Original Article

Detecting Workflow Drift before Pharmacy Algorithms Learn the Wrong Practice

Charlotte Dupuis[1]*, Hugo Perrin[1], Elise Morel[2], Thomas Martin[3]

[1]Department of Workflow Drift Detection in Pharmacy, Faculty of Pharmacy, University of Lille, Lille, France. [2]Department of AI Practice Integrity, Faculty of Pharmacy, University of Montpellier, Montpellier, France. [3]Department of Algorithm Monitoring and Correction, Faculty of Pharmacy, University of Nantes, Nantes, France.

Abstract

Artificial intelligence increasingly participates in medication-related prioritization, alerting, verification, documentation, and clinical decision support. Yet deployed algorithms may continue learning while the pharmacy work that produces their inputs, labels, feedback, and outcomes is changing. This creates a risk that temporary workarounds, redistributed responsibilities, altered documentation, changing patient populations, or unsafe local practices become represented as legitimate practice. This article defines pharmacy workflow drift as a material temporal change in the tasks, professional roles, activity sequence, patient or case population, or documentation processes through which medication-related work is performed and represented. It proposes a Multidimensional Pharmacy Workflow-Drift Framework comprising a baseline workflow contract, five-dimensional observability layer, cross-dimensional coupling assessment, consequence layer, and learning-control layer. Detection is separated from diagnosis: a statistical or process signal initiates investigation but does not confirm drift, establish causation, or authorize model updating. The framework therefore connects drift localization to medication-safety, equity, human-factors, implementation, and accountability assessments before selecting a learning state such as continued operation, enhanced surveillance, learning freeze, restricted use, controlled updating, rollback, or retirement. Validation would require technical sensitivity, false-alarm assessment, causal plausibility, workflow localization, professional usability, subgroup stability, safety evaluation, rollback capability, and recurring local requalification. The framework is an original non-empirical governance synthesis. It does not provide validated thresholds, guarantee safer decisions, allocate legal liability, or establish regulatory or deployment readiness.

Keywords: Digital pharmacy, Artificial intelligence, Pharmacy practice, Medication safety, Human–AI collaboration, Clinical decision support

Introduction

Clinical algorithms are commonly developed under assumptions about how data are generated, which patients enter a pathway, who records decisions, and how outcomes become observable. These assumptions may weaken when clinical environments, professional practices, technologies, or data-generating processes change after development [1]. In pharmacy, the resulting problem is not limited to declining discrimination or calibration. An algorithm may remain statistically stable while the work surrounding it changes sufficiently to alter the meaning of its inputs, overrides, labels, and feedback.

Clinical value therefore depends on more than model accuracy. It also depends on integration into medication-use pathways, alignment with professional work, and evaluation under the conditions in which an output is interpreted and acted upon [2]. A prediction delivered to a pharmacist before verification is not functionally equivalent to the same prediction delivered after dispensing, to another professional

Address for correspondence: Charlotte Dupuis, Department of Workflow Drift Detection in Pharmacy, Faculty of Pharmacy, University of Lille, Lille, France.
charlotte.dupuis@univ-lille.fr
Received: 07 February 2026; **Accepted:** 17 April 2026

How to cite this article: Dupuis C, Perrin H, Morel E, Martin T. Detecting Workflow Drift before Pharmacy Algorithms Learn the Wrong Practice. Arch Pharm Pract. 2026;17(2):21-9. https://doi.org/10.51847/78i64L2M3E

group, or through a documentation field introduced for administrative rather than clinical purposes.

Responsible health-machine-learning deployment consequently requires lifecycle monitoring, stakeholder oversight, and safeguards against harm [3]. However, conventional monitoring often privileges model-level indicators while treating workflow as a relatively stable deployment environment. This assumption is especially problematic for systems that learn from subsequent clinician actions, accepted recommendations, override patterns, or routinely documented outcomes. Such systems can absorb not only improved practice but also staffing-related shortcuts, alert avoidance, delayed documentation, role substitution, or temporary crisis procedures.

Machine-learning systems can produce unintended clinical and sociotechnical consequences even when their technical objectives appear to have been achieved [4]. In a pharmacy context, an algorithm may optimize conformity with observed behavior while that behavior is becoming less safe, less equitable, or less consistent with intended professional standards. "Learning from practice" is therefore not necessarily equivalent to learning correct practice.

Artificial-intelligence applications are expanding across pharmacy operations and professional activities, although evidence concerning patient-centered and practice consequences remains less mature than evidence concerning operational applications [5]. This article addresses that gap by proposing a governance framework for detecting and managing workflow drift before changed practice is incorporated into algorithmic learning. The contribution is conceptual rather than empirical: it defines a pharmacy-specific construct, specifies observable dimensions and relationships, separates detection from diagnosis, and proposes constraints on when changed workflow data may inform continued learning.

