Viewpoint

From Decision Support to Delegated Clinical Action: Governing the Autonomy Gradient of AI-Driven Clinical Decision Support Systems

Jialin Liu[1,2*], MD; Siru Liu[3*], PhD

[1]Information Center, West China Hospital of Sichuan University, Chengdu, Sichuan, China

[2]Department of Otolaryngology-Head and Neck Surgery, West China Hospital of Sichuan University, Chengdu, Sichuan, China

[3]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, United States

[*]all authors contributed equally

Corresponding Author:
Siru Liu, PhD
Department of Biomedical Informatics
Vanderbilt University Medical Center
2525 West End Ave, Suite 1475
Nashville, TN, 37212
United States
Phone: 1 615 875 5216
Email: siru.liu@vumc.org

Abstract

AI-driven clinical decision support systems can prepare executable orders, communicate with patients, and initiate multistep workflows, raising unresolved questions about delegated clinical authority. This viewpoint proposes a 5-level clinical autonomy gradient: informational support (A0), advisory decision support (A1), human-confirmed execution (A2), bounded delegated action (A3), and adaptive autonomous management (A4). Classification applies to a task-action pair within its deployment workflow. Execution authority and case-specific human authorization anchor assignment; decision influence, temporal reach, and scope adaptability describe the configuration. The action-enablement boundary (A1-A2) introduces executable action preparation requiring explicit, informed, case-specific authorization before release. The delegation boundary (A2-A3) permits execution without contemporaneous, case-specific authorization. We adapt an automation scope of practice to record permissions and operating conditions, and an autonomy safety case to link authority to deployment-specific evidence and residual risks. Proposed controls address effective human review, interacting systems, proportionate governance, and withholding, downgrading, or suspending actions when supporting conditions fail. Applications to documented systems and hypothetical scenarios illustrate classification reasoning without establishing reliability or delegated-action safety. Greater autonomy is optional; A1 or A2 may remain appropriate. Independent evaluation must assess classification agreement, feasibility, implementation burden, and effects on care.

(J Med Internet Res 2026;28:e108260) doi: 10.2196/108260

KEYWORDS

accountability; agentic AI; AI; artificial intelligence; automation scope of practice; clinical AI governance; clinical autonomy; clinical decision support systems; delegated clinical action; human oversight; patient safety

Introduction

Clinical decision support systems (CDSSs) traditionally retrieve evidence, calculate risk, generate alerts, and recommend diagnostic or therapeutic options, while clinicians interpret outputs and make decisions [1,2]. In the United States, the Food and Drug Administration (FDA) excludes certain clinical decision support functions from the device definition only when all specified criteria are met [3]. One criterion requires clinicians to be able to independently review the basis for recommendations without relying primarily on them for individual diagnostic or treatment decisions. Independent review alone does not establish nondevice status.

Generative and agentic AI extends this role through electronic health record (EHR) navigation, tool use, order preparation, patient communication, and multistep workflows [4,5]. Clinical

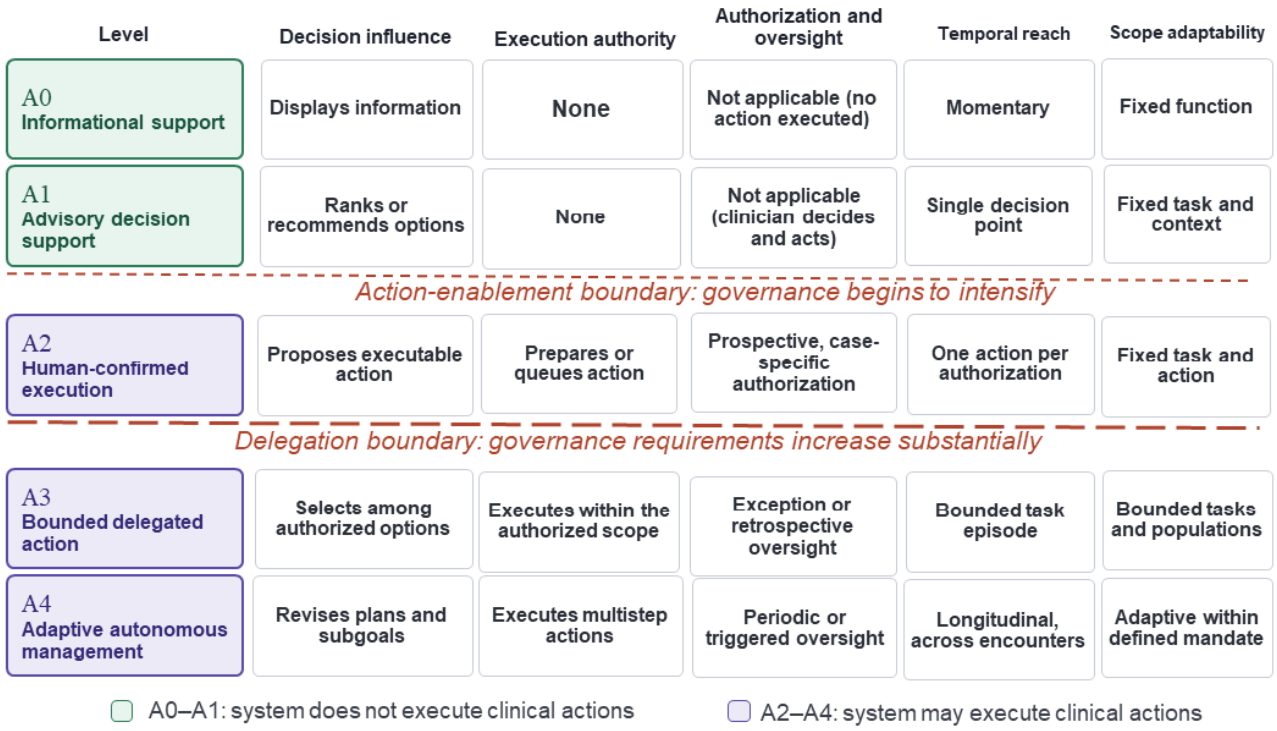
XSL•FO
RenderX

evidence remains largely exploratory, often derived from experimental or simulated settings [4]. The Medical Intelligence for Reasoning and Action (MIRA) system, for example, obtained histories, ordered and interpreted tests, and generated diagnoses and treatment plans through structured actions in a sandboxed EHR [5]. This demonstrates action-oriented capabilities in retrospective simulation, not prospective safety or effectiveness in routine care.

Governance must therefore address delegated authority as well as recommendation accuracy [6-10]. A recommendation rejected before implementation differs from an erroneous action executed before human intervention. Risk depends on model performance, workflow design, reversibility, and the capacity to detect, contain, and recover from failure [6,8-10].

We propose a clinical autonomy gradient that classifies the clinical task-action pair within its deployed workflow rather than the model alone (Figure 1). A narrative conceptual synthesis informs 5 autonomy levels and a classification procedure. Two adapted governance instruments support the proposal: an automation scope of practice (ASoP) specifying permissions and operating conditions, and an autonomy safety case linking authority to evidence, assumptions, and residual risks [11]. The framework complements lifecycle and risk-management guidance [9,12,13]. Its classification reliability, feasibility, implementation burden, and governance effectiveness require independent evaluation.

Figure 1. The clinical autonomy gradient and its governance boundaries. This proposed taxonomy classifies each task-action pair within its deployment context. Execution authority and case-specific human authorization anchor classification; decision influence, temporal reach, and scope adaptability describe its configuration. The action-enablement boundary (A1-A2) marks preparation of executable actions requiring explicit, informed, case-specific human authorization before release. The delegation boundary (A2-A3) marks execution without contemporaneous, case-specific human authorization. The deployment's ASoP specifies the authorized scope at A2 and above, with safeguards proportionate to delegated authority and clinical risk. Higher autonomy is optional; A1 or A2 may remain appropriate for a given task.



