Risk Categorization Tool

A tool for health systems (of any size) to evaluate
Life and Patient Safety risks for health AI solutions

CHAI

Coalition for Health AI

Contents

Overview

CHAI defines risk management for health AI solutions 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 your 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.

Current Focus

The Risk Categorization Tool helps health systems (of any size) evaluate Life and Patient Safety risks. Future versions will include additional domains, such as Technology, Data, and Financial risks.

Description

Tool Name: Risk Categorization Tool

Primary Risk Domain: Life and Patient Safety

Who Should Use It: Health systems (any size) and teams responsible for pre-deployment risk review. Depending on team structure, one or more members may rate one or more modifiers. We recommend that all applicable modifiers are completed.

Purpose

This tool supports early AI governance by helping identify Low, Medium, or High risk levels across several risk modifiers related to life and patient safety. The Tool does not produce a single risk score. The Tool highlights key dimensions of risk to guide further assessment, mitigation, and monitoring. Important to note: if your AI solution qualifies as Software as a Medical Device (SaMD), follow FDA regulations and guidance. This Tool applies to all other AI solutions.

What You Need Before Starting

Use CHAI's Applied Model Card or other documentation to gather details about: Intended Use and Workflow, Primary Users, Target Patient Population, and other relevant context. In particular, leverage the CHAI Applied Model Card section "Uses and Directions" for general use case information, "Ongoing Maintenance" to support Risk Modifier #06, and "AI System Facts" to support Risk Modifier #05. Lastly, for Risk Modifiers #09-11, the team(s) responsible for pre-deployment risk review will need to identify how the AI solution will be used in the context of your organization.

How to Score & Interpret Results
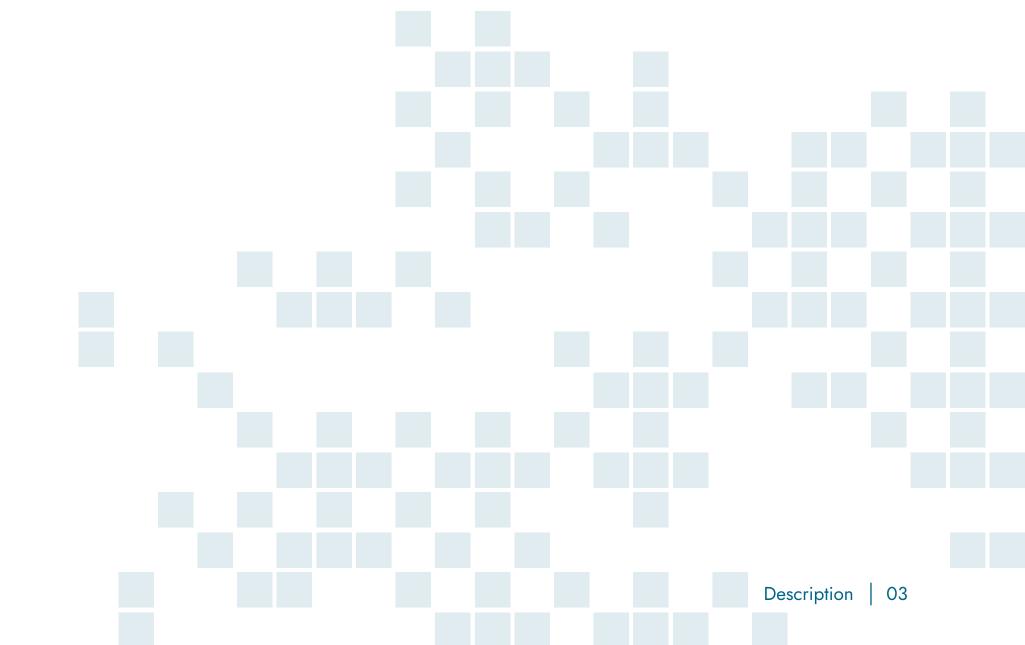
Each risk modifier is scored independently as Low, Medium, or High. If any modifier is High, a detailed risk assessment (hazard, harm, probability) is strongly recommended.

Use the results to plan mitigation strategies aligned with the risk levels of each modifier.

How to Use

Refer to the following **Example (pages 4-9)** for a step-by-step tutorial and completed illustration.

Risk Categorization Tool

Risk Domain:
Life & Patient Safety

Primary Audience:
Health systems (any size) and teams responsible for pre-deployment risk review

Use Case:
AI-assisted Patient Scheduling software (e.g., scheduling chatbot) - for outpatient, primary care clinics - simplifies booking, rescheduling, and managing patient appointments. These products allow patients to select appointment times and providers to manage their calendars, reduce no-shows, and optimize clinic workflows. Typical features include automated reminders, real-time availability updates, and integration with EHRs. Human confirms appointment once scheduled

Step 1: Align on Use Case
Review the AI solution's use case using CHAI's Applied Model Card or equivalent documentation.

