Original Article

Multimodal Pharmacy Intelligence for Linking Prescriptions, Symptoms, Laboratory Trends, Images, and Speech

Santiago Morales[1]*, Laura Rojas[2], Camila Vega[1], Diego Molina[3]

[1]Department of Multimodal AI in Pharmacy, Faculty of Pharmacy, University of Buenos Aires, Buenos Aires, Argentina. [2]Department of Cross-Modal Pharmacy Data Integration, Faculty of Pharmacy, Pontifical Catholic University of Chile, Santiago, Chile. [3]Department of Multisensory Pharmacy Intelligence, Faculty of Pharmacy, National University of La Plata, La Plata, Argentina.

Abstract

Medication decisions are rarely supported by prescriptions alone. Pharmacists may need to relate medication orders to reported symptoms, longitudinal laboratory changes, clinical images, spoken information, documentation context, and prior therapeutic events. These information forms differ in structure, reliability, timing, provenance, and clinical meaning, creating a problem that cannot be resolved through simple data concatenation or a single predictive model. This article proposes a Multimodal Pharmacy Intelligence Architecture for organizing heterogeneous patient information while preserving uncertainty and professional decision authority. The architecture contains source and identity controls, modality-specific ingestion and representation, provenance and data-quality assessment, temporal normalization, cross-modal alignment, task-specific fusion, uncertainty management, and a pharmacist-facing evidence bundle. A proposed Pharmacy Alignment Graph represents relationships among observations without treating temporal proximity as causality. Missing, stale, conflicting, or low-quality modalities remain visible rather than being silently imputed or suppressed. A Pharmacist Decision Boundary separates computational transformation from medication verification, interpretation, communication, escalation, and action. Evaluation is organized across representation integrity, alignment accuracy, missing-modality robustness, task performance, human–AI interaction, workflow consequences, medication safety, equity, and local transportability. The architecture is an original non-empirical synthesis rather than a validated clinical system. Its usefulness is conditional on task-specific evidence, semantic interoperability, prospective human-factors assessment, subgroup evaluation, incident monitoring, and governed change control. Multimodal agreement does not establish clinical truth, and model output does not independently authorize medication decisions. The proposed contribution offers a structured basis for developing and evaluating traceable multimodal decision support in pharmacy practice.

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

Introduction

Pharmacy decisions are inherently multimodal because the clinical meaning of a prescription frequently depends on information that exists outside the medication order. Dose appropriateness may depend on laboratory trajectories; an adverse-effect concern may depend on the temporal relationship between medication exposure and symptoms; adherence questions may emerge from spoken accounts that differ from structured records; and an image may contextualize a wound, device, dosage form, or visible reaction. Multimodal biomedical artificial intelligence provides a technical basis for jointly processing heterogeneous information, but combining modalities does not by itself create clinically coherent evidence [1].

Electronic health records contain structured codes, medication orders, laboratory results, free text, timestamps, and other observations generated through different workflows. Their apparent coexistence within one record can obscure differences in completeness, temporal resolution, acquisition purpose, and reliability [2]. A symptom recorded after a prescription may have begun before treatment, a laboratory result may reflect a specimen obtained earlier, and a transcribed statement may differ from the acoustic or

Address for correspondence: Santiago Morales, Department of Multimodal AI in Pharmacy, Faculty of Pharmacy, University of Buenos Aires, Buenos Aires, Argentina.
santiago.morales@agro.uba.ar
Received: 19 October 2025; **Accepted:** 27 January 2026

How to cite this article: Morales S, Rojas L, Vega C, Molina D. Multimodal Pharmacy Intelligence for Linking Prescriptions, Symptoms, Laboratory Trends, Images, and Speech. Arch Pharm Pract. 2026;17(1):13-21. https://doi.org/10.51847/SFkaknFouO

conversational context in which it was produced. Consequently, record proximity is not equivalent to clinical alignment.

The pharmacy implications extend beyond information retrieval. Artificial intelligence used in prescribing may influence which medicines patients receive and therefore affect the safety and effectiveness evidence subsequently observed in practice [3]. This creates a feedback problem: a model may not merely interpret the medication-use system but may gradually alter the data-generating process from which later models and safety evaluations learn.

Task-specific pharmacy evidence nevertheless shows that computational systems may help prioritize prescriptions carrying a higher risk of medication error [4]. Such prioritization can organize attention, but it does not establish that the system understands the patient's complete therapeutic context or can independently determine the appropriate intervention. Similarly, artificial intelligence has been investigated for filtering or prioritizing medication alerts, yet limited datasets, heterogeneous methods, and incomplete formal validation constrain claims of clinical readiness [5].

This article therefore proposes a Multimodal Pharmacy Intelligence Architecture rather than a universal multimodal predictor. Its purpose is to define how prescriptions, symptoms, laboratory trends, images, and speech may be linked while preserving provenance, temporal ambiguity, missingness, disagreement, and professional authority. The architecture is designed to produce a traceable evidence bundle for pharmacist interpretation. It does not assign the system autonomous authority to prescribe, discontinue, verify, or override treatment.

