Original Article

What Counts as Meaningful Human Oversight in AI-Assisted Medication Management?

Wouter De Smet[1]*, Liesbeth Van Dam[1], Pieter Janssens[2]

[1]Department of Human Oversight and Medication Management, Faculty of Pharmaceutical Sciences, KU Leuven, Leuven, Belgium.
[2]Department of AI-Assisted Clinical Governance, Faculty of Pharmacy, Ghent University, Ghent, Belgium.

Abstract

Artificial intelligence is increasingly positioned within medication-management activities such as prescription screening, risk prioritization, information generation, treatment selection, and clinical decision support. Yet the continued presence of a pharmacist or another healthcare professional does not necessarily constitute meaningful human oversight. A person may be formally "in the loop" while lacking the competence, time, information, organizational support, or authority required to identify and correct a hazardous output. This article develops a proposed Oversight-Capability Framework for determining when human involvement in AI-assisted medication management may be considered meaningful. The framework defines oversight through four proposed functions: detecting a reason for scrutiny, interpreting and challenging the AI-assisted output, deciding whether the output should be accepted, modified, deferred, rejected, or escalated, and implementing or coordinating an appropriate response. These functions depend on four enabling conditions—competence, time, information, and authority—and may operate prospectively, continuously, when triggered, or retrospectively. Meaningful oversight is therefore treated as a demonstrable capability to alter a potentially hazardous medication trajectory rather than as a professional title, explanation display, alert acknowledgment, or nominal override option. The framework separates model output from medication decision, technical performance from clinical usefulness, automation from professional authority, and implementation from validated benefit. Evaluation should examine erroneous recommendations, disagreement, intervention feasibility, workload, subgroup harm, escalation, and organizational learning. The proposed framework is conceptual and requires validation across medication tasks, technologies, populations, professional roles, and organizational settings. It is not a clinical recommendation, regulatory standard, validated safety mechanism, or deployment-ready protocol.

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

Introduction

Medication management is a consequential sociotechnical activity in which recommendations must be interpreted against patient characteristics, treatment history, concurrent medicines, monitoring information, professional responsibilities, and organizational constraints. Artificial intelligence may support parts of this activity, but its introduction can also create new error pathways, including inappropriate reliance, concealed uncertainty, misleading outputs, and redistribution of responsibility. Machine-learning systems can therefore produce unintended clinical consequences even when a healthcare professional remains formally involved [1].

The language of human–AI collaboration often assumes that combining computational capability with professional judgment will produce a superior result. Such complementarity is plausible but cannot be inferred from the simultaneous presence of a person and a model. It depends on which party performs which task, what information is available, how disagreement is handled, and whether the

Address for correspondence: Wouter De Smet, Department of Human Oversight and Medication Management, Faculty of Pharmaceutical Sciences, KU Leuven, Leuven, Belgium.
wouter.desmet@kuleuven.be
Received: 28 May 2025; **Accepted:** 02 September 2025

How to cite this article: De Smet W, Van Dam L, Janssens P. What Counts as Meaningful Human Oversight in AI-Assisted Medication Management? Arch Pharm Pract. 2025;16(3):71-8. https://doi.org/10.51847/w3eX398RCg

professional can meaningfully change the resulting action [2]. In medication management, this distinction is especially important because a recommendation may influence prescribing, verification, dispensing, administration, monitoring, counselling, or follow-up without formally replacing the professional responsible for the decision.

Automation bias further weakens the assumption that human review inherently provides protection. People may accept incorrect recommendations, fail to search for contradictory evidence, or reduce independent reasoning after exposure to apparently authoritative decision support. Verification can also become difficult when reviewing an output requires reconstructing complex calculations, locating missing clinical information, or understanding an unfamiliar model [3]. A nominal review step may consequently provide little protection if detecting the error demands more time, information, or expertise than the workflow permits.

