Subdomain 4.2: Risk and Impact Assessments Playbook

Guiding responsible AI adoption

Before You Start

For more about the CHAI AI Governance Framework, methods used to develop these playbooks, and information on how to use them, please refer to the CHAI AI Governance Playbooks: Introduction & Background document. Note that **baseline operational controls** were consensus-defined by Healthcare Delivery Organizations (HDOs) as "must-haves" for health AI governance. The **implementation guidance** that accompanies the baseline operational controls are **suggestions/recommendations** based on known practices. It is assumed and appropriate that implementation will differ based on organizational goals, resources, and maturity.

Disclaimer

The implementation guidance, frameworks, and examples contained herein represent generalized practices and considerations drawn from across the field of responsible AI deployment in healthcare. They are offered as one set of perspectives to inform your thinking not as the "right way" or "only way" to build, govern, or operate an AI program.

Every organization operates within its own unique context: differing patient populations, clinical priorities, regulatory environments, resource constraints, technical infrastructure, cultural norms, and stage of AI maturity. While the specific methods, timelines, and depth of implementation will appropriately vary across organizations, the underlying controls, principles, risks, and considerations raised throughout this playbook are intended to be engaged with thoughtfully and in good faith by any organization deploying AI in healthcare.

Organizations should prioritize state and federal regulatory requirements, especially for AI solutions that meet the definition of a Software as a Medical Device (SaMD), as defined by the FDA.

Not a Standard of Care. This playbook does not establish, define, or reflect a legal, regulatory, professional, or industry standard of care, nor is it intended to serve as one. It is a voluntary, aspirational, and educational resource designed to support organizational learning and dialogue. The field of healthcare AI is rapidly evolving, and reasonable organizations and professionals may differ on how to approach the issues discussed herein. The practices described may not be appropriate, feasible, or advisable for every organization, every use case, or every point in time. Adoption of any particular practice described in this playbook is entirely voluntary, and a decision to adopt, adapt, partially implement, or decline to implement any practice should not be construed as evidence of compliance with, or deviation from, any standard of care, duty, or legal obligation. This playbook is not intended to be used, and should not be used, as a benchmark, checklist, or measuring stick against which the conduct of any

organization, clinician, or other party is evaluated in any litigation, regulatory proceeding, accreditation review, or similar context.

General Informational Purpose. The information in this playbook is provided for general informational and internal operational purposes only. It is intended to support organizations in thinking through the design and stewardship of their AI programs and should be used in conjunction with from clinical leaders, technical experts, compliance officers, ethicists, patient and community representatives, and other advisors familiar with the specifics of your organization.

Not Legal Advice. Nothing in this playbook constitutes legal advice, and it should not be relied upon as a substitute for advice from a qualified attorney licensed to practice in the relevant jurisdiction. Nothing in this document creates an attorney-client relationship between the reader and CHAI, its officers, directors, employees, or contributors. The contents reflect general information and observed practices as of the date of publication and may not account for changes in law, regulation, technology, or circumstance that occur thereafter.

No Warranty. While reasonable efforts have been made to ensure accuracy, no representation or warranty is made as to the completeness, timeliness, accuracy, or suitability of the information for any particular purpose. Laws, regulations, and best practices vary by jurisdiction and evolve over time, and the application of those laws and practices to specific facts can lead to different outcomes.

Illustrative Examples. Any tools, resources, examples, or technology product described or presented are for educational and illustrative purposes only and do not represent endorsement by CHAI or guarantee specific results. Readers should consult qualified legal counsel and relevant experts for specific terms, products, or services.

Independent Judgment Required. Readers should consult with qualified legal counsel, clinical experts, and other appropriate professional advisors before taking, or refraining from taking, any action based on the information in this playbook. The decision to adopt, modify, or disregard any practice described herein rests solely with the reader and their organization. CHAI expressly disclaims any and all liability for any loss, damage, or other consequence, including but not limited to claims of negligence, breach of duty, or failure to meet any standard, arising directly or indirectly from reliance on, use of, citation to, or reference to the contents of this document.

At a Glance

Subdomain 4.2: Risk and Impact Assessments Playbook defines the approach to categorize, assess, mitigate, and monitor risks associated with AI solutions.