Modalities, Temporal Scales, and Alignment Problems

A modality is defined here as an information form with a distinguishable acquisition process, representation, error structure, temporal behavior, and clinical interpretation. Medication orders record intended therapy but may not establish dispensing, administration, adherence, or exposure. Symptoms may be patient-reported, clinician-recorded, inferred from narrative, or encoded as structured data. Laboratory values are numerical observations tied to specimen and processing times. Images preserve visual information but may contain acquisition artefacts or context unrelated to the intended task. Speech includes acoustic characteristics, spoken content, interactional cues, and a transcript, each of which may require separate interpretation.

Longitudinal electronic health-record models demonstrate that heterogeneous clinical events can be represented across time without reducing the record to a single contemporaneous snapshot [6]. However, a model-ready sequence is not automatically a clinically valid sequence. Pharmacy reasoning may require distinguishing order entry, dispensing, administration, adherence, symptom onset, specimen acquisition, result availability, and pharmacist review.

Clinical time-series research further shows that sampling decisions, outcome definitions, observation windows, and forecast horizons materially shape a modelling task [7]. Laboratory measurements may occur minutes apart in acute care or months apart in chronic therapy. Symptoms may fluctuate continuously but be documented intermittently. A valid alignment process must therefore specify which time is represented, why two observations are considered related, and how uncertainty increases as temporal separation or documentation delay grows.

Speech can carry clinically informative acoustic and prosodic features, although the meaning and performance of such signals remain population- and task-specific [8]. In pharmacy, speech should not be treated as a direct medication-effect measure. It may instead contribute to clarification of symptoms, treatment experience, cognitive or communication difficulty, or discrepancies between spoken and recorded medication use, subject to consent, language, accessibility, and recording-quality constraints.

Medical imaging and electronic-record information can be combined through early, intermediate, late, or mixed fusion strategies [9]. These approaches are not interchangeable. Early fusion may expose interactions but amplify misalignment; late fusion may preserve modality independence but miss clinically relevant relationships. The appropriate strategy depends on the pharmacy task, modality quality, and consequences of error.

Cross-system linkage additionally requires syntactic, semantic, and organizational interoperability [10]. Shared patient identity or encounter membership is insufficient when terminology, timestamps, units, source systems, and documentation purposes differ. Alignment must therefore be represented as an explicit, confidence-bounded relationship rather than an assumed property of the record.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, components, relationships, and intended uses for the article Multimodal Pharmacy Intelligence for Linking Prescriptions, Symptoms, Laboratory Trends, Images, and Speech.

Table 1. Definitions, components, relationships, and intended uses for the article Multimodal Pharmacy Intelligence for Linking Prescriptions, Symptoms, Laboratory Trends, Images, and Speech.

Component or Scientific Dimension	Problem Addressed	Inputs or Determinants	Proposed Mechanism or Relationship	Expected Contribution	Evidence Required	Failure or Uncertainty Risk	Boundary Statement
Prescription State	An order may not represent actual exposure	Order, dispensing, administration, adherence	**Proposed:** maintain separate intended, supplied, administered, and reported-use states	More precise medication context	Source concordance and task validation; prescription prioritization evidence [4]	Treating an order as confirmed use	Does not establish ingestion or therapeutic appropriateness
Symptom State	Onset and documentation may differ	Patient report, narrative, structured code, speech	**Proposed:** retain reporter, onset time, recording time, and interpretation status	Distinguishes pre-existing from treatment-emergent concerns	Agreement and temporal-linkage studies	Retrospective attribution and semantic ambiguity	Does not establish medication causality
Laboratory Trajectory	Isolated values conceal direction and timing	Specimen time, result time, units, reference ranges	**Proposed:** align trends to medication and clinical events	Supports longitudinal contextualization	Temporal and external validation [6]	Unit mismatch, delayed results, irregular sampling	A trend is not automatically a medication effect
Image Observation	Visual evidence may be detached from context	Image, acquisition metadata, interpretation	**Proposed:** preserve image provenance and task-specific linkage	Adds visual context where relevant	Image–record fusion and task validation [9]	Acquisition artefact or unrelated visual cues	Images cannot independently determine medication action
Speech Observation	Transcript alone may omit acoustic or interactional context	Audio, transcript, language, environment	**Proposed:** separate acoustic, lexical, and interpreted features	May support clarification and discrepancy detection	Population-specific validation [8]	Noise, accent bias, language mismatch, privacy risk	Speech-derived inference is not equivalent to patient intent
Temporal Alignment	Record proximity may not represent event relatedness	Multiple event and documentation times	**Proposed Temporal Anchoring Contract**	Makes the reason and confidence for linkage visible	Event-level annotation and reliability assessment [7]	False chronology or inappropriate temporal window	Temporal proximity does not establish causality
Provenance And Interoperability	Sources may differ in meaning and transformation history	Origin, terminology, units, version, transformations	**Proposed provenance object** linked to each observation	Traceability and source verification	Semantic and technical interoperability evidence [10]	Untraceable transformations or incompatible semantics	Interoperability does not prove clinical validity
Pharmacy Evidence Bundle	A fused score can conceal disagreement	Aligned observations, uncertainty, conflicts, provenance	**Original proposed synthesis:** pharmacist-facing traceable bundle	Supports review without erasing source differences	Human-factors and clinical-task studies	Automation bias or information overload	The bundle supports, but does not authorize, decisions

