Pharmacophore

ISSN-2229-5402

Journal home page: http://www.pharmacophorejournal.com

From Digital Health Signals to Medication Decisions: A Conceptual Framework for Managing Temporal Drift, Missingness, and Intervention Feedback in Artificial Intelligence–Enabled Longitudinal Pharmaceutical Care

Amelia Clarke[1]*, Carlos Mendoza[2], Hassan Al-Rashid[3], Julia Schmidt[4]

1. ***Department of Digital Health and Longitudinal Pharmaceutical Care, University of Melbourne, Melbourne, Australia.***
2. ***Department of Health Data Science and Pharmaceutical Care, Universidad de los Andes, Bogota, Colombia.***
3. ***Department of AI and Pharmaceutical Decision Science, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.***
4. ***Department of Pharmaceutical Data Science and Digital Therapeutics, University of St. Gallen, St. Gallen, Switzerland.***

ARTICLE INFO

Received:
21 May 2026
Received in revised form:
12 September 2026
Accepted:
19 September 2026
Available online:
28 October 2026

Keywords: Digital health, Artificial intelligence, Medication management, Temporal drift, Missing data, Longitudinal pharmaceutical care

ABSTRACT

Digital health technologies increasingly produce continuous or frequently repeated observations that could inform medication initiation, titration, monitoring, adherence support, and treatment review. Yet longitudinal availability does not guarantee longitudinal interpretability. A change in a wearable, mobile, or remotely acquired signal may reflect biological evolution, treatment response, altered behavior, sensor or platform change, data-quality deterioration, incomplete observation, or the downstream consequences of an earlier clinical intervention. Artificial intelligence can amplify this ambiguity when models treat incoming observations as comparable with historical training data and independent of prior actions. This article develops a conceptual account of how digital health signals should be translated into medication decisions when temporal drift, missingness, and intervention feedback coexist. The central proposition is that medication-oriented inference should distinguish changes in patient state from changes in the measurement system, observation process, model state, and intervention history before a longitudinal trajectory is treated as clinically actionable. The analysis further separates model updating from clinical escalation: recalibration, additional measurement, clinician review, model withholding, and retraining address different failure mechanisms and should not be treated as interchangeable responses to declining confidence. The resulting framework is intended to support longitudinal pharmaceutical care without assuming that repeated digital measurements are stationary, complete, or unaffected by previous therapeutic decisions.

*This is an **open-access** article distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 4.0 License, which allows others to remix, and build upon the work non commercially.*

To Cite This Article: Clarke A, Mendoza C, Al-Rashid H, Schmidt J. From Digital Health Signals to Medication Decisions: A Conceptual Framework for Managing Temporal Drift, Missingness, and Intervention Feedback in Artificial Intelligence–Enabled Longitudinal Pharmaceutical Care. Pharmacophore. 2026;17(5):74-81. https://doi.org/10.51847/kRuuDhFKso

Introduction

Digital health signals occupy an increasingly important position between patient experience and clinical action. Wearable sensors, connected devices, mobile applications, home monitoring technologies, and software-derived measures can generate observations at a frequency that conventional encounters rarely achieve. Their appeal for pharmaceutical care is evident: repeated measurements can potentially reveal emerging toxicity, loss of disease control, treatment response, adherence-related change, or the need for medication review before these developments become visible through intermittent appointments. However, the clinical meaning of a digital signal depends on more than its technical availability. A measure intended to support treatment decisions must pass through distinct stages of verification, analytical validation, and clinical validation, each tied to a specific context of use [1]. Longitudinal density therefore cannot substitute for evidence that the signal being measured is reliable, interpretable, and appropriate for the decision in question.

Corresponding Author: Amelia Clarke; Department of Digital Health and Longitudinal Pharmaceutical Care, University of Melbourne, Melbourne, Australia. E-mail: amelia.clarke@unimelb.edu.au.

The same problem appears when wearable data are transformed into digital biomarkers. Longitudinal data acquisition, processing, feature construction, contextual interpretation, and validation are interdependent, and weaknesses introduced at one stage can propagate through later inference [2]. A repeatedly measured variable may therefore acquire the appearance of stability while its underlying measurement conditions change, or appear to deteriorate while the patient remains clinically stable. For medication decisions, this distinction is consequential. A dose escalation prompted by an apparent decline, for example, may be inappropriate if the decline reflects a device or processing change; conversely, attributing a real physiological deterioration to technical instability may delay necessary treatment.