From Decision Support to Delegated Clinical Action

Conceptual Foundations and Author-Proposed Distinctions

The framework draws on distinctions among information processing, decision selection, and action implementation in human-automation research [14,15], graded clinical autonomy in dentistry [16], and clinical AI oversight and governance [6,9,17]. These sources support distinguishing types of support and delegated authority; they do not establish the 5 categories proposed here. Our synthesis separates information organization (A0) from clinical advice (A1), distinguishing forms of cognitive support without granting authority to prepare or execute clinical actions, and introduces 2 governance boundaries: preparation of executable actions requiring explicit, informed, case-specific authorization before release (A1-A2), and execution without contemporaneous, case-specific authorization (A2-A3). A3-A4 distinguishes execution within a predefined task scope from adaptive management under a broader authorized mandate, including revision of management plans and subgoals and execution of clinical actions. These are author-proposed distinctions in deployed authority, not empirically established maturity stages.

Clinical action means an implemented physical, digital, or communicative event, such as administering medication, releasing an order, changing a monitoring schedule, or sending treatment or help-seeking instructions. Action preparation

XSL•FO
RenderX

creates an executable workflow object rather than advice alone. At A2, defaults and prepopulated orders can weaken scrutiny [7]; at A3 and above, erroneous actions may propagate before review [6,8]. These differences motivate stronger safeguards at the action-enablement and delegation boundaries.

Autonomy as a Task-Workflow Property

In this framework, clinical autonomy is a property of the deployed task-workflow configuration rather than an intrinsic capability of the model. The same model may summarize records in one workflow and be authorized to initiate or modify prescribing actions in another; a simple rules-based system may also exercise substantial autonomy by initiating consequential actions without contemporaneous, case-specific human authorization. Autonomy reflects allocated operational authority, while clinical effects and risks depend on model performance, intended use, workflow integration, and deployment conditions [9,12].

Delegated action should be evaluated against real-world clinical practice, accounting for delays, omissions, clinician fatigue, inconsistency, and inequitable access, rather than an idealized standard of clinician performance. Health care already delegates selected clinical tasks to nonphysician health professionals and, under professional supervision, to support workers [18,19]. These pathways offer a governance analogy but do not establish the safety or effectiveness of AI-enabled delegation. Evaluation should therefore assess whether AI-enabled delegation is appropriately bounded, supported by context-specific evidence, monitored, and accountable [9,12]. Because deployments at the same autonomy level may differ in clinical risk, reversibility, and uncertainty, evidence requirements and safeguards should reflect both delegated authority and task-specific risk.

Defining and Classifying Clinical Autonomy

We define clinical autonomy as the degree of operational authority granted to an AI-enabled system over the selection, initiation, adaptation, or completion of clinically consequential actions, including whether they may occur without contemporaneous, case-specific human authorization. This definition separates technical capability from deployed authority and applies to a specified task-action pair, population, setting, and workflow.

Five dimensions characterize the deployment. Decision influence concerns whether the system presents information, ranks options, or selects a preferred course of action. Execution authority concerns whether it can prepare, initiate, release, or complete an action. Authorization and oversight structure concerns when case-specific human authorization is required and how monitoring, review checkpoints, exception-based intervention, and retrospective review are organized. Temporal reach concerns the duration and persistence of the system's authorized operational role, from a single action to repeated actions or longitudinal management. Scope adaptability concerns whether authority remains within predefined task and action limits or extends to revising management plans and subgoals within a broader authorized mandate. Execution authority and case-specific human authorization anchor level assignment, whereas decision influence, temporal reach, and scope adaptability describe its configuration. These dimensions are not independent scoring scales, and neither operational persistence nor parameter adjustment within a predefined task scope independently determines the autonomy level.

These dimensions characterize the deployment's autonomy configuration rather than measure clinical risk. Deployments at the same level may differ substantially in potential harm, reversibility, uncertainty, and contextual risk; evidence requirements and safeguards should therefore reflect both delegated authority and task-specific clinical risk [9,12]. Classification applies to each task-action pair within its deployment context, based on the highest authority the system is authorized to exercise for that pair, rather than assigning a universal label to a model or product.

The 5-Level Clinical Autonomy Gradient

Building on established automation and human-automation taxonomies and recent clinical AI governance frameworks, we propose the clinical autonomy gradient, summarized in Figure 1 and defined in Table 1 [6,9,14-17]. The gradient classifies deployed authority rather than stages of technological maturity. Higher autonomy is optional and requires evidence that anticipated clinical benefits outweigh the additional risks of delegation, together with safeguards that can be sustained in the deployment context. A1 or A2 may remain appropriate for a given task. Existing lifecycle, validation, and risk-management guidance applies across all levels [9,11-13]. Table 1 compares permitted actions, human authorization and oversight, proposed evidence expectations, monitoring and response, and clinical risk considerations. We propose stronger safeguards at A2 and more stringent assurance of delegated execution at A3 and A4.

XSL•FO
RenderX

Table 1. The clinical autonomy gradient for AI-driven clinical decision support systems. Evidence and monitoring expectations are proposed and have not been empirically validated. Clinical risk, patient protections, and regulatory obligations require separate assessment at every level.

Level	Permitted actions	Human authorization and oversight	Proposed evidence expectations	Monitoring and response	Clinical risk considerations
A0: informational support	Displays or organizes information; no recommendations or clinical action preparation or execution	Clinician interprets information and initiates actions	Information fidelity, provenance, usability, and subgroup performance	Monitor omissions, distortion, and data quality	Harm depends on information use and consequences of errors
A1: advisory decision support	Ranks options or recommends a clinical course; no action preparation or execution	Clinician reviews, selects, and initiates actions	Clinical validity, reviewability, automation bias, and subgroup performance	Monitor recommendation quality, clinician responses, and outcomes	Risk depends on decision, urgency, and reliance
A2: human-confirmed execution	Prepares or queues actions; execution requires case-specific authorization before release	Explicit, informed, meaningful authorization before each release	A1 evidence plus tested authorization, error interception, action integrity, and recovery	Monitor review behavior, error interception, and workload	Defaults may weaken review; harm and reversibility vary by task
A3: bounded delegated action	Executes within a predefined ASoP[a]	No case-specific authorization; monitoring, exception escalation, and timely human backup	Risk-proportionate prospective evidence; tested exclusions, abstention, escalation, and recovery	Monitor scope breaches, outcomes, and escalation; downgrade or suspend as indicated	Direct exposure to action errors; risk depends on scope and context
A4: adaptive autonomous management	Revises plans and subgoals and executes actions within a broader mandate	No case-specific authorization; system-level controls, independent oversight, and stopping criteria	Strong prospective evidence; applicable legal and regulatory authorization; ordinarily research-only	Monitor plan changes, actions, outcomes, and interactions; tested containment and recovery	Broader authority may create compound risks

[a]ASoP: automation scope of practice.

A4 permits adaptive management within a broader authorized mandate and raises unresolved questions concerning licensure, legal authority, professional accountability, patient consent, and institutional capacity [12,20-22]. Given the limited maturity of current clinical agentic AI evidence [4,5], we propose that A4 systems ordinarily remain confined to controlled research environments or tightly regulated sandboxes, with applicable institutional and regulatory authorization, predefined stopping rules, independent oversight, transparent reporting, and explicit public accountability [11,12,20]. This is a proposed governance position, not an established regulatory category.

Applying the Gradient

Classification follows a proposed 5-step procedure.

First, define the deployment. Document the task-action pair, population, setting, system version, workflow, permitted actions, time horizon, and human backup. Classify functions with different permissions separately.

Second, distinguish information from advice. If the function neither prepares nor executes clinical actions, assign A0 when it organizes or displays information without recommending a clinical course, or A1 when it ranks options or recommends a course.

Third, identify human-confirmed execution. Assign A2 when the function prepares or queues an executable action that requires explicit, informed, case-specific human authorization before release.

Fourth, distinguish bounded from adaptive delegation. For execution without contemporaneous, case-specific authorization, assign A3 when authority remains within a predefined task scope, or A4 when it permits revision of management plans and subgoals and execution of clinical actions within a broader authorized mandate.