#	Factor	Low Risk Definition	Medium Risk Definition	High Risk Definition	Team Ratings	Rationale/Evidence	Response	Action(s)	Notes & Comments
01		...pact on individual patient ...back-end functions such ...ice administrative tasks, ...alth analysis, or workflow	Indirect impact on patient care, access to care, or informational use, such as scheduling, transportation, non-clinical informational chatbots	AI solution has semi-direct involvement in patient care; such as, used by a healthcare professional as part of a broader clinical judgment; AI solution is directly involved in patient care/patient interaction	Low / Medium / High		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
02	Human In The Loop — The extent to which human oversight is involved in reviewing, verifying, or overriding the AI solution outputs before they affect patient care	AI solution output always reviewed by provider before any action taken	AI solution output has optional human in loop review by provider before any action taken	AI solution output is never reviewed by provider before an action is taken	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
03	Consequences Of Failure Or Error — The severity and likelihood of negative outcomes (e.g., morbidity, mortality) if the AI solution fails or provides incorrect information; clinical consequences are higher risk	AI solution has no direct impact and has no affect on patient harm	Errors may lead to temporary discomfort or inconvenience, with no lasting health effects (e.g., minor delays in care); Errors may result in temporary or reversible harm that requires medical intervention (e.g., prescribing the wrong medication dose that requires monitoring but does not cause long-term damage)	Errors may lead to permanent harm, permanent damage to body structure, disability, or death (e.g., AI misinterprets critical diagnostic imaging or fails to detect sepsis)	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
04	Patient Population Vulnerability — The degree to which the patient population affected is vulnerable (e.g., pediatrics, elderly, low health literacy, marginalized groups); Depends on clinical setting and presentation context	Used with patients who are noncomplex and stable	Used with patients who are medically complex but stable (e.g., patients with heart failure but on stable medication, being seen by primary care physicians)	Used with patients who are medically complex but unstable (e.g., patients with heart failure and in unstable state)	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
05	Level Of Difficulty Monitoring AI Solution Output — How robust is the AI solution's monitoring capabilities? How resource intensive will the AI solution be to monitor output and performance?; Depends on both the AI solution provider and health system capabilities	Embedded real-time monitoring and/or capability of real-time monitoring	AI solution includes partial real-time monitoring capabilities; health system still requires partial development of monitoring capabilities; periodic reports	Monitoring needs to be developed before implementation of solution; and/or manual monitoring that requires resource intensive activities	Low / Medium / High		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
06	Data Transparency — The clarity, completeness, and accessibility of the data sources and datasets used to train, test, and validate the AI solution	Health system has complete access to training data of the underlying model(s) for the AI solution; lowest level of detail for the data/datasets are shared and available (e.g., AI solution developed internally)	Health system has partial access to training data of the underlying model(s) for the AI solution; some level of detail for the data/datasets are shared and available	Health system has no access to training data of the underlying model(s) for the AI solution; no components of the data/datasets are shared or available (e.g., data provenance and data catalog/dictionary unavailable)	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
07	Clinical Level Of Care — Does the AI solution operate in a clinically sensitive or high-risk setting that requires a higher level of care (e.g., inpatient, outpatient, emergency department, etc.)	AI solution used in outpatient and non-critical settings (e.g., outpatient)	AI solution used in inpatient or urgent, but non-critical settings (e.g., inpatient)	AI solution used in life-critical settings (e.g., emergency department)	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
08	React Time — Assuming the AI output is incorrect, how quickly a decision or intervention can be made	There will be time for reaction and response planning before serious consequences of the risk	There will be limited time for reaction and response planning before serious consequences of the risk	There will be very little or no time for reaction and response planning before serious consequences of the risk	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
09	Breadth Of Potential Harm — The breadth of potential harm the AI solution could cause to patients if it performs incorrectly; Assess how broadly the AI solution is deployed across locations or institutions	Affects a single individual or a small number of patients in a limited number of settings (e.g., rare disease diagnostics, single-department pilot, one site, one clinic, or limited geographical area)	Affects a moderate number of patients (e.g., roughly half of patient population), possibly across multiple units or clinics (e.g., diabetes prediction across outpatient clinics)	Potential for widespread harm—across facilities, populations, or entire health systems (e.g., enterprise-wide triage algorithm, regional EMS AI for trauma prioritization)	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
10	Integrated Error Propagation Risk — The degree to which the AI solution's integration within the broader health IT environment increases the potential for errors to cascade across systems, workflows, and clinical decisions. This includes both the breadth of technical integration and the depth of inter-dependence, reflecting how embedded the AI solution is and how errors in one part could propagate to others.	The AI solution is functionally isolated, with minimal integration into other digital systems or workflows. Errors are unlikely to spread beyond the immediate user or use case.	The AI solution is integrated into specific modules or workflows but has limited cross-functional connections. Errors could impact related components but are unlikely to cause widespread disruptions.	The AI solution is deeply embedded across multiple systems and workflows. Its outputs are widely relied upon and shared, increasing the chance that a single point of failure could cascade across care settings, decisions, or resource allocations.	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
11	Population Sensitivity Or Disparity Risk — The risk that the AI solution could exacerbate health disparities or biases affecting sensitive populations based on race, gender, SES, etc.	There is minimal to no risk of the AI solution's output contributing to health disparities.	There is risk of the AI solution contributing to health disparities, especially if mitigation strategies are not implemented effectively or continuously evaluated.	There could be significant risk of the AI solution contributing to health disparities, such as high potential to cause harm through unequal diagnosis, treatment, or outcomes; the system could reinforce or worsen existing healthcare inequities, especially for vulnerable groups.	Low / Medium / High / N/A		Low / Medium / High	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	

Risk Categorization Tool

Risk Domain:
Life & Patient Safety

Primary Audience:
Health systems (any size) and teams responsible for pre-deployment risk review

Use Case:
AI-assisted Patient Scheduling software (e.g., scheduling chatb... ...ng, and managing patient appointments. These products allow patients to select appoi... ...ws, and optimize clinic workflows. Typical features include automated reminders, real-t... ...intment once scheduled

> **Step 2: Review Risk Modifiers**
> Assess each risk modifier individually. Use the provided definitions for Low, Medium, and High risk.