Background Context for Subdomain 4.2: Risk and Impact Assessments

This playbook outlines a flexible framework along with implementation guidance for HDOs to responsibly manage the risks of AI solutions. By focusing on five baseline operational controls, organizations of any size (from community health centers to large academic medical centers) can implement a structured approach to

categorize, assess, mitigate, and monitor risks associated with AI solutions. Risk management for health AI solutions can broadly be defined in four phases:

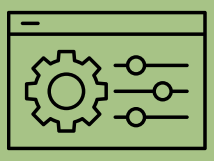
1. **Risk Categorization**: Classify AI solutions as Low, Medium, or High risk during pre-deployment. This determines the level of rigor needed for later assessments and controls.
2. **Risk Assessment**: For AI solutions with (at least) High risk, conduct a detailed analysis based on the organization's risk tolerance.
3. **Risk Mitigation**: Implement and document actions (risk mitigation controls) to reduce identified risks.
4. **Risk Monitoring:** Continuously monitor the AI solution's performance and safety over time.

Note: AI Risk & Enterprise Risk Management (ERM) Integration

AI risk categorization and assessment processes should integrate into the organization's broader Enterprise Risk Management (ERM) framework. High-risk AI solutions should be visible within existing risk registers and, where appropriate, escalated to executive or board-level oversight consistent with the organization's risk governance model. Boards are increasingly asking: "*What are our top AI risks?*" and "*Where are they reflected in ERM?*". AI risk governance is converging with enterprise risk governance.

Note: Level of Risk Categorization and Assessment

Risk categorization and assessment should be conducted at the level of the specific AI *solution* and its defined use case (e.g., ambient technology, AI-supported patient scheduling, AI-supported revenue cycle management), not at the level of the underlying foundation model. This ensures that the unique risks introduced by the specific use case context, data inputs, and workflow integration are fully evaluated.



Baseline Operational Controls for Subdomain 4.2: Risk and Impact Assessments

Conduct a Risk Categorization assessment (pre-deployment) to categorize risk in the context of the organization's AI use case scope.
Document how Risk Categorization is evaluated, recorded, retained, and reviewed.
Leverage a rigorous Risk Assessment to evaluate AI solutions that are categorized as higher risk (as defined by the organization).
Document how Risk Assessment is evaluated, recorded, retained, and reviewed.
Conduct and document an AI System Impact Assessment that includes evaluation of risks and benefits for AI solutions.

Conduct Pre-deployment Risk Categorization

Control 4.2.1: Conduct a Risk Categorization assessment (pre-deployment) to categorize risk in the context of the organization's AI use case scope.

Goal: Perform a high-level categorization of an AI use case to determine the appropriate level of oversight and rigor required for subsequent risk assessments.

Implementation Guidance

HDOs should evaluate AI solutions across specific risk domains. High-priority risk domains include components of risk that contribute to Life & Patient Safety and Technology & Data. Additional risk domains to consider include Financial, Legal, Operational, and Reputational risk. Organizations should consider their own risk domains depending on overall organizational risk tolerance. Use cases should be categorized as Low, Medium, or High risk based on the context of their intended use. Categorization should also account for how risk may stack, multiply, or attenuate based on operational context (e.g., staffing levels, shift coverage) and aggregate use (e.g., multiple AI solutions influencing the same workflow or patient population). Where contextual or cumulative factors materially change risk exposure, the categorization should be adjusted accordingly. This initial categorization acts as a triage step to prioritize resources for the most critical AI solutions.

Note: Emerging Risk Domains for Agentic AI

As agentic AI capabilities expand, HDOs should begin tracking additional risk domains beyond those listed above. These may include:

- Autonomy: The level of independent decision-making authority granted to the agent.
- Irreversibility: The difficulty or impossibility of reversing actions an agent has taken.
- Agent-to-Agent: Risks arising from interactions or hand-offs between multiple AI agents.

Within these domains, organizations should also consider risk modifiers such as level of agent decision-making authority, scope creep potential, and the presence/quality of defined stop conditions.

Risk Appetite & Tolerance Thresholds: HDOs should formally define AI-related risk appetite and tolerance thresholds across risk domains. These thresholds should be set by the designated AI governance authority (e.g., AI governance committee or equivalent body) in coordination with Enterprise Risk Management, and documented in the organization's AI policy or equivalent reference document so they can be cited consistently in decisions. Risk categorization decisions should reference these predefined tolerance levels.