Fifth, verify and record the assignment. Check workflow documentation, configuration evidence, and action logs to verify that configured permissions match documented authorization; assess whether required human controls function in practice. Record the highest authority permitted for the task-action pair, the classification rationale, and any unresolved uncertainty.

Protocol approval or session activation does not constitute contemporaneous, case-specific authorization of each subsequent action, and repeated parameter adjustments within a fixed task may remain A3. An ineffective confirmation process at A2 represents a failure of an intended safeguard rather than authorization for A3; execution without required approval is a scope breach. Where the available evidence does not establish the permitted authority, classification should remain provisional or be withheld. This procedure makes classification reasoning explicit but does not establish interrater reliability, which requires independent evaluation.

Governing the Autonomy Gradient

Proportionate Governance Across Autonomy Levels

The gradient is proposed to translate human-oversight principles into proportionate governance expectations. Requirements need not increase uniformly across levels: the 2 boundaries change the structure of authority, not merely its degree. We propose that crossing the action-enablement boundary (A1-A2) should require stronger human-factors evidence and safeguards for

XSL•FO
RenderX

action integrity, traceability, interruption, timely backup, and recovery. Crossing the delegation boundary (A2-A3), where execution no longer requires contemporaneous, case-specific authorization, should require stronger evidence supporting delegated execution, explicit accountability, monitoring, failure detection, timely human backup, recovery, and patient protections. The extent of these requirements should remain proportionate to task-specific clinical risk, uncertainty, and reversibility. Authorization should also depend on the institution's capacity to sustain the required safeguards; where that capacity is insufficient, the authorized scope should be narrowed or deployment deferred.

Relationship to Existing Governance Frameworks

Existing frameworks emphasize lifecycle risk management, transparency, monitoring, accountability, and human oversight. The National Institute of Standards and Technology (NIST) AI Risk Management Framework 1.0 organizes risk management around govern, map, measure, and manage [9]. World Health Organization guidance emphasizes human autonomy, safety, transparency, accountability, equity, and sustainability [12]. Article 14(3) of the European Union (EU) AI Act requires human oversight measures for high-risk AI systems to be commensurate with risk, level of autonomy, and context of use [20]. These frameworks inform safeguards but do not define a clinical taxonomy of delegated operational authority. The proposed clinical autonomy gradient complements them by classifying authority for each task-action pair within its deployment context while retaining existing lifecycle and risk-based requirements.

In the United States, A0 or A1 status does not establish nondevice status under the FDA nondevice clinical decision support exclusion, which requires all 4 statutory criteria to be met [3]. For A2, order preparation alone does not determine regulatory status; assessment must consider function, intended use, inputs, and whether clinicians can independently review the basis for recommendations. For A3 or A4 treatment execution, assessment should consider whether the function supports professional recommendations or directs care. A0-A4 are governance descriptors, not regulatory classifications. Failure to satisfy the exclusion does not automatically place a function within the focus of FDA's regulatory oversight, because other exclusions or enforcement-discretion policies may apply [3]. Each function requires separate regulatory assessment, alongside applicable state professional-practice law. Institutional authorization cannot substitute for required regulatory authorization.

Automation Scope of Practice

An ASoP was originally proposed for physician-AI collaboration, with clinical tasks assigned to tiers according to risk, verifiability, and required supervision [23]. We adapt it into an institutional authorization instrument for each deployment, linked to the clinical autonomy gradient; the original task tiers and our autonomy levels are not interchangeable. A0 and A1 still require documented intended use, users, populations, exclusions, limitations, and monitoring. We propose a proportionate ASoP for A2 authorization, with more stringent conditions at A3 and A4. It records institutional operational authority rather than professional licensure and supplements existing intended-use, delegation, risk-management, and change-control documentation [9,12,13,18-20,24]. Linking it to autonomy levels, downgrading, and reauthorization is a proposed extension requiring evaluation of feasibility, burden, and governance effectiveness.

The ASoP should document the following 3 areas.

First, authorized scope. Document the task-action pair, population and setting; system, model, and configuration versions; permitted data sources, tools, and downstream systems; permitted and prohibited actions; relevant medication or device classes and patient exclusions; and duration, frequency, and renewal limits.

Second, operating conditions. Specify uncertainty and abstention criteria, monitoring requirements, escalation pathways and response times, downgrade and suspension triggers, and human availability and backup requirements.

Third, governance responsibilities. Identify an accountable owner; an approving institutional authority with clinical and safety expertise; the authorization period and review schedule; and evidence required for continuation, reinstatement, or expansion. Documentation should reflect delegated authority and clinical risk and reuse existing institutional records where appropriate. Smaller organizations may combine roles if responsibilities, expertise, and escalation pathways remain explicit.

The ASoP aims to prevent autonomy creep: gradual expansion of operational authority through updates, integrations, configuration changes, workflow shortcuts, or evolving expectations. Updates do not authorize expanded scope, and an unchanged model does not establish safety under changed deployment conditions. Material changes in action scope, tool permissions, population, setting, temporal reach, oversight, or human backup capacity should trigger impact assessment and reassessment of the ASoP and its associated autonomy safety case [9,13,24]. Any expansion of delegated authority should require an explicit, evidence-supported institutional decision and corresponding ASoP revision, not technical capability or operational convenience alone.

Governing the ASoP

A clinical service lead should coauthor the ASoP with informatics, relevant clinical specialists, patient-safety staff, and the operational owner; developers provide technical specifications and evidence. A designated institutional authority should approve the authorized scope following proportionate clinical, security, human-factors, and legal review, with appropriate patient input. The operational owner maintains the scope record and coordinates monitoring, while the clinical service and governance authority reassess the safety case after material changes or safety signals. Responsibility for approval, incident response, and suspension must be explicitly assigned [17,25].

Documentation and review burden should scale with delegated authority and clinical risk. For lower-risk A0-A1 uses, a concise intended-use record may suffice, provided applicable validation

XSL•FO
RenderX

and safeguards are documented. The proposed ASoP at A2 should include evidence that case-specific authorization functions meaningfully in practice; A3-A4 require stronger monitoring, backup, and recovery assurance. Smaller institutions may combine roles and use shared templates, specialist review, and vendor or regional evidence, while verifying local applicability and retaining the expertise and institutional accountability needed for approval. If required monitoring or backup cannot be sustained, the scope should be narrowed, case-specific authorization retained or reinstated, or deployment deferred.

Worked Example of Stable Medication Renewal

Consider a hypothetical A3 service, distinct from Utah's protocol and presented as a governance illustration rather than a validated clinical rule. Its ASoP permits a single renewal for a 30-day supply, without changing the existing regimen, of a locally approved low-risk medicine for an eligible adult with verified identity, an existing prescription, an indication meeting predefined stability criteria, and current required monitoring. It excludes repeat automatic renewals, new prescriptions, dose changes, controlled or high-risk drugs, instability, pregnancy where relevant, new contraindications or interactions, and overdue follow-up. Local clinical approval and deployment-specific evidence must justify eligibility criteria, risk designations, and scope limits. The record specifies system, model, and configuration versions, pharmacy permissions, the authorization period, patient notice, and recourse.

Release is withheld for missing monitoring, conflicting data, new symptoms, suspected adverse effects, identity failure, patient-requested review, or uncertainty beyond predefined, validated operating limits. Cases are routed to a named authorized clinician within a defined response window with a specified backup pathway; delegated release pauses when required clinical coverage is unavailable. Logs and clinical audits assess eligibility checks, withheld and transmitted orders, inappropriate renewals, missed exclusions, escalation completion, adverse events, patient corrections, and subgroup differences. Missed mandatory exclusions or out-of-scope release require immediate containment and review; persistent deterioration or lost backup requires downgrade or suspension of the affected function. Restoration requires investigation, verified correction, updated documentation, and reauthorization. An A2 fallback is appropriate only if staffing and meaningful case-specific authorization before every release remain feasible.