#	Risk Modifier	Low Risk Definition	Medium Risk Definition	High Risk Definition	Team Ratings	Rationale/Evidence	Response	Action(s)	Notes & Comments
01	**Distance From Patient** How physically or operationally close the AI solution is to the patient	No direct impact on individual patient care, support back-end functions such as back office administrative tasks, population health analysis, or workflow optimization	Indirect impact on patient care, access to care, or informational use, such as scheduling, transportation, non-clinical informational chatbots	AI solution has semi-direct involvement in patient care; such as, used by a healthcare professional as part of a broader clinical judgment; AI solution is directly involved in patient care/patient interaction	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
02	**Human In The Loop** The extent to which human oversight is involved in reviewing, verifying, or overriding the AI solution outputs before they affect patient care	AI solution output always reviewed by provider before any action taken	AI solution output has optional human in loop review by provider before any action taken	AI solution output is never reviewed by provider before an action is taken	● Low ● Medium ● High ● N/A		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
03	**Consequences Of Failure Or Error** The severity and likelihood of negative outcomes (e.g., morbidity, mortality) if the AI solution fails or provides incorrect information; clinical consequences are higher risk	AI solution has no direct impact and has no affect on patient harm	Errors may lead to temporary discomfort or inconvenience, with no lasting health effects (e.g., minor delays in care); Errors may result in temporary or reversible harm that requires medical intervention (e.g., prescribing the wrong medication dose that requires monitoring but does not cause long-term damage)	Errors may lead to permanent harm, permanent damage to body structure, disability, or death (e.g., AI misinterprets critical diagnostic imaging or fails to detect sepsis)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
04	**Patient Population Vulnerability** The degree to which the patient population affected is vulnerable (e.g., pediatrics, elderly, low health literacy, marginalized groups); Depends on clinical setting and presentation context	Used with patients who are noncomplex and stable	Used with patients who are medically complex but stable (e.g., patients with heart failure but on stable medication, being seen by primary care physicians)	Used with patients who are medically complex but unstable (e.g., patients with heart failure and in unstable state)	● Low ● Medium ● High ● N/A		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
05	**Level Of Difficulty Monitoring AI Solution Output** How robust is the AI solution's monitoring capabilities? How resource intensive will the AI solution be to monitor output and performance?; Depends on both the AI solution provider and health system capabilities	Embedded real-time monitoring and/or capability of real-time monitoring	AI solution includes partial real-time monitoring capabilities; health system still requires partial development of monitoring capabilities; periodic reports	Monitoring needs to be developed before implementation of solution; and/or manual monitoring that requires resource intensive activities	● Low ● Medium ● High ● N/A		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
06	**Data Transparency** The clarity, completeness, and accessibility of the data sources and datasets used to train, test, and validate the AI solution	Health system has complete access to training data of the underlying model(s) for the AI solution; lowest level of detail for the data/datasets are shared and available (e.g., AI solution developed internally)	Health system has partial access to training data of the underlying model(s) for the AI solution; some level of detail for the data/datasets are shared and available	Health system has no access to training data of the underlying model(s) for the AI solution; no components of the data/datasets are shared or available (e.g., data provenance and data catalog/dictionary unavailable)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
07	**Clinical Level Of Care** Does the AI solution operate in a clinically sensitive or high-risk setting that requires a higher level of care (e.g., inpatient, outpatient, emergency department, etc.)	AI solution used in outpatient and non-critical settings (e.g., outpatient)	AI solution used in inpatient or urgent, but non-critical settings (e.g., inpatient)	AI solution used in life-critical settings (e.g., emergency department)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
08	**React Time** Assuming the AI output is incorrect, how quickly a decision or intervention can be made	There will be time for reaction and response planning before serious consequences of the risk	There will be limited time for reaction and response planning before serious consequences of the risk	There will be very little or no time for reaction and response planning before serious consequences of the risk	● Low ● Medium ● High ● N/A		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
09	**Breadth Of Potential Harm** The breadth of potential harm the AI solution could cause to patients if it performs incorrectly; Assess how broadly the AI solution is deployed across locations or institutions	Affects a single individual or a small number of patients in a limited number of settings (e.g., rare disease diagnostics, single-department pilot, one site, one clinic, or limited geographical area)	Affects a moderate number of patients (e.g., roughly half of patient population), possibly across multiple units or clinics (e.g., diabetes prediction across outpatient clinics)	Potential for widespread harm—across facilities, populations, or entire health systems (e.g., enterprise-wide triage algorithm, regional EMS AI for trauma prioritization)	● Low ● Medium ● High ● N/A		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
10	**Integrated Error Propagation Risk** The degree to which the AI solution's integration within the broader health IT environment increases the potential for errors to cascade across systems, workflows, and clinical decisions. This includes both the breadth of technical integration and the depth of inter-dependence, reflecting how embedded the AI solution is and how errors in one part could propagate to others.	The AI solution is functionally isolated, with minimal integration into other digital systems or workflows. Errors are unlikely to spread beyond the immediate user or use case.	The AI solution is integrated into specific modules or workflows but has limited cross-functional connections. Errors could impact related components but are unlikely to cause widespread disruptions.	The AI solution is deeply embedded across multiple systems and workflows. Its outputs are widely relied upon and shared, increasing the chance that a single point of failure could cascade across care settings, decisions, or resource allocations.	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
11	**Population Sensitivity Or Disparity Risk** The risk that the AI solution could exacerbate health disparities or biases affecting sensitive populations based on race, gender, SES, etc.	There is minimal to no risk of the AI solution's output contributing to health disparities.	There is risk of the AI solution contributing to health disparities, especially if mitigation strategies are not implemented effectively or continuously evaluated.	There could be significant risk of contributing to health disparities, such as high potential to cause harm through unequal diagnosis, treatment, or outcomes; the system could reinforce or worsen existing healthcare inequities, especially for vulnerable groups.	● Low ● Medium ● High ● N/A		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	

Risk Categorization Tool

Risk Domain:
Life & Patient Safety

Primary Audience:
Health systems (any size) and teams responsible for pre-deployment risk review

Use Case:
AI-assisted Patient Scheduling software (e.g., scheduling chatbot) - for outpatient, primary care clinics - simplifies booking, rescheduling, and manage appointments. These products allow patients to select appointment times and providers to manage their calendars, reduce no-shows, and opt workflows. Typical features include automated reminders, real-time availability updates, and integration with EHRs. Human confirms appointment once

> **Step 3: Determine Risk Level**
> Record the team's ratings, including rationale and supporting evidence, as able.

