

OPERATING CASE STUDY / OPERATIONAL EXCELLENCE

Operating Close & Workbook Governance

A clean workbook can still be an incomplete close

An anonymized project case study of a shared operating workbook: repairing formulas, reconciling monthly inputs, preserving owner accountability and defining when the management view is ready for release.

PROJECT CONTEXT / OPERATIONS, ANALYTICS & IMPLEMENTATION

BLUF / BOTTOM LINE UP FRONT

Release the close only when the full decision chain reconciles.

RECALCULATED FINDING

Replanning after the data write removed 78.26% of initial validation flags. A clean local repair still did not establish a complete shared close. Validate the input state, dependencies and owner disposition before releasing the management view.

CONTRIBUTION IN THE SOURCE RECORD

Tab-level modeling direction, repair coordination, preservation rules and narrative preparation. Registry-attested; provisional hoop-level support.

EVIDENCE STATUS

Evidence-reconciliation and proposed-governance case. Original artifacts, complete shared-master acceptance and customer impact remain unestablished.

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

EXECUTIVE ABSTRACT

Close the decision chain before releasing the monthly view

[AI] A shared operating workbook is a coordination mechanism. Its release depends on coherent inputs, valid rules, completed dependencies and accountable owner decisions.

CONTEXT | Local maintenance supported a shared management view

[HR] The register describes recurring formula repair, monthly fills and narrative preparation within a larger workbook. Maintained tabs depend on source layers and entries controlled by other owners. A completed local copy and a later shared master contain different output states.

OBJECTIVE | Define a close that leaders can rely on

[PD] Distinguish local calculation success from an authorized monthly release. Require the input cut, aggregation rule, dependency chain, exception disposition and narrative to agree on one reporting identity.

APPROACH | Reconcile workflow stages and test the validation logic

[RC] Recalculate stage-dependent validation and preserved-entry ratios. [AI] Challenge error-scan coverage, historical-oracle validity and version adoption. [PD] Specify a Dependency-Gated Close with separate computational, semantic and release gates.

FINDINGS | A clean check can answer the wrong question

[RC] Replanning after the data write removes 78.26% of the initial flags. Of preserved owner entries, 90.48% differ from their prior-period values. [AI] Neither fact establishes source correctness; scoped zero-error scans cannot certify downstream completion.

DECISION VALUE | Appoint one accountable close owner

[PD] The COO should assign a release owner to reconcile the dependency manifest and require explicit disposition of material exceptions. Preserve disputed entries, hold affected claims, and publish only the agreed cut after owner acknowledgments and independent checks.

IMPLICATIONS | Credit the operating judgment within its boundary

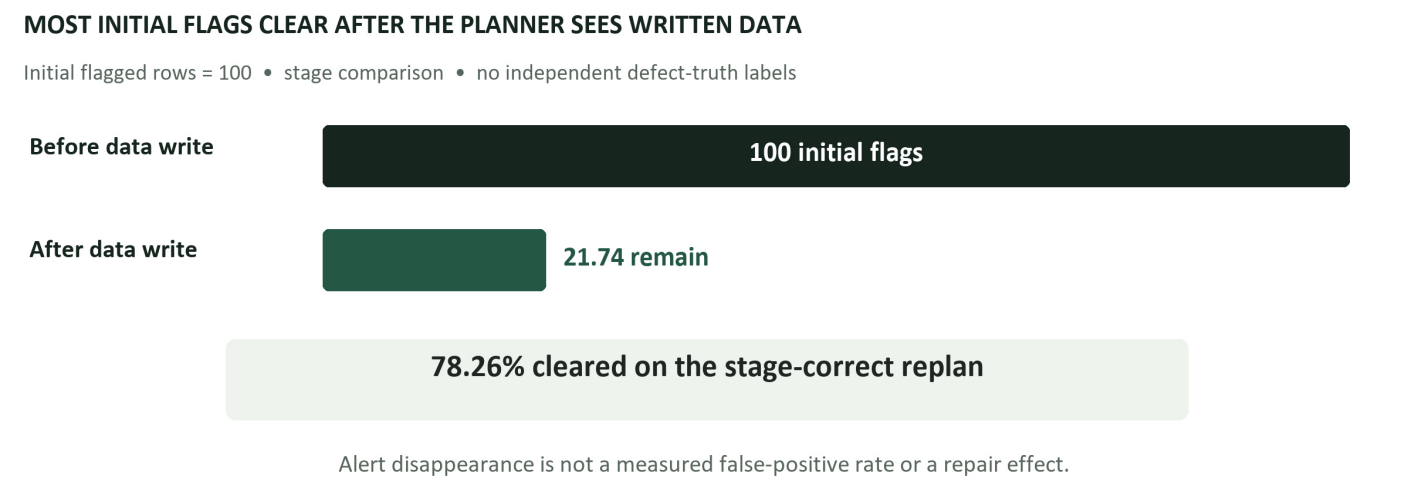
[AI] Tab-level modeling direction, repair coordination and narrative preparation receive provisional registry-level hoop support. The case establishes no end-to-end close-time saving, causal customer improvement or realized financial benefit.

S1: selected project register, M01–M14 and D01–D10 with linked evidence. HR historical record; RC recalculated; AI analytical inference; PD proposed design. Original artifacts were not independently recovered.

EXHIBIT 1 / STAGE-AWARE VALIDATION

Validate the workbook state before treating an alert as a defect

[RC] Replanning after the data write removes 78.26% of the initial flags. The comparison changes because the planner now sees a different input state.



[HR] The derived-layer planner first evaluated the workbook before the underlying monthly data had been written. A later run recovered most flagged rows. [RC] With the initial flag count indexed to 100, 21.74 remains after replanning.

What the result supports	What it cannot establish
[AI] Stage dependence	The input snapshot and order of operations affect a validation result. Every alert needs its tested state and predicate.
[AI] Truth classification	Removed flags are not automatically false positives; remaining flags are not automatically defects. Independent adjudication is still required.

[PD] Separate readiness checks before writing from correctness checks after calculation. Freeze the planned writes, apply them in an authorized copy, rebuild dependent calculations, then compare against the same declared validation contract. A changed rule or scope requires a separate reconciliation.

OPERATING DECISION

[PD] Ask the close owner to record each alert’s input version, workflow stage, predicate and disposition. Measure adjudicated residual risk and review burden rather than rewarding the largest alert count or its fastest disappearance.

[HR/RC/AI/PD] S1:D06/E03/E16. Initial flags =100; removed =78.26%; remaining =21.74%. This is a registered stage comparison, not an identified defect-reduction effect or an independent false-positive estimate.