Digital sensing research also faces broader challenges concerning standardization, representativeness, data quality, and the relationship between passive observations and clinically meaningful states [3]. These challenges become more difficult when observations are accumulated across weeks or months because the system producing the data is itself evolving. Patients change their routines, devices are replaced, software is updated, engagement fluctuates, health status changes, clinicians intervene, and medications alter the very physiological or behavioral signals being monitored. A longitudinal pharmaceutical-care model must therefore answer a more demanding question than whether a signal predicts an outcome at one point in time: it must determine whether the meaning of a signal remains sufficiently comparable across time to justify a medication decision.

The conceptual problem addressed here is consequently one of state discrimination. At any decision point, an observed digital trajectory may reflect at least four interacting processes: change in the patient, change in the measurement system, change in what is observed or missing, and change generated by previous treatment or clinical action. Artificial-intelligence systems add a fifth process because model performance, calibration, feature relationships, and uncertainty may themselves evolve after deployment. A defensible longitudinal medication decision requires these processes to be distinguished to the extent possible rather than collapsed into a single category of "signal change."

This article develops that distinction for artificial intelligence–enabled longitudinal pharmaceutical care. The analysis first identifies the major vulnerabilities that can destabilize the meaning of repeated digital measurements. It then examines temporal drift, missingness, and intervention feedback as separate but interacting mechanisms. These mechanisms provide the basis for a signal-to-medication-decision framework in which recalibration, additional measurement, clinical review, model withholding, and model updating are treated as different responses to different inferential problems. The objective is not to impose a universal algorithmic threshold for medication action. It is to clarify what must be known, monitored, or questioned before a longitudinal digital trajectory is allowed to influence pharmacotherapeutic decisions.

Challenges of Longitudinal Digital Health Evidence

One difficulty in interpreting longitudinal digital data is that genuine biological signals are often temporally structured. Wearable accelerometry, for example, can capture circadian patterns associated with age and health, demonstrating that variation across time may be biologically meaningful rather than random noise [4]. The medication-decision problem is therefore not simply to remove temporal variability but to determine whether an observed change is consistent with expected biological dynamics, clinically relevant deterioration or improvement, or a change in the way the signal is being generated. A model that interprets every departure from a prior baseline as equivalent risks confusing normal rhythmicity with treatment-relevant change.

Measurement validity creates a second vulnerability. Sensor-derived digital measures require appropriate reference measures, yet the suitability of a reference depends on the construct being measured and the intended context of use [5]. A device may perform acceptably for one analytical purpose while remaining insufficient for another. This matters longitudinally because an apparently stable algorithm can be fed by measurements whose relationship to the intended clinical construct varies across devices, settings, or periods of observation. Pharmaceutical interpretation therefore requires attention not only to model output but also to the integrity of the measurement chain beneath that output.

Wearable artificial intelligence can enhance monitoring and potentially support clinical decision-making, but its safety depends on maintaining a distinction between algorithmic inference and the medication action ultimately taken [6]. A prediction that a patient is deteriorating, for instance, is not equivalent to evidence that a specific drug should be initiated, stopped, or titrated. The latter requires pharmacological context, consideration of alternative explanations, and assessment of whether the monitored variable is sufficiently specific to the therapeutic question. This separation becomes particularly important when repeated model outputs create an impression of certainty through frequency alone.

The same constraint is visible in the regulatory treatment of digital health technologies used in drug development. Digital endpoints and measures need evidence appropriate to their defined context of use, and their acceptability is not transferable automatically between applications [7]. A measure validated for endpoint assessment in one clinical-development setting cannot therefore be assumed to support medication adjustment in routine longitudinal care. Context changes the evidentiary question: whether a measure can characterize an outcome is not identical to whether it can safely trigger therapeutic intervention.

Data provenance adds another layer. Digital biomarkers may depend on complex chains of acquisition, storage, transfer, preprocessing, and analysis, making authenticity and provenance relevant to confidence in the resulting observation [8]. In longitudinal pharmaceutical care, a discontinuity caused by a device replacement, altered data-processing route, or uncertain source should not be interpreted in the same manner as a biologically continuous trajectory. Provenance problems are not equivalent to clinical invalidity, but they weaken the basis for assuming that observations before and after a system change are directly comparable.

Finally, the digital-biomarker literature shows substantial heterogeneity in data modalities, computational methods, and intended clinical interpretations [9]. This heterogeneity argues against treating "digital health signal" as a single measurement class. Accelerometry, physiological sensing, interaction data, app-derived behavior, and other digitally acquired measures differ in their relationship to patient state and in the kinds of error or missingness they can generate. A medication-oriented framework must therefore preserve modality-specific limitations while providing a common logic for deciding when longitudinal evidence remains interpretable.