Clinical decision-support evidence similarly indicates that benefits and risks depend on usability, workflow integration, maintenance, data quality, alert design, and implementation conditions rather than technical availability alone [4]. A pharmacist who receives an alert too late, cannot inspect its basis, lacks access to relevant records, or has no practical mechanism for stopping the proposed action cannot exercise effective oversight merely by being present.

This article therefore proposes that meaningful human oversight should be judged by capability and consequence. The relevant question is not whether a human appeared somewhere in the workflow, but whether an appropriately situated person could detect a reason for scrutiny, independently evaluate the AI-assisted output, exercise decision authority, and intervene before or after harm. This is an original conceptual framework rather than an empirical validation study. Its categories and relationships are proposed unless directly grounded in the selected evidence.

Defining the Functions of Meaningful Oversight

Artificial intelligence in pharmacy encompasses heterogeneous tasks, including prescription screening, medication-information support, workflow prioritization, error detection, clinical prediction, and communication. These applications differ in autonomy, uncertainty, intended user, reversibility, and potential harm. Oversight should therefore be defined for a bounded medication-management task rather than for "AI" or "digital pharmacy" as undifferentiated categories [5].

Meaningful oversight is proposed here as the demonstrable capacity of a human or organized professional team to identify, evaluate, contest, modify, stop, escalate, or learn from an AI-assisted medication action. Merely displaying an explanation is insufficient because an explanation may be incomplete, persuasive without being reliable, or disconnected from the information required for independent verification [6]. Explanation may support oversight, but it does not itself establish comprehension, challenge, authority, or intervention.

Four proposed functions organize the concept. Detection means recognizing that an output, omission, uncertainty, workflow event, or patient circumstance requires scrutiny. Interpretation and challenge mean examining the recommendation against clinical evidence, patient context, alternative explanations, and known limitations. Decision and authorization mean determining whether the output should be accepted, modified, delayed, rejected, or referred. Intervention and escalation mean executing or coordinating an action capable of changing the medication trajectory. These functions should be connected to lifecycle accountability, monitoring, and organizational responsibility rather than treated as an isolated bedside or dispensing-stage check [7].

Human–AI performance can vary among users and cases, demonstrating that collaboration quality cannot be inferred from model accuracy alone [8]. An output that assists one professional may distract or mislead another. Meaningful oversight must therefore be evaluated as an interaction capability situated within a specific task, workflow, population, and organizational environment.

Figure 1 presents the original conceptual synthesis for oversight-capability framework for ai-assisted medication management, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