Autonomy Safety Case

We adapt structured clinical AI safety assurance [11] into an autonomy safety case organized by delegated authority. Each proposed expansion of authority should be supported by a structured, updatable, evidence-based argument for acceptable safety and a favorable benefit-risk profile for the specified task-action pair, relative to current care or a defined alternative workflow. The case should document principal claims, evidence, assumptions, residual risks, and unresolved evidence gaps across the complete sociotechnical deployment [11]. Its organization by autonomy level and linkage to ASoP-based downgrading and reauthorization are proposed extensions whose feasibility, burden, and governance effectiveness require evaluation.

The case should incorporate the associated ASoP and address clinical, human-factors, technical, security, equity, patient-protection, accountability, and recovery requirements. Table 2 summarizes the corresponding questions and how they change with delegated authority. Each claim should be linked to supporting evidence, a named accountable owner and responsible functions, and predefined conditions that would invalidate it or trigger reassessment. Where applicable, early-stage live clinical evaluations should follow DECIDE-AI (Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence) [26], and randomized trial protocols and reports should follow SPIRIT-AI (Standard Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence) and CONSORT-AI (Consolidated Standards of Reporting Trials–Artificial Intelligence) [27,28]. These guidelines support, but do not substitute for, appropriate study design, regulatory review, or deployment-specific safety assurance. Evidence expectations should reflect delegated authority, clinical risk, uncertainty, reversibility, and deployment context. A0-A1 evidence should emphasize information fidelity, clinical performance, usability, subgroup effects, and human-AI interaction. A2 additionally requires evidence of meaningful case-specific review, action traceability, interruption before release, and tested recovery. A3 should generally require prospective, deployment-specific evidence covering clinical or process outcomes, operational reliability, abstention, escalation, downstream consequences, and recovery in the intended workflow. A4 would require strong prospective evidence supporting the broader mandate, independent oversight, and applicable legal and regulatory authorization; evaluation should ordinarily remain within controlled research environments or tightly regulated sandboxes. The safety case should remain a living justification for authorization. Material changes under ASoP change-control criteria or postdeployment safety signals should trigger documented reassessment of both instruments and, where necessary, downgrade or suspension of the affected function. Expansion of delegated authority should require formal institutional reauthorization and, where applicable, regulatory review [9,13,24].

XSL•FO
RenderX

Table 2. Core elements of an autonomy safety case.

Element	Proposed evidence and governance questions	Implications of delegated authority
Clinical purpose and scope	Which task-action pair, population, setting, and time horizon are authorized? What is excluded?	Permissions and exclusions must be explicit
Performance and clinical utility	What evidence supports benefit and acceptable safety relative to current care or a defined alternative?	Evaluate downstream actions and interacting systems, beyond model benchmarks
Uncertainty and escalation	Are uncertainty, missing or conflicting data, abstention, escalation, and response times tested?	Withhold unsupported actions and obtain timely assistance
Human factors and oversight	Can personnel understand, modify, reject, or stop actions? Does review intercept errors?	A2 needs prospective authorization; A3-A4 need monitoring and timely intervention
Equity and access	Are subgroup performance, access, abstention, escalation, and outcomes assessed?	Identify unequal harms or transferred burdens
Security and permissions	Are identity, data, tools, actions, and persistent memory governed by least privilege?	Prevent unauthorized action from manipulation or excessive permissions
Traceability	Can inputs, versions, evidence, authorizations, actions, and downstream effects be reconstructed?	Support investigation and reassessment
Recovery and suspension	Are containment, continuity, recovery, and communication tested? Who suspends or restores authority?	Prevent propagation and require verified restoration
Accountability and patient rights	Who owns configuration, authorization, monitoring, harm response, notice, and recourse?	Assign responsibility and accessible human review

Effective Human Control and Dynamic Autonomy

Meaningful Human Oversight

The presence of a clinician does not itself establish meaningful oversight [6,22]. Authorization may become nominal when action volume exceeds review capacity, relevant information is unavailable, or reviewers cannot practically delay, modify, reject, or stop an action. Automation bias, interface defaults, cognitive load, and workflow pressure may further weaken scrutiny, leaving clinicians responsible without effective control [6,7,23].

One framework for meaningful oversight in medical AI identifies 4 interlocking conditions: epistemic capacity, cognitive space, decisional authority, and intervention effectiveness [6]. We propose translating these into 5 deployment requirements: adequate time and manageable cognitive workload; sufficient and appropriately presented information; task-specific competence; genuine authority to delay, modify, reject, or stop an action and, where feasible, reverse or mitigate its effects; and organizational support and protection for personnel who make justified overrides. For A2-A4 deployments, these requirements should be documented in the ASoP and supported by the autonomy safety case.

At A2, human authorization must be prospective, explicit, informed, case-specific, and meaningful because release of the executable action depends on it. Evaluation should assess review time, workload, access to relevant information, modifications, rejections, and error interception; approval rates or latency alone do not establish effective review. At A3, execution proceeds without contemporaneous, case-specific authorization; oversight combines ongoing monitoring, exception-based intervention, retrospective review, and risk-based audit sampling. The ASoP and safety case should specify and substantiate this arrangement, including action limits, eligibility checks, failure detection, escalation, human backup, and recovery. At A4, nominal human availability is insufficient; system-level controls should include an explicitly defined clinical mandate, enforceable exclusions, predefined stopping rules, independent oversight, and clear authority to suspend or deactivate the system. Across delegated deployments, the safety case should demonstrate that designated personnel can detect relevant failures and intervene within clinically appropriate time frames.

Proportionate Oversight Design and Evaluation

Oversight design should be proportionate to the authorized action and its clinical risks [6,20,23]. High-risk, irreversible, time-sensitive, or equity-sensitive actions may require hard stops, mandatory escalation, or independent secondary clinical review. Lower-risk actions may use passive cues, exception-review queues, or audit sampling where consistent with the assigned autonomy level. At A2, these measures supplement rather than replace explicit, informed, case-specific authorization before each release. Hard stops should be evaluated for delays, workarounds, excessive interruption, and other adverse workflow effects [29].

For A2 deployments, we propose measuring active review time and authorization latency, including their distributions and the proportion of reviews below a candidate task-specific time threshold; documented access to required clinical information; modification, rejection, and override rates with reasons; error interception in simulations and audited cases; and pending-review workload and response delays. Active review time should be distinguished from authorization latency: elapsed time may reflect queueing, while screen activity does not establish comprehension. Error-interception rates should be calculated as the number of erroneous proposals identified and prevented from release divided by the number of erroneous proposals actually reviewed, with proposal errors established through predefined simulation scenarios or independent clinical

XSL•FO
RenderX

adjudication. Metrics should be stratified by action risk, shift, reviewer workload, and patient group.

There is no universal review-time or override-rate threshold. Candidate thresholds should be prespecified, derived from representative task-specific human-factors testing, and assessed against error interception and clinically relevant workflow outcomes. They should serve as investigation signals rather than standalone criteria for effective review. Sustained short review combined with missing information access or poor error interception should trigger investigation and workflow redesign; low modification or override rates alone may reflect appropriate proposals. Missed mandatory exclusions, unauthorized release, or failed takeover require immediate containment of the affected function. In time-sensitive perioperative care, meaningful control requires accessible signal-quality information and dose limits, tested stop controls, and manual takeover within the clinically relevant response window; approval after dosing is not prospective A2 authorization. The objective is calibrated, demonstrably effective human control, rather than maximal human involvement.