These vulnerabilities can be organized by the point at which longitudinal meaning becomes unstable and by the medication-decision risk that follows from that instability **Table 1**.

Table 1. Longitudinal evidence failures that can misdirect medication decisions

Analytical unit	What may change across time	Why the trajectory can become misleading	Medication-decision risk	Question that must be resolved before action
Biological temporal structure	Circadian, behavioral, disease-related, or treatment-responsive state	Expected biological variability may resemble deterioration or improvement	Unnecessary escalation, premature de-escalation, or false reassurance	Is the change outside the expected temporal behavior of the patient-linked signal?
Measurement reference relationship	Device performance, reference procedure, calibration environment, acquisition conditions	Apparent continuity may conceal changing measurement validity	A technically generated shift may be treated as a therapeutic signal	Is the current measurement still comparable with the reference basis used to justify interpretation?
AI-derived signal interpretation	Feature relationships, model confidence, preprocessing, inferential context	Repeated predictions can appear authoritative despite changing input conditions	Model output may be converted too directly into medication action	Does the prediction remain valid for the current patient, setting, and therapeutic question?
Context-of-use alignment	Clinical purpose, care setting, regulatory use, treatment stage	Evidence supporting one use may not support another	A measure may be used to trigger action beyond its validated purpose	Is the contemplated medication decision within the signal's supported context of use?
Data provenance and integrity	Device identity, software path, data transfer, preprocessing origin	System changes can create discontinuities unrelated to patient state	Technical artifacts may be interpreted as medication response or failure	Can the origin and processing history of the relevant observations be established?
Modality-specific meaning	Sensor type, behavioral measure, patient interaction, physiological channel	Different modalities have different error structures and clinical specificity	Heterogeneous signals may be treated as interchangeable evidence	What does this modality measure directly, and what clinical inference remains indirect?

Note: The table is a proposed synthesis of the evidence base. It does not assign numerical probabilities or universal action thresholds.

Temporal Drift, Missingness and Intervention Feedback

The challenges above become clinically significant when they evolve after deployment. Dataset shift describes a change in the statistical relationships connecting model inputs, outcomes, and the environment in which prediction occurs. In clinical AI, those relationships are historically contingent: patient populations, documentation practices, treatment strategies, technologies, and workflows can change after a model has been developed [10]. For longitudinal pharmaceutical care, the implication is that a signal should not be interpreted solely by asking whether its value has changed. The system must also ask whether the relationship between that value and the medication-relevant outcome has changed.

Implementation itself can contribute to this instability. Longitudinal evaluations of clinical deterioration systems show that predictive performance and clinical effects cannot be separated cleanly from workflow, deployment design, and the way staff respond to model outputs [11]. A longitudinal medication-support system likewise operates inside a care process. Changes in who reviews alerts, how rapidly measurements are acted upon, or which observations are available at the point of decision may alter the practical meaning of identical model scores. Temporal drift can therefore occur at the level of the model, the data-generating system, or the clinical workflow in which the prediction is interpreted.

Medication decision support makes the feedback problem especially visible. AI approaches have been used to optimize medication-related alerts, but the resulting systems do not merely observe clinical behavior; they can alter which alerts clinicians see and how clinicians respond [12]. Once a recommendation changes prescribing, monitoring intensity, follow-up, or patient behavior, later data are no longer generated under the same conditions that produced earlier observations. This creates a conceptual dependence between prediction and future evidence. A system that ignores that dependence may incorrectly interpret treatment-generated changes as independent confirmation of its earlier inference.

Uncertainty further complicates this loop. Clinical prediction systems can produce estimates that appear precise while their calibration or uncertainty varies across tasks and settings [13]. A medication framework should therefore preserve uncertainty through the decision process instead of allowing it to disappear once a model emits a score or category. High uncertainty does not automatically mandate a particular action, but it changes which actions are defensible. In some circumstances, the appropriate response is additional measurement; in others, clinician review, temporary model withholding, or reassessment of the measurement process may be more appropriate than recalibration.

Model adaptation is one possible response to evolving data. Incremental machine-learning approaches demonstrate that clinical record processing can be updated as new information arrives rather than remaining fixed indefinitely [14]. Yet the availability of an updating mechanism does not establish that continuous updating is always desirable. If an apparent performance problem is caused by corrupted measurements, selective missingness, or a temporary workflow disruption, retraining on those data may encode the failure into the model rather than correct it. Updating should therefore follow a diagnosis of the instability rather than substitute for that diagnosis.

