JUNCTURE RESEARCH · WHITEPAPER

The Strategy Execution Gap

Why pharma loses the approved message after approval, and how to govern message execution.

JULY 2026

JUNCTURE.HEALTH

The problem: pharma defines the message once, then loses control as content scales

No industry invests more in getting a message right. Before a single promotional claim reaches a clinician, it has been shaped by brand strategy, substantiated against evidence, aligned to the approved label, and signed off by medical, legal and regulatory review. The approved message is arguably the most expensive sentence in commerce.

And then it is set loose. The same organisation that spent months defining the message spends almost nothing governing what happens to it next. The message is adapted for an email, compressed for a banner, localised for a market, summarised in a rep triggered message, rebuilt by an agency that never saw the strategy deck, and finally paraphrased by an AI system answering a clinician's question at midnight. Each step is individually reasonable. Collectively they are a controlled release into an uncontrolled system.

The scale of that release is documented. Veeva's Pulse Content Metrics Report, drawing on data from more than 350 global pharma companies, found the volume of managed content compounding at **roughly 70% per year**. At the same time, Veeva Pulse field data shows that **nearly 80% of approved content is rarely or never used** in the field. Read those two numbers together and the picture is stark: the industry produces vastly more executions of its message every year, uses a fraction of them, and has no instrument that says whether any given execution still carries the message that was approved. Volume is growing. Fidelity is unmeasured.

The audience moved too. The AMA's Physician Survey, published in March 2026, found that **81% of physicians use AI professionally**, more than double the rate when the AMA first polled doctors on health AI in 2023. Doximity's State of AI in Medicine 2026 report adds texture: among physician AI users, **69% report daily use, including 36% who use AI multiple times per day**. And the shift is not confined to chat tools. An analysis of over 130,000 health-related search queries found that **Google AI Overviews now appear in 51% of healthcare searches, double the average across industries**.

The consequence is simple to state and uncomfortable to sit with. HCP questions are moving beyond controlled channels. The question a clinician once asked a rep, a med-info line, or a search results page now goes to ChatGPT, Gemini, Perplexity, Google AI Overviews or Claude, and the answer that comes back is an execution of your message that no reviewer ever saw. The strategy execution gap is not a future risk. It is the current operating condition.

The four lifecycle friction points

The gap is not one failure. It is four distinct frictions, one at each stage of the content lifecycle, and each one leaks message fidelity in its own way. Naming them separately matters, because each friction is currently owned by a different team that believes its own step is under control.

Create: teams recreate the message instead of executing it

In most organisations the approved message lives in documents: a claims matrix in a spreadsheet, a message house in a slide, the label as a PDF. None of those is operational. So when a brand team, an agency, or a market affiliate needs an email, a sales aid, or a banner, they do not execute the message from a source. They recreate it from memory and reference, adapting tone, trimming for space, rephrasing for channel.

Every adaptation is a small translation, and small translations accumulate drift. A superiority claim softens into a vague benefit in one asset and hardens past its evidence in another. A carefully worded population statement loses its qualifier in a character-limited format. Localisation multiplies the surface: the same claim, re-derived independently in a dozen markets, exists in a dozen slightly different forms. Nobody decided to change the message. It changed because there was no single reference model to execute from, only artefacts to imitate.

Review: MLR repeatedly catches what could have been caught earlier

MLR review is the point where the organisation is most rigorous about the message, and that is precisely the problem: it is the only point. Because nothing upstream checks content against the approved message, review becomes the de facto quality gate for every mechanical failure in the creation process. Reviewers repeatedly identify the same classes of issue: a claim that wandered off the label, risk information out of proportion, a reference that no longer supports its sentence, a mandatory statement dropped in adaptation.

The cost shows up as cycle time. Veeva reports that most biopharmas take **about three weeks, on average, to deliver new content to market**, and complex materials in manual workflows take substantially longer. Tellingly, teams that restructure their review around reuse and earlier checks report **up to 57% shorter review cycles and 55% less time in review meetings**, which is a measure of how much of today's review effort is mechanical rework rather than judgment. The reviewer's scarce, accountable attention is being spent on defects the process should never have delivered to them.

