JUNCTURE RESEARCH · WHITEPAPER

# Reference Quality in Pharma Claims Libraries

A five-check framework for the citations nobody verifies.

JULY 2026

JUNCTURE.HEALTH

# Why references rot

A claims library rests on a quiet assumption: that the link between a claim and its reference, made once at authoring, stays true for the life of the claim. Everything else in the library is governed; the claim is versioned, workflowed, and scheduled for expiry. The reference is an attachment, and attachments do not age visibly. A citation recorded in 2023 looks exactly the same today, whether or not anything it asserts is still true.

The document it points at is designed to change. The European Commission's SmPC guideline states that the SmPC is the basis of information for healthcare professionals on how to use the medicinal product safely and effectively, and directs that its safety information be regularly reviewed and, if necessary, updated. The EMA's own definition carries the same phrase. A document that is the basis of information is revised whenever the information changes, and it changes often.

The volume is documented. In the US, the FDA publishes approved safety-related labeling changes in its searchable Drug Safety-related Labeling Changes database, with coverage since January 2016 across the safety sections of the Prescribing Information. A published review of FDA labeling practice puts the pace at 400 to 500 label changes every year. In the UK, the electronic medicines compendium hosts more than 14,000 regulator-checked documents and expects companies to update it within 10 days of the UK or European authority approving a change. The public registries move on a rhythm of days. The citations in a claims library were recorded once.

Version drift is only the most mechanical form of rot. Three other failure modes are endemic, and none requires anything to change after the fact; they are weak on the day they are written.

**The reference without a section pointer.** The citation reads "SmPC" or "Prescribing Information," full stop: a thirty-page document and no location within it. It is unverifiable at the speed of review; the reviewer either trusts it or re-reads the whole label, and under deadline, trust wins. A reference that cannot be checked in under a minute will, in practice, never be checked.

**The restated number.** The claim says "one in four patients." The source says 22.7%. Somewhere between label and library a value was rounded, converted, or paraphrased into something more quotable. Each restatement looks harmless; collectively they detach the library's numbers from the label's, one decimal at a time, and always in the direction that reads better.

**The publication where the label should be.** The claim is fully derivable from the label, but the citation points at a primary publication because the author had the paper at hand. The publication may state the figure differently, frame the population differently, or predate a label revision. For anything the label can support, the label is the controlling source.

The literature on citation itself says as much. A 2025 systematic review and meta-analysis covering 46 studies and 32,074 quotations found that **16.9% of quotations in medical journal articles fail to accurately reflect the source they cite, roughly half of them major errors**, and that the rate has **not meaningfully improved since the first meta-analysis a decade earlier**. That is peer-reviewed medicine, written slowly, under named authorship. A claims library assembled under launch pressure has no claim to a better base rate, and nobody is spot-checking it.

The conclusion is not that teams should try harder. It is that reference quality is auditable: linkage, substantiation, numbers, currency, and source type are all questions with checkable answers against publicly retrievable documents. What follows is the check for each.

## The five checks

Each check verifies one property, fails in a characteristic way, and produces evidence. The evidence is the point: a bare "pass" is a smaller trust problem in a new location, while a located passage, a value comparison, or a registry version is something a reviewer confirms in seconds and an auditor relies on later.

## Check 1: A reference is present

The floor. Every claim that asserts something checkable carries at least one reference, and the reference resolves: the document exists, is retrievable, and is the document the citation names.

How it fails in practice: migrations strip attachments; claims born as "common knowledge" about the brand never had a reference at all; citations point at file paths that no longer resolve. Every mature library contains claims whose reference is a rumour.

Evidence produced: for every claim, a resolved reference identity, document, edition or date, and retrieval location, or an explicit absent verdict. A claim with no reference is a finding, not a gap in the data.

## Check 2: The cited section substantiates the claim

Presence is not support. This check reads the cited section and locates the passage that substantiates the claim. Not the document level, the section level: the passage that carries the claim should actually say what the claim says.