The table's mechanisms and named constructs are proposed for empirical testing. They organize the information required for reasoning but do not establish that adding modalities improves pharmacist performance, medication safety, or patient outcomes.

Proposed Multimodal Pharmacy Architecture

Multimodal architecture requires more than placing heterogeneous inputs into one model. Representation, translation, alignment, fusion, and co-learning are distinguishable technical problems, and healthcare applications require modality combinations to be selected for the intended task rather than assumed to be universally beneficial. Interaction-aware mechanisms can represent cross-modal relationships beyond simple feature concatenation [11–13]. These mechanisms provide technical precedents, but they do not validate a pharmacy-specific configuration.

The proposed architecture contains eight layers. The source and identity layer identifies the patient, encounter, organization, consent state, and originating system. The modality-ingestion layer receives prescriptions, symptom records, laboratory observations, images, speech, transcripts, and care-context data without prematurely treating them as semantically equivalent. The quality and provenance layer records origin, units, timestamps, transformation history, completeness, and known limitations. The temporal-normalization layer distinguishes occurrence, acquisition, documentation, verification, and review times.

The modality-representation layer creates task-relevant representations while retaining access to the original observations. Separate encoders may be used for medication orders, clinical language, laboratory series, images, and speech. The alignment and fusion layer then links observations through a proposed Pharmacy Alignment Graph. Each link records the clinical event under consideration, temporal window, semantic basis, provenance, and alignment confidence. Fusion is conditional: information may be combined for one task yet remain separate for another.

The pharmacy-task layer generates bounded outputs such as prescription-review prioritization, discrepancy detection, laboratory-trend contextualization, retrieval of relevant evidence, or identification of information requiring clarification. The human decision and governance layer produces a Pharmacy Evidence Bundle that displays supporting observations, unresolved disagreement, missingness states, provenance, and uncertainty. A Pharmacist Decision Boundary separates these outputs from verification, interpretation, communication, escalation, and medication action.

Generalist medical artificial-intelligence architectures motivate reusable components that can process several modalities and tasks [14]. However, general capability does not establish authority for a particular medication decision. Each pharmacy task requires a defined intended use, excluded use, accountable professional, validation pathway, and stopping condition.

Figure 1 presents the original conceptual synthesis for multimodal pharmacy intelligence architecture, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