Measure: teams track assets, not whether the message was preserved

Content operations has metrics: assets produced, cycle times, reuse rates, channel engagement. Every one of them measures the container. None of them measures the cargo. A dashboard can say that an email was produced in fourteen days, reused a modular claim, shipped to forty thousand HCPs and earned a strong open rate, while the sentence at the heart of it quietly overstates the evidence. By every tracked metric, that asset is a success.

The Veeva Pulse finding that nearly 80% of approved content is rarely or never used is itself a measurement-gap symptom: the industry can count what it makes and where it goes, but it has no routine measure of whether the approved message survived the journey, which claims are actually reaching clinicians, and in what form.

Message preservation is the KPI nobody owns, because no system holds both the approved message and its executions in one place to compare them.

Interpret: AI answers may omit, alter or extend beyond the approved message

The newest friction point sits outside the company entirely. When a clinician asks ChatGPT, Gemini, Perplexity, Google AI Overviews or Claude about a product, the engine composes an answer from its training data and whatever sources it retrieves. It may omit material the approved message treats as inseparable, such as risk information alongside a benefit. It may alter emphasis, elevating a secondary point over the approved priority. It may extend beyond the approved message altogether, blending in third-party commentary, older data, or an adjacent indication.

None of this is malicious, and none of it passes through any submission gate. It is interpretation at scale, performed by systems that were never given the approved message as an input. For the four fifths of physicians using AI professionally, the engine's answer is a real execution of your message, delivered at the moment of clinical curiosity, and it is the one execution the current content lifecycle neither creates, reviews, nor measures.

The landscape: four controls that each solve a part, not the connection

The market has not ignored these frictions. Four control categories exist, each aimed at one of them, and each genuinely useful within its scope.

Pre-checkers. Automated compliance checks that inspect an asset before human review: claims against the label, fair balance, mandatory statements, reference substantiation. Done well, they remove mechanical failures from the reviewer's queue and shorten cycles. But a pre-checker's scope is a single asset at a single moment. It can say this asset is consistent with the label today. It cannot say how the claim inside it relates to the forty other places that claim appears, or what happens to the claim after approval.

Claims and module libraries. The modular content movement gives teams a library of pre-approved claims and content blocks to assemble from, which directly attacks the Create friction: execute from a source instead of recreating from memory. The published gains are real. But a library governs the supply of approved language, not the fidelity of what is assembled and adapted downstream, and its writ ends at the boundary of the company's own channels. A module library has no opinion about what Perplexity said this morning.

Content measurement. Analytics platforms track production, reuse, engagement and channel performance with increasing sophistication. They answer every question about the container: how fast, how many, how engaging. They do not hold the approved message as a structured object, so they cannot answer the question that matters to a medical or regulatory leader: was the message preserved?

AI-answer monitoring. The newest category tracks what AI engines say about a brand: visibility, sentiment, share of answers. As brand listening, this is valuable. But monitoring detached from the approved message is observation, not governance. Knowing that an engine mentioned your product is not the same as knowing whether the answer omitted the contraindication, altered the population, or extended beyond the indication, and without a claims-level reference model those judgments cannot be made systematically.

Each category solves its part. What none of them provides is the connection. There is no shared message model that links what is approved, what is created, what passes review, and what AI systems communicate. The pre-checker, the library, the dashboard and the monitor each hold a fragment of the message in a different shape, in a different system, owned by a different team. The gap lives between the tools, which is exactly why buying more of them, separately, has not closed it.

The model: one source of truth, and a loop that learns

Closing the gap does not require replacing any of those categories. It requires the thing they lack in common: **one approved source of truth for the message, structured so that every stage of the lifecycle can be held accountable to it.** Not a

PDF, not a slide, but an operational model of the label, the claims, the evidence behind them, the priorities among them, and the rules that govern their use.

Around that source runs a loop: **Define, Execute, Learn.** The organisation defines the message once, in structured form. Execution, by humans and by machines, is checked and traced against the definition rather than against memory. And what the organisation learns about how the message survives execution and interpretation feeds back into the next definition. Four operating steps make the loop concrete.

Ground. Structure the label, claims, evidence, priorities and rules into one reference model. This is the foundational act: converting the documents that define the message into a model that software and people can check against. Every claim carries its evidence and its boundaries. Every rule, from fair balance to mandatory statements to market-specific tolerances, is explicit and testable.