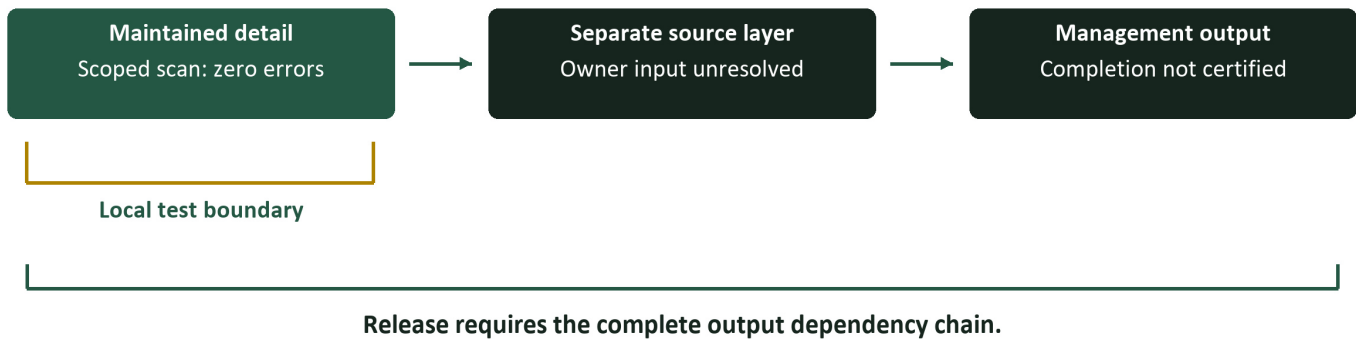
EXHIBIT 2 / DEPENDENCY COMPLETION

A clean local tab cannot certify an incomplete decision chain

[HR] The repair log records zero error cells in a scoped scan. Other source-owner dependencies and the shared management output remain unresolved.

THE RELEASE BOUNDARY EXTENDS BEYOND THE LOCAL ERROR SCAN

Registered architecture • ownership and data state determine dependency completion



[HR] Displayed detail rows do not necessarily feed the summary. Additional program, exposure and calculation layers are needed; a separate source owner supplies headline dependencies. [AI] A populated cell and a visible summary are different positions in that chain.

Control	Boundary
[AI] Formula-error scan	Detects errors in the inspected ranges. It does not prove valid references, complete inputs, correct constants or semantic consistency.
[AI] Text and zero handling	Range SUM ignores text; range AVERAGE ignores text but includes numeric zero. Incorrect types can change meaning without raising an error.

[AI] The source’s blanket claim that text zero poisons both functions is too broad. Inspect the actual formula and storage type. A missing upstream contribution, a validated observed zero and a future placeholder must also remain distinct even if their displayed values coincide.

RELEASE GATE

[PD] Trace each management output to its full dependency set. Require the steward to identify missing owners, stale constants and typed-zero states. A local scan may pass while the affected headline remains on hold.

[HR/AI/PD] S1:M03/M06/M13/E01–E07. Historical zero-error scope is limited to the maintained tab blocks. Microsoft SUM and AVERAGE documentation supports function semantics, not the authenticity of any historical formula.

EXHIBIT 3 / RULE IDENTIFICATION

A prior-period match cannot prove the current aggregation rule

[AI] Reproducing a signed historical result is a useful regression test. It cannot identify the intended rule when the oracle is wrong or several rules produce the same total.

TWO RULES CAN PASS THE SAME HISTORICAL CHECK

Synthetic example • both families in scope • reference = 100

Test input	Both-family rule	Single-family rule
Reference: A=100, B=0	100 • matches	100 • matches
Holdout: A=80, B=20	100 • includes B	80 • omits B

Historical fit cannot distinguish these rules; the holdout exposes the difference.

The data owner must still define the intended population and hierarchy.

[HR] The method selects a rule by reproducing the prior month, then applies it to the new month. The register describes duplicate leaf labels and an under-fill, but differs on whether that under-fill belonged to the historical oracle or the current resolver. [AI] Historical agreement alone cannot resolve this ambiguity.

Validation layer	Required question
[PD] Historical reproduction	Does the candidate reproduce the frozen reference, and is any tolerated difference documented?
[PD] Semantic identification	Does the rule include the intended population and ancestry, exclude overlap and remain valid after the hierarchy or period changes?

[PD] Challenge the rule with cases that distinguish alternatives: repeated leaf names, changed parent mapping, empty and inapplicable categories, and independently reconciled source totals. Preserve an explicit exception when the source cannot resolve a mapping. More agents using the same oracle do not create independent ground truth.

STEWARD ACCOUNTABILITY

[PD] Require the data owner to approve the mapping contract and the model steward to retain counterexample tests. Exact historical fit permits further validation; it does not authorize a new-period write by itself.

[HR/AI/PD] S1:D02/D03/D05/M09/E03/E08/E16. Under-fill chronology remains unresolved. The rule diagram is conceptual, not a historical replay; a recommended rollback is not verified execution.

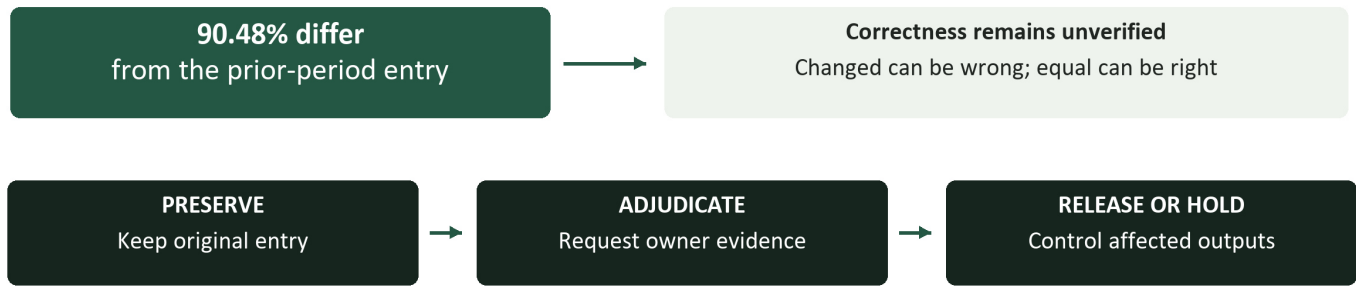
EXHIBIT 4 / OWNER-CONTROLLED EXCEPTIONS

Preserve disputed owner values and hold the dependent output

[RC] Of preserved entries, 90.48% differ from their prior-period values. Difference establishes neither current-period authenticity nor correctness.

VALUE DIFFERENCE IS A SCREENING FEATURE, NOT AUTHENTICATION

Registered comparison of preserved entries • preserve provenance; adjudicate correctness



Write authority and publication authority answer different questions.

[HR] The workflow preserves pre-existing owner entries when calculated estimates disagree. A comparison classifies most as distinct from the prior month, while a suspected cross-network copy remains in that preserved set. The record calls for escalation and withholding affected summary claims.

Separate decision	Accountable action
[PD] May the entry be replaced?	The authorized owner approves a documented correction. Preserve the original, rationale and downstream change trace.
[PD] May the output be published?	The release owner assesses materiality and holds dependent claims until the exception is resolved or explicitly bounded.

[AI] A copied value can differ from last month; a correct current entry can equal it. Equality is therefore a screening feature, not an authenticity test. Likewise, matching another network’s values raises a review question but does not independently establish the correct replacement or its financial effect.

CONTROL WITHOUT OVERWRITE

[PD] Preserve provenance, request owner evidence and quarantine the disputed publication claim. Avoid turning respect for write authority into automatic acceptance of the value, or a reviewer’s estimate into an unauthorized correction.