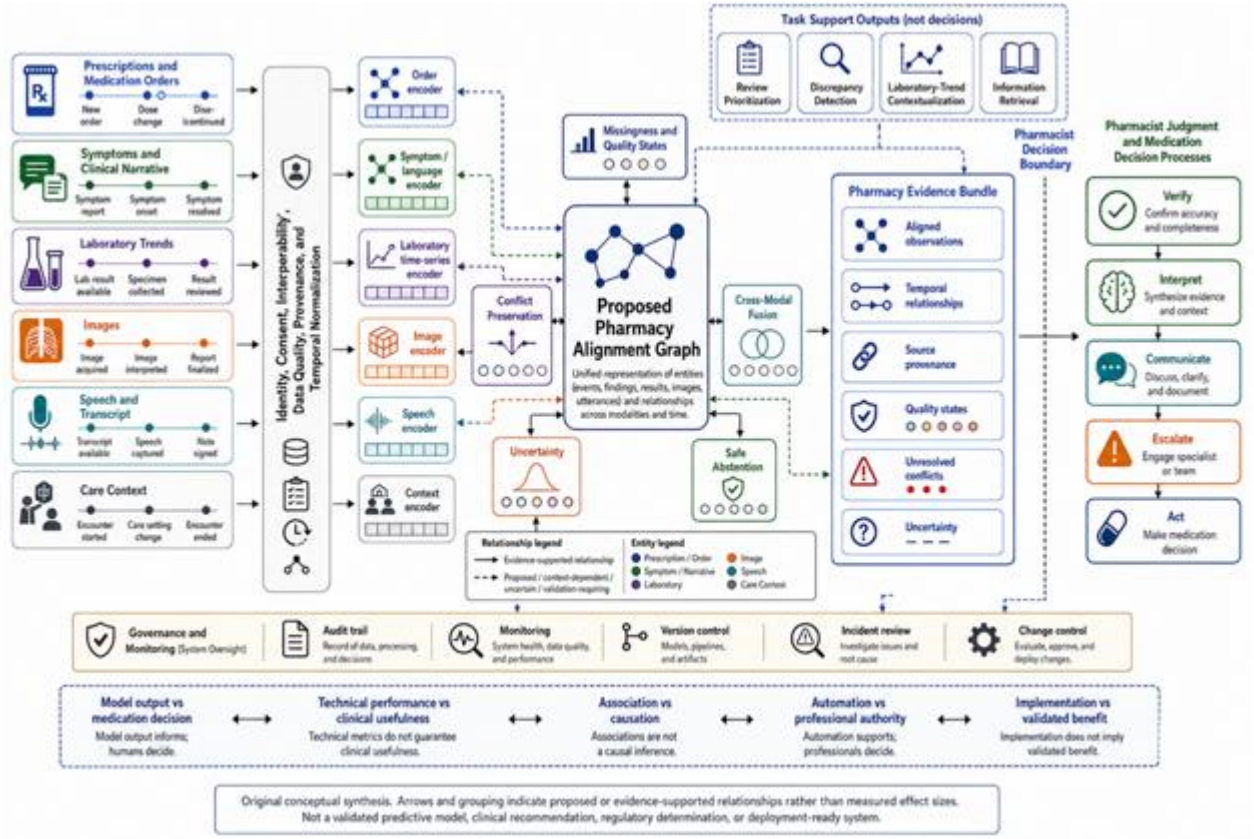

Figure 1. Multimodal Pharmacy Intelligence Architecture

The architecture additionally proposes a Safe-Abstention Gate. Abstention is not defined as model failure alone; it is a controlled response when the available information cannot support the intended task. Triggers may include unresolved patient identity, incompatible units, uncertain event linkage, missing critical modalities, severe cross-modal conflict, poor input quality, or observations outside the validated domain. The gate should specify what information is missing, whether verification is possible, and who is responsible for the next action.

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for evaluation criteria, evidence requirements, and failure modes for the article Multimodal Pharmacy Intelligence for Linking

Prescriptions, Symptoms, Laboratory Trends, Images, and Speech.

Table 2. Evaluation criteria, evidence requirements, and failure modes for the article Multimodal Pharmacy Intelligence for Linking Prescriptions, Symptoms, Laboratory Trends, Images, and Speech.

Architecture Layer or Process Stage	Core Function	Information Flow	Interaction With Other Components	Evaluation Criterion	Potential Failure Mode	Human or Organizational Responsibility	Readiness Boundary
Source And Identity	Establish patient and source context	Systems to controlled ingestion	Constrains every downstream link	Identity resolution and traceability	Wrong-patient or duplicate-record linkage	Data steward and local institution	No use when identity remains unresolved
Modality Ingestion	Preserve modality-specific observations	Native data to modality pathway	Feeds quality and provenance controls	Completeness and faithful extraction	Lost context or transformation error	System owner and modality specialist	Extraction accuracy does not establish clinical utility
Quality And Provenance	Record origin, units, time, and transformations	Source metadata to provenance object	Conditions alignment and fusion	Provenance completeness and source verifiability	Untraceable or incompatible data	Data steward and clinical informatics team	Unverifiable sources require restriction or abstention
Temporal Normalization	Distinguish event and documentation times	Timestamped observations to common event structure	Supports longitudinal reasoning	Event-order and temporal-link accuracy	False chronology or stale evidence	Clinical domain owner	Temporal order does not establish causation
Modality Representation	Create task-relevant features	Observation to representation	Preserves link to original evidence	Representation validity for intended task	Shortcut learning or loss of clinically important detail	Model developer and clinical reviewer	Encoder performance cannot authorize medication action
Alignment And Fusion	Relate observations and combine selected evidence	Representations to Pharmacy Alignment Graph	Uses provenance, quality, and task definition	Alignment reliability and controlled contribution of modalities	Inappropriate fusion or hidden conflict	Multidisciplinary validation team	Fusion requires task-specific validation [11]
Pharmacy-Task Layer	Produce bounded review support	Aligned evidence to task output	Feeds evidence bundle and abstention gate	Task performance, calibration, and error analysis	Useful metric but unsafe workflow behavior	Pharmacy clinical owner	Technical performance is not clinical usefulness
Evidence Bundle And Decision Boundary	Present evidence without transferring authority	Task output to pharmacist review	Enables verification and escalation	Interpretability, workload, reliance, and recovery	Automation bias or information overload	Pharmacist, human-factors lead, institution	Output remains advisory and context-dependent
Monitoring And Change Control	Detect altered data, models, or use	Use and incident data to governance process	May restrict or suspend components	Drift, provenance failure, incident, and version tracking	Silent degradation or uncontrolled updating	Governance body and accountable system owner	Implementation does not establish validated benefit

All readiness boundaries in **Table 2** are proposed. They indicate conditions requiring further evidence, restriction, or professional review and are not regulatory classifications or deployment thresholds.