Check. Review content against the reference model before MLR. Creation stops being imitation and starts being execution: an asset is compared with the grounded message, and drift is caught while it is cheap to fix, before it consumes a reviewer's attention. The reviewer still reviews, still judges, still signs. What changes is what arrives: fewer mechanical defects, each finding traceable to the clause it violated.

Trace. Follow the claims through the portfolio: where each claim is used, where it was adapted and how far, where it was omitted from an asset that should carry it, where it is duplicated in divergent forms. Tracing turns message preservation from an article of faith into a measurement, and it is what makes the 80% unused-content problem addressable, because you can finally see which executions of the message exist, which are live, and which have drifted.

Monitor. Compare AI-generated answers with the approved message, systematically and repeatedly, across ChatGPT, Gemini, Perplexity, Google AI Overviews and Claude. Not sentiment, but claim-level comparison: which approved claims appear in answers, which are omitted, which are altered, and where answers extend beyond the approved message. Monitoring closes the lifecycle around the one execution surface the company does not operate.

The loop matters more than any single step, because the signals feed back. What Monitor reveals about interpretation, and what Trace reveals about uptake, refine what you define: a claim the engines consistently garble may need sharper wording or better public substantiation; a claim nobody uses may not deserve its priority; a gap the engines fill with third-party content is a gap your defined message should close. Define, Execute, Learn, then define better.

Two boundaries keep the model honest. First, the reviewer stays accountable. Checks inform human judgment; they never replace it, never author claims, and never publish. Automated verdicts are inspectable and backed by a Part 11-supporting audit trail, and final sign-off stays with the named human reviewer. Second, the model is a layer beside the MLR system of record, never a replacement for it. The review workflow, the approved-asset store and the accountable decision stay exactly where they are. Governance of execution wraps around the existing process; it does not rip it out.

A worked example: Varigel

Take a fictional brand, Varigel, approved for a single narrow indication, with a known contraindication for patients on a common comorbidity medication and one headline efficacy claim backed by a defined-duration study.

Ground. The team structures the Varigel label into a reference model: the indication statement with its population boundaries, eleven approved claims each linked to evidence, the contraindication and safety statements flagged as inseparable from efficacy claims, message priorities, and the rules for fair balance and mandatory content. The work takes a focused few weeks for one brand, and for the first time the approved message exists as something other than documents.

Check. A congress wave begins: emails, rep-triggered messages, social variants, localised cuts. Each asset is checked against the model before MLR. The checks catch a subhead implying a broader population than the indication, a chart carrying an efficacy claim with no risk information in proximity, and a localisation where the contraindication line was dropped for space. Creators fix all three before a reviewer

opens anything. MLR receives cleaner assets, findings already pinned to the controlling clause, and spends its meeting on the one genuine judgment call.

Trace. Across the sixty-asset wave, tracing shows the headline efficacy claim present in fifty-one assets, adapted in nineteen of them, with three adaptations that softened the qualifier. It shows the contraindication statement missing from a third of the assets that carry the efficacy claim, mostly older pieces created before the model existed. It also shows four near-duplicate variants of the same claim doing the same job, candidates for consolidation. Message preservation now has numbers.

Monitor. A baseline run puts a public HCP-style question set to ChatGPT, Gemini, Perplexity, Google AI Overviews and Claude. The answers mostly describe Varigel accurately, but two engines routinely omit the contraindication when summarising benefits, and one extends the indication language to a neighbouring population, echoing wording from a third-party summary rather than the label.

Learn. The signals flow back. The omitted contraindication, already under-represented in the company's own assets per Trace, becomes a defined priority: inseparable pairing enforced at Check, and clear public-facing material strengthened so retrieval-based engines have the right source to draw on. The next quarterly Monitor run shows the omission rate falling and the extended-population phrasing receding. The team did not fight the engines. It repaired the message supply the engines learn from, and measured the result.

One brand, one loop, one quarter. Nothing in the walkthrough required new regulation, a platform migration, or a reorganisation. It required a shared model of the message and the discipline to hold every execution, human and machine, against it.

Implementation checklist

A programme to govern message execution can start small and compound. The sequence below is vendor-neutral; each item is achievable with disciplined process plus tooling of your choice.