[HR/RC/AI/PD] S1:M07/M08/D07/E04/E07. Distinct entries =90.48% of preserved entries; exact count and cell addresses stay private. Suspected copy and proposed correction remain adjudicated in the recovered record.

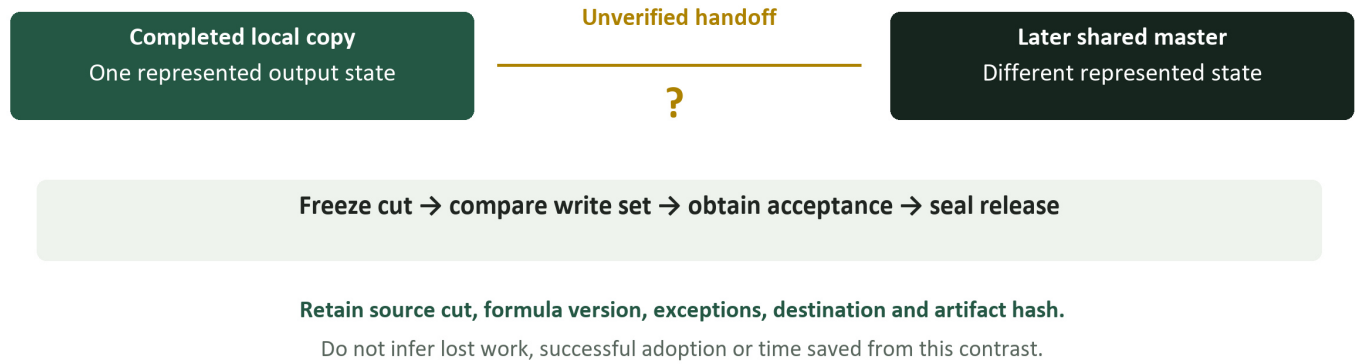
EXHIBIT 5 / SHARED-MASTER CUTOVER

Close the shared master with a controlled version cutover

[HR] A completed local fill and a later shared master contain different output states. A later file timestamp does not establish acceptance of the local work.

A COMPLETED COPY NEEDS A VERIFIED PUBLICATION TRANSACTION

Recorded state mismatch • a later timestamp does not identify the accepted source cut



[HR] The local delivery records a nonzero summary output; the later shared master carries zero in its counterpart while agreeing on a companion metric. Other fields combine different actuals cuts and incomplete dependencies. [AI] The register does not recover a verified import-and-acceptance event.

Close evidence	What must be recorded
[PD] Source and transformation	Common reporting period, extraction cut, maturity convention, mapping version, formula version and declared exceptions.
[PD] Publication transaction	Destination identity, authorized write set, before/after comparison, owner acknowledgments and the released artifact hash.

[AI] The mismatch could reflect incomplete import, subsequent revision, changed inputs or a different dependency state. It cannot be resolved by choosing the more favorable copy. The represented short fill interval also excludes or leaves uncertain source extraction, adjudication and shared-master acceptance; it is not end-to-end close time.

CLOSE OWNER ACTION

[PD] Reconcile the two lineages and complete a controlled handoff. Record unresolved fields and who may release them. Measure close time from the scheduled cycle’s start through accepted publication, using the same scope before and after any redesign.

[HR/AI/PD] S1:M06/M12/M13/E04/E06/E07/E15. The version contrast is a recorded-state discrepancy, not proof of lost work, completed adoption or operational deterioration. No cycle-time saving is identified.

EXHIBIT 6 / NARRATIVE AND CLOCK ALIGNMENT

Narrative completeness does not establish current evidence

[AI] A fully populated review can still misstate launch, maturity or performance. The narrative must inherit the released metric contract, not merely reuse an earlier paragraph.

NARRATIVE CLAIMS NEED A COMMON REPORTING IDENTITY

Proposed evidence contract • completeness, launch, performance and maturity are distinct

PLANNED START
Intended effect timing

VERIFIED LAUNCH
Execution evidence

MATURE OUTCOME
Eligible reporting cut

A populated row does not make these three clocks agree.

Metric definition + source cut + rule version + status evidence

Reopen the narrative when its approved metric basis changes.

[HR] One narrative fill is represented as complete, while the next cycle remains largely empty. The earlier narrative uses an earlier actuals cut despite subsequent restatement. Its planned start periods and performance labels come from different source fields.

Narrative claim	Evidence required
[PD] Work has launched	Verified implementation date and eligible scope; a first nonzero planned effect is not a delivery record.
[PD] Target is achieved	Matched actual, target, unit, population, period and maturity rule; label projections and partial data explicitly.

[AI] A partially filled row is not a completed row. The source’s explanation that a miss reflects execution rather than design also exceeds the evidence: plan-versus-attributed-actual comparison does not distinguish treatment effectiveness, implementation, measurement and other causes. A coefficient selected for sign and significance does not close that causal gap.

REVIEW DISCIPLINE
[PD] Link every status to the same release identity as its supporting metric. Require the operating owner to confirm delivery and the analyst to label remaining attribution uncertainty. Reopen narrative approval when the underlying cut changes.

[HR/AI/PD] S1:M05/M10/M11/D08/D09/E05/E06/E10. Narrative coverage, truth, launch and causal attribution are distinct constructs. Subsequent source revisions are not a chronological operating-outcome panel.

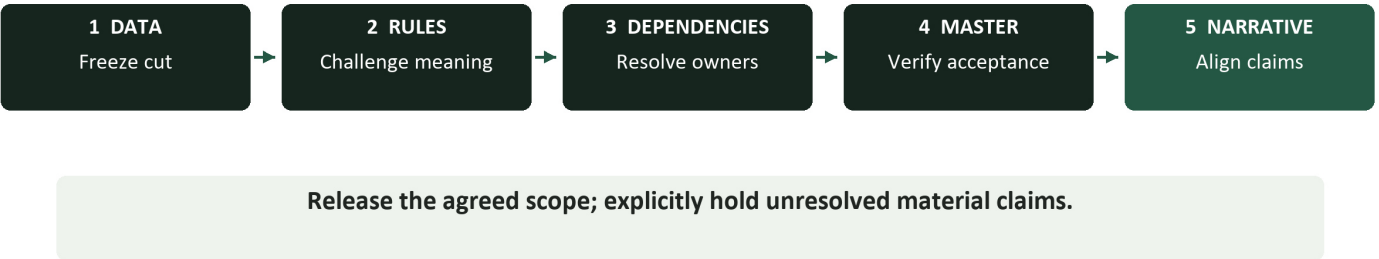
EXHIBIT 7 / DEPENDENCY-GATED CLOSE

Authorize one monthly release against an explicit integrity contract

[PD] The Dependency-Gated Close makes the management view the release unit. Passing a local check advances the workflow; it does not replace the remaining gates.

DEPENDENCY-GATED CLOSE MAKES THE RELEASE THE CONTROL UNIT

Proposed contract • one close owner coordinates stewards, row owners and publication



Success means usable, authorized evidence—not simply a workbook without errors.