Dynamic Autonomy Calibration

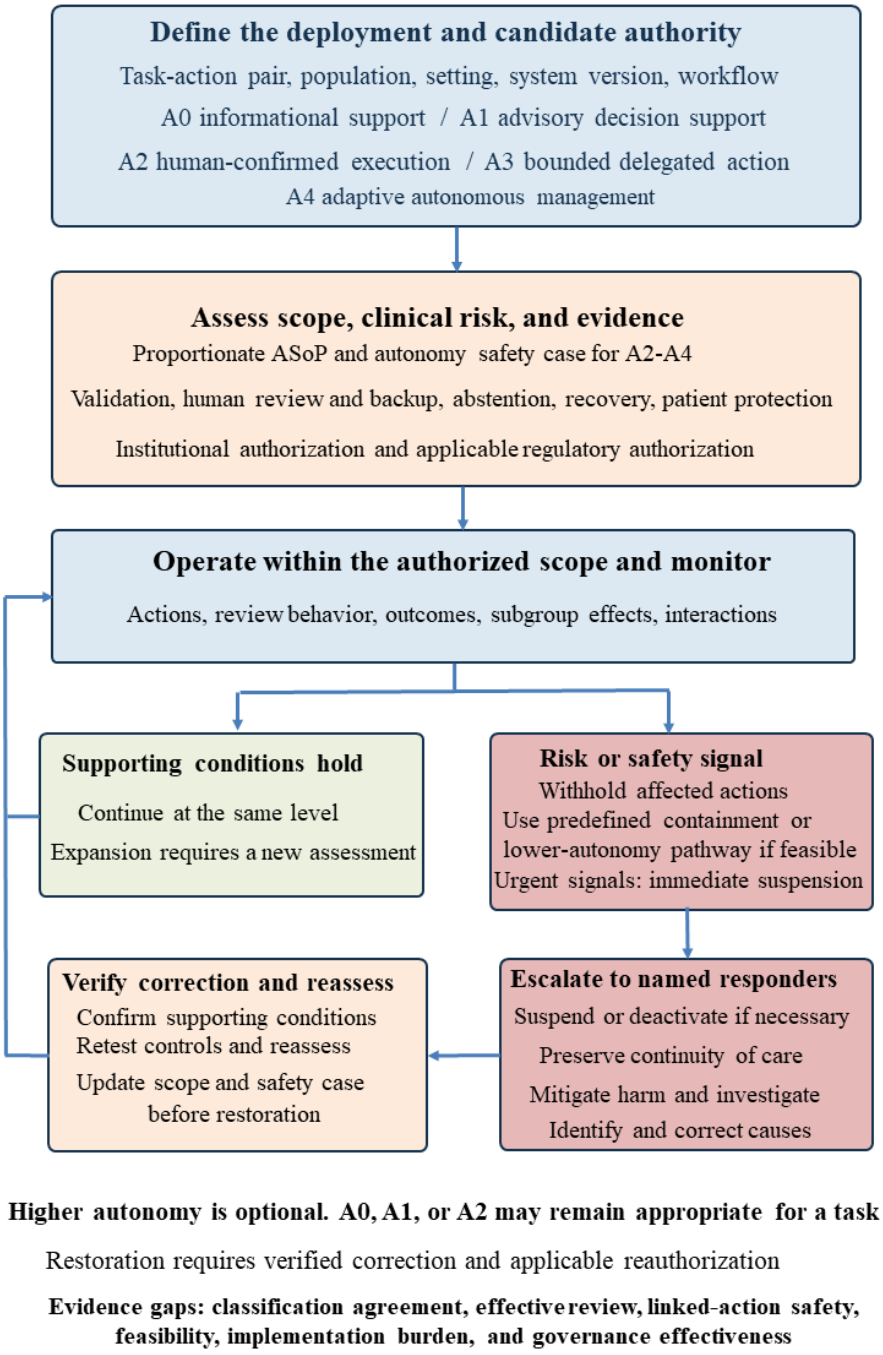
Autonomy should remain conditional rather than permanently fixed at deployment [9,13,24]. The deployment's authorized scope should remain in effect only while its supporting clinical, technical, organizational, and oversight conditions remain satisfied. Otherwise, affected actions should be withheld, routed to a predefined lower-autonomy pathway, or suspended. An A2 fallback is appropriate only when meaningful prospective review and explicit, informed, case-specific authorization remain feasible.

For A2-A4 deployments, the ASoP should specify triggers, response procedures, and responsibility for restricting or suspending authority, with automated detection where feasible. Case-level triggers may include uncertainty beyond validated limits, patient or task exclusions, and missing, conflicting, atypical, or out-of-distribution data. Deployment-level triggers may include unexpected override patterns, subgroup performance deterioration, security or system-integrity failures, excessive latency, unavailable human backup, and unvalidated system or workflow changes.

Responses should match the signal's scope and severity, contain affected actions, and escalate to named responders. Urgent signals may require immediate suspension or deactivation. Case-level fallback does not itself change the deployment's authorized classification. Figure 2 summarizes the proposed authorization and response cycle; higher autonomy is optional, and expansion requires a new assessment.

XSL•FO
RenderX

Figure 2. Proposed conditional autonomy governance. The roadmap links deployment definition, scope and evidence assessment, authorization, monitoring, containment, and reassessment. Higher autonomy is optional; A0, A1, or A2 may remain appropriate for a task. Urgent signals may require immediate suspension or deactivation. Restoration requires verified correction and applicable reauthorization and is not automatic. Case-level fallback does not itself change the deployment's authorized classification. ASoP: automation scope of practice.



Safe-State Recovery and Reauthorization

The affected function should enter a predefined safe state that preserves continuity of care and prevents actions outside the remaining authorized scope. Procedures should specify action holds or cancellations, notification of affected patients and clinicians, and activation of manual or alternative care pathways. Containment, suspension, recovery, communication, and rollback where feasible should be tested. The ASoP should identify who may suspend or deactivate the affected function or system.

Restoration requires investigation of contributing technical, workflow, and organizational factors, verified corrective action, and confirmation that supporting conditions have been reestablished. The autonomy safety case should be reassessed, with revision of the ASoP, scope, or autonomy level where necessary. Failures involving shared components or operating conditions should prompt review of other affected functions. Material changes meeting ASoP change-control criteria, and any expansion of delegated authority, should require formal institutional reauthorization and, where applicable, regulatory review [9,13,24]. Suspension remains in effect until applicable

XSL•FO
RenderX

restoration requirements are satisfied; restoration is not automatic.

Workforce Capability and Fallback Readiness

Sustained delegation of routine clinical tasks may reduce opportunities to maintain procedural familiarity, situation awareness, recognition of atypical trajectories, and takeover and recovery skills. Human-automation research supports this concern, showing that reduced involvement in task execution can impair recovery from automation failure [14,15]. However, these findings do not establish the extent of skill degradation in clinical AI deployments.

The ASoP should specify competencies to preserve, personnel responsible for assuming control, and criteria for fallback readiness. We propose periodic manual-workflow and downtime simulations, competency review, and assessment of takeover and recovery performance, with frequency and intensity proportionate to clinical risk and dependence on automation [14,15,30]. These assessments should establish whether designated personnel can evaluate and supersede AI-initiated actions and manage their consequences within clinically appropriate timeframes, including during time-sensitive scenarios. Identified readiness gaps should trigger corrective action and, where necessary, restriction or suspension of the affected function's delegated authority.

Experience with EHR downtime underscores the importance of predefined fallback procedures, staff preparation, rehearsal, and structured recovery [30]. Manual takeover and recovery must remain clinical capabilities that designated personnel can reliably exercise, supported by adequate staffing, information access, and tested procedures, without requiring continuous duplication of automated work.

Interacting Systems and Combined Workflow Risk

Individually bounded systems may create compound risks when acting on the same patient. In a hypothetical workflow, an A3 renewal agent and an A3 electrolyte agent could issue conflicting orders, exceed cumulative limits, or propagate errors from shared, stale data. The ASoP and safety case should assess the combined workflow, documenting interacting systems, shared dependencies and data provenance, action timing, and downstream permissions. Proposed safeguards include a named clinical coordinator, order reconciliation, cumulative-action limits, cross-system monitoring, and coordinated containment.

