

LIFE SCIENCES · GXP QUALITY OPERATIONS

Agentic AI Is Entering Regulated *Drug Development*

A checkpoint can clear and the batch can still be wrong.

MAUREEN DOYLE-SPARE

© 2026 Maureen Doyle-Spare.

SWIPE TO READ →

LIFE SCIENCES · GXP QUALITY OPERATIONS

Agentic AI Is Entering Regulated *Drug Development*

As autonomous workflows move into pharmacovigilance, clinical development and GxP quality operations, the control boundary shifts from validated execution to the operational meaning an agent resolves

MAUREEN DOYLE-SPARE

A checkpoint can clear and the batch can still be wrong. In agentic life-sciences workflows, the missing control can sit in the Operational Interpretation between validated source systems and regulated execution.

Agentic AI is moving into clinical trials, pharmacovigilance, quality operations, and other regulated life-sciences workflows. The significance is not simply that AI can automate more work. These are environments where a classification such as critical, serious, expected, eligible, or acceptable can determine a quarantine hold, a reporting clock, subject enrollment, or whether regulated work moves downstream.

Life sciences already has one of the most disciplined assurance environments in industry. Systems are validated, records and signatures are controlled, data integrity is protected, procedures are versioned, and quality organizations retain formal authority. Agentic systems introduce another object between those controls: the Operational Interpretation an autonomous workflow resolves before it acts.

The problem becomes visible in the actual workflows

Consider biologics batch disposition. A deviation is logged during a run. An agent can consult the approved SOP defining a critical deviation, the product specification constraining the acceptable range, prior approved dispositions from comparable events, and the in-process manufacturing data. Each source can be individually valid while the criticality boundary is not expressed identically across them. The procedure may draw the boundary in one place, the specification imply another, and historical dispositions sit slightly looser.

The agent is designed to reconcile those sources. If it synthesizes a working definition of critical that no document-controlled process approved, a deviation that should trigger quarantine can instead be classified as major and proceed under a lighter investigation pathway. The electronic record can remain complete, the signature valid, the audit trail intact, and every validated component can perform as designed. The failure sits in the meaning that set the workflow in motion.

The same structure appears in pharmacovigilance. An

© 2026 Maureen Doyle-Spare.

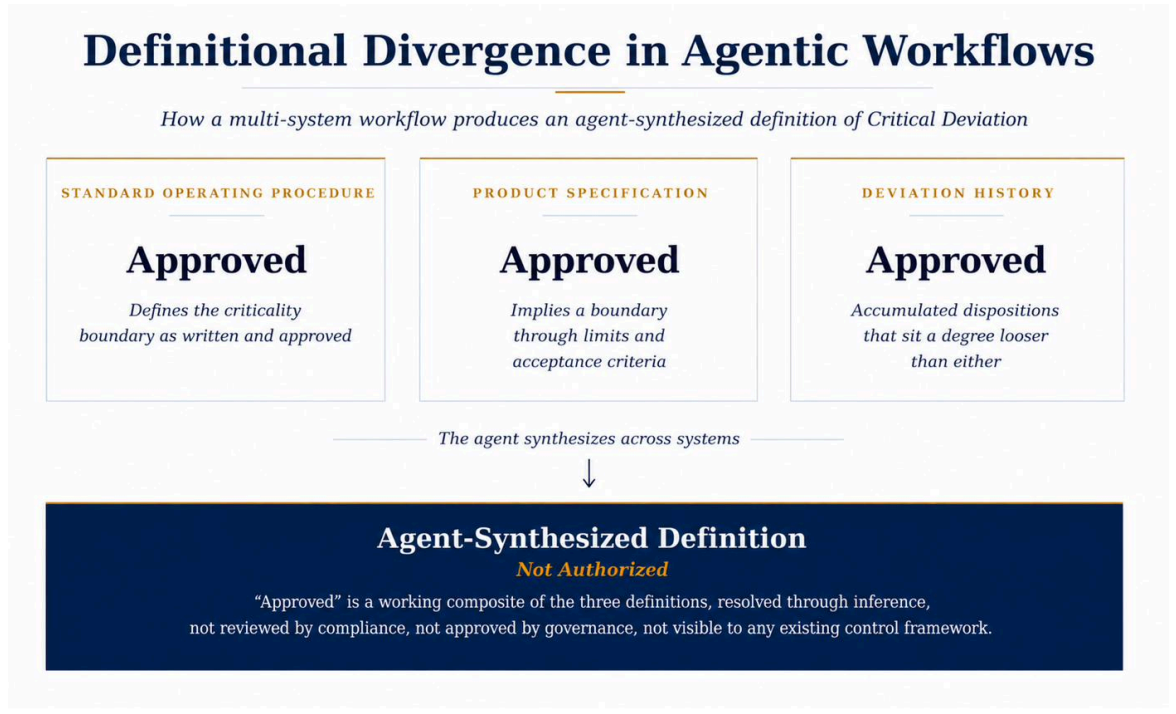


Figure 1. *Definitional Divergence in Agentic Workflows. Source: Doyle-Spare (2026)*[2].

agent triaging an adverse-event report may have to resolve seriousness and expectedness across the case narrative, safety-database conventions, the current product label, and the MedDRA coding dictionary. Those determinations are operationally consequential because they can start reporting clocks and shape regulatory obligations. In clinical operations, eligibility can be reconciled across the approved protocol, amendments, and site conventions. A subject can be enrolled under an interpretation no sponsor authorized even though each contributing artifact remains controlled.