	Risk Modifier	Low Risk Definition	Medium Risk Definition	High Risk Definition	Team Ratings	Rationale/Evidence	Response	Action(s)	Notes & Comments
01	**Distance From Patient** How physically or operationally close the AI solution is to the patient	No direct impact on individual patient care, support back-end functions such as back office administrative tasks, population health analysis, or workflow optimization	Indirect impact on patient care, access to care, or informational use, such as scheduling, transportation, non-clinical informational chatbots	AI solution has semi-direct involvement in patient care; such as, used by a healthcare professional as part of a broader clinical judgment; AI solution is directly involved in patient care/patient interaction	● Low 5 ● Medium 0 ● High 3	*Detailed rationale, artifacts, and supporting evidence, as able.*	● Low ● Medium ● High	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
02	**Human In The Loop** The extent to which human oversight is involved in reviewing, verifying, or overriding the AI solution outputs before they affect patient care	AI solution output always reviewed by provider before any action taken	AI solution output has optional human in loop review by provider before any action taken	AI solution output is never reviewed by provider before an action is taken	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
03	**Consequences Of Failure Or Error** The severity and likelihood of negative outcomes (e.g., morbidity, mortality) if the AI solution fails or provides incorrect information; clinical consequences are higher risk	AI solution has no direct impact and has no affect on patient harm	Errors may lead to temporary discomfort or inconvenience, with no lasting health effects (e.g., minor delays in care); Errors may result in temporary or reversible harm that requires medical intervention (e.g., prescribing the wrong medication dose that requires monitoring but does not cause long-term damage)	Errors may lead to permanent harm, permanent damage to body structure, disability, or death (e.g., AI misinterprets critical diagnostic imaging or fails to detect sepsis)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
04	**Patient Population Vulnerability** The degree to which the patient population affected is vulnerable (e.g., pediatrics, elderly, low health literacy, marginalized groups); Depends on clinical setting and presentation context	Used with patients who are noncomplex and stable	Used with patients who are medically complex but stable (e.g., patients with heart failure but on stable medication, being seen by primary care physicians)	Used with patients who are medically complex but unstable (e.g., patients with heart failure and in unstable state)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
05	**Level Of Difficulty Monitoring AI Solution Output** How robust is the AI solution's monitoring capabilities? How resource intensive will the AI solution be to monitor output and performance?; Depends on both the AI solution provider and health system capabilities	Embedded real-time monitoring and/or capability of real-time monitoring	AI solution includes partial real-time monitoring capabilities; health system still requires partial development of monitoring capabilities; periodic reports	Monitoring needs to be developed before implementation of solution; and/or manual monitoring that requires resource intensive activities	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
06	**Data Transparency** The clarity, completeness, and accessibility of the data sources and datasets used to train, test, and validate the AI solution	Health system has complete access to training data of the underlying model(s) for the AI solution; lowest level of detail for the data/datasets are shared and available (e.g., AI solution developed internally)	Health system has partial access to training data of the underlying model(s) for the AI solution; some level of detail for the data/datasets are shared and available	Health system has no access to training data of the underlying model(s) for the AI solution; no components of the data/datasets are shared or available (e.g., data provenance and data catalog/dictionary unavailable)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
07	**Clinical Level Of Care** Does the AI solution operate in a clinically sensitive or high-risk setting that requires a higher level of care (e.g., inpatient, outpatient, emergency department, etc.)	AI solution used in outpatient and non-critical settings (e.g., outpatient)	AI solution used in inpatient or urgent, but non-critical settings (e.g., inpatient)	AI solution used in life-critical settings (e.g., emergency department)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
08	**React Time** Assuming the AI output is incorrect, how quickly a decision or intervention can be made	There will be time for reaction and response planning before serious consequences of the risk	There will be limited time for reaction and response planning before serious consequences of the risk	There will be very little or no time for reaction and response planning before serious consequences of the risk	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
09	**Breadth Of Potential Harm** The breadth of potential harm the AI solution could cause to patients if it performs incorrectly; Assess how broadly the AI solution is deployed across locations or institutions	Affects a single individual or a small number of patients in a limited number of settings (e.g., rare disease diagnostics, single-department pilot, one site, one clinic, or limited geographical area)	Affects a moderate number of patients (e.g., roughly half of patient population), possibly across multiple units or clinics (e.g., diabetes prediction across outpatient clinics)	Potential for widespread harm—across facilities, populations, or entire health systems (e.g., enterprise-wide triage algorithm, regional EMS AI for trauma prioritization)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
10	**Integrated Error Propagation Risk** The degree to which the AI solution's integration within the broader health IT environment increases the potential for errors to cascade across systems, workflows, and clinical decisions. This includes both the breadth of technical integration and the depth of inter-dependence, reflecting how embedded the AI solution is and how errors in one part could propagate to others.	The AI solution is functionally isolated, with minimal integration into other digital systems or workflows. Errors are unlikely to spread beyond the immediate user or use case.	The AI solution is integrated into specific modules or workflows but has limited cross-functional connections. Errors could impact related components but are unlikely to cause widespread disruptions.	The AI solution is deeply embedded across multiple systems and workflows. Its outputs are widely relied upon and shared, increasing the chance that a single point of failure could cascade across care settings, decisions, or resource allocations.	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	
11	**Population Sensitivity Or Disparity Risk** The risk that the AI solution could exacerbate health disparities or biases affecting sensitive populations based on race, gender, SES, etc.	There is minimal to no risk of the AI solution's output contributing to health disparities.	There is risk of the AI solution contributing to health disparities, especially if mitigation strategies are not implemented effectively or continuously evaluated.	There could be significant risk of contributing to health disparities, such as high potential to cause harm through unequal diagnosis, treatment, or outcomes; the system could reinforce or worsen existing healthcare inequities, especially for vulnerable groups.	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls Apply organizational medium risk mitigation controls Apply organizational high risk mitigation controls	

Risk Categorization Tool

Risk Domain:
Life & Patient Safety

Primary Audience:
Health systems (any size) and teams responsible for pre-deployment risk review

Use Case:
AI-assisted Patient Scheduling software (e.g., scheduling chatbot) - for outpatient, primary care clinics - simplifies booking, rescheduling, a~~
appointments. These products allow patients to select appointment times and providers to manage their calendars, reduce no-shows, ~~
workflows. Typical features include automated reminders, real-time availability updates, and integration with EHRs. Human confirms appointm~~

Step 4: Document Risk Level

Based on the most frequently selected risk level from team ratings, assign an overall risk level (Low, Medium, High) for each risk modifier in the "Response" column.

*Based on the responses to each modifier (if you have more than one team member responding), consider toggling the "Response" field using the majority response. If there is a tie, consider discussing further to gain majority consensus or defaulting to the higher risk category.