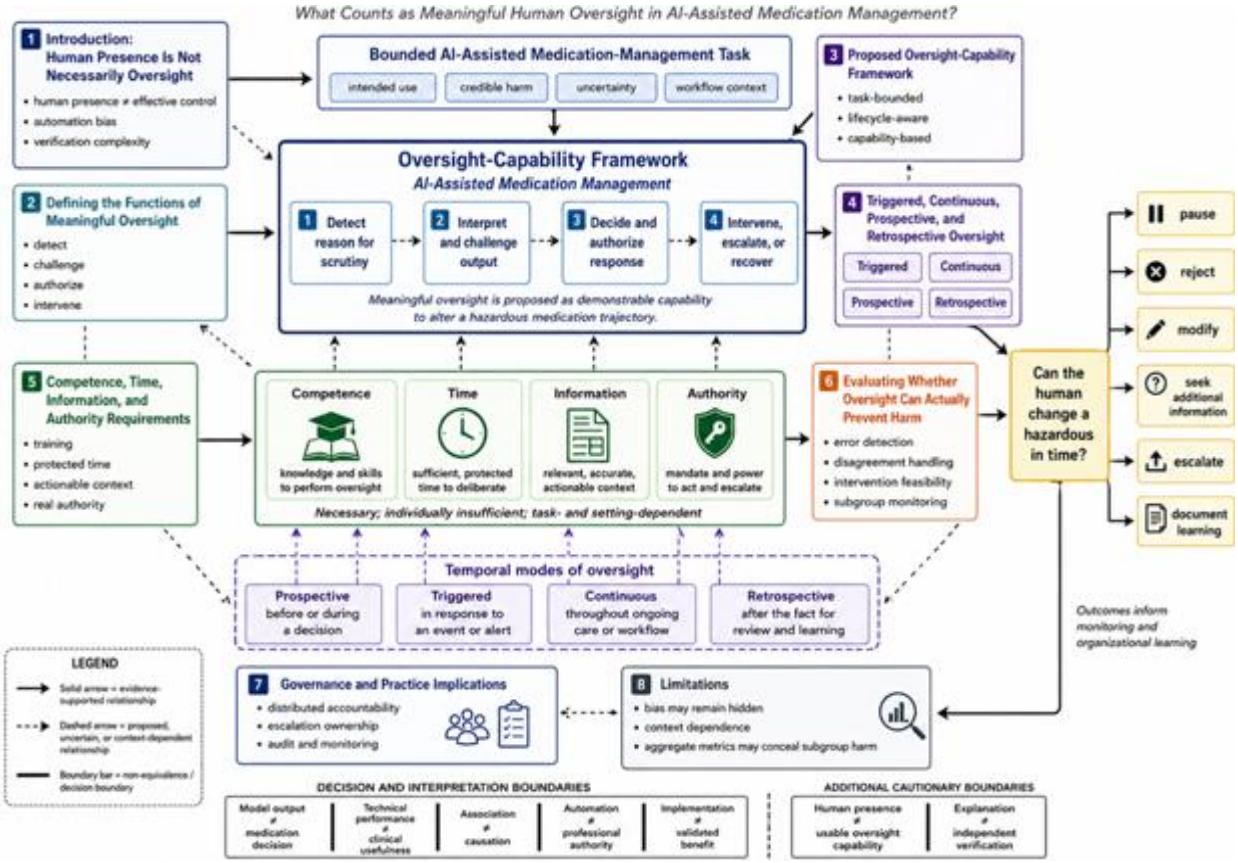

Figure 1. Oversight-Capability Framework for AI-Assisted Medication Management

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, components, relationships, and intended uses for the article What Counts as Meaningful Human Oversight in AI-Assisted Medication Management?

Table 1. Definitions, components, relationships, and intended uses for the article What Counts as Meaningful Human Oversight in AI-Assisted Medication Management?

Component or scientific dimension	Problem addressed	Inputs or determinants	Proposed mechanism or relationship	Expected contribution	Evidence required	Failure or uncertainty risk	Boundary statement
Bounded medication-management task	"AI" is too broad to define a uniform oversight requirement.	Intended use, user, patient context, workflow stage, reversibility, credible harm.	**Proposed synthesis:** oversight requirements are specified for the task and consequence rather than the technology label [5].	Prevents transfer of assumptions across dissimilar pharmacy applications.	Clear intended-use statement and workflow analysis.	Oversight may be designed for the model but not for the actual medication action.	Task definition does not establish clinical usefulness.
Human presence	A professional may remain involved without being able to detect or correct error.	Workflow position, attention, timing, access to evidence.	Presence contributes only when it enables an effective oversight function [3].	Replaces nominal involvement with observable capability.	Human-factors and workflow evidence.	Symbolic review, rubber-stamping, automation bias.	Human presence is not equivalent to meaningful oversight.
Detection	Harm cannot be prevented if no reason for scrutiny is recognized.	Alerts, uncertainty signals, patient changes, conflicting information.	**Proposed function:** identify an output, omission, or circumstance requiring examination.	Creates an opportunity for timely review.	Sensitivity to relevant hazards and manageable alert burden.	Missed signal, excessive alerts, hidden uncertainty.	Detection is not proof that the issue will be understood or corrected.