Defining Workflow Drift in Pharmacy

Concept drift traditionally concerns temporal change in relationships relevant to prediction and should be distinguished from a simple change in the distribution of observed variables [6]. Data drift may alter what the model receives, whereas concept drift may alter how those variables relate to a target. Performance drift describes a measured change in model behavior. None of these concepts, alone, identifies whether the work producing the data has changed.

Process-oriented literature shows that drift may involve control flow, data, resources, roles, and temporal behavior [7]. This broader perspective is important for pharmacy because medication-related work is distributed across people, technologies, organizational units, and time. A changed process may affect which actions occur, who performs them, their order, the cases exposed to them, and how actions are represented in digital records.

Healthcare process mining can reveal actual pathways, repeated deviations, and temporal patterns, but event logs remain partial representations of clinical work [8]. Informal consultation, tacit reasoning, interruptions, verbal clarification, and undocumented workarounds may be absent. An observed change in a digital trace may therefore reflect a real workflow change, a documentation change, or a technical extraction problem.

Drift analysis must also proceed beyond alarm generation toward identifying what changed, where it changed, and when the change occurred [9]. A monitoring signal is not a diagnosis. Abrupt changes may follow software releases or policy changes; gradual changes may reflect learning, workload adaptation, or normalization of workarounds; cyclical changes may reflect staffing or seasonal patient patterns. Each requires different investigation.

This article therefore proposes that pharmacy workflow drift is a material change over time in the tasks, professional roles, activity sequence, patient or case population, or documentation practices through which medication-related work is produced and represented, where that change can alter algorithm inputs, labels, feedback, use conditions, or consequences. The definition is sociotechnical because medication safety emerges from interactions among people, tasks, technologies, organizational conditions, and care environments [10]. Workflow drift differs from expected variation within an accepted process and from deliberately governed redesign. It also does not presume deterioration: some drift may represent improvement. The governance concern arises when a material change is unrecognized, insufficiently understood, or incorporated into learning without determining whether it reflects acceptable practice.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, components, relationships, and intended uses for the article Detecting Workflow Drift before Pharmacy Algorithms Learn the Wrong Practice.

Table 1. Definitions, components, relationships, and intended uses for the article Detecting Workflow Drift before Pharmacy Algorithms Learn the Wrong Practice.

Component or scientific dimension	Problem addressed	Inputs or determinants	Proposed mechanism or relationship	Expected contribution	Evidence required	Failure or uncertainty risk	Boundary statement

Workflow drift	Model monitoring may miss change in work itself	Tasks, roles, timing, patients, records, technologies	**Proposed:** changed work alters data generation, algorithm use, or feedback	Organizes workflow-level surveillance	Longitudinal workflow and outcome evidence	Ordinary variation may be misclassified	Not every change is harmful or algorithmically material
Task drift	Intended and performed work may diverge	Verification steps, intervention types, exception handling	**Proposed:** task substitution changes labels and exposure	Identifies what work has changed	Process observation and task-validity assessment	Hidden or informal work may be absent	A new task is not inherently unsafe
Role drift	Responsibility may move without model redesign	Profession, grade, authorization, supervision	Resource and role changes are recognized process dimensions [7]	Reveals altered interpretation and authority	Role-resolved logs and qualitative inquiry	Recorded user identity may not equal decision authority	Role change does not establish reduced competence
Sequence drift	The same actions may occur in a different order	Timing, handoffs, parallel work, bypasses	Changed ordering may alter when outputs influence decisions	Detects changed intervention opportunity	Time-stamped event and observational evidence	Retrospective entry can create false sequences	Temporal association is not causation
Population drift	The cases entering a workflow may change	Morbidity, medication complexity, language, setting, eligibility	**Proposed:** case-mix change alters both workload and learning exposure	Separates workflow from patient-distribution change	Subgroup, pathway-entry, and outcome evidence	Seasonal or policy effects may resemble drift	Population change alone is not workflow failure
Documentation drift	The digital record may change without equivalent clinical change	Templates, coding, missingness, copy-forward, delayed entry	**Proposed:** altered representation changes model-visible practice	Protects against learning documentation artifacts	Source-system and fitness-for-use validation	Documentation and task drift may be confounded	Records are not complete representations of work
Cross-dimensional coupling	One drift dimension may produce another	Shared staffing, interfaces, policies, workload	**Proposed:** role, sequence, task, population, and documentation changes interact	Supports multidimensional diagnosis	Cross-source temporal concordance	Coupling may be inferred incorrectly	Concordance strengthens but does not prove causality
Learning boundary	Detection may be mistaken for permission to adapt	Drift signal, consequences, uncertainty, authorization	**Proposed:** learning depends on diagnosis and governance approval	Prevents automatic absorption of unresolved practice	Safety, equity, workflow, and validation evidence	Freezing may preserve an already degraded model	A signal cannot directly authorize updating

The table's dimensions are proposed organizing constructs rather than validated diagnostic categories. Their observability, separability, and usefulness require empirical testing in different pharmacy settings.

