Artificial Intelligence and Medicines Security

1509 words in abstract and main text. Institutional draft; independent scholarly review pending.

Abstract

This working paper proposes a decision-centred approach to AI in medicines-security systems. It examines how forecasting, document analysis, dependency mapping and scenario support could contribute to resilience without confusing computational performance with operational benefit. Recent healthcare supply-chain scholarship is used to identify the gap between modelling and evaluated deployment, while WHO guidance provides governance context. The contribution is a proposed assurance sequence covering decision definition, data readiness, comparator selection, prospective testing, human escalation and maintenance. A distinction is made between source-grounded research assistance and operational early-warning systems. The paper proposes a limited pilot using retrospective records followed by silent prospective observation. It does not report a tested model, quantified benefit or validated shortage prediction. It concludes that AI should be judged by the quality and timeliness of authorised decisions and equitable access outcomes, with explicit reporting of failures and uncertainty.

1. The research problem

AI is frequently discussed as though a better prediction automatically creates a more resilient medicines system. That assumption overlooks the steps between an output and a response. An alert must concern the right decision, arrive early enough to matter, be understood by an accountable person and connect to a feasible action. A warning about an unavoidable interruption may still have value, but that value differs from prevention. The central question is therefore not whether AI can detect patterns. It is whether a defined application improves a defined decision under operational conditions, compared with an appropriate alternative and without creating unacceptable new dependencies.

2. Method and source boundaries

This draft combines conceptual analysis with purposively selected research and governance documents. The 2026 review by Espinosa and colleagues provides a recent examination of AI in healthcare supply chains, with its evidence search extending to May 2024 [1]. WHO’s 2021 and 2024 guidance supplies a governance foundation for health AI and large multi-modal models [2,3]. These sources were checked on 13 October 2026. No new systematic review or model experiment was performed. Healthcare supply chains include products beyond pharmaceuticals, so findings cannot be transferred to medicines without qualification. The assurance sequence proposed below is MSI’s analytical contribution and remains to be tested.

3. Evidence maturity matters

Espinosa and colleagues identify a substantial gap between proposed computational models and real-world evaluation within the literature they reviewed [1]. That bounded finding cautions against treating simulation results as demonstrated clinical or operational benefit. It does not prove that every current implementation is ineffective, particularly because the search period precedes publication. A research register should classify evidence as conceptual, simulated, retrospectively tested, prospectively observed or operationally evaluated. These categories ask different questions. A model may be technically promising while still lacking evidence that staff can use it, that an alert produces a timely response or that any benefit persists after deployment.

4. Specify the decision before the model

A suitable use-case record states the owner, decision, lead time, affected population and available actions. For example, an institution may wish to identify products requiring additional supplier investigation within a defined planning period. The output should support that investigation, not claim to predict patient harm directly. The decision record should identify the consequence of missing a signal and the cost of investigating a false alert. When those consequences differ markedly, a single accuracy statistic is inadequate. A deliberately simple baseline may be sufficient for a first deployment; complexity should be justified by useful improvement, not by the appeal of the technology.

5. Data readiness and provenance

An input catalogue should record who produced each observation, when it was available, how it was revised and whether it may be used for the intended purpose. Medicines-security data can refer to products, ingredients, manufacturing sites, suppliers or incidents; those entities must not be silently conflated. Missingness should remain visible. Training records must distinguish what was known at the decision date from information added after an event. Otherwise, apparent predictive success may result from leakage. Access rights and confidentiality are also methodological requirements: a model based on data unavailable to the operational user cannot be evaluated as though those inputs will always be present.

6. Forecasting and classification

A proposed forecasting pilot should compare performance across products and time periods, including disruptions unlike the training examples. Demand and supply signals should be evaluated separately because an observed shortage may have several explanations. Classification outputs should include uncertainty and a route to inspect the supporting record. The study must specify how outcomes are labelled and when they become observable. Changes in reporting practice can alter those labels without an equivalent change in underlying risk. The pilot should report false alerts, missed signals, calibration and useful lead time, alongside operational workload. It should not convert model confidence into an unexplained national resilience score.

7. Document analysis and generative systems

Generative AI can assist with source discovery, extraction drafts and structured summaries, but its fluent output is not evidence. A source-grounded workflow should preserve the exact document version and passage supporting each material claim. The reviewer should verify entities, dates, units and whether the passage supports an observation or an interpretation. Retrieval should not be described as verification merely because a link exists. Documents can contain irrelevant or adversarial instructions, so research systems should treat their text as data rather than authority to change the task. Sensitive supplier, patient or interview information should remain outside unapproved systems.