Accountable gate	Evidence needed to advance
[PD] Data and model stewards	Freeze the source cut, metric definitions, mappings and formulas; run stage-correct tests and independent semantic challenges.
[PD] Dependency and row owners	Complete required inputs; authenticate entries; approve corrections or record unresolved material exceptions.
[PD] Close and operating owners	Reconcile the shared master and narrative; authorize one release, disclose holds and retain the acceptance record.

[PD] Compare DGC with usual practice on fixed, held-out assignments; retain held and late tasks. Audit production prevalence before extrapolation. Guard against unsupported release, false holds and lost coverage.

[AI] Tab-level direction, repair coordination and narrative preparation have provisional registry-level hoop support. Shared-master acceptance and downstream benefit remain unestablished.

COO DECISION

[PD] Appoint the close owner, define materiality and pilot the contract. Require economically informative evidence and satisfied release guardrails before scaling.

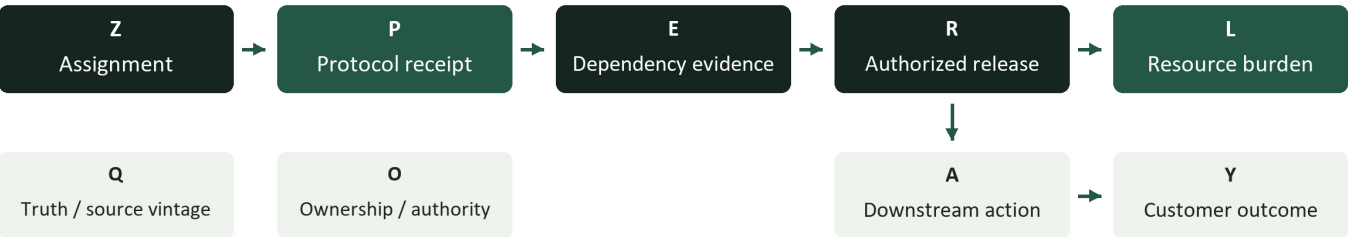
[AI/PD] Proposed DGC; S1:D01–D10/E01–E17 inform its controls. Neither the full contract nor its effectiveness is established historically. Prospective power and economic inputs are synthetic, separately disclosed planning assumptions.

Control the release before claiming an operating result

[PD] Dependency-Gated Close (DGC) treats a publication decision as the control unit. A repaired formula is one input to that decision, not its authorization.

TEST CLOSE GOVERNANCE AT THE DECISION-RESOURCE ENDPOINT

Proposed DAG • benchmark trial identifies assignment effect • arrows remain hypotheses



Additional paths: P → L (protocol effort); Q → E and L; O → R.

C: skill, workload and template difficulty → E, R and L. Condition on the frozen benchmark.

Estimate total Z → L; do not adjust away post-treatment E or R, or exclude held packets.

R → A → Y needs separate field evidence. Unexplained residuals do not identify their cause.

[PD] Z assigns DGC; P is protocol receipt; E is current-state dependency evidence; R is authorized release; L is 14-day decision-resource burden. Q represents source vintage and semantic truth; O ownership authority; C analyst skill, workload and template difficulty. R may influence downstream action A and mature customer outcome Y.

[PD] Z→P→E→R→L is the proposed governance pathway. P→L includes verification effort, so more checks can increase burden while reducing error. Q→E and Q→L capture harder source problems; C→E/R/L names confounding in historical comparisons. Randomization addresses assignment, not every mediator.

[AI] The register represents repair, preservation and escalation activities. It does not identify R→A or A→Y. Neither no-error output nor a populated narrative identifies a customer effect. Residual defects can arise through unmeasured source, rule, authority or transmission paths; their existence does not select one cause.

[PD] Estimate total Z→L without conditioning on completed assignments or approved releases. Index subsequent close cycles separately to keep feedback acyclic. Fixed deployment cost is outside L; variable DGC effort is inside it.

S1: protected register describes multiple ownership and calculation boundaries. DAG is proposed; no historical causal path is estimated.

TECHNICAL FOUNDATION • ESTIMANDS

Keep calculation success separate from release correctness

[AI] A changed value, a passing scan and an approved current-state output have different denominators and evidential meanings.

Quantity	Exact target and boundary
Formula-scan result	[PD] Error-token incidence in a declared cell/range set after a declared recalculation. A zero result excludes only scanned error types in that state.
Historical reproduction	[PD] Output distance under a candidate rule and frozen reference-period inputs. Agreement is conditional on reference validity and cannot uniquely identify the rule.
Current semantic correctness	[PD] Agreement with an independent, authorized construct specification for the current population, source cut, formula version and output.
DGC assignment benefit	[PD] Control-minus-DGC mean total variable decision-resource burden over 14 days per eligible benchmark assignment, averaged equally over 16 assignments per randomized team.
Production / customer effect	[PD] Separate estimands requiring target-population fault prevalence, workflow transport and an identified field comparison. Neither is estimated here.

[PD] Register all assignment slots before allocation; start each clock at its scheduled availability. Retain holds, withdrawals, unresolved truth and failure to submit. Use one pretrial reference cost 100, not a separate index for each team. Assignment to DGC is the intervention; receiving every check is a mediator.

[PD] Prespecify unsupported release, false hold and time-to-authorized-release as component outcomes. Incorrectly releasing an entry differs from declining to overwrite it: edit authority and publication authority are separate. A burden composite cannot excuse a breached safety guardrail.

S1: distinct-value screen, scoped recalculation, owner preservation and unresolved shared-master handoff. Definitions are proposed, not historical effect estimates.

TECHNICAL FOUNDATION • MEASUREMENT VALIDITY

Every cell needs a meaning, a clock and an accountable source

[PD] Freeze the measurement contract before testing formulas; otherwise a reproducible calculation can consistently answer the wrong question.

Metric family	Required validation
Operational rate	[PD] Define defect versus unique customer unit, eligible exposure, event/record clock, maturity cut and business-line taxonomy. Do not pool mixed-vintage monthly cells.
Program actual	[PD] Distinguish independently observed delivery from plan-scaled estimates, booked contribution, default zero and not-yet-due data. A populated actual column does not establish observation.
Narrative status	[PD] Require launch evidence, outcome cut and status rule. Partially populated differs from complete; whitespace and copied language do not establish status accuracy.
Preserved owner entry	[PD] Retain provenance, timestamp, scope and authority. Difference from a prior entry screens values; neither difference nor equality authenticates correctness.
Resource burden	[PD] Prespecify units, prices and mutually exclusive consequence codes. Include verification, rework and delay once; separate fixed deployment allocation.

[PD] Proposed audit: double-code 240 held-out records stratified by semantic family, source maturity, owner conflict and clean control. Coders work independently, blind to arm and proposed release; an authoritative panel adjudicates disagreements. Preserve natural and seeded labels and selection weights. Report confusion matrices, prevalence, raw agreement and kappa with team-bootstrap intervals.

[PD] This pilot size is an operational assumption, not a precision guarantee. Re-size from pilot prevalence and clustering. Test codebook invariance across networks and template families; high agreement alone cannot establish construct validity.

[AI] The protected record contains competing descriptions of the erroneous reference period. Preserve that chronology fork until original files and formula lineage resolve it.

S1: maturity, estimated actuals, status and reference-period conflicts. Wang–Strong motivates contextual quality; Barr et al. distinguishes available oracles from truth.