Cross-Modal Fusion, Provenance, and Temporal Reasoning

Cross-modal fusion should be treated as a task-specific operation rather than a permanent merger of all available information. A model intended to detect a discrepancy between a spoken medication history and the prescription record requires a different alignment structure from one intended to contextualize renal laboratory trends after a dose change. The contribution of each modality should therefore depend on relevance, quality, timing, provenance, and whether the modality was expected for the task.

Missing-modality research demonstrates that multimodal systems can be designed to operate when one or more expected inputs are unavailable [15]. For pharmacy use, controlled degradation should be visible. The system should state which modality is absent, whether it was necessary or optional, how the output changed, and whether the remaining evidence still supports the intended task. An imputed modality must not be displayed as an observed patient fact.

Temporal reasoning requires comparable discipline. Multi-horizon forecasting methods distinguish static information,

observed time-varying inputs, known future inputs, and outputs associated with different horizons [16]. A pharmacy architecture may adapt these distinctions by separating relatively stable patient characteristics from evolving symptoms, medication exposures, laboratory values, scheduled doses, anticipated monitoring, and future decision points. Such structuring does not imply that temporal attention weights explain causality.

Longitudinal clinical prediction also shows that risk estimates can be updated as new laboratory and clinical events arrive [17]. In pharmacy, the relevant contribution is not the transfer of a particular predictive result but the principle that evidence status changes over time. A laboratory result may strengthen, weaken, or invalidate an earlier interpretation. A symptom documented after dose modification may require reassessment when its actual onset is clarified. The architecture should therefore retain versioned evidence bundles rather than overwrite earlier reasoning without an audit trail.

Provenance determines whether aligned information can be interpreted and challenged. Data suitability is conditional on the intended question and depends on relevance, quality, and usability rather than availability alone [18]. The proposed Pharmacy Provenance Object records the originating person or system, acquisition method, terminology and units, event time, recording time, transformations, model version, and known quality limitations. Every derived representation or fused output should remain traceable to these objects.

The proposed Temporal Anchoring Contract formalizes why observations are considered related. It contains the anchor event, permissible temporal window, direction of time, clinical rationale, source dependencies, alignment confidence, and conditions that invalidate the link. For example, relating a symptom to a medication change requires more than shared encounter membership: the architecture should distinguish symptom onset, patient report, documentation, order entry, dispensing, administration, and estimated exposure.

The proposed Pharmacy Alignment Graph represents observations as nodes and their hypothesized relationships as typed edges. Edges may indicate temporal precedence, possible exposure, semantic equivalence, corroboration, contradiction, or common provenance. Uncertain links remain uncertain; conflicting evidence remains accessible to the pharmacist. The graph is not a causal model unless separately designed and validated for causal inference.

Fusion may then occur at several levels. Early fusion may combine normalized features, intermediate fusion may exchange modality representations, and late fusion may combine modality-specific outputs. A hybrid design may preserve independent modality assessments while permitting selected interactions. The architecture does not designate one approach as universally superior. The choice should reflect the intended task, missingness pattern, need for traceability, error consequences, and evidence available for validation.

The final output is a Pharmacy Evidence Bundle containing the task question, aligned observations, provenance, temporal relationships, modality-quality states, unresolved conflicts, uncertainty, and recommended verification route. "Recommended" here refers to information-handling steps—such as checking a source, clarifying timing, or requesting missing evidence—not to a clinical medication recommendation. Computational fusion remains upstream of professional judgment and cannot independently establish medication causality, therapeutic appropriateness, or decision authority.

Missing, Conflicting, and Low-Quality Modalities

Missingness is not a single state. A modality may be expected but absent, structurally unavailable, delayed, stale, unreadable, semantically ambiguous, or outside the validated domain. Because patterns of absence may themselves carry information, the architecture records why an input is missing and does not present imputed content as an observed clinical fact [19].

Conflict must likewise remain visible. Models may exploit unintended site, acquisition, or workflow cues that correlate with an outcome without representing the intended clinical phenomenon [20]. Data leakage, inappropriate partitioning, and weak external validation may further inflate apparent performance [21]. Inputs can also be corrupted or deliberately manipulated [22]. The proposed Conflict-State Taxonomy therefore distinguishes within-modality inconsistency, cross-modal contradiction, temporal incompatibility, provenance conflict, suspected artefact, and out-of-domain information.

Critical conflict activates source verification, restricted output, or safe abstention. A laboratory value that contradicts the documented clinical state may require unit verification, specimen-time review, or repeat testing. A spoken medication history that differs from the prescription record may require patient clarification rather than automatic reconciliation. Agreement among modalities can increase consistency, but it does not establish clinical truth or medication causality.

Evaluation across Technical and Clinical Tasks

Evaluation must progress from responsible problem formulation and transparent prediction-model reporting to assessment of the complete artificial-intelligence intervention, including input handling, error states, output presentation, and human interaction [23–25]. The proposed evaluation pathway includes representation integrity, temporal-alignment reliability, missing-modality robustness, calibration and abstention, pharmacist task performance, workflow effects, medication-safety consequences, subgroup performance, and local transportability. No individual layer is sufficient to establish overall value.