Population heterogeneity makes the problem more difficult because aggregate performance can remain apparently stable while subgroup behavior changes. Longitudinal evaluation of clinical prediction has shown that fairness can drift over time and that model maintenance can affect groups differently [15]. In medication-oriented systems, this creates a requirement to distinguish overall temporal stability from clinically relevant subgroup stability. A model that remains calibrated on average may still become unreliable for a patient subgroup whose data distribution, treatment patterns, or measurement conditions have changed.

Missingness belongs within the same logic. In longitudinal care, absence of data is not simply an empty cell between two observations. It can coincide with changes in device use, patient engagement, care setting, measurement availability, or clinical state. Operational data-quality work in multicentre clinical datasets demonstrates the need for active, continuing assurance rather than an assumption that incoming information is uniformly complete and comparable [16]. The interpretive consequence is important: when the reason for an unobserved interval is uncertain, a model should not treat the surrounding trajectory as though continuous observation had occurred. Nor should imputation or interpolation be allowed to create false confidence about a medication-relevant trend.

Temporal drift, missingness, and intervention feedback therefore interact. A medication change can alter physiology; altered physiology can change the digital signal; the intervention can also change monitoring intensity or patient engagement; missing observations may then occur nonuniformly; and the model may encounter a distribution different from its development environment. The central analytical task is to identify which part of this chain has changed sufficiently to invalidate direct comparison with prior observations.

This distinction produces different monitoring responses. A suspected measurement-system discontinuity calls first for verification of acquisition and comparability, whereas evidence of model calibration drift may justify recalibration. Unexplained gaps may require additional measurement rather than prediction from sparse data. High-stakes discordance between model output and the broader clinical picture should trigger clinician review. When the system cannot establish that the current input distribution remains within an interpretable range, temporary model withholding may be safer than forcing a prediction. Retraining becomes appropriate only when the instability is sufficiently characterized and when new data represent the target deployment environment rather than the failure state itself.

The resulting response logic is summarized in **Table 2**. The rules are intentionally qualitative because the approved evidence does not justify universal numerical thresholds for pharmaceutical action.

Table 2. Proposed monitoring and response rules for unstable longitudinal medication evidence

Observed longitudinal condition	Competing explanations that must be distinguished	Preferred initial response	When escalation is warranted	Why automatic retraining is inappropriate
Abrupt signal change without parallel clinical evidence	Biological change; device or platform discontinuity; acquisition error	Verify measurement continuity and obtain corroborating information	Clinical review when the signal is clinically consequential or discordance persists	Retraining could absorb a technical artifact as if it were biology
Gradual degradation in prediction reliability	Patient/population evolution; calibration drift; workflow change	Reassess calibration and deployment conditions	Recalibration or model revision when degradation is reproducible and relevant to the target setting	The source of degradation may not be a learnable population shift
Intermittent or prolonged data absence	Reduced engagement; device non-use; infrastructure failure; care-context change; unobserved clinical change	Characterize the observation gap and seek additional measurement when feasible	Withhold or downgrade model-supported action when missingness materially weakens interpretation	Training on selectively observed data may reinforce the observation bias
Rising model uncertainty with otherwise plausible inputs	Model unfamiliarity; changed feature relationships; insufficient information	Preserve uncertainty and seek corroborating evidence	Clinical review or model withholding when the medication consequence is high stakes	Updating before identifying the cause can convert uncertainty into unstable confidence
Stable aggregate performance with subgroup deterioration	Differential population shift; changing treatment patterns; unequal measurement quality	Perform subgroup-sensitive monitoring	Targeted recalibration, validation, or withholding for affected groups	Global retraining may conceal or worsen subgroup-specific failure
Signal change following a	Intended treatment response; adverse effect; behavior	Interpret the new trajectory relative to intervention timing	Reassessment when post-intervention signals conflict with the clinical	Post-intervention data are not independent

medication or clinical intervention	change; altered monitoring intensity	and expected mechanisms	course or treatment objective	observations from the pre-intervention state
Persistent discordance between digital evidence and the clinical record	Measurement failure; model shift; incomplete clinical context; atypical patient response	Prioritize clinician review and evidence reconciliation	Temporary model withholding when discordance cannot be resolved	More training data do not resolve an unidentified source of incompatibility

Note: These are proposed mechanism-matched response rules. They are not validated universal thresholds and do not replace patient-specific pharmacological judgment.