TECHNICAL FOUNDATION • RULE IDENTIFICATION

A matching reference period does not identify the right rule

[AI] Passing a historical fixture is necessary for an agreed regression test; it is insufficient evidence that the candidate rule expresses the intended population.

[PD] Synthetic counterexample: reference inputs A=100 and B=0 give 100 under both $f=A+B$ and $g=A$. Holdout inputs A=80 and B=20 yield 100 and 80. Both rules pass the reference; only the specification that includes both families establishes which is valid. Historical fit cannot supply missing semantic authority.

Challenge	Prespecified distinguishing test
Family omission / overlap	[PD] Activate each authorized leaf separately; test aggregate identities and duplicate membership. Freeze disjointness rules before comparing totals.
Month / denominator mapping	[PD] Perturb source period and eligible exposure independently. Check sign, units, time alignment and expected scaling; inspect dependency lineage.
Actual-versus-plan meaning	[PD] State the zero-delta and positive-delta behavior required by the owner. Test whether a plan-scaled quantity is a constructed estimate rather than an independent actual.
Cell type semantics	[PD] Test numeric 0, text"0", blank and unavailable states through the actual functions used. SUM and AVERAGE do not have identical treatment of text and numeric zero.

[AI] Microsoft documents that SUM over references ignores text, while AVERAGE includes numeric zeros and ignores text/empty references. A blanket claim that text zero necessarily corrupts SUM is therefore unsupported. Direct arithmetic, denominators and downstream conversions still require specific tests.

[PD] Hold out entire dependency-template families from rule development. Parameter variants of the same copied template are not independent evidence of generalization. Record every candidate rule and failed challenge, not only the surviving rule.

S1: signed-period gate and formula semantics. Barr et al.2015; Panko 2006; Microsoft SUM/AVERAGE documentation. Counterexample is synthetic, not an internal fixture.

Revalidate the current state before committing the write

[AI] No-clobber preservation controls one write behavior. It does not resolve the accuracy of the preserved value or a concurrent owner edit.

DGC stage	Required release evidence
Freeze	[PD] Record authorized scope, source cuts, dependency versions and destination version. Snapshot the original; retain ownership for every proposed write.
Stage inputs	[PD] Validate source eligibility and write the approved data layer into an isolated candidate state. Keep unresolved evidence explicit.
Recalculate / challenge	[PD] Rebuild dependencies, test semantic invariants and recompute payload-dependent exceptions. Run current-state checks after prerequisite writes.
Adjudicate	[PD] Preserve disputed owner entries; obtain authorized evidence or hold affected claims. Escalation includes an owner, deadline and materiality rule.
Commit / accept	[PD] Compare destination version against the frozen version, apply only the authorized write set, and verify accepted output/hash. Changed version triggers re-read and re-adjudication.

[AI] Most represented pre-write flags disappeared after the data layer changed. This shows state dependence of that comparison, not automatic correctness of the resulting payload. Retaining a stale exception list can waste review; clearing it without fresh semantic checks can hide defects.

[PD] RFC9110 conditional-request semantics motivate a version-match precondition against lost updates; this is a design analogy, not a claim that the workbook implements HTTP transactions. Microsoft full rebuild updates calculation dependencies but cannot approve business meaning.

[PD] Define rollback and partial-release boundaries before execution. A verified local copy and a later shared master need an acceptance record; timestamp order alone does not establish adoption, loss or successful publication.

S1: data-write/replan order, protected ownership and publication ambiguity. RFC9110 §13.1.1; Microsoft CalculateFullRebuild; ICAEW2024 principles 17–20.

TECHNICAL FOUNDATION • VARIATION & SELECTION

Mixed measurement states do not establish operating change

[RC] The private audit computes 37 ordered screens without manufacturing missing months. No screen supports an operating-improvement or stability classification.

[PD] For a commensurate sequence $x_1 \dots x_n$, $MR_t = |x_t - x_{t-1}|$ for $t=2 \dots n$. Tentative individual limits are $\text{mean}(x) \pm 2.66 \text{ mean}(MR)$; this audit uses moving-range upper limit $3.268 \text{ mean}(MR)$. Limits require meaningful order, a comparable process phase and an appropriate rational subgroup.

Recorded sequence	Interpretation retained
Monthly operational cells	[AI] Tentative formulas are reproducible, but mixed maturity and source vintages invalidate a stable-process reading. A crossing would not identify the repair effect.
Plan / projection paths	[AI] Mechanical shape checks only. Management-set paths are not realizations of the process they target.
All-zero error scans	[AI] Degenerate dispersion in a scoped detection instrument; not proof that semantic error incidence is zero.
Within-run flag comparison	[AI] Different payload states, not successive independent process observations. Clearing flags measures comparison-state change.
Sparse endpoint pairs	[AI] Missing intermediate observations stay missing. Two-point limits cannot justify a trend or stable process claim.

[PD] No commensurate binary count/denominator panel supports a Laney p' chart; rounded rates cannot create one. Preserve source-version alternatives rather than averaging them.

[AI] No fixed worst-N treatment cohort and follow-up comparison is supplied for this workbook-governance case. Upstream targeting does not identify a workbook effect. A future selected-site outcome study must freeze selection, estimate reliability and compare untreated reversion; no historical regression-to-the-mean adjustment is estimable here.

S1 and private ordered-screen audit; NIST individuals-chart formulas; Barnett et al. on regression to the mean. Screen count describes analytical coverage, not independent histories.

TECHNICAL FOUNDATION · CONTRIBUTION & MODERATORS

Attribute the design role without assigning unmeasured outcomes

[HR] The register describes a bounded tab stewardship and review role inside a wider shared reporting system. The original workbooks and approvals remain unauthenticated in this package.

Contribution claim	Process-tracing grade and test
Directed scoped repair and fill	[HR] Provisional hoop: registry narratives and artifact locators are consistent with direction. Recover contemporaneous requests, versions and sign-offs; copied attribution remains an alternative.
Specified conventions and exceptions	[HR] Provisional hoop: represented instructions identify scope, preservation and escalation choices. Original dated rule decisions would strengthen this claim.
Owned the shared release	[AI] Unestablished: ownership of selected tabs does not establish authority over cross-owner calculations or final publication.
Delivered customer or cash benefit	[AI] Unestablished: no controlled outcome linkage, mature field panel or financial realization record supports attribution.

[AI] No smoking-gun or doubly decisive grade is assigned. Multiple rows copied from the same underlying narrative are not independent corroboration. A named file author or last modifier can support artifact association but cannot alone discriminate design, implementation and approval.

[PD] Expected moderators include dependency fan-out, frequency of source revisions, ownership fragmentation, template reuse and operator expertise. DGC should help where output chains cross boundaries and accepted rules exist. It cannot resolve an undefined construct by adding more checks.

[PD] Prespecify moderator interactions and retain class-specific uncertainty. Complex work may cost more to verify even when defect escapes fall. Avoid ranking small networks or teams from mixed-denominator totals; the trial should evaluate the governance package before claiming business-line superiority.