How it fails in practice: the right document, the wrong section, a pharmacodynamics section cited for a tolerability claim. Or the section pointer is missing entirely, which makes the check impossible to run and is recorded as its own finding. Or the section discusses the topic but says something narrower than the claim: the claim generalises, the source qualifies.

Evidence produced: the substantiating passage, quoted, with its location, or a finding that no substantiating passage exists in the cited section. Found, the claim becomes verifiable by anyone in under a minute; not found, the reviewer knows exactly what is broken.

## Check 3: The numbers match

Every value in the claim, percentages, absolute figures, durations, dose strengths, appears in the cited section and matches. Exact matches pass. Rounded, converted, or derived values are flagged: not necessarily wrong, but no longer the label's number, and someone accountable should decide whether "one in four" is an acceptable rendering of 22.7% in a promotional context.

How it fails in practice: rounding that flatters. Percentages derived from event counts the source never presents as a percentage. A composite-endpoint number attached to a claim about a single component. A value that appears nowhere in the cited section because it came from a congress slide that never entered the label.

Evidence produced: a value-by-value comparison, each value marked exact, rounded, derived, or absent. Rounding is a finding with a severity; absence is a failure. The table is the audit artifact: it shows the numbers were checked, which is precisely what nobody can currently show.

## Check 4: The label version is current

The citation points at the current published version of the label. This is checkable because currency is public: the FDA's SrLC database publishes approved safety labeling changes searchable by drug and date, the EMA publishes current product information for centrally authorised medicines, and the emc carries the current UK documents with a 10-day expectation on updates. A currency check is simply a comparison between the version the citation records and the version the registry shows.

How it fails in practice: the citation was recorded against the label as it stood at claim creation, and the label has since been revised, sometimes in the very section the claim cites. At 400 to 500 safety-related label changes a year in the US alone, superseded citations accumulate as arithmetic, not negligence. The quieter failure is the citation that records no version at all: no date, no revision, just "the SmPC." It can never be verified as current by anyone, so **a citation that records no version is itself a finding**, prior to any question of drift.

Evidence produced: the version or date the citation records, the current version in the relevant registry, and a verdict: current, superseded, or version not recorded.

## Check 5: The evidence hierarchy is respected

For any claim the label can support, the label is the source that should be cited. The SmPC and the prescribing information are the regulator-agreed basis of information for the product; a primary publication is an input to that document, not a substitute for it. This check classifies each claim as label-derivable or genuinely beyond-label, and verifies that label-derivable claims cite the label.

How it fails in practice: a publication cited for a figure that sits in the label, because the paper was on the author's desk. A congress abstract substantiating a claim the label states more carefully. A publication that frames the number for a study population where the label frames it for the approved population, and the claim quietly inherits the wider frame.

Evidence produced: the classification, the verdict, and for each publication-cited-where-label-should-be finding, the label section that should replace it. The hierarchy is a preference order, not a ban: publications remain the legitimate source for what the label does not carry, and the check records that explicitly rather than penalising it.

# Scoring and rollup

Five checks per reference produce more signal than a review meeting can consume, so the results compress into a per-reference verdict with three bands.

**Solid.** Reference present and resolving, substantiating passage located, values exact, version current, hierarchy respected. Nothing for a human to do; the citation would survive an audit today, and the evidence pack proves it.

**Review.** The reference substantiates the claim but something needs an accountable eye: a rounded value, a version one revision behind in an untouched section, a

hierarchy question, missing version metadata on an otherwise sound citation. Judgment calls; the band exists so humans spend their time exactly here.

**Fix.** The reference is absent, does not resolve, points at a section that does not substantiate the claim, contains numbers the source does not, or cites a superseded version in a revised section. No judgment required to know something is wrong; judgment is required to decide what replaces it.

Band assignment should be conservative: anything ambiguous falls down a band, not up. A system that flatters its own library converts an unexamined assumption into a certified one.