The exposure also varies by modality. Small-molecule products often operate against tighter numeric specifications. Complex biologics, cell therapies, and gene therapies introduce more process-inherent variability and qualitative judgment. The interpretive space is therefore not uniform across life sciences. It can be widest precisely where the science is most variable and the consequence of a misjudged boundary is highest.

Validation governs the system. Runtime governance governs the interpretation.

Computer system validation, computer software assurance, 21 CFR Part 11 controls, ALCOA+ data-integrity principles, quality risk management, and pharmaceutical quality systems remain essential. They govern software behavior, records, data, procedures, changes, signatures, and regulated outcomes. They were built around a long-standing assumption: the behavior under assurance was specified before execution.

Agentic systems change that assumption when they determine operational meaning at runtime. Historically, a LIMS, manufacturing execution system, electronic batch record, safety database, or quality management system supplied validated information and a qualified professional decided what that information meant before regulated action occurred. Agentic deployment can move part of that interpretive responsibility into the workflow itself.

© 2026 Maureen Doyle-Spare.

That is the distinction captured by Agentic Workflow Drift[2]. It is not model drift and it does not require a malfunction. It occurs when a reasoning-driven system synthesizes signals across validated sources into an Operational Interpretation that no controlled process authorized, and execution proceeds under that interpretation. A cleared checkpoint confirms that the workflow executed. It does not necessarily confirm that the workflow executed against the meaning the organization authorized.

A cleared checkpoint confirms that the workflow executed. It does not necessarily confirm that the workflow executed against the meaning the organization authorized.

Human accountability does not require human resolution of every case

The answer is not to put a person back into every decision. That would often recreate the bottleneck agentic automation is intended to remove. The institution instead needs to define which decision classes may execute autonomously, which authoritative meaning applies to them, how much semantic divergence is tolerable, and which cases must be held or escalated.

The Semantic Control Plane[1] provides that runtime structure. The institution maintains an authorized Reasoning Baseline for the decision class. The Runtime Semantic State exposes the Operational Interpretation driving the pending action. The Semantic Deviation Index measures divergence from the Baseline. The Deterministic Gate applies the institution's calibrated threshold before execution authority is emitted. The resulting RSSR and Semantic Audit Trail preserve the governance evidence needed to reconstruct the decision.

The security boundary is different from ordinary drift

Most semantic divergence in a regulated workflow will not be malicious. Procedures, specifications, historical dispositions, labels, reference dictionaries, and site practices can differ for legitimate operational reasons. An agent built to complete the task may treat those differences as signals to reconcile rather than a reason to stop.

Deliberate manipulation is a different class of problem. Agentic Workflow Subversion and Semantic Layer Integrity Attack become relevant when an adversary conditions the artifacts or context from which the agent resolves meaning. In a life-sciences setting, the target could be a release boundary, a seriousness determination, expectedness, or trial eligibility. No record necessarily has to be forged and no validated system has to be modified. The attack can operate on the semantic environment the agent treats as ground truth.

Life sciences is becoming a proving ground for governed autonomy

The current movement of agentic AI into clinical development, pharmacovigilance, and quality operations makes the governance problem concrete. These functions combine controlled systems, scientific judgment, regulatory definitions, and consequences that can reach patients. They are exactly where an institution needs more than a record that an agent acted.

The next maturity step is the ability to show which

© 2026 Maureen Doyle-Spare.



Figure 2. *The Pre-Execution Assurance Protocol. Source: Doyle-Spare (2026)*[1].

authorized definition governed a consequential decision, what interpretation the agent resolved, how far it diverged, what authorization decision was made, and what evidence was preserved. Validation establishes that the system is fit for intended use. Runtime semantic governance establishes that the meaning under which the system is about to act remains institutionally authorized.

Life sciences does not need to replace GxP validation to govern agentic AI. It needs to extend assurance to the interpretive object validation was never designed to supervise. As autonomous workflows move deeper into regulated drug development, governing the agent is necessary. Governing the interpretation it acts on is the next control boundary. ◆

References

1 *Doyle-Spare, Maureen. Agentic AI Systems Governance: A Runtime Reference Architecture for the Reasoning Layer and the Semantic Control Plane in Regulated Financial Institutions. Zenodo, 2026. doi.org/10.5281/zenodo.20749051*

2 *Doyle-Spare, Maureen. Agentic Workflow Drift in Life Sciences: Extending the Reasoning-Layer Risk Taxonomy to GxP-Regulated Pharmaceutical and Biotechnology Operations. Zenodo, 2026. doi.org/10.5281/zenodo.21223427*

Originally published in *The Agentic AI Governance Playbook* on Substack, August 17, 2026. maureendoylespare.substack.com

© 2026 Maureen Doyle-Spare.



Governing the interpretation it acts on is the next control boundary.

THE AGENTIC AI GOVERNANCE PLAYBOOK

Download the research

Papers, briefs and reference architecture at maureendoylespare.com

Visit the Playbook

WEEKLY NEWSLETTER

Follow the newsletter

The Agentic AI Governance Playbook Weekly Newsletter

Follow Newsletter

INTELLECTUAL PROPERTY: © 2026 Maureen Doyle-Spare. All rights reserved. No part of this article, including its text, figures, tables, and images, may be reproduced or distributed without the author's prior written permission.