

Model validation memo — Transaction Monitoring Scenario Suite v6

AML — transaction monitoring scenarios

Transaction monitoring scenario validation

MODEL

Transaction Monitoring Scenario Suite v6

ENGAGEMENT

Sample — transaction monitoring annual
validation

PREPARED

2026-10-07

PREPARED BY

AlphaIQ CertifyIQ, computed by the product

RECOMMENDATION



Inconclusive — the evidence is insufficient

9 sections · 5 traced figures · 0 finding(s) · 7 test(s) named as not performed

SYNTHETIC SAMPLE

Every figure in this report was computed by CertifyIQ from a synthetic book — 6,580 alerts across five scenarios, with a reviewed below-the-line sample. No real institution, model or customer is described. The book is `aml-alert-log.csv`; how it was generated, and the defects planted in it, are documented in `aml-alert-log-dgp.json`.

AT A GLANCE · SYNTHETIC SAMPLE

Transaction Monitoring Scenario Suite v6

AML — transaction monitoring scenarios · Transaction monitoring scenario validation

Inconclusive — the evidence is insufficient

No row in this pack was judged against a threshold: all 5 were reported without one, and each is listed as such in the results table. No defect was demonstrated and none was ruled out, so the honest conclusion is that the evidence is insufficient rather than that the model passed. Declaring the institution's own thresholds against this engagement is what turns these rows into judgements.

9

SECTIONS

5TRACED
FIGURES**0**

MATERIAL

0

NOTABLE

0

OBSERVATIONS

7NOT
PERFORMED

Supervisory frame: this memo is built to "AML Model Validation (FFIEC)", published under US · FFIEC / FinCEN. No supervisory frame is declared for this engagement, so the frame above was chosen by the model's family rather than by its supervisor. That is not a statement that the two agree: nobody has recorded who supervises this model, and a reader in another jurisdiction should treat the section structure, the citations and the recommendation vocabulary as US's.

Evidence manifest: none sealed. The documents behind each figure are named where the figure is; the set of documents on record when this memo was written has not been attested by anyone, so it cannot be cited here.

How to read this sample. Every number below was computed by CertifyIQ from a synthetic book (aml-alert-log.csv), through the same resolver, metric pack and report builder a signed-in engagement uses. Each figure carries its own trace — where it came from and what it was judged against. Tests that could not be performed are named with the reason, never left out. Where no threshold was declared, a metric is reported but not judged, because this report does not invent an institution's tolerances.

CONTENTS

1 Purpose and scope**2** Scenario inventory and risk-based coverage**3** Alert volume and distribution**4** Threshold tuning: above- and below-the-line testing**5** Alert disposition and SAR conversion**6** Segmentation stability**7** Findings**8** Third-party dependencies and residual risk**9** Not performed

1 Purpose and scope

This memo validates Transaction Monitoring Scenario Suite v6 against the transaction-monitoring alert log extracted on 2026-08-31: 6,580 row(s) read, 6,580 alert(s) kept across 5 declared scenario(s), of which 4,823 were reviewed to a disposition and 1,757 were raised or sampled and never reviewed. Every rate in this memo is computed over the 4,823 reviewed alert(s); the 1,757 unreviewed alert(s) are counted where volume is counted and judged by nothing, because an alert nobody dispositioned carries no outcome to judge.

The log was declared, not defaulted: 5 scenario threshold(s) with their band widths, each cited to the institution's own tuning document; SAR labels "sar-filed" and cleared labels "cleared", "closed-no-action" in the institution's own vocabulary; and a segmentation reference window ending 2026-05-04, cited to Model risk policy s. 8.2 — the segmentation reference window is the H1 book. Any figure whose input was not declared appears as a named not-computable row in section 8 rather than as a default this platform chose on the institution's behalf.

This validation carries no replica of the monitoring system under test and reads no ongoing-monitoring series: the figures below are measured directly from the declared alert log against the institution's own declared claims, so the recommendation is not gated on a replication step this family's evidence chain has no equivalent of. It does not validate sanctions screening, KYC risk rating or the data feeding the monitoring system, and it does not grade scenario coverage against the institution's risk assessment — each is named in section 8 with the reason, and the documentation obligations among them are counted in the engagement's completeness checklist, not here.