Proposed Signal-to-Medication-Decision Framework

The central requirement for longitudinal pharmaceutical inference is to avoid treating every incoming digital observation as though it were generated under unchanged conditions. Health-system experience with clinical AI shows that implementation depends on workflow, governance, validation, and organizational context in addition to model performance [17]. The proposed framework therefore begins by separating five states that can evolve independently: the patient state, measurement-system state, observation state, model-inference state, and intervention state. A medication decision is defensible only when the available evidence permits sufficient discrimination among these states for the consequence under consideration.

Uncertainty should remain visible throughout this translation. Clinical AI research demonstrates that uncertainty proxies and calibration behavior vary across models and tasks [18]. Accordingly, a model output should not be converted directly into an action threshold. When uncertainty rises, the relevant question is whether the uncertainty arises from an unfamiliar patient trajectory, degraded measurement, sparse observation, altered feature relationships, or unresolved clinical discordance. The answer determines whether the next step should be additional measurement, clinician review, recalibration, temporary withholding, or another response.

Postdeployment evidence is equally important because development-time validation cannot establish indefinite longitudinal stability [19]. Monitoring should therefore evaluate not only predictive performance but also changes in input distributions, measurement provenance, data availability, subgroup behavior, and relationships between predictions and subsequent interventions. This transforms model maintenance from periodic technical servicing into a clinical inference problem.

Available evidence on dataset shift further indicates that detection and mitigation methods are heterogeneous and that no single correction strategy is universally preferable [20]. The framework consequently rejects automatic retraining as the default response to instability. Retraining is appropriate only after the changed process has been characterized sufficiently to establish that new observations represent the deployment environment that the model should learn, rather than corrupted measurements, selective observation, temporary workflow disruption, or intervention-generated states.

Empirical work also shows that harmful shifts can be detected and, in particular settings, mitigated by targeted adaptation [21]. The proposed framework therefore uses mechanism-matched response as its governing rule: identify which state has changed, assess whether that change compromises medication-relevant interpretation, and select the least assumptive corrective action capable of restoring defensible inference. **Figure 1** summarizes this architecture.

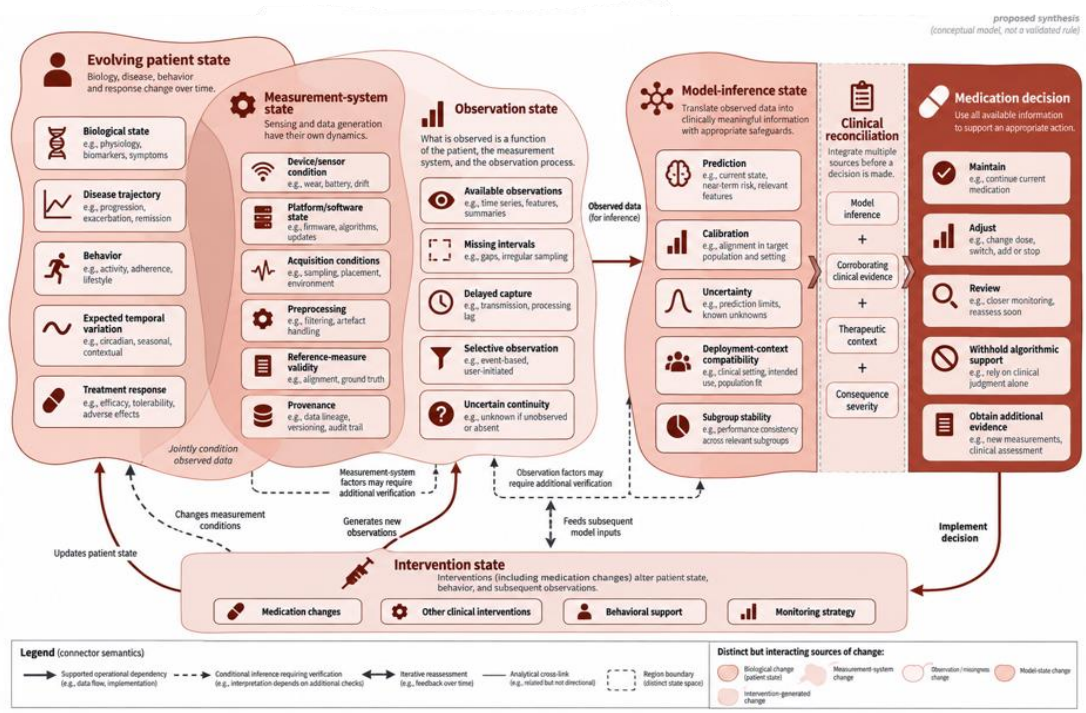


Figure 1. When a longitudinal digital signal can legitimately enter a medication decision