	Risk Modifier	Low Risk Definition	Medium Risk Definition	High Risk Definition	Team Ratings	Rationale/Evidence	Response	Action(s)	Notes & Comments
01	**Distance From Patient** How physically or operationally close the AI solution is to the patient	No direct impact on individual patient care, support back office administrative tasks, such as back office administrative tasks, population health analysis, or workflow optimization	Indirect impact on patient care, access to care, or informational use, such as scheduling, transportation, non-clinical informational chatbots	AI solution has semi-direct involvement in patient care; such as, used by a healthcare professional as part of a broader clinical judgment; AI solution is directly involved in patient care/patient interaction	● Low 5 ● Medium 0 ● High 3	*Detailed rationale, artifacts, and supporting evidence, as able.*	● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
02	**Human In The Loop** The extent to which human oversight is involved in reviewing, verifying, or overriding the AI solution outputs before they affect patient care	AI solution output always reviewed by provider before any action taken	AI solution output has optional human in loop review by provider before any action taken	AI solution output is never reviewed by provider before an action is taken	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
03	**Consequences Of Failure Or Error** The severity and likelihood of negative outcomes (e.g., morbidity, mortality) if the AI solution fails or provides incorrect information; clinical consequences are higher risk	AI solution has no direct impact and has no affect on patient harm	Errors may lead to temporary discomfort or inconvenience, with no lasting health effects (e.g., minor delays in care); Errors may result in temporary or reversible harm that requires medical intervention (e.g., prescribing the wrong medication dose that requires monitoring but does not cause long-term damage)	Errors may lead to permanent harm, permanent damage to body structure, disability, or death (e.g., AI misinterprets critical diagnostic imaging or fails to detect sepsis)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
04	**Patient Population Vulnerability** The degree to which the patient population affected is vulnerable (e.g., pediatrics, elderly, low health literacy, marginalized groups); Depends on clinical setting and presentation context	Used with patients who are noncomplex and stable	Used with patients who are medically complex but stable (e.g., patients with heart failure but on stable medication, being seen by primary care physicians)	Used with patients who are medically complex but unstable (e.g., patients with heart failure and in unstable state)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
05	**Level Of Difficulty Monitoring AI Solution Output** How robust is the AI solution's monitoring capabilities? How resource intensive will the AI solution be to monitor output and performance?; Depends on both the AI solution provider and health system capabilities	Embedded real-time monitoring and/or capability of real-time monitoring	AI solution includes partial real-time monitoring capabilities; health system still requires partial development of monitoring capabilities; periodic reports	Monitoring needs to be developed before implementation of solution; and/ or manual monitoring that requires resource intensive activities	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
06	**Data Transparency** The clarity, completeness, and accessibility of the data sources and datasets used to train, test, and validate the AI solution	Health system has complete access to training data of the underlying model(s) for the AI solution; lowest level of detail for the data/datasets are shared and available (e.g., AI solution developed internally)	Health system has partial access to training data of the underlying model(s) for the AI solution; some level of detail for the data/datasets are shared and available	Health system has no access to training data of the underlying model(s) for the AI solution; no components of the data/ datasets are shared or available (e.g., data provenance and data catalog/ dictionary unavailable)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
07	**Clinical Level Of Care** Does the AI solution operate in a clinically sensitive or high-risk setting that requires a higher level of care (e.g., inpatient, outpatient, emergency department, etc.)	AI solution used in outpatient and non-critical settings (e.g., outpatient)	AI solution used in inpatient or urgent, but non-critical settings (e.g., inpatient)	AI solution used in life-critical settings (e.g., emergency department)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
08	**React Time** Assuming the AI output is incorrect, how quickly a decision or intervention can be made	There will be time for reaction and response planning before serious consequences of the risk	There will be limited time for reaction and response planning before serious consequences of the risk	There will be very little or no time for reaction and response planning before serious consequences of the risk	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
09	**Breadth Of Potential Harm** The breadth of potential harm the AI solution could cause to patients if it performs incorrectly; Assess how broadly the AI solution is deployed across locations or institutions	Affects a single individual or a small number of patients in a limited number of settings (e.g., rare disease diagnostics, single-department pilot, one site, one clinic, or limited geographical area)	Affects a moderate number of patients (e.g., roughly half of patient population), possibly across multiple units or clinics (e.g., diabetes prediction across outpatient clinics)	Potential for widespread harm—across facilities, populations, or entire health systems (e.g., enterprise-wide triage algorithm, regional EMS AI for trauma prioritization)	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
10	**Integrated Error Propagation Risk** The degree to which the AI solution's integration within the broader health IT environment increases the potential for errors to cascade across systems, workflows, and clinical decisions. This includes both the breadth of technical integration and the depth of inter-dependence, reflecting how embedded the AI solution is and how errors in one part could propagate to others.	The AI solution is functionally isolated, with minimal integration into other digital systems or workflows. Errors are unlikely to spread beyond the immediate user or use case.	The AI solution is integrated into specific modules or workflows but has limited cross-functional connections. Errors could impact related components but are unlikely to cause widespread disruptions.	The AI solution is deeply embedded across multiple systems and workflows. Its outputs are widely relied upon and shared, increasing the chance that a single point of failure could cascade across care settings, decisions, or resource allocations.	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
11	**Population Sensitivity Or Disparity Risk** The risk that the AI solution could exacerbate health disparities or biases affecting sensitive populations based on race, gender, SES, etc.	There is minimal to no risk of the AI solution's output contributing to health disparities.	There is risk of the AI solution contributing to health disparities, especially if mitigation strategies are not implemented effectively or continuously evaluated.	There could be significant risk of contributing to health disparities, such as high potential to cause harm through unequal diagnosis, treatment, or outcomes; the system could reinforce or worsen existing healthcare inequities, especially for vulnerable groups.	● Low ● Medium ● High		● Low ● Medium ● High ● N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	

Risk Categorization Tool

Risk Domain: Life & Patient Safety

Primary Audience: Health systems (any size) and teams responsible for pre-deployment risk review

Use Case:
AI-assisted Patient Scheduling software (e.g., scheduling chatbot) - for outpatient, primary care clinics - simplifies booking, rescheduling, and managing patient appointments. These products allow patients to select appointment times and providers to manage their calendars, reduce no-shows, and optimize clinic workflows. Typical features include automated reminders, real-time availability updates, and integration with EHRs. Human confirms appointment once scheduled

Step 5: Address Risk Modifiers
Apply organizational risk mitigation controls for each risk modifier

Note: CHAI has not yet developed a Risk Assessment — use your organization's processes.