Aggregated across the library, the bands become a **reference-health profile**: the share of claims in each band, cut by brand, market, claim type, and age. It does two things a spot-check audit cannot. It turns "we should look at our citations" into a ranked worklist, fix-band claims in the highest-traffic assets first, then review-band items in order of exposure. And it trends: a rising solid share is demonstrable governance, where an annually audited library spends eleven months of the year unverified at full speed.

One caution: the rollup is triage, not compliance. A high solid share is not a certification, just a management instrument that shows where the weak citations are and which way the trend runs.

## A worked example: Varigel

Varigel is a fictional brand with a single approved indication; any resemblance to a real product is unintended. Three claims from its library, through the checks.

**Claim A: "One in four patients achieved clear or almost clear skin by week 12."** The reference is present and points at the efficacy section of the current label. Check 2 passes: the substantiating passage is located, and the section reports exactly this endpoint. Check 3 fails: the label reports 22.7%, and "one in four" is 25%, a restatement that rounds 2.3 points in the brand's favour. Verdict: review, with the

value comparison attached. The accountable reviewer decides whether the rounding stands; either way, the decision is now recorded against the evidence.

**Claim B: "Varigel is well tolerated in long-term use."** The reference is present and cites section 5.1 of the SmPC, pharmacodynamic properties. Check 2 fails: the cited section describes mechanism of action and contains no substantiating passage on tolerability at any duration. A plausible history: the citation was copied from an adjacent claim, and nobody has read section 5.1 since. Verdict: fix. The check notes that the safety experience section is where a tolerability claim would have to be substantiated, and leaves whether a reworded claim is supportable there to a human.

**Claim C: a health-economics statement citing a peer-reviewed adherence study.** Check 5 passes: the claim is genuinely beyond-label, the label carries no adherence data, so a publication is the right kind of source. Check 4 records **"currency not applicable"**: a journal article is fixed at publication and has no current version to drift from. A system that reported a version-currency verdict on a journal article would be inventing a check it cannot run; recording non-applicability honestly is evidence the framework knows what it is checking. Verdict: solid, with the annotation.

Three claims: one fix, one review, one solid. Nobody re-read the label end to end, and the reviewer's attention landed on exactly the two claims that needed it, each arriving with its evidence.

## The operating model

A one-time audit produces a snapshot that starts rotting the day it is delivered. The framework holds its value only run continuously, on three triggers.

**On claim creation.** The five checks run before the claim enters review, so a weak citation never enters the library at all. This is also where **suggest-the-reference** belongs: given the draft claim text, the system proposes candidate sections from the indexed label and reference set, the author picks, and the check validates the pick. The no-pointer citation dies here, at authorship, the only cheap place to kill it.

**On label revision events.** The registries that make check 4 possible also make it event-driven: a watch on the FDA's SrLC feed, EMA product information, and the emc turns a label revision into a trigger. Every claim citing the revised label is re-checked, changed sections first, and the results re-rank the worklist. This converts "the label changed" from an anxiety into a finite, ordered list of claims.

**On a cadence.** A standing sweep, monthly or quarterly, backstops what the event feeds miss: links that stopped resolving, documents moved in a migration, a publication retracted or corrected. The cadence is the safety net, not the primary mechanism; if it is doing most of the catching, the event wiring is broken.

Two operational properties decide whether any of this survives contact with a real content team. The first is **self-service reference ingest**: adding a source must be as simple as dropping a PDF or pasting a URL, with the system extracting the text, indexing the sections, and recording version metadata at ingest. If adding a reference requires a services ticket, the indexed library will trail the real one permanently, and every check downstream inherits the gap. The second is governance: the checks produce findings, never decisions. What happens to a flagged claim, reword, re-reference, retire, is a human call by a named, accountable reviewer. Where records and signatures are regulated, the trail behind those calls should be Part 11-supporting, and the sign-off stays with the reviewer, not the system.

## Implementation checklist

Vendor-neutral, in rollout order. Most items cost process, not software.