8. Human accountability and escalation

WHO’s guidance places ethical governance and accountability at the centre of health AI [2,3]. In this proposed operational framework, every application needs an owner who can challenge the output, pause its use and explain the decision taken. Escalation rules should cover disagreement, missing evidence, unexpected behaviour and system outage. Staff need a usable alternative when the application fails. Human review should be designed as substantive judgement rather than routine approval of automated suggestions. The organisation should record where the model affected a decision and where it was rejected. That record supports learning and prevents accountability from becoming diffuse between supplier, software and institution.

9. Security and dependency risks

An AI application can improve one part of a system while creating a new single point of failure. Concentrated vendors, opaque model updates, inaccessible logs or dependence on one external service may weaken continuity. A deployment assessment should consider data access, model portability, cyber response, outage tolerance and contractual exit. Those questions are distinct from predictive validity, but both determine whether the application is usable during stress. The institution should test recovery without the model and define which tasks can continue manually. An early-warning system that becomes unavailable during the event it is intended to support has failed an essential capability test.

10. A prospective pilot design

The proposed pilot begins with a defined product group and documented historical observations. Stage one builds an auditable baseline using simple rules. Stage two compares the candidate model using temporal holdout data and predefined metrics. Stage three operates silently alongside normal practice, recording alerts without changing clinical or procurement decisions. Only after independent review should a controlled operational phase be considered. The evaluation should measure whether alerts reach authorised owners, whether actionable options exist and how staff workload changes. Any later outcome study would need a credible comparison and attention to confounding. No stage should be represented as complete before evidence has been collected.

11. Equity, adoption and workforce

Efficiency alone cannot determine success. A model may prioritise products with richer data while overlooking small populations or fragmented supply pathways. Evaluation should therefore ask whose risks become visible and whose remain unobserved. Accessibility, training and staff capacity affect whether outputs are acted upon. A warning dashboard can consume attention without improving decisions if it is not integrated into existing work. The proposed adoption record includes the target user, training, authority, response resources and feedback channel. Changes should be reviewed with relevant practitioners and patient perspectives, rather than assuming that one workflow fits every institution.

12. Limitations and research agenda

The framework has not been tested and cannot establish that its sequence improves outcomes. Selected sources do not cover every AI technique or implementation, and rapid technical development can outpace published evaluation. Retrospective data may be incomplete, while prospective interruption outcomes may be rare. These constraints argue for precise, small studies with honest negative findings. Future work should investigate the relationship between technical performance, useful lead time and decision quality. Medicines-security AI should be assessed as a maintained institutional capability. Its contribution is credible when the evidence shows what improved, for whom, under which conditions and at what cost.

Appendix: proposed AI use-case assurance record

Record the intended decision, user, owner, relevant medicine group, planning horizon and permitted actions. Document source rights, version dates, missingness and the point at which each input becomes available. State the baseline comparator and the outcomes used for evaluation. Predefine how false alerts, missed signals, calibration, lead time and staff workload will be assessed. Record which populations or product categories have weak data coverage and how that affects use.

Before any operational phase, document human escalation, pause authority, monitoring, outage procedures and a practical non-model alternative. Keep a log of accepted and rejected outputs and the reasons for decisions. Specify how changes to data sources, model versions or operational conditions trigger review. Independent assessment should challenge evidence quality, generalisability and whether an apparent technical improvement changes an authorised response. This record is a proposed assurance instrument, not a validated certification checklist. It does not remove the need for appropriate data protection, clinical, regulatory and organisational review.

References and publication record

1. Espinosa, O., Basto, S., Ordóñez, A. et al. (2026). A review of artificial intelligence in healthcare supply chains: untapped potential? BMC Health Services Research 26, 742. doi:10.1186/s12913-026-14559-2.
https://link.springer.com/article/10.1186/s12913-026-14559-2
Accessed 13 October 2026.

2. WHO (2021). Ethics and governance of artificial intelligence for health.
https://www.who.int/publications/i/item/9789240029200
Accessed 13 October 2026.

3. WHO (2024). Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models.
https://www.who.int/publications/i/item/9789240084759
Accessed 13 October 2026.

Authorship, funding, interests and AI use

Working-paper draft v0.1, 13 October 2026. Institutional research draft prepared with generative-AI assistance and direct source checks. Human scholarly review, publication-specific authorship, funding and conflict declarations remain pending. No original interviews or empirical study were conducted. Correspondence: contact@medicines-security.net.

The founder has interests in advisory and life-sciences ventures. Publication-specific declarations and accountable authorship must be completed before final scholarly release. Cited authors and institutions are sources, not partners or endorsers.

Draft citation

Medicines Security Institute (2026). Artificial Intelligence and Medicines Security. MSI Working Papers 02, draft v0.1, 13 October.