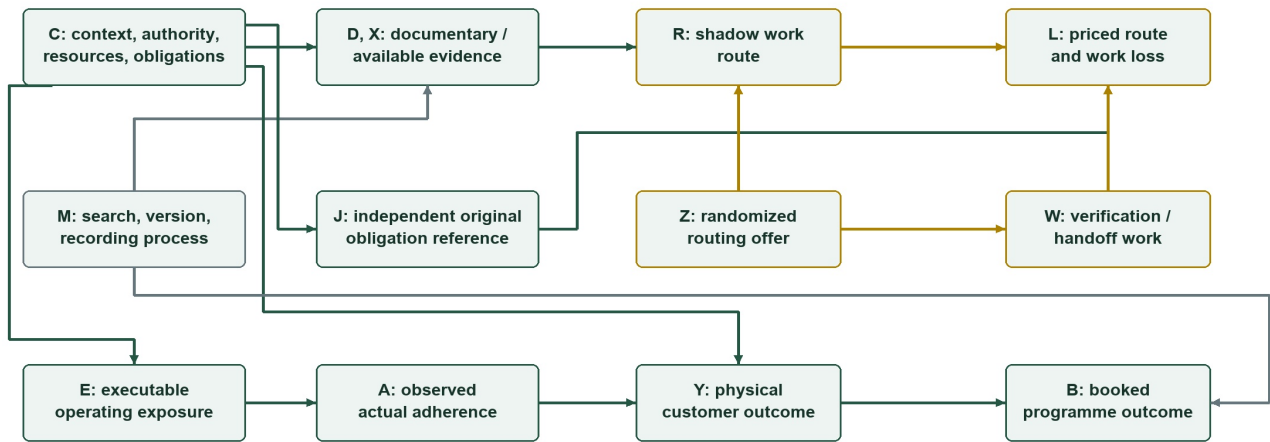
[HR] Selected source register S1; original recovery and historical coordinates are private. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical causal impact is not established.

TECHNICAL ANALYSIS

A standard informs action; its existence does not identify execution

[PD] The proposed Evidence-to-Action Contract (EAC) separates admissible instructions, feasible work, observed execution and customer outcomes. Its first test is a shadow routing benchmark.

EAC: DOCUMENTS, ROUTING, EXECUTION AND OUTCOMES HAVE DIFFERENT EVIDENCE



P: programme plan is a comparison basis for B; it is not an execution observation.
S: common preassignment request selection may depend on B and management attention.
J depends on original context C. Independent adjudication does not mean independence from truth.
The lower operating chain is a separate extension. No shadow R-to-Y effect is asserted.

[PD] C is preassignment context: current authority, scope, tools, resources and obligations. D is documentary evidence, mediated by search and version process M. X is the evidence available to a routing rule. Randomized Z changes route R and verification work W. Independently adjudicated J assesses the original admissible obligations; loss L compares R with J and prices work. J depends on C, not on the reviewer's chosen route.

[PD] The lower chain is a separate operating extension: executable exposure E, actual adherence A and context C can affect physical outcome Y. Recording process M turns Y into booked outcome B. A programme plan P supplies a comparison denominator, not a treatment or execution observation. The shadow benchmark has no asserted R-to-Y arrow.

[AI] A poor-booking trigger can select requests S through B and management attention. Freeze a common original frame before randomization. The benchmark identifies a routing-policy effect for that frame, not why historical programmes missed plan. Conditioning on postassignment acceptance, successful verification or completed handoff would change that estimand.

Basis: EAC theory of change and shadow/operating distinction [PD/AI]. Historical quantities remain private.

TECHNICAL ANALYSIS

The same shortfall can require four different repairs

[AI] Independent synthetic law: booked gain = 100 x exposure x adherence x efficacy x capture. Efficacy is relative to the invented plan. The construction grants a valid measurement model that the historical source does not establish.

World	Exposure	Adherence	Efficacy	Capture	Booked
A	0.5	0.8	1	1	40
B	1	0.4	1	1	40
C	1	1	0.4	1	40
D	1	1	1	0.4	40

[AI] Making adherence perfect while holding the other factors fixed gives booked gains of 50, 100, 40 and 40. Full adherence already exists in C and D. In D the underlying physical gain is 100, but only 40 is recorded. Fixing capture changes the booking, not the physical outcome. Identical attainment therefore does not identify either adherence or the value of adherence pressure.