Key Tools & Resources

CHAI Risk Categorization Tool (v2)

- *The Risk Categorization Tool is for health systems (any size) and teams responsible for pre-deployment risk review. This tool supports early AI governance by helping identify low, medium, or high-risk levels across several risk modifiers related to the risk domains (1) Life & Patient Safety and (2) Technology & Data. The tool highlights key dimensions of risk to guide further assessment, mitigation, and monitoring.*
- *Use the Risk Categorization Tool (v2)* as a multi-stakeholder, consensus-defined tool to help operationalize the organization's pre-deployment categorization of risk.

CHAI Risk Mapping: CHAI Risk Categorization Tool (v2) mapped to NIST AI RMF and EU AI Act

- *CHAI developed a detailed information mapping between the CHAI Risk Categorization Tool (v2), National Institute of Standards and Technology (NIST) AI Risk Management Framework (RMF), and European Union (EU) AI Act.*

Implementation Guiding Questions

Below are standard questions that HDOs may ask to guide discussion during the initial categorization of risk for **control 4.2.1**:

- How operationally or physically close is the AI output to the patient (e.g., back-office administration vs. clinical diagnosis support)?
- To what extent is a human involved in reviewing, verifying, or overriding AI outputs before they affect care?
- What is the potential severity of negative outcomes, such as morbidity or mortality, if the AI provides incorrect information?
- Is the solution used for patients in sensitive groups, such as pediatrics, older adults (aged 65 and older), or populations underrepresented in training data?
- Does the system rely on identified protected health information (PHI) or de-identified data with residual re-identification potential?
- What type(s) of AI is the solution using (large language model, regression model)?
- How likely is the AI output to directly impact patient care or patient care decisions?
- How and where is the data stored for the purpose of the intended use case? How is the data stored or deleted after its intended use? Is the data subject to backups, replication, or other retention measures that should be accounted for in the risk categorization?
- Does the AI solution touch PHI within the proposed workflow?
- What is the level of difficulty in detecting unintended results, negative effects, or model drift?
- To what extent does the organization fully control the AI and its outputs (e.g., does it rely on inputs from outside the organization or send outputs to outside actors)?
- What is the technical load and integration burden of deploying this AI solution (e.g., interface dependencies, data pipeline changes, downstream system impacts, and ongoing maintenance overhead)?
- How sensitive is the information being generated or inferred by the AI (e.g., predictions of risk for developing a chronic condition or other inferences that could carry downstream consequences for the patient)?
- **Important**: How deeply is the AI embedded in the health IT environment, and could an error in this solution cascade or propagate to other systems and workflows?

Document and Retain the Risk Categorization

Control 4.2.2: Document how Risk Categorization is evaluated, recorded, retained, and reviewed.

Goal: Create a transparent, auditable trail for why specific AI solutions were assigned their respective risk levels.

Implementation Guidance

Standardizing the categorization of use cases is essential for consistency. All documentation regarding these decisions should be retained and reviewed periodically to account for changes in the AI's performance or its environment. HDOs should keep record of the specific evidence and rationale (e.g., vendor model card information, vendor intake questionnaires, or other supplemental AI solution documentation to use as evidence for the team's risk categorization decision).

For AI features/capabilities increasingly embedded by vendors directly into systems already in use, HDOs should define what documentation they require vendors to provide (e.g., model/solution cards, transparency reports, among several other potential requirements) to support ongoing risk categorization and auditability. CHAI recommends establishing a vendor AI Intake process as part of procurement. Additional detail and resources about the AI intake process is in the Subdomain 4.1: Responsible AI Lifecycle Management and Use Playbook.

Illustrative Example

HDOs may classify all administrative AI use cases that do not access patient data as "Low Risk" by default. The criteria for this categorization, and supporting evidence from the AI solution, should be formally documented to ensure transparency and auditability.

Implementation Guiding Questions

Below are standard questions that HDOs may ask to guide discussion for **control 4.2.2**:

- Who is the designated official or governance body responsible for reviewing and signing off on the risk categorization, and what is the defined scope of their authority and accountability for this decision? Organizations should elevate this designation formally, specifying decision-making rights and reporting obligations.
- Are the known limitations of the AI system, such as specific populations where the model is less accurate, documented during the categorization process? In addition, has the organization documented the methods and effort used to identify those limitations (e.g., vendor disclosures requested, internal testing performed, literature reviewed) so that gaps in knowledge are visible?
- How often will the categorization be revisited to ensure it still reflects the current use case scope (e.g., review cadence)?

Leverage Risk Assessment for Higher Risk AI Solutions

Control 4.2.3: Leverage a rigorous Risk Assessment to evaluate AI solutions that are categorized as higher risk (as defined by the organization).