S1: role boundaries and unauthenticated locators. Collier process-tracing logic distinguishes necessary evidence from discriminating proof; proposed moderators are untested.

TECHNICAL FOUNDATION • PROSPECTIVE EVALUATION

Randomize teams while holding the evidence task constant

[PD] A parallel sandbox trial tests DGC against current close practice before production rollout. Randomize at the smallest independent team/operator unit that can prevent learning spillovers.

[PD] Freeze 16 eligible assignments per team before allocation. Hold out dependency-template families from development; balance a fixed benchmark across arms. Include seeded semantic faults, adjudicated naturally occurring faults and clean assignments. Keep their origins explicit. Independent teams may not share solutions, tooling changes or adjudication answers during follow-up.

Trial element	Preregistered contract
Primary ITT	[PD] Equal-team mean of equal-assignment 14-day variable burden; control minus DGC. Retain every slot from scheduled availability, including noncompletion and owner holds.
Reference panel	[PD] Blind adjudicators to assignment and release choice. Use semantic specifications, independently recovered lineage and owner authority; disagreement is not automatically a false release.
Unresolved evidence	[PD] Apply prespecified terminal resource penalties and report truth-uncertainty bounds separately. Never infer correctness from non-detection or drop unresolved work.
Guardrails	[PD] Unsupported release, false holds, unauthorized edits, coverage and delay thresholds. Predeclare multiplicity and the decision rule for any safety breach.

[PD] Estimate the cluster-level difference with prespecified variance diagnostics and randomization sensitivity. The planning model assumes independent Gaussian team means and equal variance conditional on the fixed benchmark; check this in a blinded pilot. Heavy tails, unequal task counts or shared-template random effects require revised power.

[PD] Assignment, receipt and fidelity stay distinct. A stepped crossover is avoided because verification learning persists. Production deployment and customer impact require separate transport evidence and field evaluation; sandbox cost performance does not identify either.

Proposed cluster experiment, not historical identification. Campbell et al.2012 cluster-trial reporting guidance; finite-sample design calculations are disclosed on the following model pages.

TECHNICAL FOUNDATION • FRAMEWORK & LITERATURE

Dependency-Gated Close makes four claims testable

[PD] DGC combines semantic specification, state-aware validation, bounded authority and verified acceptance. Its propositions concern release behavior, not assumed customer savings.

Proposition	Mechanism and disconfirming evidence
P1 • Distinguishing tests	[PD] Held-out dependency-family challenges reduce semantic escapes relative to reference-fit-only checks. Contradicted if the upper confidence endpoint for the randomized reduction falls below its minimum useful threshold.
P2 • Current-state gates	[PD] Post-prerequisite validation reduces stale exceptions without increasing missed defects. Contradicted by a prespecified escape-risk breach or a precise absence of the intended review reduction.
P3 • Authority plus versions	[PD] Owner adjudication with version-checked commit reduces unauthorized overwrite and stale release. Contradicted if monitored failure rates exceed the preregistered guardrail.
P4 • Net resource value	[PD] DGC reduces total variable burden enough to cover fixed deployment. Scale only if the lower 95% confidence endpoint clears break-even; an upper endpoint below it contradicts the investment claim.

[AI] The source is consistent with Panko’s distinction between superficial checks and logic inspection, and Barr et al.’s oracle problem. It does not empirically establish DGC’s propositions. Burnett et al. support assertion-based specification, but their cited system lacks temporal operators; workflow ordering here is a proposed extension.

[AI] Campbell’s measurement-reactivity argument predicts pressure to optimize visible scan counts; Wang–Strong explains why accuracy alone misses fitness for use. Tucker–Edmondson motivates feeding exceptions into corrected routines rather than celebrating repeated workarounds.

[PD] Inconclusive intervals neither confirm nor falsify the economic proposition. Report the remaining decision uncertainty. A lower burden with a breached unsupported-release guardrail fails the joint investment-and-safety contract.

Primary literature informs testable mechanisms; no prestige-based validation is claimed. DGC is an authored proposed framework whose incremental value remains to be tested.

TECHNICAL FOUNDATION · LIMITATIONS & REPRODUCTION

The evidence boundary determines what can be published

[AI] This case establishes a governance problem and a reproducible interpretation of the supplied register. It does not authenticate unavailable original files or estimate a historical treatment effect.

Limitation	Consequence for use
Source and chronology	[AI] Original formulas, cached states and approvals are unavailable; source descriptions disagree on the erroneous reference period. Preserve competing versions and recover lineage before operational sign-off.
Measurement	[AI] Actual labels mix observed, estimated and default states; source maturity and taxonomy differ. No homogeneous historical rate panel or true semantic-error denominator is supplied.
Identification / contribution	[AI] No counterfactual identifies repair, publication or customer effects. Personal role passes only provisional evidence tests; outcome attribution remains unestablished.
Economics / transport	[PD] Trial costs, variance, effect and audit priors are invented assumptions. Fixed benchmark prevalence is not production prevalence; resource release is not cash realization.
Confidentiality	[PD] Public historical quantities are indexed or relative; raw values and source locators remain in the private annex. Synthetic fixtures are explicitly identified.

[PD] Reproduce in order: verify the register hash; re-extract rows; rebuild source-version calculations; run ordered screens; execute the synthetic model; regenerate pages. Retain hashes, units, checks and independent reviews.

[PD] Before live use, obtain the authoritative current workbook, accepted metric dictionary, dependency graph, owner permissions and destination version. Repeat independent holdout tests and verify the exact published state. Successful reproduction of this case is not permission to modify a shared artifact.

[AI] The novel interpretation survives without an assumed historical gain: local correctness, current-state correctness and authority to release are distinct control requirements.

Consolidated limitations constrain every historical, causal and economic statement. Public calculations contain synthetic inputs and protected summaries; originals remain a required operational gate.

PROPOSED DECISION MODEL · RESOURCE ECONOMICS

A complete resource endpoint makes break-even explicit

[PD] All inputs below are invented planning assumptions. The benchmark has no authenticated resource-cost history and provides no estimate of production savings.

[PD] Let B be control-minus-DGC mean 14-day variable burden per eligible assignment, in common reference-cost index points. The clock starts at original scheduled availability, including late, withheld and unsubmitted assignments. Include ordinary review, new variable protocol effort, rework, unique consequences and prepriced hold/delay costs once; retain unresolved assignments through terminal categories and missing-data bounds.

Accounting contract	Planning assumption
Reference pool	[PD] Common independently frozen baseline=100 per eligible benchmark assignment.
Variable benefit B	[PD] Assumed 20 index points; protocol-induced variable work is already subtracted through the outcome.
Fixed allocation K	[PD] Deployment/training allocation 6 per assignment in the SAME pool/horizon; never a total setup bill.
Net / break-even	[PD] Net=B-K; break-even B=6. Realization and non-overlap factors equal 1 by the resource accounting contract; no cash claim.

Assumed B	K=4	K=6	K=9
6	2	0	-3
12	8	6	3
20	16	14	11

[PD] Grid entries are net resource-value index points. Assumed B20 gives net 14 at K6. Finance must price real released capacity and distinguish it from cash before deployment. If prices, avoidability or exclusivity change, reconstruct both arms and K consistently; do not discount net benefit while leaving new costs untouched.