If risk modifier #01 is Low risk, apply your organization's Low risk mitigation controls; if any risk modifier is rated High, conduct a rigorous risk assessment for that modifier.

#	Risk Modifier	Low Risk Definition	Medium Risk Definition	High Risk Definition	Team Ratings	Rationale/Evidence	Response	Action(s)	Notes & Comments
01	**Distance From Patient** — How physically or operationally close the AI solution is to the patient	No direct impact on individual patient care, support back office administrative tasks, such as population health analysis, or workflow optimization	Indirect impact on patient care, access to care, or informational use, such as scheduling, transportation, non-clinical informational chatbots	AI solution has semi-direct involvement in patient care; such as, used by a healthcare professional as part of a broader clinical judgment; AI solution is directly involved in patient care/patient interaction	Low 5 / Medium 0 / High 3	*Detailed rationale, artifacts, and supporting evidence, as able.*	Low	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	*Additional notes and comments*
02	**Human In The Loop** — The extent to which human oversight is involved in reviewing, verifying, or overriding the AI solution outputs before they affect patient care	AI solution output always reviewed by provider before any action taken	AI solution output has optional human in loop review by provider before any action taken	AI solution output is never reviewed by provider before an action is taken	Low 6 / Medium 2 / High 0		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
03	**Consequences Of Failure Or Error** — The severity and likelihood of negative outcomes (e.g., morbidity, mortality) if the AI solution fails or provides incorrect information; clinical consequences are higher risk	AI solution has no direct impact and has no affect on patient harm	Errors may lead to temporary discomfort or inconvenience, with no lasting health effects (e.g., minor delays in care); Errors may result in temporary or reversible harm that requires medical intervention (e.g., prescribing the wrong medication dose that requires monitoring but does not cause long-term damage)	Errors may lead to permanent harm, permanent damage to body structure, disability, or death (e.g., AI misinterprets critical diagnostic imaging or fails to detect sepsis)	Low 3 / Medium 5 / High 0		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
04	**Patient Population Vulnerability** — The degree to which the patient population affected is vulnerable (e.g., pediatrics, elderly, low health literacy, marginalized groups); Depends on clinical setting and presentation context	Used with patients who are noncomplex and stable	Used with patients who are medically complex but stable (e.g., patients with heart failure but on stable medication, being seen by primary care physicians)	Used with patients who are medically complex but unstable (e.g., patients with heart failure and in unstable state)	Low 0 / Medium 2 / High 6		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
05	**Level Of Difficulty Monitoring AI Solution Output** — How robust is the AI solution's monitoring capabilities? How resource intensive will the AI solution be to monitor output and performance?; Depends on both the AI solution provider and health system capabilities	Embedded real-time monitoring and/or capability of real-time monitoring	AI solution includes partial real-time monitoring capabilities; health system still requires partial development of monitoring capabilities; periodic reports	Monitoring needs to be developed before implementation of solution; and/or manual monitoring that requires resource intensive activities	Low / Medium / High / N/A		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
06	**Data Transparency** — The clarity, completeness, and accessibility of the data sources and datasets used to train, test, and validate the AI solution	Health system has complete access to training data of the underlying model(s) for the AI solution; lowest level of detail for the data/datasets are shared and available (e.g., AI solution developed internally)	Health system has partial access to training data of the underlying model(s) for the AI solution; some level of detail for the data/datasets are shared and available	Health system has no access to training data of the underlying model(s) for the AI solution; no components of the data/datasets are shared or available (e.g., data provenance and data catalog/dictionary unavailable)	Low / Medium / High / N/A		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
07	**Clinical Level Of Care** — Does the AI solution operate in a clinically sensitive or high-risk setting that requires a higher level of care (e.g., inpatient, outpatient, emergency department, etc.)	AI solution used in outpatient and non-critical settings (e.g., outpatient)	AI solution used in inpatient or urgent, but non-critical settings (e.g., inpatient)	AI solution used in life-critical settings (e.g., emergency department)	Low 1 / Medium 4 / High 2		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
08	**React Time** — Assuming the AI output is incorrect, how quickly a decision or intervention can be made	There will be time for reaction and response planning before serious consequences of the risk	There will be limited time for reaction and response planning before serious consequences of the risk	There will be very little or no time for reaction and response planning before serious consequences of the risk	Low 4 / Medium 0 / High 0		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
09	**Breadth Of Potential Harm** — The breadth of potential harm the AI solution could cause to patients if it performs incorrectly; Assess how broadly the AI solution is deployed across locations or institutions	Affects a single individual or a small number of patients in a limited number of settings (e.g., rare disease diagnostics, single-department pilot, one site, one clinic, or limited geographical area)	Affects a moderate number of patients (e.g., roughly half of patient population), possibly across multiple units or clinics (e.g., diabetes prediction across outpatient clinics)	Potential for widespread harm—across facilities, populations, or entire health systems (e.g., enterprise-wide triage algorithm, regional EMS AI for trauma prioritization)	Low 0 / Medium 3 / High 1		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
10	**Integrated Error Propagation Risk** — The degree to which the AI solution's integration within the broader health IT environment increases the potential for errors to cascade across systems, workflows, and clinical decisions. This includes both the breadth of technical integration and the depth of inter-dependence, reflecting how embedded the AI solution is and how errors in one part could propagate to others.	The AI solution is functionally isolated, with minimal integration into other digital systems or workflows. Errors are unlikely to spread beyond the immediate user or use case.	The AI solution is integrated into specific modules or workflows but has limited cross-functional connections. Errors could impact related components but are unlikely to cause widespread disruptions.	The AI solution is deeply embedded across multiple systems and workflows. Its outputs are widely relied upon and shared, increasing the chance that a single point of failure could cascade across care settings, decisions, or resource allocations.	Low 0 / Medium 3 / High 1		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	
11	**Population Sensitivity Or Disparity Risk** — The risk that the AI solution could exacerbate health disparities or biases affecting sensitive populations based on race, gender, SES, etc.	There is minimal to no risk of the AI solution's output contributing to health disparities.	There is risk of the AI solution contributing to health disparities, especially if mitigation strategies are not implemented effectively or continuously evaluated.	There could be significant risk of contributing to health disparities, such as high potential to cause harm through unequal diagnosis, treatment, or outcomes; the system could reinforce or worsen existing healthcare inequities, especially for vulnerable groups.	Low 0 / Medium 4 / High 0		Low / Medium / High / N/A	Apply organizational low risk mitigation controls / Apply organizational medium risk mitigation controls / Apply organizational high risk mitigation controls	