[AI] A zero product is equally compatible with zero exposure, zero execution, zero efficacy or zero capture when the other factors equal one. Repetition does not rank those explanations without likelihood assumptions and independent process measurements. An uncited expected-benefit range remains a scenario, not an empirical bound.

[AI] A separate relabeling example moves category A from 40 to 25 and category B from 60 to 75 while the total stays 100. That coding operation produces a favorable named-bucket movement without reducing total defects. This example does not prove that an actual programme was merely relabeling or that a concurrent intervention had zero effect.

[PD] Diagnose the missing factor before choosing an action: verify eligible exposure, observe required execution, test the causal mechanism, and reconcile the booking basis. Programme rate and unit attainment also require a common exposure bridge; one ratio cannot stand in for the other.

Basis: exact source-independent counterexample.py/json [AI]. All values are hypothetical and are not scaled source values.

TECHNICAL ANALYSIS

Count eligible opportunities, then verify each action gate

[HR/AI] The mapping combines defect rows, mechanism labels, policy identifiers and multiple regional taxonomies. These are different counting units. A repeated link can be useful evidence without adding another covered opportunity.

Gate	Evidence to establish or track
Identity and version	Retrieved content, effective date, approved owner and preserved source hash
Applicable scope	Geography, business model, activity and population match
Authorized executable action	Permitted task, capable role, tools, access and available capacity
Observable eligible grain	Unique opportunity and required action; valid site or person measurement level
Mechanism and outcome evidence	Record the current indication; track subsequent execution and mature outcomes separately before causal credit

[AI] Synthetic coverage example: four eligible mechanisms have exposure weights 10, 20, 30 and 40. Five link rows cover only three distinct mechanisms. Only the first is verified executable; the second has an inapplicable document, the third an expired one, and the fourth is unknown. Known executable exposure is 10%; if the unknown fourth passes all gates it becomes 50%. Link-row counts cannot identify either quantity.

[HR/AI] A failed search is a bounded statement about that search and corpus. Repeated passes sharing a corpus are not independent demonstrations of absence. Blank applicability cells need their documented convention. A later denied fetch does not prove access was denied when an earlier analysis was written; a later successful fetch does not authenticate every prior version.

[PD] Store a versioned many-to-many relation from defect, context and valid time to candidate actions. Resolve duplicates, overrides and incompatible instructions before release. Unknowns route to verification. A valid local containment step and a capability obligation may coexist; a governance referral needs an accepted recipient and return path.

Basis: registry mapping and version limits [HR/AI]; synthetic coverage and proposed EAC gates [AI/PD].

TECHNICAL ANALYSIS

Define routing success before attributing customer benefit

[PD] A correct action assignment means that the route matches admissible obligations under the original applicable rule and evidence. It does not mean a person caused the historical defect.

[PD] The primary shadow estimand is $\tau_L = E[L(\text{usual}) - L(\text{EAC})]$ per original eligible request, including verification, rejected handoffs, mixed obligations and unresolved requests. Net value subtracts fixed incremental cost K . The safety estimand is $\tau_F = \Pr(\text{unsupported local assignment under usual}) - \Pr(\text{unsupported local assignment under EAC})$, on the same full original denominator. Positive values favor EAC; the safety gate requires superiority rather than a conditional accuracy among answered cases.

[PD] Lock the shadow answer at original intake +48 hours. Complete blinded adjudication by intake +14 days, reconstructing the version, authority and evidence available at lock. Later policy changes cannot retroactively authorize a route. Reference disagreement remains visible. Accepted governance ownership means a recorded acknowledgment and defined next action, not merely a department name.

[PD] A later operating study needs a different estimand: all-original customer promise failure and resource cost under an implemented action policy. Document availability, action receipt and enactment should be measured separately, following Bellg's fidelity distinctions. Proctor's implementation outcomes likewise do not become customer outcomes through a label change.

[AI/PD] For an outcome bounded between 0 and B , unresolved reference fractions u_0 and u_1 can widen a known-only contrast by as much as $B(u_0 + u_1)$. Apply each endpoint's actual denominator and support; uncertain obligation membership needs reassignment bounds. Dropping unknowns or grading only accepted cases can manufacture apparent success.

[PD] Compare the complete workflow with the existing shortcut, a low-cost verified-scope rule and every feasible fixed route. All comparators obey the same known-blocker and authority rules. Beating a weak shortcut alone does not justify paying for a complex workflow.