Clinical Implementation and Monitoring Implications

Medication-oriented monitoring should begin from the recognition that digital interventions can themselves change medication behavior. Evidence from cardiovascular care indicates that eHealth interventions can influence adherence, although effectiveness varies across intervention types and contexts [22]. A change in an adherence-related signal after a digital intervention therefore cannot be treated as independent confirmation of the model that prompted that intervention.

Medication applications also differ in the functions through which they attempt to support adherence, including reminder, educational, monitoring, and organizational components [23]. Consequently, engagement with an application is not equivalent to medication ingestion or pharmacological response. Longitudinal interpretation should identify what the observed digital behavior actually represents before it is allowed to influence therapeutic decisions.

Remote patient monitoring extends the same issue across repeated home measurements. Its value depends on integration between incoming data, clinical review, escalation pathways, and the broader care infrastructure [24]. Increasing the frequency of observation may improve temporal visibility, but it also creates more opportunities for false alarms, selective missingness, unresolved discordance, and intervention-generated feedback. Monitoring density should therefore be matched by interpretive capacity.

Pharmaceutical decisions additionally require domain-specific therapeutic context. Digital technologies may support antimicrobial stewardship and other medication decisions by combining clinical, microbiological, behavioral, and monitoring information, but they do not eliminate the need to interpret those data within prescribing constraints and patient-specific circumstances [25]. The relevant question is not whether the algorithm has detected change, but whether the change is sufficiently attributable, clinically consequential, and pharmacologically coherent to justify action.

Prospective telemonitoring studies also illustrate that monitoring and intervention form a coupled system: detected changes can trigger clinical actions that subsequently alter the trajectory being monitored [26]. Once such an intervention occurs, the post-intervention period should be treated as a new inferential state rather than as a continuation of untreated baseline conditions. This is particularly important when medication adjustment changes symptoms, physiology, behavior, monitoring intensity, or healthcare contact.

Feedback-loop research provides the final caution. When actions influence which future observations become available, biases can be reinforced if the resulting data are interpreted as though they were generated independently of prior decisions [27]. Longitudinal pharmaceutical care should therefore preserve intervention history as part of the model's interpretive context. **Figure 2** represents this temporal dependence and the points at which reassessment is required.