Early clinical evaluation should be staged and context-sensitive, examining actual user behaviour, workflow disruption, inappropriate reliance, and recovery from error [26]. Technical evaluation should determine whether modalities are represented faithfully and aligned to the correct events. Clinical-task evaluation should determine whether pharmacists identify, interpret, and resolve relevant medication problems more appropriately than with the available comparator. Human-factors evaluation should examine workload, attention allocation, source inspection, override behaviour, and comprehension of uncertainty.

Progression should stop when provenance is unverifiable, critical inputs are absent, temporal links cannot be defended, users cannot recognize failure, subgroup harm is plausible, or outputs alter medication action without adequate prospective evidence. Passing a technical gate does not imply clinical usefulness, medication safety, equity, or readiness for implementation.

Implementation and Human-Factors Implications

Clinical decision support operates within workflows shaped by usability, alert design, data quality, and risk management [27]. Local configuration, infrastructure, communication, and professional expertise may further change how an artificial-intelligence output is interpreted and acted upon [28].

Figure 2 presents the original conceptual synthesis for cross-modal alignment, provenance, missingness, and temporal reasoning, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

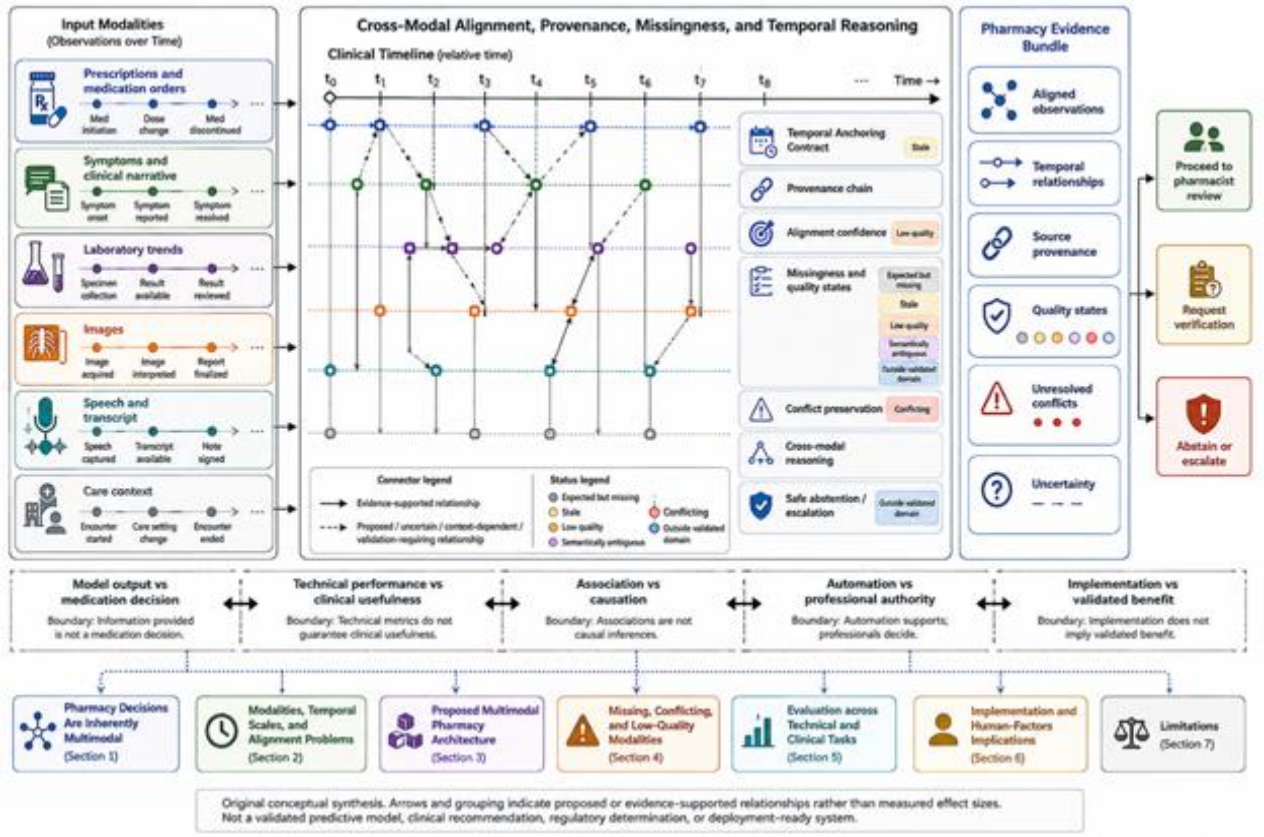

Figure 2. Cross-Modal Alignment, Provenance, Missingness, and Temporal Reasoning

Human–AI performance may improve or deteriorate according to interface design, professional expertise, and reliance behaviour [29]. Pharmacists must therefore be able to inspect original sources, challenge an alignment, identify missing information, request clarification, and reject the system output. An override should be recorded for monitoring but should not automatically become a training label without governed review.