Interaction among A3 systems does not itself establish A4. A4 classification requires authority to revise management plans and subgoals and execute clinical actions within a broader authorized mandate. The proposed safeguards require evaluation; shared failures or material changes should trigger coordinated reassessment of affected scopes and safety cases.

Accountability and Patient Protection

Prospective Accountability Allocation

Accountability should reflect each actor's control over risk-producing decisions and capacity to prevent or respond to harm, rather than rest solely with the clinician nearest to the patient [21,22]. We propose allocating responsibilities to developers and vendors for technical design, performance characterization, updates, and safeguards; health systems for procurement, configuration, integration, and workflows; and institutional leaders for authorization, evidence requirements, and oversight resources [9,12,25]. Clinical responsibilities should reflect clinicians' time, information, competence, authority, and practical opportunity to intervene, without displacing applicable professional duties [6,21-23].

At A0-A1, clinicians retain primary responsibility for case-specific interpretation and decisions, while other actors remain accountable for matters within their control. A2 authorization does not absolve responsibility for the design and conditions supporting meaningful review. A3 requires explicit accountability for delegated execution, monitoring, escalation, and recovery; A4 additionally requires accountability for plan and subgoal revision within the broader mandate [6,9,12,21,22,25]. Responsibility should not rest primarily with clinicians lacking practical control over delegated actions. Legal liability remains jurisdiction- and fact-specific and cannot be inferred from autonomy level alone.

Before deployment, the ASoP and safety case should assign accountable owners and operational responsibilities for validation, configuration, authorization, monitoring, incident response, patient notification, recourse, and remediation, including compensation and regulatory reporting where applicable [9,12,20,25]. These allocations should preserve each actor's responsibility for matters within their control and be reassessed after material changes meeting ASoP change-control criteria.

Patient Disclosure and Proportionate Consent

Patients should generally be informed when AI materially shapes, directly communicates, or executes a clinically consequential action, including actions without direct patient interaction [12,31,32]. Disclosure should identify the system's role and authorized actions, whether human review precedes release or oversight relies on monitoring and escalation, and how patients can obtain human review, correct inaccurate information, or request reconsideration [12,32]. The ASoP can support accessible explanations. Disclosure should precede the relevant decision or action when patients can meaningfully express preferences or choose an alternative; otherwise, it should occur as soon as practicable, subject to applicable legal requirements.

We propose assessing the need for AI-specific consent according to delegated authority, clinical risk, reversibility, departure from standard care, and applicable legal and research requirements [12,31-33]. Autonomy level alone neither determines the need for separate consent nor replaces consent for the underlying intervention. Routine, low-risk automation within standard care may not require separate AI-specific consent when transparent notice, meaningful recourse, and human review are available, subject to applicable law and institutional policy. We propose seeking explicit AI-specific consent when delegation materially changes risks, alternatives, or human involvement, and offering an alternative pathway where clinically and operationally feasible [31,32]. Research use must satisfy applicable consent

XSL•FO
RenderX

requirements, including lawful exceptions or duly authorized waivers [12,31,32].

Consent should not substitute for evidence, monitoring, or institutional accountability. Accessible communication, reasonable accommodations, and human review pathways should address language, digital literacy, disability, and other participation barriers, including the needs of children and patients requiring legally authorized representatives [12].

Contestability, Correction, and Recourse

We propose incorporating contestability, correction, and recourse into deployment safety and governance, consistent with patient-protection principles [12,32]. Patients, caregivers, and surrogate decision-makers may provide safety-relevant information unavailable to the system, such as new symptoms, inaccurate medication histories, or changed or misunderstood preferences. Patients should be able to challenge AI-influenced decisions or actions, request timely human review, and correct relevant data through ASoP-defined pathways with named responsible owners. Responses should withhold potentially harmful pending actions where necessary and activate predefined containment or feasible lower-autonomy fallback pathways. Corrections should propagate to affected queued actions and downstream systems, with reversal or mitigation where clinically and technically feasible.

Patient-initiated corrections and disputes should inform monitoring, audit, and safety-case reassessment. Serious safety concerns or recurring disputes, particularly when disproportionately concentrated in specific patient groups, should trigger investigation and, where warranted, reassessment of the ASoP, authorized scope or autonomy level, or continued deployment. Response windows should reflect potential harm, urgency, reversibility, and whether execution could precede human review. Recourse should support prevention and mitigation of harm and review of completed actions across affected systems.

Applying the Gradient to Clinical Scenarios

Overview

These applications combine documented systems with hypothetical scenarios. Classifications are author-conducted analytical illustrations based on reported or explicitly assumed permissions, not independent classification exercises or empirical validation. Safeguards absent from the sources are presented as proposed governance conditions.

Stable Medication Renewal

Recommending renewal is A1; preparing an executable order requiring case-specific authorization is A2; releasing it within a predefined scope without such authorization is A3. Publicly available documentation describes Doctronic's Utah pilot as remaining in phase 1, requiring licensed-practitioner authorization before each renewal is transmitted [34]. This is analytically consistent with A2 but does not establish meaningful review in practice. Phase 2, which had not yet been implemented as of June 20, 2026, would permit direct submission to a pharmacist after the Utah Office of Artificial Intelligence Policy approved each medication group. Without case-specific practitioner authorization, renewal transmission would support A3 classification, without establishing safety or classifying pharmacy dispensing [34]. Documented safeguards include identity and prescription verification, medication restrictions, monthly reporting, and staged approval. Proposed expansion conditions include deployment-specific evidence, validated exclusions, escalation, backup, and recovery. Missing monitoring, changed clinical status, or failed backup should trigger withholding and human review.

Chemotherapy Dosing

Calculating body surface area from specified inputs and retrieving protocols without recommending treatment may be A0; recommending dose adjustments is A1; preparing an executable regimen requiring qualified-clinician authorization is A2. The 2024 American Society of Clinical Oncology and Oncology Nursing Society standards support multiple verification steps for antineoplastic therapy, not validation of these autonomy categories [35]. A1 or A2 may remain appropriate.

Outpatient Dyspnea

Record summarization is A0, differential diagnosis is A1, and preparation of testing or referral orders requiring authorization is A2. We propose excluding delegated diagnostic disposition and treatment initiation because incomplete context and potentially serious causes make delayed recognition consequential [7,36]. Automatic urgent-review routing or team notification could be A3 for that bounded action; recommending escalation remains A1. Evaluation should address missed alerts, false positives, queue overload, and assessment delays. Different task-action pairs within one encounter may therefore occupy different levels.

Protocolized Inpatient Potassium or Magnesium Replacement

Conventional protocols provide a delegation analogue, not evidence of AI-directed ordering safety [37,38]. Initiating replacement without case-specific approval within predefined limits would be A3. We propose prospective evaluation and an ASoP defining eligibility, laboratory and renal-function criteria, relevant urine output, interacting medications, arrhythmia risk, routes, doses, repeat testing, and monitoring. Severe or rapidly changing abnormalities should trigger withholding and escalation; A1 or A2 fallback requires adequate staffing and oversight.

Perioperative Closed-Loop Infusion

A randomized multicenter study evaluated automated propofol and remifentanil adjustment guided by a bispectral-index target [39]. Adjustment within this predefined control task is analytically consistent with A3; parameter adaptation alone does not establish A4. Trial evidence applies to the studied conventional controller and setting. For AI-enabled implementations, we propose tested responses to signal failure, discordant physiology, dose-limit breaches, and lost takeover capacity that preserve safe anesthesia and continuity of care. These conditions are not established by the cited trial.

XSL•FO
RenderX

MIRA EHR-Sandbox Simulation

MIRA performs structured actions using retrospective patient cases in a sandbox [5]. Test ordering without case-specific approval is analytically consistent with A3. Authority to revise management plans and subgoals and execute clinical actions within a broader mandate would require assessment against A4 criteria; planning capability alone is insufficient. Simulated permissions do not establish clinical authorization. Patient-care deployment would require prospective evaluation, enforceable scope limits, legal assessment, escalation, and tested recovery.