- **Inventory the library's citations as citations.** Count claims with no reference, references with no section pointer, and references with no version metadata. This baseline is usually the argument for everything that follows.
- **Require a section pointer on every new citation, starting today.** Zero cost, immediate effect: every citation created from now on is verifiable in under a minute.

- **Record version metadata at citation time.** Document date or revision, and the registry it came from. Treat "no version recorded" as a finding, not a formatting preference.
- **Index the labels.** Ingest the current label for each brand and market into a searchable, versioned store, self-service, so ingest never becomes the bottleneck.
- **Run the five checks on one brand first.** Calibrate the verdict bands on a library someone knows well, and let the false-positive argument happen on familiar ground.
- **Work the worklist by band and traffic.** Fix-band claims in the most-used assets first; review-band items in order of exposure. Resist the urge to boil the library.
- **Wire the registry watch.** FDA SrLC, EMA product information, emc, whichever registries govern your markets: a label revision should become a re-check event, not a memo.
- **Add suggest-the-reference at claim creation.** The cheapest point in the entire lifecycle to raise reference quality is the moment the claim is first written.
- **Keep the reviewer accountable.** Findings route to named humans; decisions are theirs; the trail is Part 11-supporting where regulation requires it.
- **Report reference health quarterly.** The band distribution and its trend, to medical affairs leadership. Direction of travel, not a certification.

## Where a tool fits

Everything above can be run manually: a label PDF, a spreadsheet, and patience will execute all five checks on a small library, and doing it by hand once is the best way to believe the findings. What does not scale by hand is continuity, re-checking hundreds of claims on every label revision, forever. Juncture implements this framework as part of its Content Intelligence layer: claims and references ingest self-service, the five checks run at creation, on revision events, and on a cadence, and reviewers receive evidence-backed findings while keeping the accountable decision on every claim. The framework itself is deliberately tool-agnostic: if this paper only changes what your team requires of a citation, it has done its job.

## The takeaway

A claims library that governs its claims and trusts its citations is governed on one side. The reference is where the claim touches the evidence, and it is the only part of the library that a regulator, a reviewer, and a machine can all check against the same public document. Five checks, presence, substantiation, numbers, currency, hierarchy, make that trust auditable; three bands turn the results into a worklist; three triggers keep it true continuously. The label will keep changing several hundred times a year. The only question is whether your library notices.

## Sources


1. European Commission, "A Guideline on Summary of Product Characteristics (SmPC)," Revision 2, September 2009 (the SmPC as the basis of information for healthcare professionals; safety information regularly reviewed and updated). health.ec.europa.eu
2. European Medicines Agency, glossary entry "Summary of product characteristics" (SmPCs form the basis of information for healthcare professionals on safe and effective use). ema.europa.eu
3. US Food and Drug Administration, "Drug Safety-related Labeling Changes (SrLC) Database Overview" (approved safety labeling changes since January 2016; sections tracked; searchable by drug, date range, and labeling section). fda.gov
4. Kircik L, Sung JC, Stein-Gold L, Goldenberg G, "United States Food and Drug Administration Product Label Changes," Journal of Clinical and Aesthetic Dermatology, 2016 (400 to 500 label changes every year). pubmed.ncbi.nlm.nih.gov
5. electronic medicines compendium (emc), "About emc" (more than 14,000 regulator-checked documents; companies should update the emc within 10 days of an approved change). medicines.org.uk

6. Baethge C, Jergas H, "Systematic review and meta-analysis of quotation inaccuracy in medicine," Research Integrity and Peer Review, 2025 (46 studies, 32,074 quotations; 16.9% of quotations incorrect, about half major errors). pmc.ncbi.nlm.nih.gov

7. Jergas H, Baethge C, "Quotation accuracy in medical journal articles, a systematic review and meta-analysis," PeerJ, 2015 (the earlier meta-analysis; error rates broadly unimproved a decade on). pubmed.ncbi.nlm.nih.gov