Basis: proposed original-request ITT, shadow clocks and uncertainty protocol [PD]; no historical treatment effect claimed.

TECHNICAL ANALYSIS

Validate execution and the reference independently

[PD] A retrieved document, an executable instruction and an observed action are separate measurements. Each needs a defined population, unit, clock, owner and failure mode.

[PD] Sample original local, governance, mixed and verification requests across document versions, regions, programmes and evidence-quality strata. Include requests with missing or denied source access. Use known sampling probabilities and population weights. Two coders blinded to treatment and later outcomes assess identity, scope, authority, feasibility and original obligations; a separate panel adjudicates disputes from primary records.

[PD] Report class prevalence, weighted confusion matrices, agreement and kappa with cluster-bootstrap intervals. A simple random sample of 400 gives an approximate 95% worst-case agreement margin of 4.9 percentage points; this is not kappa power and does not account for clustering or rare unsupported assignments. Validate the rare safety class with additional sampled evidence and correct weights.

[PD] Test content validity with operational and policy owners, criterion validity against the independent reference, and measurement invariance across versions and contexts. Audit false absence from search, stale applicability, hidden resource constraints and falsely recorded recipient acceptance. A shared reference error can make coders agree while every route is wrong.

[PD] For an operating extension, observe whether eligible work actually occurred with the prescribed content, frequency and duration, rather than infer fidelity from booked benefit. Carroll supplies the implementation-fidelity distinction. Glasgow's RE-AIM dimensions keep reach and adoption at their proper levels; neither is a customer-effect estimate.

[HR/AI] The source contains site-level metrics and different regional standards. A join cannot authorize individual attribution where the measure was never designed for that level. Distinguish package, unit, event, programme and policy rows; retain valid line items while preventing duplicate opportunity denominators.

Basis: proposed validation and sampling [PD]; source grain restrictions [HR/AI]; accessed primary literature scopes recorded privately.

TECHNICAL ANALYSIS

A changing plan is not a stable compliance process

[RC/AI] Final inventory: 11 source/sensitivity screens, 19 separate version/basis comparisons, 7 design-shape audits and 10 cross-sectional profiles. No stable operating phase or eligible Laney panel is established.

[RC/AI] Two fitted source representations flag the last X point and the fourth moving range, but encode the same changing-plan credit series. They are not independent signals. The short observed monthly credit sequence stays within tentative limits under all eight rounding corners. Two retrospective tiny-prefix probes flag terminal X/MR; one has zero baseline moving range. These do not provide valid stable-baseline detection.

[PD] Tentative X limits are mean plus or minus 2.66 times mean moving range; the MR upper limit is 3.267 times that mean. Keep version comparisons, unordered profiles and invented plans outside process diagnosis. Negative-start normalizations reverse direction; zero baselines are unindexable. No binary counts or exposures are invented to compute a Laney adjustment.

[AI] Under an assumed linear conditional-mean model, expected untreated follow-up is $m + \rho \times (x - m)$, where x is the selected baseline value, m the population mean and ρ assumed persistence. Values 0.4, 0.7 and 0.9 imply 60%, 30% and 10% reversion of the selected excess. These are hypothetical sensitivities, not estimated explanations for source movements.

[PD] Prespecify modifiers measured before assignment: applicable version, region, baseline evidence quality, recipient capacity, programme maturity and obligation mix. Report interactions with uncertainty while retaining pooled ITT. Do not stratify the primary effect on realized successful verification, recipient acceptance or completed work.

[PD] A later nonrandom operational rollout needs a separate group-time ATT and mature original cohorts. Show pretrend plots, equivalence and sensitivity checks; a nonsignificant pretrend does not establish parallel trends. Concurrent programmes, reporting changes and shared governance capacity constrain identification and transport.

Basis: final evidence_spc.json, 47 vectors [RC/AI]; proposed selection, moderation and future evaluation [PD].

TECHNICAL ANALYSIS

Separate analytical participation from accepted operating change

[HR/AI] The register describes a policy-to-defect mapping, machine-assisted research and later citation corrections. It does not independently authenticate sole invention, a deployed lookup, accepted recommendations or realized benefit.