Governance should assign responsibility for validation, professional training, incident review, version control, monitoring, and restriction of use [30]. Responsibility cannot be dispersed merely because information passes through several models or organizations. The institution remains responsible for defining the authorized task, while pharmacists retain authority over medication interpretation and action.

Table 3 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for governance responsibilities, implementation conditions, and monitoring requirements for the article Multimodal

Pharmacy Intelligence for Linking Prescriptions, Symptoms, Laboratory Trends, Images, and Speech.

Table 3. Governance responsibilities, implementation conditions, and monitoring requirements for the article Multimodal Pharmacy Intelligence for Linking Prescriptions, Symptoms, Laboratory Trends, Images, and Speech.

Governance Domain	Responsible Actor	Required Authority	Documentation Requirement	Monitoring Indicator	Escalation Trigger	Corrective Action	Accountability Boundary
Data Stewardship	Data steward	Verify source access, semantics, and permitted use	Provenance and suitability record [18]	Missing, incompatible, or stale inputs	Untraceable critical source	Correct, restrict, or abstain	Data availability is not clinical validity
Clinical Ownership	Pharmacy lead	Define intended, excluded, and escalation uses	Task and Pharmacist Decision Boundary statement [27]	Unsafe reliance, unusual overrides, or delayed review	Use beyond the authorized task	Suspend use and review workflow	Medication authority remains professional
Technical Change Control	Model owner	Control model, terminology, and interface versions	Versioned change log [30]	Drift, failure, or altered modality mix	Unreviewed material change	Revalidate, restrict, or roll back	Technical control does not confer clinical authority
Human-Factors Oversight	Human-factors lead	Evaluate interface and workflow interaction	Usability, workload, and reliance record [28]	Confusion, excessive workload, or missed conflict	Users cannot detect or recover from error	Redesign interface or restrict use	Acceptability does not establish effectiveness
Medication-Safety And Equity Oversight	Safety and equity leads	Review incidents and subgroup performance	Safety, incident, and subgroup dossier [30]	Disparate errors or emerging safety signal	Plausible avoidable harm	Investigate, constrain, or stop use	Aggregate accuracy is insufficient
Institutional Governance	Accountable multidisciplinary body	Authorize restricted evaluation and monitoring	Conditions, responsible actors, and review dates [26]	Breach of conditions or unresolved incident	Evidence no longer supports continued use	Pause, withdraw, or requalify	Local authorization is not universal readiness

The responsibilities in **Table 3** are proposed governance assignments rather than universal legal or regulatory allocations. Local professional, institutional, technical, and jurisdictional requirements remain controlling.

Limitations

Post-hoc explanation and aggregate performance cannot independently establish reliability or equity because proxy bias and subgroup underdiagnosis may persist [31–33]. Generalisability must be established for the intended task, population, institution, workflow, and decision rather than presumed from prior validation [34].

The architecture is conceptual and may omit relevant modalities, care settings, organizational constraints, or failure pathways. Its additional complexity may also offer less value than a simpler system for narrowly defined tasks. Each proposed construct therefore requires empirical comparison, and validation of one component cannot be transferred automatically to the complete architecture.

Conclusion

Multimodal pharmacy intelligence should not be defined as the accumulation of more patient data or the replacement of pharmacist judgment by a general model. The proposed architecture links prescriptions, symptoms, laboratory trends, images, speech, and care context through explicit provenance, temporal anchoring, modality-quality states, conditional fusion, conflict preservation, and safe abstention. Its principal output is a traceable Pharmacy Evidence Bundle separated from medication action by the Pharmacist Decision Boundary.

The architecture offers testable constructs for technical, clinical, human-factors, organizational, safety, equity, and governance evaluation. Its usefulness remains conditional on semantic interoperability, task-specific validation, transparent failure handling, prospective workflow assessment, subgroup evaluation, incident monitoring, and governed change control.

Multimodal agreement does not establish causality, explanation does not establish trustworthiness, and implementation does not establish benefit. The contribution is therefore a framework for disciplined development and evaluation rather than a validated clinical system, autonomous medication-decision mechanism, regulatory standard, or recommendation for routine deployment.