Proposed Multidimensional Drift Framework

Figure 1 presents the original conceptual synthesis for multidimensional pharmacy workflow-drift framework, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

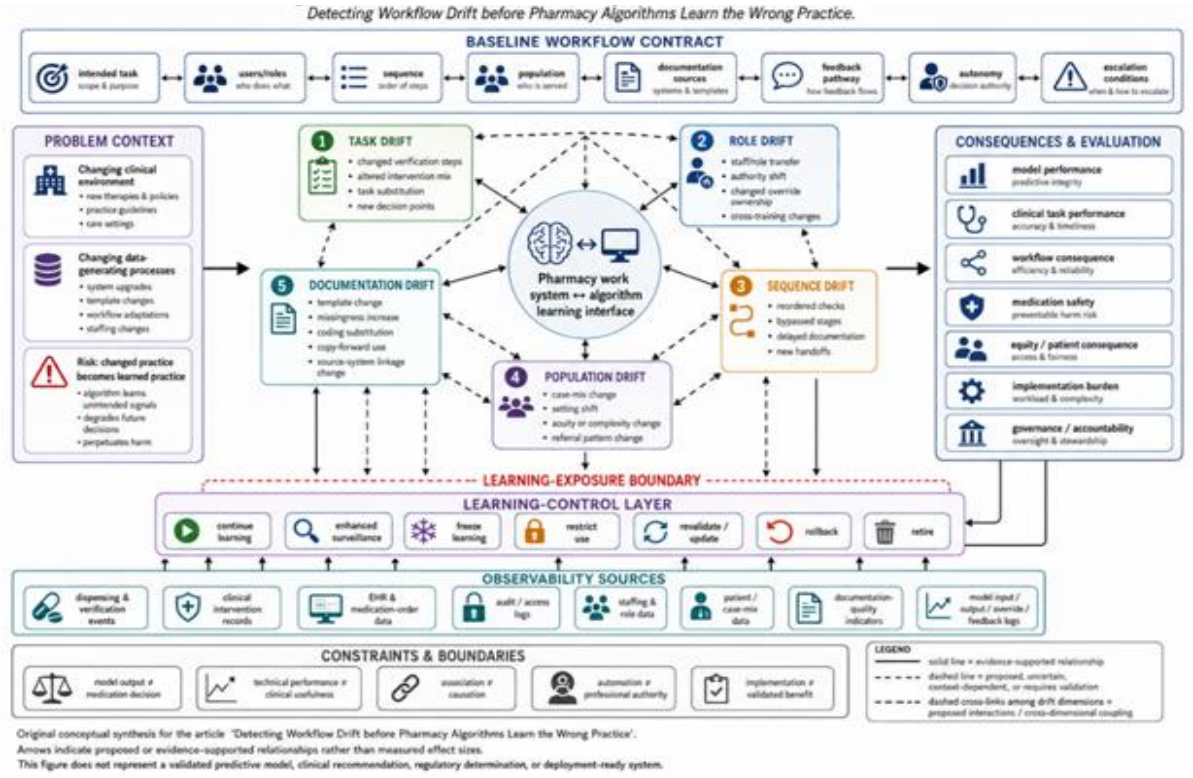

Figure 1. Multidimensional Pharmacy Workflow-Drift Framework

The framework places a pharmacy work system–algorithm learning interface at its center. This reflects the human-factors principle that work-system conditions, care processes, outcomes, transitions, and feedback should be considered together rather than as isolated technical elements [11]. The central interface is surrounded by a baseline workflow contract, five drift dimensions, an observability layer, a consequence layer, and a learning-control layer.

The baseline workflow contract is a proposed, versioned description of the algorithm's intended task, authorized users, patient population, expected sequence, documentation sources, feedback pathway, autonomy level, and escalation conditions. It is not a legal contract. Its purpose is to make assumptions about work explicit enough to compare intended and observed practice.

The five-dimensional observability layer assesses task, role, sequence, population, and documentation change. Human–AI performance depends on complementarity and appropriate division of work rather than the substitution of an isolated algorithm for professional judgment [12]. Role drift is therefore not merely a change in username or staffing category; it may alter who interprets an output, who possesses authority to act, and who creates the feedback later treated as ground truth.