Risk Categorization Tool

Example

Risk Domain:
Life & Patient Safety

Primary Audience:
Health systems (any size) and teams responsible for pre-deployment risk review

Use Case:
AI-assisted Patient Scheduling software (e.g., scheduling chatbot) - for outpatient, primary care clinics - simplifies booking, rescheduling, and managing patient appointments. These products allow patients to select appointment times and providers to manage their calendars, reduce no-shows, and optimize clinic workflows. Typical features include automated reminders, real-time availability updates, and integration with EHRs. Human confirms appointment once scheduled

Step 6: Complete All Modifiers
Ensure all risk modifiers (1—11) are reviewed, categorized, and documented.

Risk Modifier	Low Risk Definition	Medium Risk Definition	High Risk Definition	Team Ratings	Rationale/Evidence	Response	Action(s)	Notes & Comments
01 Distance From Patient — How physically or operationally close the AI solution is to the patient	No direct impact on individual patient care, support back-end administrative tasks, such as back office administrative tasks, population health analysis, or workflow optimization	Indirect impact on patient care, access to care, or informational use, such as scheduling, transportation, non-clinical informational chatbots	AI solution has semi-direct involvement in patient care; such as, used by a healthcare professional as part of a broader clinical judgment; AI solution is directly involved in patient care/patient interaction	Low 5; Medium 0; High 3	*Detailed rationale, artifacts, and supporting evidence, as able.*	**Low**; Medium; High; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*
02 Human In The Loop — The extent to which human oversight is involved in reviewing, verifying, or overriding the AI solution outputs before they affect patient care	AI solution output always reviewed by provider before any action taken	AI solution output has optional human in loop review by provider before any action taken	AI solution output is never reviewed by provider before an action is taken	Low 6; Medium 2; High 0	*Detailed rationale, artifacts, and supporting evidence, as able.*	**Low**; Medium; High; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*
03 Consequences Of Failure Or Error — The severity and likelihood of negative outcomes (e.g., morbidity, mortality) if the AI solution fails or provides incorrect information; clinical consequences are higher risk	AI solution has no direct impact and has no affect on patient harm	Errors may lead to temporary discomfort or inconvenience, with no lasting health effects (e.g., minor delays in care); Errors may result in temporary or reversible harm that requires medical intervention (e.g., prescribing the wrong medication dose that requires monitoring but does not cause long-term damage)	Errors may lead to permanent harm, permanent damage to body structure, disability, or death (e.g., AI misinterprets critical diagnostic imaging or fails to detect sepsis)	Low 4; Medium 4; High 0	*Detailed rationale, artifacts, and supporting evidence, as able.*	Low; **Medium**; High; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*
04 Patient Population Vulnerability — The degree to which the patient population affected is vulnerable (e.g., pediatrics, elderly, low health literacy, marginalized groups); Depends on clinical setting and presentation context	Used with patients who are noncomplex and stable	Used with patients who are medically complex but stable (e.g., patients with heart failure but on stable medication, being seen by primary care physicians)	Used with patients who are medically complex but unstable (e.g., patients with heart failure and in unstable state)	Low 0; Medium 2; High 6	*Detailed rationale, artifacts, and supporting evidence, as able.*	Low; Medium; **High**; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*
05 Level Of Difficulty Monitoring AI Solution Output — How robust is the AI solution's monitoring capabilities? How resource intensive will the AI solution be to monitor output and performance?; Depends on both the AI solution provider and health system capabilities	Embedded real-time monitoring and/or capability of real-time monitoring	AI solution includes partial real-time monitoring capabilities; health system still requires partial development of monitoring capabilities; periodic reports	Monitoring needs to be developed before implementation of solution; and/or manual monitoring that requires resource intensive activities	Low 0; Medium 0; High 0	*Abstain b/c don't have this use case information*	Low; Medium; High; **N/A**	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	
06 Data Transparency — The clarity, completeness, and accessibility of the data sources and datasets used to train, test, and validate the AI solution	Health system has complete access to training data of the underlying model(s) for the AI solution; lowest level of detail for the data/datasets are shared and available (e.g., AI solution developed internally)	Health system has partial access to training data of the underlying model(s) for the AI solution; some level of detail for the data/datasets are shared and available	Health system has no access to training data of the underlying model(s) for the AI solution; no components of the data/datasets are shared or available (e.g., data provenance and data catalog/dictionary unavailable)	Low 0; Medium 0; High 0	*Abstain b/c don't have this use case information*	Low; Medium; High; **N/A**	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	
07 Clinical Level Of Care — Does the AI solution operate in a clinically sensitive or high-risk setting that requires a higher level of care (e.g., inpatient, outpatient, emergency department, etc.)	AI solution used in outpatient and non-critical settings (e.g., outpatient)	AI solution used in inpatient or urgent, but non-critical settings (e.g., inpatient)	AI solution used in life-critical settings (e.g., emergency department)	Low 1; Medium 4; High 2	*Detailed rationale, artifacts, and supporting evidence, as able.*	Low; **Medium**; High; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*
08 React Time — Assuming the AI output is incorrect, how quickly a decision or intervention can be made	There will be time for reaction and response planning before serious consequences of the risk	There will be limited time for reaction and response planning before serious consequences of the risk	There will be very little or no time for reaction and response planning before serious consequences of the risk	Low 4; Medium 0; High 0	*Detailed rationale, artifacts, and supporting evidence, as able.*	**Low**; Medium; High; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*
09 Breadth Of Potential Harm — The breadth of potential harm the AI solution could cause to patients if it performs incorrectly; Assess how broadly the AI solution is deployed across locations or institutions	Affects a single individual or a small number of patients in a limited number of settings (e.g., rare disease diagnostics, single-department pilot, one site, one clinic, or limited geographical area)	Affects a moderate number of patients (e.g., roughly half of patient population), possibly across multiple units or clinics (e.g., diabetes prediction across outpatient clinics)	Potential for widespread harm—across facilities, populations, or entire health systems (e.g., enterprise-wide triage algorithm, regional EMS AI for trauma prioritization)	Low 0; Medium 3; High 1	*Detailed rationale, artifacts, and supporting evidence, as able.*	Low; **Medium**; High; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*
10 Integrated Error Propagation Risk — The degree to which the AI solution's integration within the broader health IT environment increases the potential for errors to cascade across systems, workflows, and clinical decisions. This includes both the breadth of technical integration and the depth of inter-dependence, reflecting how embedded the AI solution is and how errors in one part could propagate to others.	The AI solution is functionally isolated, with minimal integration into other digital systems or workflows. Errors are unlikely to spread beyond the immediate user or use case.	The AI solution is integrated into specific modules or workflows but has limited cross-functional connections. Errors could impact related components but are unlikely to cause widespread disruptions.	The AI solution is deeply embedded across multiple systems and workflows. Its outputs are widely relied upon and shared, increasing the chance that a single point of failure could cascade across care settings, decisions, or resource allocations.	Low 0; Medium 3; High 1	*Detailed rationale, artifacts, and supporting evidence, as able.*	Low; **Medium**; High; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*
11 Population Sensitivity Or Disparity Risk — The risk that the AI solution could exacerbate health disparities or biases affecting sensitive populations based on race, gender, SES, etc.	There is minimal to no risk of the AI solution's output contributing to health disparities.	There is risk of the AI solution contributing to health disparities, especially if mitigation strategies are not implemented effectively or continuously evaluated.	There could be significant risk of contributing to health disparities, such as high potential to cause harm through unequal diagnosis, treatment, or outcomes; the system could reinforce or worsen existing healthcare inequities, especially for vulnerable groups.	Low 0; Medium 4; High 0	*Detailed rationale, artifacts, and supporting evidence, as able.*	Low; **Medium**; High; N/A	Apply organizational low risk mitigation controls; Apply organizational medium risk mitigation controls; Apply organizational high risk mitigation controls	*Additional notes and comments*