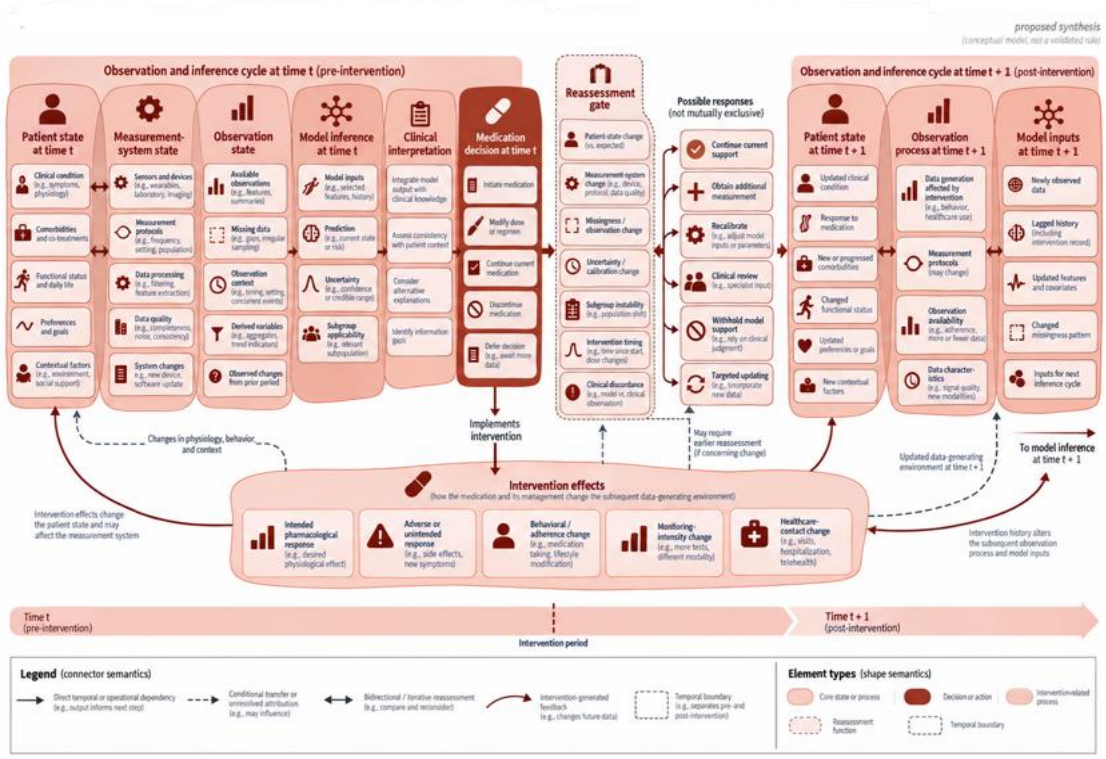


Figure 2. How medication action changes the evidence used for the next decision

Conclusion

Longitudinal digital evidence can support medication decisions only when repeated measurements remain interpretable in relation to the processes that generated them. Temporal density does not establish continuity of meaning. A trajectory may

change because the patient has changed, because the measurement or observation system has changed, because relevant data have become selectively unavailable, because the model's relationships have drifted, or because an earlier medication decision has altered the subsequent state being observed.

The central implication is that these possibilities should not be collapsed into a single technical category of model drift. Their distinction changes the appropriate response. Measurement discontinuity may require verification; uncertain observation may require additional data; calibration failure may justify recalibration; unresolved clinical discordance may require clinician review or temporary model withholding; and sufficiently characterized distributional change may justify targeted updating. Retraining is therefore one response among several, not the default consequence of instability.

Artificial intelligence–enabled pharmaceutical care becomes more defensible when medication decisions are treated as interventions within the longitudinal data-generating process rather than terminal outputs from it. The strongest use of digital health signals is consequently conditional: they should influence medication action only to the extent that patient change can be distinguished adequately from measurement change, missingness, model instability, and the feedback produced by previous therapeutic decisions.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, et al. Verification, analytical validation, and clinical validation (V3): The foundation of determining fit-for-purpose for biometric monitoring technologies (BioMeTs). NPJ Digit Med. 2020;3:55. doi:10.1038/s41746-020-0260-4
2. Daniore P, Nittas V, Haag C, Bernard J, Gonzenbach R, von Wyl V. From wearable sensor data to digital biomarker development: Ten lessons learned and a framework proposal. NPJ Digit Med. 2024;7(1):161. doi:10.1038/s41746-024-01151-3
3. Akre S, Seok D, Douglas C, Aguilera A, Carini S, Dunn J, et al. Advancing digital sensing in mental health research. NPJ Digit Med. 2024;7(1):362. doi:10.1038/s41746-024-01343-x
4. Shim J, Fleisch E, Barata F. Circadian rhythm analysis using wearable-based accelerometry as a digital biomarker of aging and healthspan. NPJ Digit Med. 2024;7(1):146. doi:10.1038/s41746-024-01111-x
5. Bakker JP, McClenahan SJ, Fromy P, Turner S, Peterson BT, Vandendriessche B, et al. A hierarchical framework for selecting reference measures for the analytical validation of sensor-based digital health technologies. J Med Internet Res. 2025;27:e58956. doi:10.2196/58956
6. Mahajan A, Heydari K, Powell D. Wearable AI to enhance patient safety and clinical decision-making. NPJ Digit Med. 2025;8(1):176. doi:10.1038/s41746-025-01554-w
7. Erdemli G, Murphy T, Walinsky S. Regulatory considerations for successful implementation of digital endpoints in clinical trials for drug development. NPJ Digit Med. 2025;8(1):142. doi:10.1038/s41746-025-01513-5
8. Mahajan A, Powell D. Improving authenticity and provenance in digital biomarkers: The case for digital watermarking. NPJ Digit Med. 2025;8(1):31. doi:10.1038/s41746-024-01374-4
9. Qi W, Zhu X, Wang B, Shi Y, Dong C, Shen S, et al. Alzheimer's disease digital biomarkers multidimensional landscape and AI model scoping review. NPJ Digit Med. 2025;8(1):366. doi:10.1038/s41746-025-01640-z
10. Lea AS, Jones DS. Mind the gap—machine learning, dataset shift, and history in the age of clinical algorithms. N Engl J Med. 2024;390(4):293-5. doi:10.1056/NEJMp2311015
11. van der Vegt AH, Campbell V, Mitchell I, Malycha J, Simpson J, Flenady T, et al. Systematic review and longitudinal analysis of implementing artificial intelligence to predict clinical deterioration in adult hospitals: What is known and what remains uncertain. J Am Med Inform Assoc. 2024;31(2):509-24. doi:10.1093/jamia/ocad220
12. 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
13. Acharya A, Shrestha S, Chen A, Conte J, Avramovic S, Sikdar S, et al. Clinical risk prediction using language models: Benefits and considerations. J Am Med Inform Assoc. 2024;31(9):1856-64. doi:10.1093/jamia/ocae030
14. Streiffer C, Saini D, Whitehead G, Daniel J, Garzon-Mrad C, Kavanaugh L, et al. The incremental design of a machine learning framework for medical records processing. J Am Med Inform Assoc. 2024;31(10):2236-45. doi:10.1093/jamia/ocae194
15. 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