The consequence layer separates model performance from clinical task performance, workflow consequence, medication safety, patient experience, equity, implementation burden, and accountability. This separation is important because medication-alert optimization remains limited by incomplete external validation and routine implementation evidence [13]. A model can produce technically improved prioritization without establishing that the resulting work is safer, more equitable, or professionally usable.

The learning-control layer governs whether changed data may influence future behavior. Alert effects can depend on interaction design and the clinical role receiving the alert [14]. Consequently, an apparent improvement in acceptance or override rates may reflect changed placement, responsibility, or workload rather than improved clinical appropriateness. The proposed framework treats such changes as diagnostic questions rather than direct learning targets.

Conceptual layout: Baseline workflow assumptions feed the central work system–algorithm interface. Five surrounding drift dimensions supply candidate signals. Signals enter a consequence-assessment boundary before reaching the learning-control layer. Feedback arrows from algorithm outputs to professional actions and subsequent documentation indicate that the system can reshape the data from which it later learns. Dashed cross-links indicate proposed,

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for evaluation criteria, evidence requirements, and failure modes for the article Detecting Workflow Drift before Pharmacy Algorithms Learn the Wrong Practice.

Table 2. Evaluation criteria, evidence requirements, and failure modes for the article Detecting Workflow Drift before Pharmacy Algorithms Learn the Wrong Practice.

Architecture layer or process stage	Core function	Information flow	Interaction with other components	Evaluation criterion	Potential failure mode	Human or organizational responsibility	Readiness boundary
Baseline workflow contract	Make intended work and learning assumptions explicit	Intended task, users, sequence, population, records, feedback	Provides comparator for all drift dimensions	Completeness, version control, local interpretability	Baseline normalizes unsafe or obsolete practice	Pharmacy service owner with informatics review	**Proposed:** no readiness without an agreed baseline
Observability layer	Detect potentially material temporal change	Workflow events, staffing, patient mix, documentation, model logs	Supplies signals to triage, not directly to learning	Sensitivity, timeliness, false-alarm burden	Partial logs or extraction errors create misleading alarms	Data and workflow monitoring functions	Detection alone cannot establish drift
Five-dimension classifier	Localize task, role, sequence, population, or documentation change	Candidate signals plus contextual evidence	Supports cross-dimensional assessment	Distinguishability and reproducibility	Overlapping dimensions are forced into one category	Multidisciplinary drift-review team	Categories require empirical validation
Cross-dimensional coupling	Identify interacting changes	Temporally aligned signals from independent sources	Connects role, task, sequence, population, and records	Cross-source concordance and plausible mechanism	Coincident changes are interpreted causally	Clinical, operational, and technical investigators	Association cannot authorize corrective learning
Consequence layer	Separate technical from clinical and organizational effects	Model, task, safety, equity, workload, patient evidence	Determines seriousness and permissible response	Domain-specific validity and subgroup assessment	Stable aggregate performance conceals harm	Medication-safety, pharmacy, and equity functions	Technical stability is not clinical readiness
Human–AI authority boundary	Preserve professional decision responsibility	Output, explanation, override, escalation, final action	Constrains automation and interpretation of feedback	Role clarity, usability, escalation reliability	Automation is treated as professional authority	Authorized clinician and governance body	Model output is not a medication decision
Learning-control layer	Select whether and how learning may continue	Diagnosis, consequence assessment, uncertainty, authorization	Governs freeze, restriction, update, rollback, or retirement	Traceability and evidence proportionality	Temporary workarounds become training targets	Designated lifecycle-governance authority	**Proposed:** no update without explicit authorization
Requalification loop	Test the system after material change	Updated workflow, model, subgroup, safety, and usability evidence	Returns approved changes to monitored operation	Local validity and rollback capability	Performance recovery is mistaken for restored workflow fit	Independent validation and release functions	Implementation does not establish validated benefit

The framework does not rank the five dimensions or assume that all must be monitored with equal intensity. Monitoring scope should be conditional on the algorithm's task, autonomy, feedback pathway, medication-safety consequences, and the organization's capacity to investigate signals.

Detecting Task, Role, Sequence, Population, and Documentation Drift

Detection should combine statistical, process, organizational, and qualitative evidence. Empirical drift-detection studies show that performance depends on feature selection, sample characteristics, and the detection method used [15]. A pharmacy organization should therefore validate candidate signals locally rather than import thresholds from another technology or setting. Detection should also complement, not replace, monitoring of clinical and operational consequences.

Task drift may be reflected in changed intervention categories, new verification steps, altered exception handling, task substitution, or a shift from clinical assessment toward administrative completion. Role drift may appear when actions move among pharmacists, technicians, physicians, nurses, automated services, or supervisory levels. Relevant indicators include role-resolved event logs, authorization records, assignment patterns, escalation routes, and qualitative reports from staff. These indicators remain incomplete when actual authority differs from the identity recorded in the system.