Goal: Conduct a deeper-dive analysis for AI solutions identified as higher risk, focusing on hazards, harms, and the implementation of mitigation controls.

Implementation Guidance

For higher risk use cases, a detailed analysis of the probability and magnitude of harm is recommended. This includes evaluating risk modifiers such as population sensitivity, disparity risk, data provenance verifiability (with models whose training data provenance is unknown or unverifiable carrying a higher default risk regardless of intended use), and the sufficiency of information about the solution's training data. Organizations should implement risk mitigation controls (i.e., specific actions taken to reduce identified hazards) and then re-evaluate the risk level to determine if the residual risk is acceptable. Each organization may have a different definition of "High Risk", based on risk tolerance and other organizational variables. The goal of this control is not to establish a definition for "High Risk", but to ensure that when high risk is determined, organizations conduct the appropriate level of rigor and risk assessment.

Note

Per *ISO 14971:2019: Medical devices — Application of risk management to medical devices*, definitions of hazards, harms, and related risk assessment terms are included:

- Harm: injury or damage to the health of people, or damage to property or the environment
- Hazard: potential source of *harm*
- Hazardous situation: circumstance in which people, property or the environment is/are exposed to one or more *hazards*
- Severity: measure of the possible consequences of a *hazard*
- Risk: combination of the probability of occurrence of *harm* and the *severity* of that *harm*
- Risk Control: process in which decisions are made and measures implemented by which *risks* are reduced to, or maintained within, specified levels
- Residual Risk: *risk* remaining after *risk control* measures have been implemented

HDOs should familiarize themselves with these terms and discuss how to operationalize within internal workflows.

Note

In addition to *ISO 14971:2019*, HDOs should evaluate whether the AI solution:

- Meets criteria for U.S. Food and Drug Administration (FDA)-regulated Software as a Medical Device (SaMD);
- Implicates Conditions for Coverage (CfCs) & Conditions of Participation (CoPs) | CMS related to clinical decision-making or patient safety;
- Introduces privacy obligations under the Health Insurance Portability and Accountability Act (HIPAA).

HDOs should familiarize themselves with applicable regulatory frameworks and integrate this analysis into the formal risk assessment.

Higher risk categorization should trigger defined governance actions, including: formal committee review, enhanced post-deployment monitoring, documented residual risk acceptance by designated governance authority, and defined re-review triggers (e.g., after a system update, incident, or drift detection).

Additionally, HDOs should define explicit conditions under which an AI solution will *not* proceed to deployment, including:

- Unmitigated high-severity harm potential that cannot be reduced through available controls;
- Vendor refusal to provide certain AI Intake/procurement documentation (e.g., model cards);
- Residual risk exceeding predefined organizational tolerance thresholds.
- A continuously adaptive or self-calibrating AI system for which the vendor cannot provide defined governance procedures specifying: (a) the criteria under which a calibration change constitutes a governance notification event, (b) the disclosure mechanism for such events, and/or (c) whether such changes trigger re-risk-categorization.

Decisions to halt deployment under these conditions should be made by the designated AI governance authority (e.g., AI governance committee), and organizations should define a documented appeal or reconsideration pathway (e.g., who may submit an appeal, what new evidence or mitigations would warrant re-review, and the body responsible for the final determination).

Key Tools & Resources

ISO 14971:2019: Medical devices — Application of risk management to medical devices

- *This standard establishes a systematic framework for identifying hazards associated with medical devices, estimating and evaluating the associated risks, controlling these risks, and monitoring the effectiveness of the controls throughout the entire device lifecycle.*

Security Risk Assessment (SRA) Tool (v3.6)

- *The Office of the National Coordinator for Health Information Technology (ONC), in collaboration with the U.S. Department of Health and Human Services (HHS) Office for Civil Rights (OCR), developed a downloadable Security Risk Assessment (SRA) Tool. The target audience of this tool is medium- and small-sized providers.*

NIST AI Risk Management Framework (AI RMF) & Playbook

- *A comprehensive framework for managing AI risks, including a structured playbook with actionable guidance across the AI lifecycle.*

International Medical Device Regulators Forum (IMDRF) Software as a Medical Device (SaMD) Clinical Risk Categorization Framework

- *An international framework for categorizing the clinical risk of Software as a Medical Device.*

Failure Modes and Effects Analysis (FMEA)