Claim	Provisional process-tracing assessment
Participation in the mapping and analytical framing	Registry-level hoop: derivative artifact descriptions and metadata support an association
Every classification was personally invented	Unestablished: machine research, inherited taxonomies and partner inputs remain alternatives
Later identifiers and source scope were corrected	Registry-level hoop: reported correction supports a lead; original fetch and version records remain unrecovered
The recommendations changed operations or outcomes	Unestablished: no authenticated adoption, implementation or counterfactual outcome record

[AI] Creator and last-modifier fields, even if recovered, record file association rather than exclusive intellectual authorship. Repeated statements sharing a source lineage are not independent corroboration. A missing distribution or approval record does not prove no adoption occurred elsewhere; it limits the supported claim.

[HR/AI] The documented lookup is a proposed integration. Naming an owner function does not show an accepted handoff, and a policy author is not automatically the implementer. An infrastructure deadline can be monitored in a weekly review without implying its outcome can change in that week.

[PD] Close the record with original drafts, version history, assignments, source-query logs, decision comments, recipient acceptance and change tickets. Preserve the analyst's defensible contribution: making policy, action and evidence dependencies inspectable. Separate that contribution from the present case's new counterexample and from any future evaluation result.

Basis: derivative registry and source-recovery limits [HR/AI]; Collier process-tracing logic. All historical contribution grades remain provisional.

TECHNICAL ANALYSIS

Test the route before attaching operating value

[PD] Randomize independent review-capacity domains to the existing document-presence shortcut or EAC. Freeze original requests and the reference rule before assignment; keep every request through the declared adjudication horizon.

[PD] The proposed finite law distinguishes local-only work, governance-only work, genuinely mixed obligations and unresolved evidence. Reviewers observe fallible scope, capability and recipient signals, never the hidden reference state. Shared domain evidence quality creates dependence. Common known blockers force verification in every arm; undetected insufficiency remains an error risk.

[PD] Price unsupported local assignment, unnecessary governance, omitted obligations, verification delay and returned handoffs separately. Add fixed and variable verification costs once. Discount favorable components for uncertain attribution or maturity while charging adverse components in full; a net gain can conceal additional harm in another category. Customer savings and avoided programme payments are outside this shadow endpoint.

[PD] Compare the minimum detectable loss reduction with the economic threshold K at each tested allocation. Distinguish detecting zero-effect departure at break-even from establishing a lower confidence bound above break-even. Require the safety contrast to clear its own gate. A joint-power lower bound of zero is uninformative, not zero actual power.

[PD] Exact finite moments and an explicitly approximate planning rule will be checked against independent generative simulations at selected actual null, economic and safety boundaries. Report Monte Carlo uncertainty and changed nuisance laws. No exhaustive minimum-allocation or uniform finite-coverage claim is needed. The final numerical law and validation are presented on the following design pages.

[PD] Estimate information value from an explicit prior, feasible actions, observation likelihood and payoff table, then subtract evidence cost. Strong and colleagues inform that separation. A simpler admissible rule may dominate the complete workflow; preserve that result. A favorable benchmark would justify a bounded test, not automatic live routing or a claim of physical benefit.

Basis: proposed shadow evaluation, power and economics [PD]. Model parameters are wholly hypothetical; final numerical audit is separately reproducible.

TECHNICAL ANALYSIS

EAC supplies falsifiable routing claims and explicit limits

[PD] These propositions are untested. They concern admissible work and evidence quality; none assumes that an outcome miss reveals intent or individual fault.

Proposition and boundary	Observable falsification
P1: scope and feasibility checks reduce unsupported local assignments under common reference quality	No prespecified reduction on the original request denominator, or effect disappears under blinded reference audit
P2: direct fidelity observation discriminates shortfall mechanisms beyond document and booking data	Execution measurements add no prespecified predictive discrimination in held-out contexts
P3: versioned contextual joins reduce wrong-scope routes without hiding unknowns	Wrong-scope errors fail to decline, or apparent improvement relies on dropping unresolved opportunities
P4: accepted mixed obligations reduce omission and returned work when recipient capacity exists	Priced omission/rework fails to improve, or extra governance work exceeds the decision gain

[AI] EAC extends implementation research with a decision contract between documentary evidence and permitted work. Deming cautions against tampering; Campbell and Goodhart against proxy-driven behavior; Ittner-Larcker-Randall against an untested measure-to-outcome chain. Tucker-Edmondson motivates retaining local containment and systemic work together. Composite-indicator guidance requires explicit denominators and weights; a high link count cannot compensate for a failed authority gate.