Sequence drift concerns changes in order, timing, handoffs, parallel work, or bypassed stages. Time-stamped medication orders, verification events, dispensing records, interventions, acknowledgements, and overrides may reveal altered pathways. However, delayed or retrospective documentation

can create a digital sequence that differs from the sequence of actual care.

Population drift concerns changes in the patients or cases entering an algorithm-supported pathway. Candidate indicators include care setting, age, morbidity, medication complexity, language, service eligibility, referral source, and formulary exposure. Population change may alter workload and model inputs without constituting workflow drift. It becomes workflow-relevant when it changes which tasks are performed, how responsibilities are distributed, or whether existing processes remain appropriate.

Documentation drift requires particular caution because EHR data quality is multidimensional and dependent on fitness for the intended use [16]. Changes in missingness, template adoption, code substitution, copy-forward, free-text practices, delayed entry, or source-system linkage can alter the apparent prevalence of clinical actions without equivalent changes in practice. Monitoring should therefore compare data sources, documentation timing, and extraction pipelines before interpreting record-level changes as clinical change.

Workflow fragmentation and documentation burden can also redistribute where, when, and by whom information is recorded [17]. A new documentation field may improve traceability, create duplicate work, or shift recording to a person who did not make the underlying decision. Conversely, disappearance of a coded intervention may reflect migration to free text rather than cessation of the intervention. Documentation signals should therefore be investigated alongside staffing, workload, policy, interface, and observational evidence.

Contextual analysis of clinical decision support can reveal whether an apparent problem is associated with workflow configuration, staffing, education, or system design rather than the predictive model itself [18]. Explainability methods may assist localization, but they do not establish causation. The proposed detection process consequently treats every signal as one of four preliminary possibilities: genuine workflow drift, expected variation, documentation or technical artifact, or unresolved change. Only structured diagnosis can determine whether the algorithm's learning state should change.

Diagnosis, Response, and Learning Constraints

A detected signal should initiate diagnosis rather than adaptation. The causal source of a distributional change should be examined before selecting an update because associations learned under one data-generating process may not remain valid after that process changes [19]. Diagnosis should ask what changed, when and where it changed, whether the change was planned, and how it affected algorithm inputs, outputs, labels, overrides, or observed outcomes.

Continual monitoring should connect detection to investigation, controlled updating, and quality-improvement governance [20]. The proposed pathway is signal → triage → localization → contextual diagnosis → consequence assessment → containment → learning-state decision. Relevant explanations include service redesign, staffing pressure, shortages, policy or formulary changes, interface modification, extraction failure, documentation migration, population change, and emergent workaround.

Maintenance also requires user reporting, disablement, rollback, and retirement pathways rather than metric surveillance alone [21]. The proposed learning states are continued operation, enhanced surveillance, learning freeze, restricted inference, controlled update and revalidation, rollback, or retirement. Adaptive clinical decision support requires transparent oversight and controlled change [22]. Accordingly, no drift signal should directly authorize learning; labels created during unresolved high-consequence disruption should be quarantined where feasible; target redefinition should be explicit; and resumption of learning should require accountable authorization.

These states and constraints are proposed. Their thresholds, comparative effects, workload, and medication-safety consequences require local empirical validation.

Drift-Evaluation and Governance Requirements

Evaluation must distinguish detection performance from usefulness in live medication-related work. Early clinical evaluation of AI decision support should examine workflow integration, human factors, errors, and failures rather than model performance alone [23]. The framework therefore requires assessment of signal validity, temporal sensitivity, false-alarm burden, localization, professional usability, clinical task performance, medication safety, subgroup stability, and rollback reliability.

Trustworthy healthcare AI evaluation also encompasses fairness, universality, traceability, usability, robustness, and explainability across the lifecycle [24]. Bias assessment should examine data, modeling, implementation, and subgroup consequences rather than relying on one aggregate fairness measure [25]. Continual-learning systems additionally require predefined change-control and monitoring boundaries [26].

The proposed transition gates are: signal validity, localization, causal plausibility, consequence, evidence sufficiency, accountable authorization, and postchange requalification. Passing these gates would document a reasoned decision; it would not independently establish effectiveness, regulatory compliance, or unrestricted deployment readiness.

Implementation Implications

Implementation should anticipate complexity, nonadoption, abandonment, scale-up, and sustainability rather than treating technical installation as completion [27]. Context assessment should consider the intervention, inner and outer settings, involved individuals, and implementation process [28]. Workflow fit, trust, training, technical integration, and organizational resources may facilitate or obstruct routine AI use [29]. Operational deployment also requires explicit infrastructure for triggering, data flow, feedback, monitoring, evaluation, and controlled release [30].