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

References

1. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. Nat Med. 2022;28(9):1773-84. doi:10.1038/s41591-022-01981-2
2. Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: A systematic review. J Am Med Inform Assoc. 2018;25(10):1419-28. doi:10.1093/jamia/ocy068
3. Dickinson H, Teltsch DY, Feifel J, Hunt P, Vallejo-Yagüe E, Virkud AV, et al. The unseen hand: AI-based prescribing decision support tools and the evaluation of drug safety and effectiveness. Drug Saf. 2024;47(2):117-23. doi:10.1007/s40264-023-01376-3
4. Corny J, Rajkumar A, Martin O, Dode X, Lajonchère JP, Billuart O, et al. A machine learning-based clinical decision support system to identify prescriptions with a high risk of medication error. J Am Med Inform Assoc. 2020;27(11):1688-94. doi:10.1093/jamia/ocaa154
5. 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
6. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. NPJ Digit Med. 2018;1:18. doi:10.1038/s41746-018-0029-1
7. Harutyunyan H, Khachatrian H, Kale DC, Ver Steeg G, Galstyan A. Multitask learning and benchmarking with clinical time series data. Sci Data. 2019;6:96. doi:10.1038/s41597-019-0103-9
8. Kiyoshige E, Ogata S, Kwon N, Nakaoku Y, Hayashi C, Blaylock N, et al. Developing and testing AI-based voice biomarker models to detect cognitive impairment among community dwelling adults: A cross-sectional study in Japan. Lancet Reg Health West Pac. 2025;59:101598. doi:10.1016/j.lanwpc.2025.101598
9. Huang SC, Pareek A, Seyyedi S, Banerjee I, Lungren MP. Fusion of medical imaging and electronic health records using deep learning: A systematic review and implementation guidelines. NPJ Digit Med. 2020;3:136. doi:10.1038/s41746-020-00341-z
10. Lehne M, Sass J, Essenwanger A, Schepers J, Thun S. Why digital medicine depends on interoperability. NPJ Digit Med. 2019;2:79. doi:10.1038/s41746-019-0158-1
11. Baltrušaitis T, Ahuja C, Morency LP. Multimodal machine learning: A survey and taxonomy. IEEE Trans Pattern Anal Mach Intell. 2019;41(2):423-43. doi:10.1109/TPAMI.2018.2798607
12. Krones F, Marikkar U, Parsons G, Szmul A, Mahdi A. Review of multimodal machine learning approaches in healthcare. Inf Fusion. 2025;114:102690. doi:10.1016/j.inffus.2024.102690
13. Chen RJ, Lu MY, Wang J, Williamson DFK, Rodig SJ, Lindeman NI, et al. Pathomic Fusion: An integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. IEEE Trans Med Imaging. 2022;41(4):757-70. doi:10.1109/TMI.2020.3021387
14. Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ, et al. Foundation models for generalist medical artificial intelligence. Nature. 2023;616(7956):259-65. doi:10.1038/s41586-023-05881-4
15. Liu J, Capurro D, Nguyen A, Verspoor K. Attention-based multimodal fusion with contrast for robust clinical prediction in the face of missing modalities. J Biomed Inform. 2023;145:104466. doi:10.1016/j.jbi.2023.104466
16. Lim B, Arık SÖ, Loeff N, Pfister T. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. Int J Forecast. 2021;37(4):1748-64. doi:10.1016/j.ijforecast.2021.03.012
17. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. Nature. 2019;572(7767):116-9. doi:10.1038/s41586-019-1390-1
18. Shang N, Weng C, Hripcsak G. A conceptual framework for evaluating data suitability for observational studies. J Am Med Inform Assoc. 2018;25(3):248-58. doi:10.1093/jamia/ocx095
19. Che Z, Purushotham S, Cho K, Sontag D, Liu Y. Recurrent neural networks for multivariate time series with missing values. Sci Rep. 2018;8:6085. doi:10.1038/s41598-018-24271-9
20. Badgeley MA, Zech JR, Oakden-Rayner L, Glicksberg BS, Liu M, Gale W, et al. Deep learning predicts hip fracture using confounding patient and healthcare variables. NPJ Digit Med. 2019;2:31. doi:10.1038/s41746-019-0105-1
21. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell. 2021;3(3):199-217. doi:10.1038/s42256-021-00307-0
22. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. Science. 2019;363(6433):1287-9. doi:10.1126/science.aaw4399
23. 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
24. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385. doi:10.1136/bmj-2023-078378
25. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. Nat Med. 2020;26(9):1364-74. doi:10.1038/s41591-020-1034-x
26. 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
27. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: Benefits, risks, and strategies for success. NPJ Digit Med. 2020;3:17. doi:10.1038/s41746-020-0221-y
28. Farič N, Hinder S, Williams R, Ramaesh R, Bernabeu MO, van Beek E, et al. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: A qualitative study. J Am Med Inform Assoc. 2024;31(1):24-34. doi:10.1093/jamia/ocad191
29. Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human-computer collaboration for skin cancer recognition. Nat Med. 2020;26(8):1229-34. doi:10.1038/s41591-020-0942-0
30. Labkoff S, Oladimeji B, Kannry J, Solomonides A, Leftwich R, Koski E, et al. Toward a responsible future: Recommendations for AI-enabled clinical decision support. J Am Med Inform Assoc. 2024;31(11):2730-9. doi:10.1093/jamia/ocae209
31. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3(11). doi:10.1016/S2589-7500(21)00208-9
32. 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
33. 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
34. Futoma J, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. Lancet Digit Health. 2020;2(9). doi:10.1016/S2589-7500(20)30186-2