Limitations and identifying assumptions

[AI/PD] Originals and a linked execution/outcome panel remain unrecovered. Search recall, historical applicability, exposure and causal efficacy are unknown. The shadow benchmark assumes a stable authorized reference, independent domains, calibrated signal errors and defensible loss weights; capacity spillovers, reference errors or changed case mix can overturn its ranking. No historical customer effect or realized saving is identified. Recruitment feasibility, reference validity and the best simpler comparator must be resolved before commissioning a live evaluation.

Basis: proposed EAC propositions [PD], source and identification limits [AI]; six fresh primary sources and continuing methods literature.

PROPOSED ROUTING BENCHMARK / 1 OF 4

One docket can need two kinds of work

[PD] EAC tests an evidence-supported work recommendation. Local execution, governance repair, both and verification are distinct routes; none labels a person noncompliant.

[PD] Randomize independent domains with 20 frozen dockets each, conditionally independent given shared evidence grade (70% / 30%). Original reference states are local-only, governance-only, mixed and unresolved (30% / 30% / 25% / 15%). Both arms retain the same eligibility law; 8% known blockers always verify. Lock shadow decisions at intake + 48 hours; hidden day-14 adjudication reconstructs that original authority/evidence state.

Base signal	Stronger grade	Weaker grade
Local admissibility: sensitivity / false positive	92% / 6%	72% / 22%
Governance need: sensitivity / false positive	90% / 12%	65% / 30%
Recipient evidence: supported / unsupported need	90% / 65%	60% / 45%

[PD] Lookup routes by document presence. EAC recommends both when both checks qualify; local-only if local evidence qualifies without an accepted handoff; governance-only if only that handoff qualifies; verify if neither does. Signals are conditionally independent by state and grade. No rule sees truth; recipient evidence can be wrong. Mixed cases receive both routes in 55.7% of base EAC dockets versus none under lookup.

[PD] Charge errors, omitted work and waiting

Illustrative resource loss: unsupported local assignment 9; unnecessary governance 3; omitted obligation 4 each, or 2 each while verifying; verification wait 1; dual-route coordination 0.4; returned handoff 0.75. Common lookup costs 0.12; full checks add 0.55 and setup 0.18 per original docket, including blockers.

[PD] Entire law and prices are invented. No live execution, sanctions, promise changes or remedy delays. Research adjudication is separately budgeted.

PROPOSED ROUTING BENCHMARK / 2 OF 4

Size for break-even, not the largest effect

[PD] At 32 domains per arm the assumed full effect is easy to detect, but beneficial detection at break-even is only 9.8%. That design cannot reliably distinguish a small worthwhile gain from zero.

[PD] Require a lower directional bound on resource gain above setup cost K and a positive reduction in unsupported local recommendations, each per all original dockets. Exact finite moments feed fixed, law-derived normal radii at nominal 97.5% directionally. MDE uses 80% beneficial-tail detection at a two-sided 5% zero threshold. These are approximate bounds, not exact coverage.

Direct design	Per arm	MDE / K	Detection at K
initial	32	0.735 / 0.18	9.8%
base	640	0.165 / 0.18	86.2%
Scope stress	640	0.165 / 0.18	86.7%
Corroborated	256	0.261 / 0.32	93.1%

[PD] All four full-effect joint normal floors exceed 99.9%; this does not establish superiority to simpler rules. Base expected lower gain is 2.241, above K = 0.18. The base allocation needs 1,280 independent domains and 25,600 dockets; source evidence establishes neither recruitment nor variance. Calibrate first.

Finite check: estimate [95% MC interval]	Base	Corroborated
Economic false clearance at break-even	2.30% [1.99, 2.65]	1.99% [1.70, 2.32]
Unsupported-local superiority at null	2.49% [2.17, 2.85]	2.51% [2.19, 2.88]

[PD] Eight selected scenarios use 8,000 finite trials each. Base/corroborated full joint success is 8,000/8,000, with 99.95% Wilson lower bounds. Null zero-test rejection is 5.25% / 5.15%. Selected boundary checks are conditional calibration, not a guarantee over unknown reference errors or adverse response families.

[PD] Resource units per original docket. Lambda is a per-docket response mixture conditional on shared grade, not observed compliance. No sample minimum claimed.