Table 3 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for governance responsibilities, implementation conditions, and monitoring requirements for the article Detecting Workflow Drift before Pharmacy Algorithms Learn the Wrong Practice.

Table 3. Governance responsibilities, implementation conditions, and monitoring requirements for the article Detecting Workflow Drift before Pharmacy Algorithms Learn the Wrong Practice.

Governance domain	Responsible actor	Required authority	Documentation requirement	Monitoring indicator	Escalation trigger	Corrective action	Accountability boundary
Workflow baseline	Pharmacy service owner	Approve intended task and roles	Versioned workflow contract	Departure from intended work	Material unexplained change	Review and revise baseline	Baseline approval does not validate safety
Technical surveillance	Informatics or model team	Inspect system and data pipelines	Model, data, and release logs	Input, output, or feedback change	Persistent or high-consequence signal	Investigate, freeze, or roll back	Technical staff do not authorize clinical use alone
Clinical diagnosis	Pharmacist-led multidisciplinary team	Examine workflow meaning	Drift investigation record	Task, role, or sequence change	Plausible medication consequence	Contain and assess locally	Association does not establish causation
Safety and equity	Safety and equity functions	Require restriction or further evidence	Safety and subgroup assessment	Harm, access, or subgroup instability	Credible unequal or unsafe effect	Restrict, redesign, or revalidate	Aggregate stability is insufficient
Learning authorization	Lifecycle-governance body	Approve update, resumption, or retirement	Decision rationale and uncertainty	Completion of transition gates	Insufficient evidence or failed validation	Continue freeze, roll back, or retire	Development does not confer authorization
Frontline reporting	Pharmacy workforce	Report mismatch and workaround	Structured incident or concern record	Recurrent usability or workflow problem	Repeated or severe reports	Investigate configuration and work system	Reports initiate review but do not prove drift
Release and recovery	Independent validation and release functions	Release or withdraw versions	Validation and rollback record	Local performance and workflow fit	Failed postchange evaluation	Roll back or withdraw	Implementation does not establish validated benefit

Figure 2 presents the original conceptual synthesis for drift detection, diagnosis, response, and learning constraints, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