- *A structured methodology for identifying, prioritizing, and mitigating potential failure points in AI systems. FMEA is widely used in healthcare quality and safety contexts and can be adapted for AI risk assessment workflows.*

Implementation Guiding Questions

Below are standard questions that HDOs may ask to guide discussion for **control 4.2.3**:

- Harm Identification: What specific clinical, operational, or legal harms could emerge from the use of this AI solution?
- Data Sufficiency: Is the training and testing data representative of the specific patient populations the organization serves? Organizations should define quantitative subgroup performance disparity thresholds that trigger escalation or deployment pause when thresholds aren't met.
- Security Resilience: How exposed is the model to security vulnerabilities, such as data poisoning or adversarial examples, based on its technical deployment?
- Mitigation Strategy: If a high-risk of bias or error is identified, what specific controls (e.g., algorithmic adjustments, increased human oversight, or restricted use cases) are being put in place?
- For vendor-provided or co-developed AI systems, have contractual risk allocation, indemnification provisions, and/or incident notification requirements been reviewed and found acceptable?
- What plan is in place to temporarily or permanently remove the AI in the event that becomes necessary?

Document and Retain the Risk Assessment

Control 4.2.4: Document how Risk Assessment is evaluated, recorded, retained, and reviewed.

Goal: Ensure that the rigorous risk assessments and their associated mitigation strategies are formally documented and kept current throughout the AI lifecycle.

Implementation Guidance

Risk management is a continuous, iterative process. Documentation should include the results of testing against predefined metrics, identified data gaps, and the effectiveness of implemented risk mitigation controls. This documentation should be updated whenever the system undergoes significant changes, such as retraining or a shift in use case setting. There are a few key items that HDOs should keep in mind when documenting and conducting their risk assessment(s):

- Residual Risk Acceptance Sign-Off: Residual risk acceptance for higher risk AI systems should require documented sign-off by the designated governance authority, consistent with organizational risk governance policies. Acceptance decisions should be time-bound and subject to re-evaluation at defined intervals or when material changes occur.
- Risk and Mitigation Control Libraries: Organizations may consult established control libraries to inform the development of AI-specific risk mitigation controls. Recommended references, such as the NIST AI Risk Management Framework (AI RMF), CHAI's Risk Categorization Tool, and/or third-party governance platforms that developed their own risk and mitigation control libraries are listed in the Tools & Resources section of this playbook. This work is still in early stages, but risk mitigation libraries are increasingly useful as HDOs assemble a standard set of mitigation controls that can be applied to an AI solution once appropriate risk categorization and assessment have been completed.
- Measuring Effectiveness of Risk and Mitigation Controls: Organizations should define how they will measure whether implemented risk and mitigation controls are achieving their intended effect. Approaches may include: tracking AI output override rates; monitoring incident reports and near-miss events; reviewing performance drift indicators; conducting periodic stakeholder feedback surveys; and comparing pre/post-mitigation performance metrics against predefined thresholds.

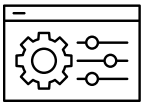
Implementation Guiding Questions

Below are standard questions that HDOs may ask to guide discussion for **control 4.2.4**:

- Version Control & Change Management: Is there a clear version history that documents when the risk assessment was conducted, what assumptions were made, and what system changes (e.g., retraining, new integrations, expanded user groups) triggered reassessment?
- Testing Evidence: Are validation results, performance metrics, and known limitations formally linked to the identified hazards and corresponding mitigation controls?
- Incident Response: Is there a documented plan for detecting and responding to output anomalies, drift, or security breaches?
- Residual Risk Acceptance: Has the designated governance authority explicitly reviewed and formally accepted the documented residual risk after mitigation controls were applied?
- Accountability & Escalation Authority: Who is the designated governance authority (and under what predefined conditions) to pause, restrict, or fully deactivate the system if performance degrades or emergent safety risks arise?

- Review Cadence: Is there a defined timeline (e.g., quarterly, annually, post-incident) for formally re-reviewing the risk assessment to ensure it remains accurate in the current clinical and technical environment?
- Risk and Mitigation Control Effectiveness: How will the organization measure whether implemented risk and mitigation controls are working as intended?

Conduct and Document an AI System Impact Assessment



Control 4.2.5: Conduct and document an AI System Impact Assessment that includes evaluation of risks and benefits for AI solutions.

Goal: Evaluate the overall balance of risks and benefits for AI solutions, considering the impact on all relevant stakeholders.