- **Name an owner for the approved message model.** Not a committee. One accountable owner, typically in medical affairs or content strategy, for the structured message as an asset.
- **Start with one brand and one label.** A single product with a stable label gives a clean baseline and a fast learning loop. Resist the portfolio-wide programme on day one.
- **Ground the message.** Structure the label, claims, evidence, priorities and rules into one reference model. Record claim boundaries and inseparable pairings, such as benefit statements and their risk context, explicitly.
- **Make the rules testable.** Rewrite review guidance as checks that pass or fail: mandatory statements, fair balance proximity, population boundaries, reference substantiation. If a rule cannot be tested, sharpen it until it can.
- **Put the check before MLR, not inside it.** Content is compared with the reference model when it is finished and before a reviewer opens it. Findings go back to creators first.
- **Keep the reviewer accountable.** No automated approval, no machine-authored claims, no auto-publishing. Human sign-off, supported by a Part 11-supporting audit trail, validated under your own SOPs.
- **Baseline your traceability.** Map where each priority claim currently appears across live assets, in what form, and where required statements are missing. Expect surprises; that is the point.
- **Add a message-preservation metric.** Alongside reuse and engagement, report the share of live assets whose claims match the approved model, and trend it. What is reported gets owned.
- **Baseline the AI answer layer.** Put a realistic, public HCP-style question set to ChatGPT, Gemini, Perplexity, Google AI Overviews and Claude, and compare answers with the approved message at claim level: present, omitted, altered, extended.
- **Set a monitoring cadence.** Engines change continuously; a single audit decays fast. Re-run the question set on a fixed cadence and trend omission, alteration and extension over time.

- **Close the loop on a schedule.** Quarterly, bring Trace and Monitor findings into the message definition: reword what the machines garble, prioritise what is under-supplied, retire what nobody uses.
- **Leave the system of record alone.** The MLR workflow and the approved-asset store stay where they are. Everything above is a layer beside them, which is what keeps adoption cheap and reversible.

The takeaway

The strategy execution gap is not a content problem or an AI problem. It is a governance asymmetry: pharma governs message definition with more rigour than any industry and message execution with almost none. Content volume compounding at 70% a year widened the gap. Four fifths of physicians using AI professionally pushed its far edge outside the company's walls entirely. The response is not more definition and not more volume. It is one structured source of truth for the approved message, and a Define, Execute, Learn loop that grounds the message, checks content before review, traces claims through the portfolio, and monitors what the machines say, with every signal feeding back.

This is the layer **Juncture** is built to be: the approved message grounded once, and creation, review, measurement and AI monitoring held to it, beside the systems you already run. How the loop works in practice is documented at **how it works**. But the argument of this paper does not depend on any vendor. Whoever builds your loop, build the loop. The approved message is the most expensive sentence your organisation produces. It deserves governance that survives its own approval.

Sources

1. Pharmaphorum, "Utilising content metrics to improve digital engagement," summarising the Veeva Pulse Content Metrics Report (70% compound annual growth in managed content volume; data across 350+ global pharma companies). **pharmaphorum.com**

2. Veeva, "Veeva Pulse Report Finds Content-Driven Engagement Lags Despite Proven Boost to Treatment Adoption" (nearly 80% of approved content rarely or never used; Pulse data drawn from over 600 million HCP interactions annually). veeva.com
3. American Medical Association, "More than 80% of physicians use AI professionally," AMA Physician Survey, published March 2026 (81% of physicians use AI professionally, more than double the 2023 rate). ama-assn.org
4. Doximity, "State of AI in Medicine Report 2026" (among physician AI users, 69% report daily use, including 36% multiple times per day). doximity.com
5. WebFX, "AI Overviews in Healthcare" (analysis of 130,000+ health-related queries; AI Overviews appear in 51% of healthcare searches, double the cross-industry average). webfx.com
6. Veeva, "Getting Started With Modular Content in Pharma" (most biopharmas take about three weeks, on average, to deliver new content to market; cites the Veeva Pulse Content Metrics Report). veeva.com
7. Veeva, "Future-Proof Your MLR Review and Approval Process" (teams with optimised review workflows report up to 57% shorter review cycles and 55% less time in review meetings). veeva.com