PROPOSED ROUTING BENCHMARK / 3 OF 4

Passing the lookup test is not enough

[PD] Poor scope and recipient evidence still lets EAC beat the lookup shortcut. It nevertheless loses to simpler alternatives. More observations cannot repair that decision.

All-cost rule	Base	Scope stress	Corroborated
Lookup	4.84	4.84	4.84
Full EAC	2.66	4.35	2.55
Scope-only, matched evidence	3.16	4.03	3.40
Always verify	3.32	3.32	3.32
Always governance	3.64	3.64	3.64
Always local / always both	6.13 / 5.71	6.13 / 5.71	6.13 / 5.71

[PD] Full checks cost 0.55 + K 0.18; corroborated checks 0.95 + K 0.32. Matched scope-only checks cost 0.25 + 0.08, or 0.65 + 0.22 when corroborated. They share the full rule's local-assignment error rate within context. Always-verify/governance have zero unsupported local assignments but can omit useful local work. All fixed rules respect known blockers.

[PD] Prices require independent owner calibration. With deferral priced 0.5 per omitted obligation, always-verify costs 1.67 versus base EAC 2.32 and corroborated 2.31. No benchmark validation can validate the preferred price. A confirmatory study must include the best admissible simple comparator.

[PD] Signed realization, with every added burden retained

Net = a * h * favorable category reductions - adverse category increases - extra verification - K. Here a is avoidability and h the non-overlap share. Split changes by reference state, grade and cost category before discounting; do not net out extra waiting or returned work.

a * h	Base net	Stress net	Corroborated net
0.40	0.18	-0.68	-0.04
0.75	1.34	-0.0021	1.32
1.00	2.18	0.48	2.29

[PD] Hypothetical resource units/docket, no cash savings. Private grid also varies K by 0.75/1/1.5. Reprice and re-power before any purchasing decision.

PROPOSED ROUTING BENCHMARK / 4 OF 4

Buy evidence for a defined next decision

[PD] Research value concerns whether to retain the bounded lookup, verify current scope, or repair the ownership/capability basis. It is separate from routing loss and customer outcomes.

Research action / utility	Executable scope	Mixed duties	Missing basis
Retain bounded lookup	2	0	-3
Verify current scope	5	2	-2
Repair ownership/capability	-2	6	8

[PD] State priors are 40% / 35% / 25%. Without evidence, repair has expected utility 3.30; perfect information adds at most 2.80. A three-result study reports the true state with probability q and splits errors equally. All 27 contingent rules are evaluated.

Signal accuracy q	Study cost	EVSI	Net EVSI
55.0%	0.35	0.6625	0.3125
75.0%	1.00	1.6125	0.6125
90.0%	2.20	2.3250	0.1250

[PD] The middle study has the highest net value under these invented inputs. Fund bounded calibration of prices, reference quality, current authorization, executable local work and handoff evidence. The shadow result cannot authorize operational routing.

[PD] Missing reference can overturn apparent clearance

Day 14 is an audit deadline, not ultimate truth. At 2% unknown or wrong labels in each original arm, worst-case resource-contrast uncertainty is +/-0.593 and local-assignment uncertainty +/-4 percentage points. That erases the stressed expected lower resource bound of 0.547. Retain all dockets and paid work; re-adjudicate at day 28 against the original 48-hour state.

[PD] The bound uses score support 14.82 and original all-docket denominators; uncertain state membership needs separate validation. Later field evaluation must measure actual work, delay and customer outcomes under unchanged authority. Shadow scores establish none of these effects.

[PD] Separate research-utility units. No double counting with operational value; independent hidden reference must be validated, not assumed from a later label.

REPRODUCIBILITY / READER PROTOCOL

Reproduce the evidence route before trusting the action

[AI] EAC separates documentary coverage, diagnostic inference, authorized action and mature outcome evidence. Each requires a reproducible record.

Reproduction layer	Required reconstruction
Source and contribution	Verify the exact register cells, source-recovery attempts and authorship evidence. Preserve artifact creation, policy authority, adoption and operating delivery as separate claims.
Taxonomy and coverage	Fix the eligible mechanisms and versioned context. Reconstruct many-to-many policy relationships without duplicating episodes, conflating document count with coverage or erasing unresolved matches.
Attainment and measurement	Recompute rate and unit attainment on compatible exposure. Audit category transfers, denominator changes, zero bookings and the differences between observed results and unsupported planning ranges.
Routing and reference	Generate the explicit hypothetical states and fallible evidence. Reproduce each permitted route, paid verification step, independent mature reference and error classification.
Value and comparison	Price every adverse route and required work. Compare the full workflow against simpler policies, recalculate break-even detectability and evaluate separate information value.