16. Fu A, Shen T, Roberts SB, Liu W, Vaidyanathan S, Marchena-Romero KJ, et al. Optimizing the efficiency and effectiveness of data quality assurance in a multicenter clinical dataset. J Am Med Inform Assoc. 2025;32(5):835-44. doi:10.1093/jamia/ocaf042
17. Poon EG, Lemak CH, Rojas JC, Guptill J, Classen D. Adoption of artificial intelligence in healthcare: Survey of health system priorities, successes, and challenges. J Am Med Inform Assoc. 2025;32(7):1093-100. doi:10.1093/jamia/ocaf065
18. Savage T, Wang J, Gallo R, Boukil A, Patel V, Safavi-Naini SAA, et al. Large language model uncertainty proxies: Discrimination and calibration for medical diagnosis and treatment. J Am Med Inform Assoc. 2025;32(1):139-49. doi:10.1093/jamia/ocae254
19. Koski E, Das A, Hsueh PYS, Solomonides A, Joseph AL, Srivastava G, et al. Towards responsible artificial intelligence in healthcare—getting real about real-world data and evidence. J Am Med Inform Assoc. 2025;32(11):1746-55. doi:10.1093/jamia/ocaf133
20. Silva GFDS, Barcellos Filho FN, Wichmann RM, da Silva Junior FC, Chiavegatto Filho ADP. Strategies for detecting and mitigating dataset shift in machine learning for health predictions: A systematic review. J Biomed Inform. 2025;170:104902. doi:10.1016/j.jbi.2025.104902
21. Subasri V, Krishnan A, Kore A, Dhalla A, Pandya D, Wang B, et al. Detecting and remediating harmful data shifts for the responsible deployment of clinical AI models. JAMA Netw Open. 2025;8(6):e2513685. doi:10.1001/jamanetworkopen.2025.13685
22. Miao Y, Luo Y, Zhao Y, Liu M, Wang H, Wu Y. Effectiveness of eHealth interventions in improving medication adherence among patients with cardiovascular disease: Systematic review and meta-analysis. J Med Internet Res. 2024;26:e58013. doi:10.2196/58013
23. Wang T, Huang YM, Chan HY. Exploration of features of mobile applications for medication adherence in Asia: Narrative review. J Med Internet Res. 2024;26:e60787. doi:10.2196/60787
24. Paul MM, Khera N, Elugunti PR, Ruff KC, Hommos MS, Thomas LF, et al. The state of remote patient monitoring for chronic disease management in the United States. J Med Internet Res. 2025;27:e70422. doi:10.2196/70422
25. Rawson TM, Zhu N, Galiwango R, Cocker D, Islam MS, Myall A, et al. Using digital health technologies to optimise antimicrobial use globally. Lancet Digit Health. 2024;6(12):e914-25. doi:10.1016/S2589-7500(24)00198-5
26. Yun S, Comín-Colet J, Calero-Molina E, Hidalgo E, José-Bazán N, Cobo Marcos M, et al. Evaluation of mobile health technology combining telemonitoring and teleintervention versus usual care in vulnerable-phase heart failure management (HERMeS): A multicentre, randomised controlled trial. Lancet Digit Health. 2025;7(5):100866. doi:10.1016/j.landig.2025.02.006
27. Aikens RC, Chen JH, Baiocchi M, Simard JF. Feedback loop failure modes in medical diagnosis: How biases can emerge and be reinforced. Med Decis Making. 2024;44(5):481-96. doi:10.1177/0272989X241248612