Implementation Guidance

Distinguishing the Three Assessment Types:

Assessment Type	Scope	Applies To
Risk Categorization	High-level triage; assigns Low/Medium/High risk category	All AI solutions (pre-deployment)
Risk Assessment	Deep-dive hazard, harm, and mitigation assessment	Higher-risk AI solutions (threshold defined by the organization)
AI System Impact Assessment	Broader system-level consequences; risk and benefit across stakeholders	Higher-risk AI solutions (threshold defined by the organization)

While risk assessments focus on identifying and mitigating specific hazards, an AI System Impact Assessment evaluates the broader, system-level consequences of deploying the AI solution. This includes examining how the AI solution impacts and aligns with organization culture, value, patient needs, and financial incentives. An impact assessment should consider both intended and unintended consequences, such as shifts in decision-making patterns, changes in patient-provider interactions, or downstream financial incentives among other examples.

Not all AI solutions will require a formal impact assessment; however, organizations should define clear internal thresholds (e.g., risk level, scale of deployment, patient-facing functionality) that trigger one. Specifically, organizations should articulate when a full impact assessment is required (typically for higher-risk, patient-facing, or broadly deployed solutions) versus when a lighter-weight review is sufficient, and document the criteria that distinguish the two paths. The decision to conduct an impact assessment should align with the organization's risk tolerance and governance maturity. The AI System Impact Assessment is intended to align with, and may extend,

existing Privacy Impact Assessment (PIA) and Data Protection Impact Assessment (DPIA) processes so that organizations can build on established workflows.

HDOs should periodically evaluate cumulative AI risk across all deployed systems, particularly where multiple AI solutions influence the same workflow, patient population, or workforce staff (e.g., clinicians, nurses, or other employees who participate in multiple AI-supported workflows simultaneously). Portfolio-level risk should be reviewed at defined governance intervals (e.g., semi-annually or annually) and reported to the designated governance authority.

Note: A Brief Operational Process for Portfolio-level Risk Review

Included below are a few steps HDOs may take when conducting a portfolio-level risk review:

(1) Define Scope: Which AI systems are included based on shared workflow or patient population touch points?

(2) Develop a Core Assessment Question(s): e.g., "Does the combined operation of these systems create risks not present in any individual system, including interaction effects, compounding errors, or population-level automation bias?"

(3) Establish a Recommended Cadence: e.g., at least annually as a standing governance committee agenda item

(4) Create a Decision Tool: e.g., a simple interaction matrix template showing AI systems on both axes with cells indicating workflow intersection type.

While the risk-benefit analysis component of the Impact Assessment is presented here within the Subdomain 4.2: Risk and Impact Assessments Playbook, organizations may also consider integrating this analysis with procurement decision-making processes. Cross-referencing this assessment with the Subdomain 4.4: Third Party Management Playbook is encouraged to ensure alignment between risk-benefit and vendor selection decisions.

Key Tools & Resources

ISO/IEC 42005:2025: Information technology — Artificial intelligence (AI) — AI system impact assessment

- *This standard supports transparency, accountability, and trust in AI by helping organizations identify, evaluate, and document potential impacts throughout the AI system lifecycle.*

Microsoft Responsible AI Impact Assessment Template

- *The Responsible AI Impact Assessment Template is the product of a multi-year effort at Microsoft to define a process for assessing the impact an AI system may have on people, organizations, and society.*

Implementation Guiding Questions

Below are standard questions that HDOs may ask to guide discussion for **control 4.2.5**:

- Risk-Benefit Profile: Does the clinical or operational benefit of the AI solution outweigh the potential risks identified in earlier assessments?
- Stakeholder Impact: How does this AI affect different groups, including those who may not interact with it directly but are impacted by its outputs?

- Bias and Disparity Considerations: Does the impact assessment specifically address how the system might exacerbate or reduce existing health disparities?
- Accessibility: Is the AI's output presented in a way that is comprehensible and usable for the clinicians or patients receiving the information?
- Does the deployment of this AI solution create or alter financial incentives that could influence clinical behavior, documentation practices, or care resource allocation?
- Have potential sources of bias in the AI system been explicitly identified, documented, and addressed as part of the impact assessment (separate from, but complementary to, risk categorization and assessment)?

Real-World Example

Explore the below real-world example of how an HDO is successfully approaching risk and impact assessments. This example offers valuable insights and practical strategies that you can adapt to your own organization.