Interpretation and challenge	Explanations may persuade without supporting independent verification.	Output rationale, clinical records, alternatives, limitations, uncertainty.	**Proposed function:** compare the AI-assisted output with independent clinical and contextual evidence [6].	Supports contestability and calibrated reliance.	Comprehension, verification, and disagreement studies.	Explanation bias, incomplete context, inability to reconstruct reasoning.	Explanation is not equivalent to verification.
Decision and authorization	Review has little value when the reviewer cannot alter the decision.	Professional scope, accountability, escalation rules, organizational policy.	**Proposed function:** accept, modify, defer, reject, or refer the output.	Links professional judgment to an actionable decision.	Evidence that users possess and exercise real authority.	Nominal discretion, unclear responsibility, deference to automation.	Model output is not the medication decision.
Intervention and escalation	Recognizing an error does not prevent harm unless action is feasible.	Time, communication channels, staffing, access to prescribers and records.	**Proposed function:** change, pause, stop, or escalate the medication trajectory.	Connects oversight to potential risk control.	Timeliness and intervention-feasibility evidence.	Delayed response, unavailable decision-maker, irreversible action.	Override access is not equivalent to feasible intervention.
Lifecycle accountability	Point-of-use review cannot address every design, data, or monitoring failure.	Development, validation, deployment, updating, incident review.	Oversight is distributed across users and organizations throughout the system lifecycle [7].	Prevents the frontline professional from becoming the sole safety backstop.	Role allocation, monitoring, audit, and change-control evidence.	Responsibility gaps and organizational diffusion.	Professional oversight does not replace developer or institutional accountability.
Human–AI interaction	Model performance may not predict combined performance.	User expertise, case difficulty, interface, reliance, disagreement.	Interaction effects are context-dependent and must be tested [8].	Directs evaluation toward combined task performance.	Comparative human-only, AI-only, and human–AI evidence where appropriate.	Performance degradation for some users or cases.	Technical performance is not equivalent to clinical usefulness.

Proposed Oversight-Capability Framework

The proposed framework begins with a bounded AI-assisted medication-management task. The task should specify the intended user, information inputs, decision affected, timing, reversibility, plausible error pathways, and potential consequences. Medication-risk prioritization illustrates why this boundary matters: a system may identify prescriptions requiring review, but the safety contribution depends on whether the prioritized cases are clinically meaningful and whether a reviewer can investigate and act on them [9]. Risk identification should not be described as harm prevention unless the downstream review and intervention pathway has also been evaluated.

The framework then asks whether four enabling conditions are present. Competence concerns the knowledge and skills required to understand the medication task, recognize model limitations, evaluate uncertainty, and challenge an output. Time concerns whether scrutiny can occur before the relevant decision becomes effectively irreversible. Information concerns access to the patient, treatment, evidence, model, and workflow context needed for an independent judgment. Authority concerns real permission and practical power to alter, delay, reject, or escalate the action. These conditions are proposed as necessary but individually insufficient.

When the enabling conditions are adequate, the four oversight functions may be enacted: detection, interpretation and challenge, decision and authorization, and intervention or escalation. The sequence is not necessarily linear. A reviewer may return for additional information, seek another professional opinion, escalate uncertainty, or initiate retrospective analysis. The framework therefore represents a revisable capability pathway rather than a fixed algorithm.

Oversight also extends beyond the moment at which an output is displayed. Responsible machine-learning practice requires attention to problem formulation, data provenance, validation, deployment, monitoring, and system change [10]. The professional using the system cannot independently repair biased training data, inappropriate objectives, distribution shift, inaccessible audit records, or unsafe organizational incentives. Point-of-use oversight must therefore be embedded within lifecycle governance.

A further boundary separates technical performance from clinical task performance. A model may classify, rank, or generate text accurately under a benchmark while failing to improve professional decisions or workflow outcomes. Clinical impact additionally depends on external validity, integration, usability, prospective evaluation, and the consequences of errors [11]. Accordingly, the proposed framework evaluates not only whether a model produces a technically correct output, but whether the human–AI system enables an appropriate medication action under real conditions.