Precision Medicine

Reviews of perioperative bioinformatics and precision pain management provide context for personalized care, not evidence supporting delegated AI treatment [40,41]. Genomic or biomarker-based advice is A1; preparing an executable individualized regimen requiring authorization is A2. Evidence maturity, assay validity, subgroup applicability, delayed harms, consent, and reversibility require separate assessment. Personalization alone does not justify greater authority; limited evidence may favor research use or retained case-specific authorization.

Limitations and Research Priorities

Limitations

This viewpoint proposes a conceptual governance framework linking 5 autonomy levels to an institutional scope-of-practice record and an evidence-based safety justification. Neither the classification nor its governance benefits have been empirically validated. Analytical applications illustrate reasoning without establishing reliability or improved care. Hybrid workflows may complicate classification, and clinical risk may vary within the same action type. The 5 dimensions describe deployment configurations rather than independent scoring scales. Documented oversight may also diverge from practice: nominal confirmation can undermine A2 safeguards without altering the scope of authorization or legitimizing action taken without case-specific human approval.

The proposal draws primarily on United States and EU governance concepts; professional authority, institutional responsibility, and liability may differ elsewhere [9,12,20]. Resource-constrained settings may lack the assumed monitoring and governance capacity, while potentially benefiting from carefully bounded delegation; context-sensitive adaptation and evaluation are therefore necessary [12]. Autonomy levels also omit safety-relevant properties such as model opacity, data quality, environmental volatility, cybersecurity, and organizational safety culture, which require separate assessment [9,12].

Research and Policy Priorities

First, multidisciplinary refinement should be followed by independent empirical evaluation. A consensus process involving patients and relevant clinical, technical, governance, and legal stakeholders could refine definitions and safeguards, drawing on DECIDE-AI's modified Delphi approach [26]. Consensus would not validate the taxonomy. Independent reviewers should classify prespecified workflow dossiers, reporting agreement, chance-adjusted reliability estimates with CIs, disputed assignments, and missing information. Evaluation should include hybrid workflows and alternative category structures.

Second, comparative implementation studies should distinguish classification performance from governance effectiveness. Outcomes should include unauthorized or inappropriate actions, downstream harms, escalation and recovery, workload, subgroup effects, and implementation costs. Studies should test whether A2 review indicators predict error interception and whether proportionate documentation is feasible, following applicable DECIDE-AI, SPIRIT-AI, or CONSORT-AI guidance [26-28].

Third, health systems could extend AI inventories into registries of task-action pairs, versions, authorization status, supporting evidence, accountable owners, monitoring, and incidents [9,25]. Evaluation should assess whether these registries detect uncontrolled expansion of permissions, expose shared dependencies, and support coordinated containment.

Fourth, institutional and regulatory review should distinguish advice, action preparation, and delegated execution without treating autonomy levels as regulatory categories. Permission to recommend an action does not imply permission to prepare or execute it. Material changes should trigger institutional reassessment and applicable regulatory review [3,20,24].

Fifth, payment and liability policies should support meaningful review, monitoring, patient recourse, and recovery, without incentivizing removal of necessary safeguards. Policy development should consider actors' control over risks while preserving jurisdiction- and fact-specific legal assessments [21,22].

Finally, international comparisons should evaluate the task-action approach across legal, institutional, and resource settings [3,9,12,13,20,24]. A shared vocabulary need not imply uniform authorization thresholds, which remain dependent on applicable law, professional requirements, clinical risk, and the capacity to sustain safeguards.

Conclusion

As AI-enabled CDSSs move toward executable and delegated actions, governance must specify authority for each task-action pair within its workflow. We propose the clinical autonomy gradient, an adapted ASoP, and an autonomy safety case to define permissions, link authority to deployment-specific evidence, and organize oversight and accountability. Authorization should remain conditional on effective oversight, patient protections, and the capacity to restrict or suspend authority when supporting conditions fail. Accountability should reflect actors' operational control, while legal liability requires separate assessment. A1 or A2 may remain appropriate for a given clinical task; greater autonomy is optional and requires task- and context-specific justification. Independent evaluation is needed to establish classification reliability, feasibility, implementation burden, and effects on care before governance effectiveness can be claimed.

XSL•FO
RenderX

Funding

The authors declared no financial support was received for this work.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Generative AI Use

Following preparation of the initial draft, ChatGPT 5.5 (OpenAI) was used for language editing. The authors reviewed and, where appropriate, revised all AI-assisted text before approving the final manuscript, for which they assume full responsibility.

Authors' Contributions

JL and SL conceptualized and designed the viewpoint, conducted the literature analysis, and prepared the initial draft. All authors reviewed and revised the work for important intellectual content and approved the final version.

Conflicts of Interest

None declared.

References

1. 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. [FREE Full text] [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
2. Middleton B, Sittig DF, Wright A. Clinical decision support: a 25 year retrospective and a 25 year vision. Yearb Med Inform. 2016;25(Suppl 1):S103-S116. [FREE Full text] [doi: 10.15265/IYS-2016-s034] [Medline: 27488402]
3. Clinical decision support software: guidance for industry and Food and Drug Administration staff. US Food and Drug Administration. 2026. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software [accessed 2026-06-20]
4. Collaco BG, Haider SA, Prabha S, Gomez-Cabello CA, Genovese A, Wood NG, et al. The role of agentic artificial intelligence in healthcare: a scoping review. NPJ Digit Med. 2026;9(1):345. [FREE Full text] [doi: 10.1038/s41746-026-02517-5] [Medline: 41832341]
5. Ferber D, Hilgers L, Höper C, Kinny-Köster B, Eckardt J, Egger-Heidrich K, et al. Towards autonomous medical artificial intelligence agents. Nature. 2026;655(8125):1282-1291. [doi: 10.1038/s41586-026-10675-5] [Medline: 42310457]
6. van de Sande D, Economou-Zavlanos N, van Genderen ME. Meaningful oversight of medical AI beyond human in the loop. NPJ Digit Med. 2026;9(1):569. [FREE Full text] [doi: 10.1038/s41746-026-02971-1] [Medline: 42493536]
7. Khera R, Simon MA, Ross JS. Automation bias and assistive AI: risk of harm from AI-driven clinical decision support. JAMA. 2023;330(23):2255-2257. [doi: 10.1001/jama.2023.22557] [Medline: 38112824]
8. Sittig DF, Singh H. Recommendations to ensure safety of AI in real-world clinical care. JAMA. 2025;333(6):457-458. [doi: 10.1001/jama.2024.24598] [Medline: 39602298]
9. Tabassi E. Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg, Maryland. National Institute of Standards and Technology; 2023.
10. Yang Y, Liu J, Liu S. A futures framework for clinical AI governance: anticipating emerging risks, shifting roles, and regulatory challenges. J Med Internet Res. 2026;28:e96152. [FREE Full text] [doi: 10.2196/96152] [Medline: 42372265]
11. Festor P, Jia Y, Gordon AC, Faisal AA, Habli I, Komorowski M. Assuring the safety of AI-based clinical decision support systems: a case study of the AI Clinician for sepsis treatment. BMJ Health Care Inform. 2022;29(1):e100549. [FREE Full text] [doi: 10.1136/bmjhci-2022-100549] [Medline: 35851286]
12. Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva, Switzerland. World Health Organization; 2021.
13. Good machine learning practice for medical device development: guiding principles. International Medical Device Regulators Forum. 2025. URL: https://tinyurl.com/vp9tfsww [accessed 2026-06-20]
14. Parasuraman R, Sheridan TB, Wickens CD. A model for types and levels of human interaction with automation. IEEE Trans Syst Man Cybern A Syst Hum. 2000;30(3):286-297. [doi: 10.1109/3468.844354] [Medline: 11760769]
15. Endsley MR, Kaber DB. Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics. 1999;42(3):462-492. [doi: 10.1080/001401399185595] [Medline: 10048306]
16. Shujaat S. A six-level clinical autonomy framework for artificial intelligence in dentistry. Front Oral Health. 2026;7:1836492. [FREE Full text] [doi: 10.3389/froh.2026.1836492] [Medline: 42453389]
17. Wang A, Freeman S, Magrabi F. Governance for safe and responsible AI in healthcare organisations: a scoping review of frameworks. NPJ Digit Med. 2026;9(1):516. [FREE Full text] [doi: 10.1038/s41746-026-02679-2] [Medline: 42062453]