Example #1: AI Governance at Tampa General Hospital (TGH) - Risk Review Process

The AI Governance Committee at TGH evaluates all solutions containing Artificial Intelligence capabilities through one of three review categories: (1) Full Product Review, (2) Expedited Review, or (3) Exempt Review.

Evaluation Criteria for the 3 Review Categories:

(1) Full Product Review – discussion and vote

- **Clinical management impact**: The AI drives, suggests, or changes diagnosis /treatment/triage (i.e., not just information display).
- **Automation/agentic behavior**: Any tool that does not rely on human-in-the-loop design.
- **High severity if wrong**: Failure could plausibly cause patient harm (morbidity/mortality), major privacy/security incident, or major reputational/legal risk.
- **Sensitive data**: Uses PHI at scale, free-text notes, audio, images, or data types with re-identification risk.
- **Novel/insufficient evidence**: New model, new workflow, or limited validation in similar settings.

Minimum packet: use case description, workflow map, data flows, vendor model card, validation evidence, human-in-the-loop controls, downtime/rollback plan, monitoring metrics (override rate, unsafe recommendation rate, etc.)

(2) Expedited Review – consent agenda

ALL OF THE FOLLOWING MUST BE TRUE:

- **Supportive decision support only**: Provides information/support but does **not** drive management or decision making, AND user remains the decision-maker.
- **Low-to-moderate severity if wrong** and clear mitigations (required user sign-off, audit logs, easy rollback).
- **Limited data exposure:** minimal or no PHI.
- Prior evidence of safe use in similar health systems/workflows.

Minimum packet: use case description, workflow map, data flows, vendor model card, validation evidence, human-in-the-loop controls, downtime/rollback plan, monitoring metrics (override rate, unsafe recommendation rate, etc.)

(3) Exempt Review – no vote needed

ALL OF THE FOLLOWING MUST BE TRUE:

- **No patient data** (or truly de-identified with minimal re-ID risk).
- **No impact on patient care/clinical decisions**.
- **Commodity functionality** (i.e., generic meeting transcription not used for clinical documentation).

Minimum packet: use case description, data flows, vendor model card, validation evidence, downtime/rollback plan

Based on which category of review a tool requires, TGH pulls in different AI Governance Assessment Team subject matter experts (SMEs) to assess the solution. The TGH Assessment team is made up of SMEs from 6 different evaluation domains: (1) Cybersecurity, (2) Utility & Validity, (3) Compliance, (4) Risk, (5) Ethics, (6)

Educational Impact. Below is the AI Governance Tool Evaluation Matrix, which outlines which domains are required to assess solutions based on their review category.

	Cybersecurity	Utility & Validity	Compliance	Risk	Ethics	Educational Impact
Full Review *discussion & vote*	●	●	●	●	●	●
Expedited *consent agenda*	●	●	●	●	●	●
Exempt *no vote required*	●	●				

● Required assessment domain □ Not required for this review category

Full Product Review:

- Requires assessments from all 6 evaluation domains
- Discussion period allotted for any questions or concerns
- Committee votes to approve or deny solution at monthly Governance meeting

Expedited Review:

- Requires assessments from all 6 evaluation domains
- No discussion period for Expedited tools.
- Committee consent agenda votes to approve or deny solution at monthly Governance meeting

Exempt Review:

Tools & Resources

Included below are tools and resources that may be helpful for developing and operationalizing your own risk and impact assessments.

CHAI Risk Categorization Tool (v2)

- *The Risk Categorization Tool (v2) is for health systems (any size) and teams responsible for pre-deployment risk review. This tool supports early AI governance by helping identify low, medium, or high-risk levels across several risk modifiers related to the risk domains (1) Life & Patient Safety and (2) Technology & Data. The tool highlights key dimensions of risk to guide further assessment, mitigation, and monitoring.*

CHAI Risk Mapping: CHAI Risk Categorization Tool (v2) mapped to NIST AI RMF and EU AI Act

- *CHAI developed a detailed information mapping between the CHAI Risk Categorization Tool (v2), National Institute of Standards and Technology (NIST) AI Risk Management Framework (RMF), and European Union (EU) AI Act.*

CHAI Partner Ecosystem

- *Collaborate with trusted partners delivering excellence in AI governance, testing, data management, and advisory services - all aligned with the CHAI AI Governance Framework.*

ISO 14971:2019: Medical devices — Application of risk management to medical devices