Medication-direction error detection by large language models provides a relevant example. Such systems may identify potentially problematic instructions, but the output

remains an informational contribution rather than the medication decision itself [12]. A meaningful oversight pathway would specify who examines the flag, what source information is available, how uncertainty is communicated, how false-positive and false-negative outputs are managed, who contacts the prescriber or patient, and who authorizes the final wording. The framework therefore treats model flagging, professional review, medication authorization, and implementation as separate stages.

Responsible translation consequently requires lifecycle governance connecting problem formulation, local validation, workflow integration, and post-deployment monitoring [7, 10, 11]. The proposed architecture does not claim that every stage must be controlled by one professional or that all medication applications require identical evidence. Instead, it provides a structure for asking whether responsibility, information, and intervention capacity are sufficient for the risk and claim being made.

Triggered, Continuous, Prospective, and Retrospective Oversight

Oversight may operate through different temporal modes. These modes are proposed categories rather than validated levels of safety, and more than one mode may be required for a single medication-management application.

Triggered oversight begins when an event creates a reason for scrutiny. Triggers may include an AI-generated alert, low-confidence output, data inconsistency, unexpected dose, contraindication, patient deterioration, disagreement between systems, or professional concern. A trigger is meaningful only when it reaches an appropriate reviewer, communicates why attention is required, and supports an independent verification pathway. Experimental evidence shows that users may remain susceptible to incorrect AI advice, meaning that notification alone cannot be treated as protection [13]. Trigger design should therefore be evaluated for both missed hazards and inappropriate reliance.

Continuous oversight refers to human attention or supervisory capability maintained throughout an AI-assisted process. It may be relevant where outputs change dynamically, decisions occur repeatedly, or deterioration must be detected over time. Continuous exposure to AI, however, is not equivalent to continuous oversight. Human–AI effects can vary with expertise and case conditions, and assistance may improve some judgments while impairing others [14]. Continuous oversight requires sustainable attention, information access, and corrective capacity rather than the permanent visibility of a dashboard or recommendation.

Prospective oversight occurs before a medication action is finalized or executed. It may include reviewing a proposed prescription, checking an AI-generated instruction, assessing a treatment recommendation, or authorizing an automated prioritization outcome. AI recommendations can influence treatment selections, including choices involving medicines, which makes reliance, disagreement, and reversibility important evaluation targets [15]. Prospective oversight is most protective when the reviewer has sufficient information and authority before the decision becomes operationally difficult to reverse.

Retrospective oversight occurs after an output, decision, or medication action. It may involve reviewing overridden alerts, investigating errors, examining subgroup performance, auditing recurrent recommendations, identifying drift, or revising local governance. Retrospective oversight can reveal patterns that are not visible in individual cases and can support organizational learning [7]. It cannot, however, reverse every completed harm and should not substitute for prospective or triggered intervention when immediate action is feasible.

The modes should therefore be combined according to risk, uncertainty, reversibility, and workflow. A high-risk prescription-screening system may require prospective review of selected cases, triggered escalation for uncertainty, continuous monitoring of workflow burden, and retrospective analysis of errors and overrides. A lower-consequence administrative prioritization tool may require a different combination. Lifecycle monitoring remains necessary because performance and use can change after implementation [10].

No temporal mode is inherently meaningful. Triggered oversight fails when alerts are ignored or misleading; continuous oversight fails when attention is unsustainable; prospective oversight fails when review is rushed or powerless; and retrospective oversight fails when findings do not lead to corrective action. The proposed test is whether the selected combination gives appropriately competent and authorized people a realistic opportunity to detect, challenge, change, and learn from hazardous AI-assisted medication actions.

Competence, Time, Information, and Authority Requirements

Meaningful oversight requires more than a formally designated reviewer. The proposed framework treats competence, time, information, and authority as necessary enabling conditions whose adequacy depends on the medication task, potential harm, workflow, and degree of uncertainty. None is individually sufficient. A knowledgeable pharmacist cannot intervene meaningfully without time or authority, while an authorized professional cannot evaluate an output safely without relevant competence and information.