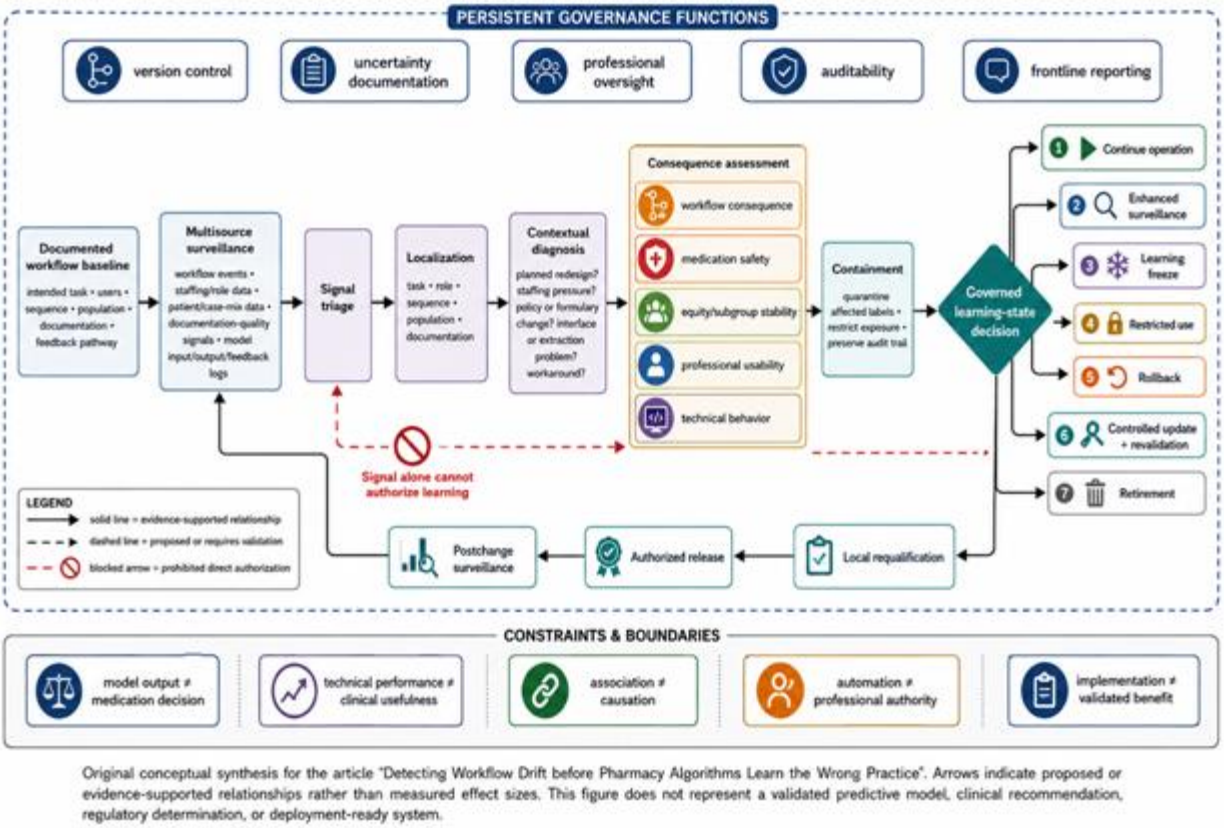

Figure 2. Drift Detection, Diagnosis, Response, and Learning Constraints

Limitations

A workflow baseline may itself encode inappropriate practice. Proxy targets can reproduce structural inequity even when they appear operationally reasonable [31]. Aggregate performance may also conceal systematic underperformance in underserved groups [32]. EHR-derived learning can encode selection, access, missingness, clinician behavior, and documentation biases [33], while subgroup consequences may change after deployment through fairness drift [34].

The framework is further limited by incomplete event logs, unrecorded communication, overlapping drift dimensions, low-volume tasks, and uncertain causal attribution. Planned improvement may resemble harmful drift, whereas freezing learning may preserve an already degraded model. The framework therefore remains a testable governance proposition rather than a validated monitoring standard.

Conclusion

Pharmacy algorithms do not learn from an abstract clinical reality; they learn from work as selected, sequenced, performed, and documented within particular organizations. When that work changes, an algorithm may absorb temporary workarounds, redistributed authority, documentation artifacts, altered populations, or unsafe practice while conventional performance indicators remain reassuring.

The proposed Multidimensional Pharmacy Workflow-Drift Framework makes this risk governable by connecting a versioned workflow baseline to task, role, sequence, population, and documentation surveillance. It separates detection from diagnosis, technical performance from clinical usefulness, model output from professional authority, and implementation from validated benefit. Its central governance rule is that observed change cannot directly authorize learning.

The framework offers a basis for empirical research on signal validity, workflow localization, causal investigation, medication-safety consequences, subgroup stability, professional usability, learning restraint, requalification, and rollback. It does not establish universal thresholds, legal responsibility, regulatory acceptability, clinical effectiveness, or deployment readiness. Its contribution is a structured means of asking whether an algorithm is learning from appropriate practice before changed practice becomes embedded in future decisions.