2 Scenario inventory and risk-based coverage

The institution declared 5 scenario(s). For each, the threshold production alerts at or above, the width of the band sampled either side of it, and the tuning document the pair is cited to — as declared, unaltered:

cross-border-high-risk: threshold 15,000, band width 3,000, cited to AML scenario inventory v6, s. 3.4 — cross-border, high-risk jurisdiction. 399 alert(s) in the log.

dormant-account-reactivation: threshold 20,000, band width 4,000, cited to AML scenario inventory v6, s. 3.5 — reactivation of a dormant account. 458 alert(s) in the log.

rapid-movement: threshold 50,000, band width 10,000, cited to AML scenario inventory v6, s. 3.2 — rapid movement of funds. 568 alert(s) in the log.

structuring-cash: threshold 10,000, band width 2,000, cited to AML scenario inventory v6, s. 3.1 — cash structuring, EUR 10,000 aggregate. 1,983 alert(s) in the log.

wire-round-amount: threshold 25,000, band width 5,000, cited to AML scenario inventory v6, s. 3.3 — round-amount wire activity. 3,172 alert(s) in the log.

Whether this inventory covers the institution's businesses, products, services and customers as its BSA/AML risk assessment describes them is not graded here. It is the risk-based-coverage obligation the FFIEC BSA/AML Examination Manual places on the institution, recorded in the engagement's completeness checklist against the institution's own rationale document; an alert log shows which scenarios ran, not which risks exist. The inventory rationale for each threshold is likewise the institution's document to supply, and this section states the citations it declared rather than judging them.

3 Alert volume and distribution

6,580 alert(s) across 5 scenario(s); 4,823 reviewed, 1,757 not. Per scenario (all alerts / reviewed / SARs): cross-border-high-risk 399 / 362 / 52; dormant-account-reactivation 458 / 353 / 0; rapid-movement 568 / 547 / 193; structuring-cash 1,983 / 1,695 / 51; wire-round-amount 3,172 / 1,866 / 10.

Concentration is reported as a Herfindahl index of alert shares across the declared scenarios, in [1/5, 1]: 1/5 would be volume spread evenly, 1 would be a single scenario raising everything. No authority publishes a tolerance and the institution declared none, so the figure is reported and not judged.

Alert volume concentration across scenarios (Herfindahl index)

MEASURED

0.34

Alert volume concentration across scenarios (Herfindahl index): 0.34 measured on 6,580 alerts carrying 306 filed SARs. Benchmark source: 3 NYCRR §504.3(a)(1)–(3) — scenarios matched to the institution's risk assessment, which is what an examiner reads a concentrated alert distribution against. All alerts counted, reviewed or not. No instrument bounds the index and none is shipped.

4 Threshold tuning: above- and below-the-line testing

Above- and below-the-line testing looks at the alerts in the band immediately either side of each declared threshold: above it, whether the alerts the threshold admits are productive; below it, whether reportable activity sits just outside what the threshold catches. Per scenario, read off the declared threshold and band width:

cross-border-high-risk (line at 15,000, band 3,000): above the line [15,000, 18,000] — 21 row(s), 19 reviewed, 7 SAR(s); below the line [12,000, 15,000] — 67 row(s), 67 reviewed, 7 SAR(s).

dormant-account-reactivation (line at 20,000, band 4,000): above the line [20,000, 24,000] — 93 row(s), 75 reviewed, 0 SAR(s); below the line [16,000, 20,000] — 40 row(s), 40 reviewed, 0 SAR(s).

rapid-movement (line at 50,000, band 10,000): above the line [50,000, 60,000] — 80 row(s), 78 reviewed, 27 SAR(s); below the line [40,000, 50,000] — 39 row(s), 39 reviewed, 1 SAR(s).

structuring-cash (line at 10,000, band 2,000): above the line [10,000, 12,000] — 1,555 row(s), 1,316 reviewed, 33 SAR(s); below the line [8,000, 10,000] — 70 row(s), 70 reviewed, 0 SAR(s).

wire-round-amount (line at 25,000, band 5,000): above the line [25,000, 30,000] — 378 row(s), 217 reviewed, 2 SAR(s); below the line [20,000, 25,000] — 52 row(s), 52 reviewed, 0 SAR(s).