Time includes both chronological opportunity and usable cognitive capacity. A reviewer must receive the output early enough to examine it before the medication action becomes

difficult or impossible to reverse. The workflow must also permit sustained attention rather than requiring rapid acknowledgment amid competing demands. Workload, task complexity, and repeated alerts can contribute to alert fatigue and reduce attention to decision-support messages [16]. The existence of a review interval therefore does not demonstrate that adequate review is realistically possible. Evaluation should examine interruption, queue length, competing responsibilities, response delay, and whether high-risk cases can be prioritized without creating new blind spots.

Information means access to the contextual evidence needed to evaluate the AI-assisted output independently. Depending on the task, this may include indication, dose, organ function, laboratory values, allergy history, concomitant therapy, adherence, treatment response, previous adverse effects, prescriber intent, model inputs, missing-data indicators, and known limitations. Medication-related alerts are frequently overridden, and the appropriateness of those decisions varies according to alert type and clinical context [17]. Oversight should therefore not be assessed by alert acceptance or override frequency alone. The relevant question is whether the reviewer received sufficiently specific, timely, and actionable information to distinguish a useful warning from a misleading or irrelevant one.

Competence includes medication-domain expertise, understanding of the intended task, recognition of uncertainty, knowledge of common AI failure modes, and the ability to compare the output with independent evidence. Professional registration or general familiarity with digital systems does not establish these capabilities. Existing educational programs for healthcare professionals vary considerably in content and assessment, indicating that AI competence cannot be presumed from exposure to technology [18]. Task-specific education should be paired with observable assessment, such as identifying erroneous outputs, articulating limitations, seeking missing information, and selecting appropriate escalation routes. The framework does not prescribe a universal curriculum or competency threshold.

Authority means more than theoretical discretion. The reviewer must possess practical and organizational power to pause, reject, modify, defer, or escalate an AI-assisted medication action. Authority may be undermined by unclear responsibility, inaccessible prescribers, rigid automation, staffing limitations, managerial pressure, or fear of punitive consequences. It should therefore be evaluated through actual workflow permissions and responses to disagreement. An override button is not meaningful authority when its use is discouraged, undocumented, delayed, or incapable of changing the action.

Human intervention capacity is consequently constrained by workload, alert burden, and competence rather than professional presence alone [16-18]. The proposed framework treats failure of any enabling condition as a potential break in the oversight pathway. However, the conditions should not be converted into unsupported universal scores. Their adequacy must be tested for the specific medication task, setting, user group, and foreseeable harm.

Evaluating Whether Oversight Can Actually Prevent Harm

The central evaluative question is whether human oversight can change a hazardous medication trajectory under realistic conditions. Model accuracy, explanation availability, reviewer satisfaction, alert acknowledgment, and agreement with AI may provide useful information, but none alone establishes harm prevention. Evaluation should reconstruct the pathway from model output to signal, human interpretation, decision authority, intervention, and resulting consequence.

Clinical evaluation reports should describe the AI intervention, intended use, human interaction, error handling, and outcomes sufficiently to determine what the professional actually did and whether that action influenced the result [19]. Studies should distinguish errors generated by the AI, errors introduced through interaction, errors detected by the human, and errors remaining after review. Reporting that a human approved an output is insufficient unless the reviewer's information, timing, options, and response are also described.

Prospective protocols should specify the intended user, workflow position, data inputs, human role, escalation pathway, anticipated errors, and potential harms before claims about meaningful oversight are made [20]. The protocol should define what counts as successful detection, appropriate disagreement, effective intervention, delayed action, and unresolved uncertainty. Where oversight is expected to prevent harm, the evaluation design should include outcomes capable of testing that claim rather than relying solely on technical or usability proxies.

Early live clinical evaluation should examine actual use, workflow adaptations, safety events, human factors, and deviations from intended practice [21]. Relevant measures may include whether reviewers recognize incorrect recommendations, whether correct outputs are unnecessarily rejected, whether escalation occurs in time, whether workload changes, whether users become over-reliant, and whether the system creates new medication-error pathways. Evidence should also identify which users and cases benefit or are disadvantaged.