Protected history and independent constructions

[AI] Historical values, indices, source identities and version mappings remain in the private annex. Public numerical illustrations are source-independent constructions. Their logical conclusions survive a common positive rescaling when all related quantities and prices transform consistently.

Evidence needed after a successful shadow test

[AI] Better routing under an agreed loss function does not demonstrate better operating outcomes. A deployment changes workloads, incentives and behavior. It requires accepted ownership, feasible capacity and an evaluation of the completed action package and mature customer outcome.

DECISION HANDOFF

[PD] Preserve uncertainty, verify scope and charge for unresolved work. Expand the routing workflow only when its additional value survives a strong comparator and the measurement contract supports the proposed use.

Private source maps, executable calculations, explicit hypothetical models and hash-bound reviews support reproduction. They do not authenticate missing originals or demonstrate historical impact.

RESEARCH REFERENCES / POLICY AND IMPLEMENTATION

Distinguish a written standard from executable practice

Primary research informs interpretation and prospective design. It does not authenticate historical files, establish adoption or supply a missing operating effect.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, 2, 40. doi:10.1186/1748-5908-2-40. Full primary article: adherence and dose are measured against intended delivery; a document link or booked benefit is not an adherence observation.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011; online 2010). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(2), 65-76. doi:10.1007/s10488-010-0319-7. Full primary article: distinguish feasibility, fidelity and penetration from service outcomes; declare each measure's referent and level.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6, 42. doi:10.1186/1748-5908-6-42. Full primary article, COM-B development: assess capability and opportunity alongside motivation. The framework does not identify an individual's intent from an outcome shortfall.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, 4, 50. doi:10.1186/1748-5908-4-50. Full primary article, readiness constructs: test resources, compatible workflows and access to knowledge before treating a listed standard as executable.

Bell, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., Ogedegbe, G., Orwig, D., Ernst, D., & Czajkowski, S.; NIH Behavior Change Consortium Treatment Fidelity Workgroup (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology*, 23(5), 443-451. doi:10.1037/0278-6133.23.5.443. Original abstract only: distinguish design, training, delivery, receipt and enactment in the proposed validation protocol.

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health*, 89(9), 1322-1327. doi:10.2105/AJPH.89.9.1322. Original abstract and metadata accessed: distinguish reach, efficacy, adoption, implementation and maintenance at their appropriate levels. No healthcare effect size is transferred.

Moreau, L., and Missier, P., eds. (2013). PROV-DM: The PROV Data Model. W3C Recommendation, 30 April. Official full specification: entities, activities, agents, derivation and revision support traceable evidence. EAC separately proposes applicability, expiry and authorization gates; provenance is not a correctness guarantee.

Primary scholarly or official sources; access scope is documented privately. Methods and conceptual frameworks do not transfer empirical effects across settings.

RESEARCH REFERENCES / MEASUREMENT AND EVALUATION

Preserve the population and test the complete workflow

Primary research informs interpretation and prospective design. It does not authenticate historical files, establish adoption or supply a missing operating effect.

NIST/SEMATECH. e-Handbook of Statistical Methods, section 6.3.2.2: Individuals Control Charts. Official method. Tentative limits summarize ordered programme-booking observations only after scope and maturity checks. A within-limit attainment series does not identify adherence or a stable operating phase.

Laney, D. B. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531–537. Primary publisher abstract. Overdispersion adjustment requires actual binary counts, exposure and rational subgroups. Programme-plan ratios and document counts are not eligible operating attribute panels.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37–46. Primary publisher record. Apply blinded agreement and kappa to current-version, applicability and executable-route labels. Shared coder agreement can coexist with a 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. Selecting poorly performing programmes creates regression-to-the-mean risk. Use an explicitly assumed persistence sensitivity and a concurrent comparison, not recovery alone as evidence of compliance improvement.