Synthetic accounting model; no internal financial inputs. Full cost taxonomy, weights and all computations are retained in the reproducibility model.

PROPOSED DECISION MODEL · CLUSTERED BENCHMARK

Size the evaluation for the investment threshold

[PD] Randomize independent teams equally to DGC or current close practice; learning rules out crossover. Each team completes 16 preregistered held-out assignments. No availability is established.

[PD] Fix template-family holdouts and balance their difficulty across arms. Conditional on this benchmark, assume task SD 18 and exchangeable ICC 0.20: $\text{Var}(\text{team mean})=18^2[0.20+0.80/16]=81$. Independent Gaussian team means and equal variances give a pooled t test with $df=2n-2$ for n teams per arm.

Total teams	MDE at 80%	P(LCB>6)	Joint gate
24	10.771	95.3%	Fail
48	7.435	>99.9%	Fail
74	5.942	>99.9%	Pass
96	5.200	>99.9%	Pass

[PD] MDE uses two-sided alpha 0.05 against zero. The economic column is the probability that the lower endpoint of a two-sided 95% t interval exceeds 6 when the true synthetic benefit is 20. Require both $MDE \leq 6$ and economic power $\geq 80\%$. Values integrate the noncentral Student distribution; they are conditional planning results.

[PD] The smallest model-passing design is 74 teams (1184 assignments);96 teams is a buffered candidate, conditional on pilot and recruitment. Neither is a deployment recommendation. These finite powers are below 1. Teams supply independent replication.

GUARDRAILS

[PD] Unsupported release, false hold, owner override, completeness and delay require preregistered margins and multiplicity control. Primary burden power does not size these tests; a prespecified safety breach blocks release.

Proposed parallel cluster experiment on a fixed benchmark. Gaussian/equal-variance assumptions are unvalidated; neither historical precision nor production effectiveness is claimed.

PROPOSED DECISION MODEL · SENSITIVITY & TRANSPORT

A challenge-set result needs a production bridge

[PD] Seeded faults test robustness at a chosen mixture. They do not establish the frequency or cost of faults in routine monthly close.

Team-mean SD	K=4	K=6	K=9
6.0	74	34	18
9.0	162	74	34
12.0	286	128	58

[PD] Entries are the smallest even total-team counts satisfying both planning gates at assumed benefit 20. Task SD 12/18/24 gives team SD 6/9/12 under the same 16-task, ICC 0.20 contract. A higher fixed cost raises break-even: detecting that larger effect is easier, although proving benefit above it becomes harder. The joint gate checks both.

[PD] At 96 teams, ICC 0.40 produces MDE 6.88, failing the 6-point precision gate. Doubling tasks to 32 under ICC 0.20 gives MDE 4.93. Additional tasks cannot remove persistent team variation; correlations are hypothetical.

[PD] For transport, stratify faults by semantic family, owner conflict, source maturity and clean control. Estimate target prevalences π_c from an independent representative audit and conditional resource effects B_c . Target benefit is $\Sigma \pi_c B_c$ only with overlap, comparable workflows and stable class-specific prices. Recalculate uncertainty, power and K at target volume.

NO PREVALENCE SHORTCUT

[PD] Keep seeded and naturally observed cases labeled. Include clean cases to measure false holds. If target prevalence or unresolved truth is unknown, publish bounds and collect the missing audit; do not extrapolate the benchmark average.

All grid and covariance inputs are synthetic. Effects conditional on benchmark templates do not identify generalization to unseen dependency families or production prevalence.

PROPOSED DECISION MODEL · INFORMATION VALUE

Provenance evidence can change the owner decision

[PD] A separate hypothetical docket asks whether a preserved entry is correct. Prior probabilities are 0.70 correct and 0.30 wrong; this is not an observed error rate.

Authorized action	Correct entry	Wrong entry
Release preserved entry	0	-40
Owner-approved replacement	-16	-2
Hold for owner adjudication	-4	-6

[PD] Payoffs are net resource utilities, including each action’s review/delay/replacement cost. The prior-optimal action is an owner-adjudication hold, value−4.6. An independent provenance audit reports correct/wrong with assumed equal sensitivity and specificity. The audit does not itself confer write permission.

Audit accuracy	EVSI	Audit cost	Net EVSI
65%	0.000	0.250	-0.250
85%	0.850	0.500	0.350
95%	2.870	1.000	1.870

[PD] $EVSI = E[\max_a E(U_a | \text{signal})] - \max_a E(U_a)$. The low-information audit never changes the hold decision. At 85% accuracy, a correct signal favors release; at 95%, a wrong signal favors owner-approved replacement. Perfect information is worth 4.000, bounding every scenario.

[PD] Net EVSI subtracts only incremental information acquisition; replacement and hold costs are already in the payoff table. These values concern a separate decision docket and cannot be added to trial benefits describing the same prevented consequences. Validate priors, diagnostic likelihoods and loss weights before commissioning an audit.

Wholly hypothetical Bayesian decision model with two states and three authorized actions. No historical decision value, customer benefit or cash saving is estimated.

Read each validation result against the property it actually tests

[AI] The case examines an operating close through the evidence available in a project register. Registered repair and fill work remains distinct from original-artifact authentication, publication acceptance and business impact.

Claim class	Meaning in this case
[HR] Historical record	A represented statement, artifact state or decision in the selected register. Its underlying original files were not independently authenticated.
[RC] Recalculated	Arithmetic from registered quantities, protected comparisons or a declared synthetic example. Recalculation does not upgrade source reliability.
[AI] Analytical inference	An implication constrained by scope, versions, missing dependencies or unresolved truth. It is not identified historical causation.
[PD] Proposed design	The Dependency-Gated Close, future measurement or evaluation, and synthetic planning assumptions. These are not measured project results.

[AI] The main contribution is a stricter release decision. The recorded stage comparison removes 78.26% of flags after data writing, while the preserved-entry comparison shows numerical difference without establishing validity. These results change what the checks can authorize.

[AI] Keep syntax, semantics and release readiness separate. A formula can calculate while referencing the wrong cohort; an owner entry can be preserved while its dependent claim is withheld; a local delivery can be complete while the shared master remains unreconciled.

[PD] Reproduction chain: frozen register → cell-linked ledger → version and stage calculations → protected exhibits → separately labeled proposed model → independent arithmetic and artifact review. Preserve unknown mappings and unresolved chronology rather than assigning a favorable interpretation.

DISTRIBUTION

[PD] Share the anonymized case PDF. Keep the separate evidence annex, identity mapping, raw values, internal source locations and calculation package private.

S1: supplied project register. Public historical measures are indexed or relative. Synthetic illustrations and prospective models are labeled; neither references nor internal review certify historical impact.

REFERENCES / PRIMARY METHODS AND SUBSTANTIVE LITERATURE

Test the property, the dependency and the oracle

These sources support the analytical interpretation and proposed controls. They do not authenticate the source workbook or establish a historical operational benefit.

Microsoft. SUM function. Official function documentation, checked September 2026. Range-based SUM ignores text. A text-stored value can be omitted without an error; historical behavior still requires the actual formula and data type.