The proposed framework can be falsified when a reviewer does not receive a meaningful signal, cannot verify the output, lacks adequate time or information, cannot alter the action, or repeatedly fails to identify hazardous recommendations. It also fails when review creates additional delay, over-reliance, inequity, or diffusion of responsibility. The principal test is not whether a human interacted with the

system, but whether the interaction produced a timely and justifiable change when change was needed.

Demonstrating prevention requires an appropriate comparative design. Agreement with AI, explanation viewing, override availability, or user confidence should remain secondary or process measures unless their relationship to safer medication decisions has been established. Evaluation should be proportionate to intended use: low-consequence administrative support may not require the same evidence as systems influencing dose, contraindication management, treatment selection, or dispensing authorization.

Governance and Practice Implications

Meaningful oversight is an organizational property as well as an individual professional capability. Institutions deploying AI-assisted medication-management systems remain responsible for defining intended use, validating local performance, assigning roles, maintaining infrastructure, monitoring changes, and responding to incidents. Ethical challenges associated with bias, opacity, and changing professional roles cannot be transferred entirely to the person who receives the final output [22].

Responsibility is distributed across developers, vendors, deploying organizations, managers, clinical teams, pharmacy professionals, and monitoring bodies throughout the system lifecycle [23]. Developers influence objectives, data representation, interface design, and error visibility. Organizations determine procurement, staffing, workflow integration, access to records, escalation processes, and monitoring. Professionals contribute contextual judgment and medication expertise. Meaningful oversight requires these responsibilities to be explicit rather than concealed behind a generic assertion that a human remains in control.

Professional authority must also be preserved in practice. Algorithmic decision support can reshape autonomy and responsibility even when it is formally advisory [24]. A pharmacist may appear to retain discretion while organizational incentives, interface defaults, time pressure, or inaccessible escalation routes make disagreement difficult. Governance should therefore evaluate whether professionals can document uncertainty, reject a recommendation, request additional evidence, contact another decision-maker, and stop an unsafe action without unreasonable obstruction.

Organizations should define escalation ownership, downtime arrangements, audit access, incident-review procedures, change control, and responsibility for recurrent error. Monitoring should include overrides, disagreements, missed alerts, false reassurance, workflow adaptations, delayed interventions, subgroup performance, and changes in system use. These arrangements should support learning rather than simply attributing adverse outcomes to individual users.

The framework is not a legal allocation of liability or a regulatory standard. Local responsibilities will vary with jurisdiction, professional scope, technology, and organizational structure. Its governance contribution is a proposed test: an institution should not claim meaningful human oversight when the people assigned to review outputs lack the resources, information, authority, or organizational protection needed to perform that role.

Limitations

Human oversight cannot reliably compensate for every technical or organizational failure. Biased data, unsafe objectives, opaque error modes, poor calibration, and inappropriate deployment can remain difficult for reviewers to detect [25]. The framework therefore does not position pharmacists or other professionals as universal safety backstops.

Oversight capability is also context-dependent. Performance and human interaction may change across institutions, populations, technologies, interfaces, and medication tasks, limiting assumptions of generalizability [26]. Local evaluation remains necessary before decision-use or safety claims are made.

Aggregate results may conceal systematic harm affecting underserved groups [27]. Evaluation should therefore examine subgroup-specific error, detectability, intervention, and consequence. Performance can remain unsafe under bias, distribution shift, or underserved-population error even when a human reviewer is present [25-27]. These limitations constrain the framework's interpretation and reinforce its status as a proposed, empirically testable synthesis.

Conclusion

Meaningful human oversight in AI-assisted medication management should not be defined by human presence, professional title, explanation access, alert acknowledgment, or a nominal override mechanism. This article proposes that oversight becomes meaningful only when appropriately situated people possess the capability to detect a reason for scrutiny, interpret and challenge an AI-assisted output, make an authorized decision, and intervene or escalate in time.