ACKNOWLEDGMENTS: None
CONFLICT OF INTEREST: None
FINANCIAL SUPPORT: None
ETHICS STATEMENT: None

References

1. Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, Zittrain J, et al. The clinician and dataset shift in artificial intelligence. N Engl J Med. 2021;385(3):283-6. doi:10.1056/NEJMc2104626
2. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195. doi:10.1186/s12916-019-1426-2
3. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: A roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337-40. doi:10.1038/s41591-019-0548-6
4. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA. 2017;318(6):517-8. doi:10.1001/jama.2017.7797
5. Hatzimanolis J, Riley B, El-Den S, Aslani P, Zhou J, Chaar BB. Applications of artificial intelligence in current pharmacy practice: A scoping review. Res Social Adm Pharm. 2025;21(3):134-41. doi:10.1016/j.sapharm.2024.12.007
6. Lu J, Liu A, Dong F, Gu F, Gama J, Zhang G. Learning under concept drift: A review. IEEE Trans Knowl Data Eng. 2019;31(12):2346-63. doi:10.1109/TKDE.2018.2876857
7. Sato DMV, de Freitas SC, Barddal JP, Scalabrin EE. A survey on concept drift in process mining. ACM Comput Surv. 2022;54(9):189:1-189:38. doi:10.1145/3472752
8. De Roock E, Martin N. Process mining in healthcare—an updated perspective on the state of the art. J Biomed Inform. 2022;127:103995. doi:10.1016/j.jbi.2022.103995
9. Adams JN, van Zelst SJ, Rose T, van der Aalst WMP. Explainable concept drift in process mining. Inf Syst. 2023;114:102177. doi:10.1016/j.is.2023.102177
10. Carayon P, Wooldridge A, Hose BZ, Salwei M, Benneyan J. Challenges and opportunities for improving patient safety through human factors and systems engineering. Health Aff (Millwood). 2018;37(11):1862-9. doi:10.1377/hlthaff.2018.0723
11. Carayon P, Wooldridge A, Hoonakker P, Schoofs Hundt A, Kelly MM. SEIPS 3.0: Human-centered design of the patient journey for patient safety. Appl Ergon. 2020;84:103033. doi:10.1016/j.apergo.2019.103033
12. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7
13. Graafsma J, Murphy RM, van de Garde EMW, Karapinar-Çarkit F, Derijks HJ, Hoge RHL, et al. The use of artificial intelligence to optimize medication alerts generated by clinical decision support systems: A scoping review. J Am Med Inform Assoc. 2024;31(6):1411-22. doi:10.1093/jamia/ocae076
14. Hussain MI, Reynolds TL, Zheng K. Medication safety alert fatigue may be reduced via interaction design and clinical role tailoring: A systematic review. J Am Med Inform Assoc. 2019;26(10):1141-9. doi:10.1093/jamia/ocz095
15. Kore A, Abbasi Bavil E, Subasri V, Abdalla M, Fine B, Dolatabadi E, et al. Empirical data drift detection experiments on real-world medical imaging data. Nat Commun. 2024;15(1):1887. doi:10.1038/s41467-024-46142-w
16. Lewis AE, Weiskopf N, Abrams ZB, Foraker R, Lai AM, Payne PRO, et al. Electronic health record data quality assessment and tools: A systematic review. J Am Med Inform Assoc. 2023;30(10):1730-40. doi:10.1093/jamia/ocad120
17. Moy AJ, Hobensack M, Marshall K, Vawdrey DK, Kim EY, Cato KD, et al. Understanding the perceived role of electronic health records and workflow fragmentation on clinician documentation burden in emergency departments. J Am Med Inform Assoc. 2023;30(5):797-808. doi:10.1093/jamia/ocad038
18. Liu S, McCoy AB, Peterson JF, Lasko TA, Sittig DF, Nelson SD, et al. Leveraging explainable artificial intelligence to optimize clinical decision support. J Am Med Inform Assoc. 2024;31(4):968-74. doi:10.1093/jamia/ocae019
19. Subbaswamy A, Saria S. From development to deployment: Dataset shift, causality, and shift-stable models in health AI. Biostatistics. 2020;21(2):345-52. doi:10.1093/biostatistics/kxz041
20. Feng J, Phillips RV, Malenica I, Bishara A, Hubbard AE, Celi LA, et al. Clinical artificial intelligence quality improvement: Towards continual monitoring and updating of AI algorithms in healthcare. NPJ Digit Med. 2022;5(1):66. doi:10.1038/s41746-022-00611-y
21. Davis SE, Embí PJ, Matheny ME. Sustainable deployment of clinical prediction tools—a 360° approach to model maintenance. J Am Med Inform Assoc. 2024;31(5):1195-8. doi:10.1093/jamia/ocae036
22. Petersen C, Smith J, Freimuth RR, Goodman KW, Jackson GP, Kannry J, et al. Recommendations for the safe, effective use of adaptive CDS in the US healthcare system: An AMIA position paper. J Am Med Inform Assoc. 2021;28(4):677-84. doi:10.1093/jamia/ocaa319
23. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33. doi:10.1038/s41591-022-01772-9
24. Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ. 2025;388. doi:10.1136/bmj-2024-081554
25. Wang HE, Landers M, Adams R, Subbaswamy A, Kharrazi H, Gaskin DJ, et al. A bias evaluation checklist for predictive models and its pilot application for 30-day hospital readmission models. J Am Med Inform Assoc. 2022;29(8):1323-33. doi:10.1093/jamia/ocac065
26. Vokinger KN, Feuerriegel S, Kesselheim AS. Continual learning in medical devices: FDA's action plan and beyond. Lancet Digit Health. 2021;3(6). doi:10.1016/S2589-7500(21)00076-5
27. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res. 2017;19(11). doi:10.2196/jmir.8775
28. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci. 2022;17(1):75. doi:10.1186/s13012-022-01245-0
29. Wenderott K, Krups J, Weigl M, Wooldridge AR. Facilitators and barriers to implementing AI in routine medical imaging: Systematic review and qualitative analysis. J Med Internet Res. 2025;27. doi:10.2196/63649
30. Corbin CK, Maclay R, Acharya A, Mony S, Punnathanam S, Thapa R, et al. DEPLOYR: A technical framework for deploying custom real-time machine learning models into the electronic medical record. J Am Med Inform Assoc. 2023;30(9):1532-42. doi:10.1093/jamia/ocad114
31. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447-53. doi:10.1126/science.aax2342
32. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nat Med. 2021;27(12):2176-82. doi:10.1038/s41591-021-01595-0
33. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. JAMA Intern Med. 2018;178(11):1544-7. doi:10.1001/jamainternmed.2018.3763
34. Davis SE, Dorn C, Park DJ, Matheny ME. Emerging algorithmic bias: Fairness drift as the next dimension of model maintenance and sustainability. J Am Med Inform Assoc. 2025;32(5):845-54. doi:10.1093/jamia/ocaf039