

Inspection-Ready AI in Pharmacovigilance

Governance, lifecycle assurance and transaction-level traceability

Artificial intelligence can improve the speed and consistency of pharmacovigilance. It does not reduce the accountability of the marketing authorisation holder (MAH), including where activities are outsourced. The inspection question is whether AI-supported work can be understood, controlled, monitored and evidenced in practice. [1, 2, 7, 8]

Three connected levels of control

Level	Inspection question	Representative evidence
Governance	Is the use known, bounded, classified by risk and owned?	AI inventory; context of use; risk and supplier assessments; oversight by the qualified person responsible for pharmacovigilance (QPPV).
Lifecycle	Was the configured service assessed, tested, released, monitored and changed under control?	Requirements; datasets; contextual tests; approvals; monitoring; change and deviation records.
Transaction	Can a real decision be traced from original source to final disposition?	Input; deployed configuration; output; attributable reviewer action; audit trail; downstream action.

Timely retrieval and explanation cut across all three levels: an MAH should be able to produce coherent records, reconcile them across systems and demonstrate how the controls worked. This four-part framing is an author proposal, not an official EU or EMA inspection checklist. [2, 3, 4]

The risk-based premise

Classify the AI-enabled service by what it does and the consequences of error, not by labels such as “assistive.” A drafting aid differs from a service that extracts case data, excludes literature or routes deadline-sensitive reports. Existing EU pharmacovigilance obligations remain controlling; the EMA reflection paper calls for risk-based management of AI in its actual context of use. [1, 5, 6]

A risk-based AI scope model

Maintain an inventory spanning purpose-built AI systems, embedded vendor functions, outsourced activities and relevant model or cloud dependencies. Define each use case by its inputs, outputs, intended and prohibited use, affected products and languages, accountable roles, deployed configuration, fallback and monitoring. The broad inventory and proportionate assurance depth are proposed governance practices, not prescribed EU inventory fields. [5, 6]

Evidence should follow risk

Assure the configured AI-enabled service and workflow, not just the model or a supplier demonstration. Test against the decision, operating population, interfaces, reviewer behaviour and consequences of error. Supplier qualification may be leveraged after documented assessment; MAH-specific assurance may still be needed. [5-7]

Use case	Principal risk	Priority evidence
ICSR intake / extraction	Missed safety information	Source traceability; false negatives; language coverage; review.
Literature screening	Missed relevant article	Recall; exclusions; challenge cases; manual fallback.
Duplicate detection	False merge or missed duplicate	Representative tests; confirmation; audit trail.
Coding / translation	Altered clinical meaning	Segmented performance; correction analysis; reviewer control.
Signal prioritisation	Delayed or missed signal	Completeness; calibration; triage rationale; monitoring.
Regulatory drafting	Incorrect or untraceable content	Source verification; expert approval; version history.

When a service may exclude or deprioritise safety information, measure missed targets and test the fallback. Overall accuracy alone can conceal rare, serious errors or weak performance in particular languages and sources; report sensitivity, critical errors and relevant subgroup results against an adjudicated reference standard. These metric priorities are risk-based author recommendations. [5, 6, 9]

Human oversight must be measurable

A review step is not proof of effective control. Reviewers need source visibility, training, sufficient time and authority to correct, escalate or stop unsafe use. Assess whether they catch plausible safety-relevant errors; sample accepted outputs as well as overrides, and monitor workload and correction patterns. This operational testing is a proposed way to evidence human oversight. [5, 6, 8]

Reproducibility versus reconstruction

Exact technical reproducibility means regenerating the same output under the same inputs, model, parameters, prompt, retrieval state and environment. Hosted, updated or non-deterministic systems may not support this. Defensible historical reconstruction instead uses contemporaneous records to establish what actually happened; it does not promise byte-for-byte replay. This distinction is an author proposal informed by traceability principles. [2, 6, 10]

For an inspector-selected transaction, link the original source and receipt time to the deployed application release; model or provider version or release identifier; prompt-template and retrieval-source versions; rules and thresholds; output presented; authenticated, attributable reviewer actions; final disposition; and downstream update, submission or escalation. Link these to the approved production release and contemporaneous change, limitation and monitoring records. This evidence chain is an author recommendation. [2, 3, 5, 6, 10]

The eight evidence packages of the inspection-ready operating model

- | | |
|--|--|
| 1. AI system map and inventory | Which uses and dependencies exist? Inventory, process map and PSMF cross-reference. |
| 2. Context, governance and risk | Who owns each use and its stop decision? Context, risk assessment and QPPV oversight. |
| 3. Supplier and data governance | Can vendor and data controls be evidenced? Contracts, data flows and inspection access. |
| 4. Assurance and release | Was the actual service fit for use? Requirements, tests and approval. |
| 5. Operational control | Is quality stable? Segmented metrics, drift, exceptions and workload. |
| 6. Configuration and change | Which state ran when? Version register, impact assessment and rollback. |
| 7. Deviations and CAPAs | What was affected and corrected? Deviations and corrective and preventive actions (CAPAs). |
| 8. Transaction reconstruction | Can a decision be shown end to end? Source-to-disposition evidence export. |

Conclusion

Inspection readiness is not a vendor certificate or generic AI policy. It is the demonstrated ability to show that material AI use is known, bounded, assessed, monitored, owned and linked to an accountable final decision. If AI materially influences a regulated pharmacovigilance decision, that influence should be reconstructable.

Selected references

- [1] European Parliament and Council. Directive 2001/83/EC of 6 November 2001 on the Community code relating to medicinal products for human use, as amended.
- [2] European Commission. Commission Implementing Regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities, as amended by Commission Implementing Regulation (EU) 2025/1466 of 22 July 2025.
- [3] European Medicines Agency and Heads of Medicines Agencies. Guideline on good pharmacovigilance practices (GVP), Module III – Pharmacovigilance inspections (Rev 2). 4 September 2026; effective 10 September 2026.
- [4] European Medicines Agency and Heads of Medicines Agencies. Guideline on good pharmacovigilance practices (GVP), Module II – Pharmacovigilance system master file (Rev 2). 28 March 2017.
- [5] European Medicines Agency. Reflection paper on the use of Artificial Intelligence (AI) in the medicinal product lifecycle. 9 September 2024.
- [6] European Medicines Agency and U.S. Food and Drug Administration. Guiding principles of good AI practice in drug development. January 2026.
- [7] European Medicines Agency and Pharmacovigilance Inspectors Working Group. Human pharmacovigilance inspections: questions and answers, electronic-system validation/qualification Q&A. EMA webpage; undated.
- [8] European Medicines Agency and Heads of Medicines Agencies. Guideline on good pharmacovigilance practices (GVP), Module I – Pharmacovigilance systems and their quality systems. 22 June 2012.
- [9] European Medicines Agency and Heads of Medicines Agencies. Guideline on good pharmacovigilance practices (GVP), Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). 28 July 2017.
- [10] Medicines and Healthcare products Regulatory Agency. “GXP” Data Integrity Guidance and Definitions. Revision 1, March 2018. UK supporting guidance.

EVIDENCE BASIS | Adapted from the supplied research dossier (cut-off: 23 September 2026). The three-level model, cross-cutting retrieval capability and eight evidence packages are author proposals, not an official inspection checklist. Sources include EU law, GVP guidance, EMA reflection material and EMA/FDA principles; MHRA guidance is not EU law. Informational only; not legal advice.