XSL•FO
RenderX

18. Mirhoseiny S, Geelvink T, Martin S, Vollmar HC, Stock S, Redaelli M. Does task delegation to non-physician health professionals improve quality of diabetes care? Results of a scoping review. PLoS One. 2019;14(10):e0223159. [FREE Full text] [doi: 10.1371/journal.pone.0223159] [Medline: 31603900]
19. Shore CB, Maben J, Mold F, Winkley K, Cook A, Stenner K. Delegation of medication administration from registered nurses to non-registered support workers in community care settings: a systematic review with critical interpretive synthesis. Int J Nurs Stud. 2022;126:104121. [FREE Full text] [doi: 10.1016/j.ijnurstu.2021.104121] [Medline: 34896760]
20. European Parliament, Council of the European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Publications Office of the European Union. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj [accessed 2026-06-21]
21. Habli I, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. Bull World Health Organ. 2020;98(4):251-256. [FREE Full text] [doi: 10.2471/BLT.19.237487] [Medline: 32284648]
22. Cavalcante Siebert L, Lupetti ML, Aizenberg E, Beckers N, Zgonnikov A, Veluwenkamp H, et al. Meaningful human control: actionable properties for AI system development. AI Ethics. 2023;3(1):241-255. [doi: 10.1007/s43681-022-00167-3]
23. Pan H, Liu J, Liu S. Enhancing physician resilience to generative AI: multilevel framework for shared authority, verification, and skill preservation. J Med Internet Res. 2026;28:e88058. [FREE Full text] [doi: 10.2196/88058] [Medline: 42341297]
24. Marketing submission recommendations for a predetermined change control plan for artificial intelligence-enabled device software functions: guidance for industry and Food and Drug Administration staff. US Food and Drug Administration. 2025. URL: https://tinyurl.com/2s7n2v2u [accessed 2026-06-21]
25. Loufek B, Vidal D, McClintock DS, Lifson M, Williamson E, Overgaard S, et al. Embedding internal accountability into health care institutions for safe, effective, and ethical implementation of artificial intelligence into medical practice: a Mayo Clinic case study. Mayo Clin Proc Digit Health. 2024;2(4):574-583. [FREE Full text] [doi: 10.1016/j.mcpdig.2024.08.008] [Medline: 40206521]
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):924-933. [doi: 10.1038/s41591-022-01772-9]
27. Rivera SC, Liu X, Chan A-W, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nat Med. 2020;26(9):1351-1363. [FREE Full text] [doi: 10.1038/s41591-020-1037-7] [Medline: 32907797]
28. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-1374. [FREE Full text] [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]
29. Powers EM, Shiffman RN, Melnick ER, Hickner A, Sharifi M. Efficacy and unintended consequences of hard-stop alerts in electronic health record systems: a systematic review. J Am Med Inform Assoc. 2018;25(11):1556-1566. [FREE Full text] [doi: 10.1093/jamia/ocy112] [Medline: 30239810]
30. Lyon R, Jones A, Burke R, Baysari MT. What goes up, must come down: a state-of-the-art electronic health record downtime and uptime procedure in a metropolitan health setting. Appl Clin Inform. 2023;14(3):513-520. [FREE Full text] [doi: 10.1055/s-0043-1768995] [Medline: 37406674]
31. Rose SL, Shapiro D. An ethically supported framework for determining patient notification and informed consent practices when using artificial intelligence in health care. Chest. 2024;166(3):572-578. [FREE Full text] [doi: 10.1016/j.chest.2024.04.014] [Medline: 38788895]
32. Hurley ME, Lang BH, Kostick-Quenet KM, Smith JN, Blumenthal-Barby J. Patient consent and the right to notice and explanation of AI systems used in health care. Am J Bioeth. 2025;25(3):102-114. [doi: 10.1080/15265161.2024.2399828] [Medline: 39288291]
33. Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical considerations of using ChatGPT in health care. J Med Internet Res. 2023;25:e48009. [FREE Full text] [doi: 10.2196/48009] [Medline: 37566454]
34. Utah Department of Commerce, Office of Artificial Intelligence Policy. Doctronic AI regulatory mitigation agreement. Utah Department of Commerce. 2025. URL: https://commerce.utah.gov/ai/regulatory-relief/authorized-ai-pilots/doctronic/ [accessed 2026-06-20]
35. Siegel RD, LeFebvre KB, Temin S, Evers A, Barbarotta L, Bowman RM, et al. Antineoplastic therapy administration safety standards for adult and pediatric oncology: ASCO-ONS standards. JCO Oncol Pract. 2024;20(10):1314-1330. [doi: 10.1200/op.24.00216]
36. Berliner D, Schneider N, Welte T, Bauersachs J. The differential diagnosis of dyspnea. Dtsch Arztebl Int. 2016;113(49):834-845. [FREE Full text] [doi: 10.3238/arztebl.2016.0834] [Medline: 28098068]
37. Pearson DJ, Sharma A, Lospinoso JA, Morris MJ, McCann ET. Effects of electrolyte replacement protocol implementation in a medical intensive care unit. J Intensive Care Med. 2018;33(10):574-581. [doi: 10.1177/0885066616679593] [Medline: 27881698]
38. O'Neill K, Bloomer MJ. An integrative review of potassium replacement protocol use in critical care: development, use and critical care nurse autonomy. Intensive Crit Care Nurs. 2023;79:103524. [FREE Full text] [doi: 10.1016/j.iccn.2023.103524] [Medline: 37598503]

XSL•FO
RenderX

39. Liu N, Chazot T, Hamada S, Landais A, Boichut N, Dussaussoy C, et al. Closed-loop coadministration of propofol and remifentanil guided by bispectral index: a randomized multicenter study. Anesth Analg. 2011;112(3):546-557. [doi: 10.1213/ANE.0b013e318205680b] [Medline: 21233500]
40. Jahangeerbasha A, E R, Suchitra MR, Kumar V R H, P R, Ravindran C, et al. Bioinformatics in anaesthesia: a new era of personalised and predictive perioperative care. Indian J Clin Anaesth. 2026;13(1):34-44. [doi: 10.18231/j.ijca.14586.1766463788]
41. Parthasarathy S, Krubaa P, Kadaksha P, Geetha K, Jahangeerbasha A, Suchitra M, et al. Emerging frontiers in precision pain management in anaesthesia: an integrative narrative review of biomaterials, gene editing, and microbiome modulation. Indian J Clin Anaesth. 2026;13(3):412-425. [doi: 10.18231/j.ijca.16562.1781243123]

Abbreviations

ASoP: automation scope of practice
CDSS: clinical decision support system
CONSORT-AI: Consolidated Standards of Reporting Trials–Artificial Intelligence
DECIDE-AI: Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence
EHR: electronic health record
EU: European Union
FDA: Food and Drug Administration
MIRA: Medical Intelligence for Reasoning and Action
NIST: National Institute of Standards and Technology
SPIRIT-AI: Standard Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence

Edited by J Khuntia; submitted 29.Jul.2026; peer-reviewed by A Jahangeerbasha, M Bhatt; comments to author 30.Sep.2026; revised version received 10.Oct.2026; accepted 04.Nov.2026; published 24.Nov.2026

Please cite as:
Liu J, Liu S
From Decision Support to Delegated Clinical Action: Governing the Autonomy Gradient of AI-Driven Clinical Decision Support Systems
J Med Internet Res 2026;28:e108260
URL: https://www.jmir.org/2026/1/e108260
doi: 10.2196/108260
PMID:

©Jialin Liu, Siru Liu. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 24.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX