

QA Sampling for AI-Assisted Review (2026): A Defensible Approach + Tool Shortlist

<https://counterbench.ai/guides/qa-sampling-ai-assisted-review-2026> · Last updated: 2026-03-08

QUICK ANSWER

A defensible QA sampling plan is bucketed (by output type), randomized, tied to a simple per-batch rule, and logged with an error taxonomy plus stop/adjust thresholds—otherwise you can't detect systemic failure modes or explain what you verified.

BENCH-TESTED CHECKLIST

- Step 1: define the AI task in one sentence (what it does and doesn't do).
- Step 2: define error types that matter (critical/material/minor).
- Step 3: sample by bucket (responsive/non-responsive/privilege/hot/low priority).
- Step 4: use a repeatable rule (fixed count or percent per batch for high-risk buckets).
- Step 5: randomize the sample (avoid cherry-picking).
- Step 6: review with a short QA checklist (text matches, cite-backs, privilege indicators).
- Step 7: apply stop/adjust/proceed thresholds and document changes.
- Step 8: log everything (batch, bucket, sample size, errors, decision, changes).

Get templates + the full workflow: <https://counterbench.ai/guides/qa-sampling-ai-assisted-review-2026>