The pooled below-the-line SAR rate above is SARs over reviewed rows across every below-the-line band. A non-zero rate is reportable activity the current thresholds exclude — an observation for the institution's tuning decision, which is recorded in the document each threshold is cited to, not made here. Per-scenario 95% intervals on each band's conversion are produced by the engine's ATL/BTL reading when the institution declares its bootstrap draw count and its source; this run carries no such declaration, so no interval is printed and none is implied.

Scenario tuning effectiveness (share of declared scenarios producing a SAR)

MEASURED

80.00%

Scenario tuning effectiveness (share of declared scenarios producing a SAR): 80.00% measured on 4,823 alerts carrying 306 filed SARs. Benchmark source: FFIEC BSA/AML Examination Manual, suspicious activity reporting — a monitoring system's filtering criteria and thresholds; 3 NYCRR §504.3(a)(8), ongoing analysis of the relevance of scenarios and thresholds. No authority publishes a floor for the productive share: judged against the institution's declared appetite, or reported unjudged.

Below-the-line SAR rate (reviewed alerts just under the declared thresholds)

MEASURED

2.99%

Below-the-line SAR rate (reviewed alerts just under the declared thresholds): 2.99% measured on 4,823 alerts carrying 306 filed SARs. Benchmark source: 3 NYCRR §504.3(a)(5), end-to-end testing; FFIEC BSA/AML Examination Manual, threshold validation. Pooled over the reviewed rows in each declared below-the-line band (this platform); per-scenario 95% intervals stay in the engine's ATL/BTL report, which needs declared draws. No authority publishes a tolerance.

5 Alert disposition and SAR conversion

Of 4,823 reviewed alert(s), 306 were dispositioned as a SAR filing and 4,517 were reviewed and cleared, in the institution's own labels as declared. 1,757 alert(s) carry no disposition and are outside this section.

Conversion is reported against no threshold: the institution declared no conversion appetite for this run, and no authority publishes one. The figure is stated so a reviewer can read it against the institution's own history; it is not graded.

The protocol under which investigators reached these dispositions — who reviews, what escalates, what results in a filing — is the alert-disposition obligation the institution establishes and the completeness checklist records; this memo reports the dispositions the log carries and does not audit how they were reached.

SAR conversion efficiency (SARs filed over alerts reviewed)

MEASURED 6.34%

SAR conversion efficiency (SARs filed over alerts reviewed): 6.34% measured on 4,823 alerts carrying 306 filed SARs.

Benchmark source: Definitional — precision of reviewed alerts against the institution's own dispositions, as this platform computes it; alerts never reviewed are excluded from the denominator and counted in the census. "Peer band 2–6%" is not shipped: a conversion band is one institution's risk appetite.

6 Segmentation stability

Alert volume by declared customer segment (4 segment(s) seen) is compared between the reference window ending 2026-05-04, cited to Model risk policy s. 8.2 — the segmentation reference window is the H1 book, and the period after it, as a population stability index over segment shares. Alerts are counted whether or not reviewed — volume is what moves when a segment's behaviour or the system's coverage of it changes.

The index is reported against no threshold: the institution declared no stability tolerance for this run. Credit's conventional 0.10 / 0.25 bands are NOT applied here — they were published for score distributions, not for segment shares, and importing them would be a shipped default this pack refuses to carry.

Segmentation stability (PSI of alert volume across declared segments)

MEASURED 0.29

Segmentation stability (PSI of alert volume across declared segments): 0.29 measured on 6,580 alerts carrying 306 filed SARs.