- *This standard establishes a systematic framework for identifying hazards associated with medical devices, estimating and evaluating the associated risks, controlling these risks, and monitoring the effectiveness of the controls throughout the entire device lifecycle.*

ISO/IEC 42005:2025: Information technology — Artificial intelligence (AI) — AI system impact assessment

- *This standard supports transparency, accountability, and trust in AI by helping organizations identify, evaluate, and document potential impacts throughout the AI system lifecycle.*

Microsoft Responsible AI Impact Assessment Template

- *The Responsible AI Impact Assessment Template is the product of a multi-year effort at Microsoft to define a process for assessing the impact an AI system may have on people, organizations, and society.*

Security Risk Assessment (SRA) Tool (v3.6)

- *The Office of the National Coordinator for Health Information Technology (ONC), in collaboration with the U.S. Department of Health and Human Services (HHS) Office for Civil Rights (OCR), developed a downloadable Security Risk Assessment (SRA) Tool. The target audience of this tool is medium- and small-sized providers.*

NIST AI Risk Management Framework (AI RMF) & Playbook

- *A comprehensive framework for managing AI risks, including a structured playbook with actionable guidance across the AI lifecycle.*

International Medical Device Regulators Forum (IMDRF) Software as a Medical Device (SaMD) Clinical Risk Categorization Framework

- *An international framework for categorizing the clinical risk of Software as a Medical Device.*

Failure Modes and Effects Analysis (FMEA)

- *A structured methodology for identifying, prioritizing, and mitigating potential failure points in AI systems. FMEA is widely used in healthcare quality and safety contexts and can be adapted for AI risk assessment workflows.*

Example Template: Risk & Impact Assessment Phases as Governance Decision Gates

- *An illustrative example that highlights how an organization may approach establishing a formal decision gate within the organization's governance framework, including a defined decision required to advance, a named accountable owner, and documented evidence of completion.*

Next Steps

Links below will connect you to other AI governance playbooks in this series.

Attribution

CHAI AI Governance Playbooks were developed through the contributions and collaboration of more than 150 organizations across the healthcare ecosystem. CHAI extends its sincere appreciation to all participating organizations and contributors for their time, expertise, and thoughtful feedback throughout the development process. We would also like to provide special recognition to the individuals below who requested attribution by name for their contributions to these efforts:

- Josh Hjelmstad, MHA, AVIA Health
- Sandra Gregg, DNP, MHA, RN, PMP, LSSBB, CCMP, HCD, Booz Allen Hamilton, Inc.
- Vidhi Vinay Kothari, Carnegie Mellon University
- Stephen McLaughlin, PhD, Children's National Hospital
- Ryan Marshall Felder, PhD HEC-C, Cleveland Clinic
- Po-Hao Chen, MD, MBA, Cleveland Clinic Foundation
- Katherine Mulready, JD MPS, cliexa
- Holden Groves, MD, Columbia University/NewYork-Presbyterian
- Shakira J. Grant, MBBS, MSCR, CROSS Global Research & Strategy
- Namrata Rastogi, MBBS MRCGP MS, Enigma Ventures
- Aleocidio Sette Balzanelo, MD, MBA, FOLKS
- Rocco Orlando, MD, Hartford HealthCare
- Michael Blumental, HealthLeap
- Romanus M. Joseph, Innovative Health Accountable Care Organization (ACO)
- Rebecca G. Mishuris, MD, MS, MPH, Mass General Brigham
- Nasibeh Zanjirani Farahani, PhD, CSM, Mayo Clinic
- Jason Gillman, MD, FIDSA, MCG Health
- Douglas Graham, Individual Contributor
- Nicole Petroski, Mercy Health Services
- Haris Shuaib, Newton's Tree
- Rémy Ibarcq, Individual Contributor
- Srikar Nallan, MS, Stanford Healthcare
- Jaimie Weber, MD, Tampa General Hospital
- R. Brandon Hunter, MD, FAAP, Texas Children's Hospital
- Angela L. Lipscomb, JD, LL.M, The University of Texas MD Anderson Cancer Center
- Christopher H. Lee, MBA, BSN, RN-BC, UCLA Health
- Abby Steele, Individual Contributor

- Shiba Kuanar, University of Minnesota
- Joshua Miller, JD, CHC, University of Rochester Medicine
- Anne Travis, MD, MSc, Wolters Kluwer
- Michael L. Shaw, JD, ZS

License