The proposed Oversight-Capability Framework links these functions to competence, time, information, and authority and distinguishes prospective, continuous, triggered, and retrospective oversight. It separates model output from medication decision, technical performance from clinical usefulness, automation from professional authority, and implementation from validated benefit. Its evaluative focus is whether the human–AI arrangement can alter a hazardous medication trajectory without creating unacceptable new errors, delays, inequities, or responsibility gaps.

The framework also places oversight within lifecycle governance. Professionals cannot independently correct

deficient data, inappropriate objectives, distribution shift, unsafe organizational incentives, or inaccessible monitoring. Developers, vendors, deploying institutions, managers, and healthcare professionals therefore retain distinct and overlapping responsibilities.

The contribution is conceptual rather than validated. Its functions, relationships, temporal modes, and failure conditions require empirical testing across medication tasks, professional groups, patient populations, technologies, and organizational settings. It should not be interpreted as a clinical recommendation, regulatory determination, universal standard, or deployment-ready protocol. Its intended contribution is to replace the weak question of whether a human remains involved with the more demanding question of whether the human and organization possess demonstrable capability to recognize, contest, and change an unsafe medication action.

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

References

1. 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
2. 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
3. Lyell D, Coiera E. Automation bias and verification complexity: A systematic review. J Am Med Inform Assoc. 2017;24(2):423-31. doi:10.1093/jamia/ocw105
4. 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
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. 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
7. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc. 2020;27(3):491-7. doi:10.1093/jamia/ocz192
8. 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
9. 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
10. 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
11. 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
12. Pais C, Liu J, Voigt R, Gupta V, Wade E, Bayati M. Large language models for preventing medication direction errors in online pharmacies. Nat Med. 2024;30(6):1574-82. doi:10.1038/s41591-024-02933-8
13. Gaube S, Suresh H, Raue M, Merritt A, Berkowitz SJ, Lermer E, et al. Do as AI say: Susceptibility in deployment of clinical decision-aids. NPJ Digit Med. 2021;4(1):31. doi:10.1038/s41746-021-00385-9
14. Kiani A, Uyumazturk B, Rajpurkar P, Wang A, Gao R, Jones E, et al. Impact of a deep learning assistant on the histopathologic classification of liver cancer. NPJ Digit Med. 2020;3:23. doi:10.1038/s41746-020-0232-8
15. Jacobs M, Pradier MF, McCoy TH Jr, Perlis RH, Doshi-Velez F, Gajos KZ. How machine-learning recommendations influence clinician treatment selections: The example of the antidepressant selection. Transl Psychiatry. 2021;11(1):108. doi:10.1038/s41398-021-01224-x
16. Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R, et al. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. BMC Med Inform Decis Mak. 2017;17(1):36. doi:10.1186/s12911-017-0430-8
17. Nanji KC, Seger DL, Slight SP, Amato MG, Beeler PE, Her QL, et al. Medication-related clinical decision support alert overrides in inpatients. J Am Med Inform Assoc. 2018;25(5):476-81. doi:10.1093/jamia/ocx115
18. Charow R, Jeyakumar T, Younus S, Dolatabadi E, Salhia M, Al-Mouaswas D, et al. Artificial intelligence education programs for health care professionals: Scoping review. JMIR Med Educ. 2021;7(4). doi:10.2196/31043
19. Liu X, Rivera SC, 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
20. Rivera SC, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. Nat Med. 2020;26(9):1351-63. doi:10.1038/s41591-020-1037-7
21. 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
22. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. N Engl J Med. 2018;378(11):981-3. doi:10.1056/NEJMp1714229
23. Morley J, Machado CCV, Burr C, Cowls J, Joshi I, Taddeo M, et al. The ethics of AI in health care: A mapping review. Soc Sci Med. 2020;260:113172. doi:10.1016/j.socscimed.2020.113172
24. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. J Med Ethics. 2020;46(3):205-11. doi:10.1136/medethics-2019-105586
25. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf. 2019;28(3):231-7. doi:10.1136/bmjqs-2018-008370
26. 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
27. 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