Benchmark source: 3 NYCRR §504.3(a)(1)–(3) — scenarios matched to the institution's customer segmentation. Computed only over a DECLARED segment column, alert-date column and cited reference window; refuses by name when any is absent, because inventing either would decide the segmentation this metric exists to review. No authority publishes a tolerance, and credit's PSI bands are not applied to segment shares.

7 Findings

No findings were raised. Every metric judged against a declared claim returned within it; rows reported without a claim were not judged and raise nothing.

Against the findings register: 0 finding(s) open.

8 Third-party dependencies and residual risk

No third-party dependency is declared for this engagement. That is a statement by the institution that the model embeds no vendor or open-source component it cannot validate inside — not an absence of the question.

9 Not performed

The following checks were not performed, each with the reason. A validation report's most dangerous property is the test it quietly omitted, so this register is part of the memo rather than an appendix to it. Entries that name a documentation obligation are established by the institution and recorded in the engagement's completeness checklist; they are listed here so the reader of this memo knows they exist and knows this memo did not establish them.

NOT PERFORMED — AND WHY

The detection scenario inventory documented with the rationale for each threshold and parameter

3 NYCRR §504.3(a)(4) and (a)(6) require the scenarios, thresholds, parameters and assumptions to be documented with their rationale. This pack CONSUMES the inventory — every threshold and band arrives declared and cited to the institution's tuning document, and an uncited one is refused — but whether the rationale in that document is adequate is a reading of the document, not an arithmetic over the log.

Ongoing analysis of the continued relevance of scenarios, rules, thresholds and assumptions

3 NYCRR §504.3(a)(8) and the Wolfsberg Group's Statement on Effective Monitoring for Suspicious Activity (2024) require a standing process. This pack's tuning-effectiveness and below-the-line rows are evidence such a review would use over ONE period; that the review happens, at what interval, and what it changed is governance, evidenced by its minutes rather than by a figure.

The documented protocol for investigating alerts and deciding which result in a filing

3 NYCRR §504.3(a)(7) and 31 CFR §1020.320 require the protocol and the filing. The dispositions this pack reads are that protocol's OUTPUT, in the institution's declared vocabulary; the protocol itself — who decides, how it is documented, how a decision not to file is recorded — is a document, and the two-stage case funnel it implies is not carried by this log.

Identification of every data source feeding the monitoring system and validation of the integrity, accuracy and quality of that data

3 NYCRR §504.3(c)(1)–(3). An alert log evidences what the monitoring system emitted, not what it was fed: a transaction that never reached the system raises no alert and leaves no trace here. Source-to-system reconciliation is a data-lineage exercise over the institution's feeds, which the ETL reconciliation performs for a file this platform loads and nothing performs for the feeds it does not.

Scenarios based on the institution's risk assessment and matched to its businesses, products, services and customers

3 NYCRR §504.3(a)(1)–(3) and 31 CFR §1020.210. Whether the declared inventory covers the risks the institution's own assessment names is a mapping of one document onto another — the typology-coverage judgement named as a reading and not an arithmetic — and this pack does not perform it.

Independent validation of the monitoring system under model risk management — conceptual soundness, ongoing monitoring and outcomes analysis

The Interagency Statement on Model Risk Management for Bank Systems Supporting BSA/AML Compliance (Board of Governors, FDIC, NCUA, OCC and FinCEN; 9 April 2021) applies the three MRM elements to these systems. This pack is the outcomes-analysis element over one log; conceptual soundness and ongoing monitoring are the standard checks every family carries and the governance around them, not rows of this pack.

Sanctions and watchlist screening effectiveness

A different control with a different evidence shape — name matching against a list, not a threshold against a value. Nothing in an alert log from a transaction monitoring system evidences it, and a pack that appeared to cover it would leave the gap unexamined. corrected the deck line that sold it.



AlphaIQ CertifyIQ — independent validation of models and the decisions they drive. This sample was produced by the product itself; nothing in it was written by hand.

Memo content hash 44c99e7131e25e13ee858759767dc42e8569bf3ccd4e39d5b26288a9097a791f

This document is a recommendation, not a certification. A certificate is issued only on attested evidence, which a synthetic sample does not have.