Microsoft. AVERAGE function. Official function documentation, checked September 2026. Range text and empty cells are ignored, while numeric zeros are included. Identical display strings can therefore imply different averaging populations.

Burnett, M., Cook, C., Pendse, O., Rothermel, G., & Summet, J. (2002). End-User Software Engineering with Assertions. Oregon State University TR02-60-05. Author full text. Supports explicit cell specifications and dependency assertions; its mechanism does not express workflow sequence.

Panko, R. R. (2006). Recommended Practices for Spreadsheet Testing. EuSpRIG, 73–84; author manuscript posted 2007. Supports logic inspection beyond visible error values. Its general testing recommendations do not measure this case's repair effectiveness.

Barr, E. T., Harman, M., McMin, P., Shahbaz, M., & Yoo, S. (2015). The Oracle Problem in Software Testing: A Survey. IEEE Transactions on Software Engineering, 41, 507–525. Author repository full text. Agreement with a reference is only as valid as the oracle and specification.

Hermans, F., Pinzger, M., & van Deursen, A. (2011). Supporting Professional Spreadsheet Users by Generating Leveled Dataflow Diagrams. ICSE, 451–460. Author full text. Dependency visibility supports comprehension; a populated display row need not feed the decision output.

Roy, S., Hermans, F., & van Deursen, A. (2017). Spreadsheet Testing in Practice. SANER, 338–348. Author institutional record and manuscript. Describes testing and maintenance practices; this case's release contract is a separate proposed operating design.

Microsoft. Application.CalculateFullRebuild method. Official VBA documentation. Recalculates open workbooks and rebuilds their formula dependencies. Rebuilding calculation order does not identify the intended metric, validate source truth or authorize publication.

Original research, author manuscripts and official technical guidance. Applicability and access limits are retained in the private review record. No external study supplies a project effect or a verified historical cost saving.

REFERENCES / PRIMARY METHODS AND SUBSTANTIVE LITERATURE

Keep decision measures aligned with their operating meaning

These sources support the analytical interpretation and proposed controls. They do not authenticate the source workbook or establish a historical operational benefit.

Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33. Author-hosted text. Data quality is contextual; local syntax is insufficient for a complete management decision.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37–46. Original article. Use chance-corrected agreement alongside prevalence, raw disagreement and independent truth assessment; agreement does not authenticate owner entries.

Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Original occasional-paper record, ERIC. A rewarded indicator can diverge from the intended outcome. Treating fewer alerts or faster publication as success creates that design risk; gaming is not established here.

Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance Implications of Strategic Performance Measurement in Financial Services Firms. *Accounting, Organizations and Society*. Author institutional summary. Operational measures need demonstrated links to strategy; no such project-specific impact link is identified here.

Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures. *California Management Review*. Author working paper. Immediate repair and systematic organizational learning differ. A recurring ownership or interface failure requires more than another local patch.

NIST/SEMATECH. e-Handbook of Statistical Methods: Individuals Control Charts. Official formula guidance. Tentative limits describe ordered arithmetic; mixed maturities, revised definitions and deterministic plans do not establish a stable historical operating process.

The W. Edwards Deming Institute. The Funnel Experiment. Institutional explanation of tampering with routine variation. Supports validating an action trigger before responding; this case does not identify a stable source process or a treatment-induced signal.

OECD / European Commission JRC (2008). Handbook on Constructing Composite Indicators: Methodology and User Guide. Institutional methods guidance. Any decision-burden aggregate needs explicit components, weights and sensitivity. Preserve unweighted release and delay guardrails.

Original research, author manuscripts and official technical guidance. Applicability and access limits are retained in the private review record. No external study supplies a project effect or a verified historical cost saving.

REFERENCES / PRIMARY METHODS AND SUBSTANTIVE LITERATURE

Govern the release and evaluate the policy

These sources support the analytical interpretation and proposed controls. They do not authenticate the source workbook or establish a historical operational benefit.

W3C (2013). PROV Model Primer. Official provenance guidance. Versioned entities, activities and responsible agents support traceable handoffs. Provenance alone proves neither numerical correctness nor authorized shared-master acceptance.

Fielding, R., Nottingham, M., & Reschke, J., eds. (2022). RFC9110: HTTP Semantics, §13.1.1, If-Match. Official conditional-update standard. The version-match principle informs proposed release controls; this case does not claim the historical workbook implemented HTTP concurrency protection.

Hernán, M. A., & Robins, J. M. Causal Inference: What If. Author-hosted book and current materials. Treatment, eligibility, assignment and follow-up must precede protocol-induced holding, completion or publication decisions.

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44, 823–830. Tests discriminate specific contribution hypotheses. Registry association and repeated memory lineage provide limited support, not independently authenticated authorship or adoption.

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. Selection on extreme prior burden can produce apparent recovery without policy effect.

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*. Values information through changed decisions under stated priors and likelihoods; no positive historical program effect is required.

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). CONSORT 2010 Statement: Extension to Cluster Randomised Trials. *BMJ*, 345, e5661. Original reporting guidance. Specify cluster assignment, recruitment, attrition, intracluster correlation and the analysis level; the benchmark's power is separately derived.

Original research, author manuscripts and official technical guidance. Applicability and access limits are retained in the private review record. No external study supplies a project effect or a verified historical cost saving.

FINAL REFEREE REPORT / INTERNAL, AGENT-ASSISTED REVIEW

The case turns local validation into an accountable release decision

[AI] Author-led review and independent parallel agents challenge the evidence, methods and executive decision. This is internal review, not external academic or client certification.

Major issue	Resolution or remaining gate
Meaning of validation	Stage-dependent flag removal is not a false-positive estimate. Scoped zero-error scans do not certify semantic or end-to-end correctness. Official function documentation replaces an overbroad text-zero explanation.
Oracle and owner truth	Historical-oracle versus current-resolver chronology remains unresolved. Numerical difference does not prove a current owner entry. Suspected corrections remain scenarios pending authorization.
Adoption and contribution	Local completion and later shared-master state are reconciled as competing versions. Tab-level direction and repair work receive provisional registry-level hoop support; publication acceptance and outcomes remain unestablished.
Prospective evaluation	Independent covariance, Student-t, exact-information and simulation checks pass 105 tests. Benchmark effect and cost share one horizon; guardrail sizing and production transport remain prerequisites.

Partner critique: would the COO act on page 1?

[AI] Yes: assign a close owner, require a dependency manifest and hold the affected management claim until its material exceptions are resolved. The 78.26% stage comparison and the gap between numerical difference and verified entries make the approval change concrete.

Positive contribution within the evidence boundary

[AI] The register describes directed formula repair, explicit aggregation rules, owner-entry preservation and narrative preparation. The case retains those contributions while testing their limits; it does not equate sophisticated tooling or repeated checks with established business impact.

REVIEW DISPOSITION

[AI] Numerical, method and artifact review support this anonymized operating-governance case and its proposed benchmark model. Historical customer effect, end-to-end close-time saving, production effectiveness and realized financial value remain unestablished.

Internal author-led and independent agent review. Numerical reproduction, issue closure, source-boundary checks and page-by-page artifact reviews are retained in the private package.