Hudgens, M. G., & Halloran, M. E. (2008). Toward Causal Inference with Interference. *Journal of the American Statistical Association*, 103, 832–842. Original manuscript. Shared reviewers and governance recipients can transmit routing practices or congest handoffs across dockets. Randomize independent review-capacity domains and measure cross-domain leakage.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, 225, 200–230. Author manuscript. A later nonrandom operational rollout requires separately defined group-time ATT, eligible controls and defensible parallel trends. Shadow routing improvement cannot identify historical programme effects.

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. Value evidence only when a specified likelihood can change a feasible routing or verification decision. Subtract acquisition costs and compare with fixed rules; do not add information value to operating benefit.

Primary scholarly or official sources; access scope is documented privately. Methods and conceptual frameworks do not transfer empirical effects across settings.

RESEARCH REFERENCES / OPERATING GOVERNANCE

Connect accountability to mechanism and authority

Primary research informs interpretation and prospective design. It does not authenticate historical files, establish adoption or supply a missing operating effect.

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44, 823–830. Primary publication. Grade the registered role, historical correction and accepted decision separately. File metadata and derivative memory cannot establish sole invention, accepted handoff or implementation.

Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures. *California Management Review*, 45(2), 55-72. Author manuscript. Repeated local workarounds may leave system problems unresolved. EAC keeps a mixed local action and accepted governance obligation open together rather than forcing a single blame category.

The W. Edwards Deming Institute. The Funnel Experiment. Official exposition. Routine variation does not justify repeated corrective pressure. Diagnose whether the standard, opportunity, implementation or recording changed before altering the operating response.

Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Original institutional record. Targeted metrics can alter behavior and classification. Retain all-cause customer outcomes and independent fidelity observations beside booked programme credits.

Goodhart, C. A. E. (1992). The Objectives for, and Conduct of, Monetary Policy in the 1990s. Reserve Bank of Australia Annual Conference. Original author paper, section 1(b)(i). A targeted proxy can lose its relationship to the objective. Policy-link coverage and booked benefits are not substitutes for executable reach or customer outcomes.

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. Test the measure-to-action-to-outcome chain. A named policy and a low programme attainment ratio do not identify the causal mechanism or the return to adherence pressure.

OECD / European Commission JRC (2008). Handbook on Constructing Composite Indicators. Official handbook. Aggregation depends on definitions, denominators and weights. Report distinct eligible mechanisms and exposure-weighted executable reach; repeated policy aliases cannot multiply coverage.

Primary scholarly or official sources; access scope is documented privately. Methods and conceptual frameworks do not transfer empirical effects across settings.

FINAL REFEREE REPORT / INTERNAL AGENT-ASSISTED REVIEW

Accept the diagnostic correction; test the routing contract

[AI] This author-led, agent-assisted review is not external peer review. Acceptance concerns the evidence reconciliation and proposed evaluation, not historical implementation or financial benefit.

Major issue	Disposition in the revised case
Adherence inferred from plan miss	Exact constructions reproduce the same booked result under different rollout, adherence, efficacy and capture. A plan gap cannot identify the mechanism.
Coverage treated as capability	Fix the eligible mechanism denominator, separate source relationships and preserve current applicability, executability, authority and uncertainty.
Search absence treated as nonexistence	Bound the finding to the searched corpus and access. Parallel queries can share blind spots; active programs can coexist with missing dedicated guidance.
Reported gain treated as benefit	Retain category transfers and differing exposure bases. Other programs’ targets, pilot results and cost frames do not establish this project’s outcome.
Automation treated as implementation	The historical lookup remains a design. Proposed routes require evidence refresh, accepted ownership, paid verification and comparison with simple alternatives.

Partner critique: can a COO act on page one?

[AI] Yes: appoint accountable policy and measurement owners, verify executability and stop using document presence as a causal diagnosis. Authorize a bounded routing test with a strong comparator before expanding automation or assigning outcome accountability.

Remaining conditions and minor issues

[AI] Original-source authentication, coverage crosswalks, reference validity, evidence-error rates, loss prices, verification capacity and transport across operating modes remain unresolved. Future tests must preserve mixed and missing cases. Internal review cannot supply historical adoption or a missing causal effect.

REVIEW DISPOSITION / PASS WITHIN STATED SCOPE

[AI] Accept the analytical correction and falsifiable design. Invalid authority, incompatible measurement, insufficient break-even detectability or failure against a simpler policy requires revision before reliance.

Internal evidence, numerical, methodological, executive and actual-page reviews. No external endorsement, consulting rank or verified historical effectiveness is claimed.