Risk Categorization Tool

Risk Domain:
[Life & Patient Safety]

Primary Audience:
Health systems (any size) and teams responsible for pre-deployment risk review

Use Case:

Risk Modifier	Low Risk Definition	Medium Risk Definition	High Risk Definition	Team Ratings	Rationale/Evidence	Response	Action(s)	Notes & Comments
01 Distance From Patient How physically or operationally close the AI solution is to the patient	No direct impact on individual patient care, support back-end functions such as back office administrative tasks, population health analysis, or workflow optimization	Indirect impact on patient care, access to care, or informational use, such as scheduling, transportation, non-clinical informational chatbots	AI solution has semi-direct involvement in patient care; such as, used by a healthcare professional as part of a broader clinical judgment; AI solution is directly involved in patient care/patient interaction	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
02 Human In The Loop The extent to which human oversight is involved in reviewing, verifying, or overriding the AI solution outputs before they affect patient care	AI solution output always reviewed by provider before any action taken	AI solution output has optional human in loop review by provider before any action taken	AI solution output is never reviewed by provider before an action is taken	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
03 Consequences Of Failure Or Error The severity and likelihood of negative outcomes (e.g., morbidity, mortality) if the AI solution fails or provides incorrect information; clinical consequences are higher risk	AI solution has no direct impact and has no affect on patient harm	Errors may lead to temporary discomfort or inconvenience, with no lasting health effects (e.g., minor delays in care); Errors may result in temporary or reversible harm that requires medical intervention (e.g., prescribing the wrong medication dose that requires monitoring but does not cause long-term damage)	Errors may lead to permanent harm, permanent damage to body structure, disability, or death (e.g., AI misinterprets critical diagnostic imaging or fails to detect sepsis)	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
04 Patient Population Vulnerability The degree to which the patient population affected is vulnerable (e.g., pediatrics, elderly, low health literacy, marginalized groups); Depends on clinical setting and presentation context	Used with patients who are noncomplex and stable	Used with patients who are medically complex but stable (e.g., patients with heart failure but on stable medication, being seen by primary care physicians)	Used with patients who are medically complex but unstable (e.g., patients with heart failure and in unstable state)	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
05 Level Of Difficulty Monitoring AI Solution Output How robust is the AI solution's monitoring capabilities? How resource intensive will the AI solution be to monitor output and performance?; Depends on both the AI solution provider and health system capabilities	Embedded real-time monitoring and/or capability of real-time monitoring	AI solution includes partial real-time monitoring capabilities; health system still requires partial development of monitoring capabilities; periodic reports	Monitoring needs to be developed before implementation of solution; and/or manual monitoring that requires resource intensive activities	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
06 Data Transparency The clarity, completeness, and accessibility of the data sources and datasets used to train, test, and validate the AI solution	Health system has complete access to training data of the underlying model(s) for the AI solution; lowest level of detail for the data/datasets are shared and available (e.g., AI solution developed internally)	Health system has partial access to training data of the underlying model(s) for the AI solution; some level of detail for the data/datasets are shared and available	Health system has no access to training data of the underlying model(s) for the AI solution; no components of the data/datasets are shared or available (e.g., data provenance and data catalog/dictionary unavailable)	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
07 Clinical Level Of Care Does the AI solution operate in a clinically sensitive or high-risk setting that requires a higher level of care (e.g., inpatient, outpatient, emergency department, etc.)	AI solution used in outpatient and non-critical settings (e.g., outpatient)	AI solution used in inpatient or urgent, but non-critical settings (e.g., inpatient)	AI solution used in life-critical settings (e.g., emergency department)	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
08 React Time Assuming the AI output is incorrect, how quickly a decision or intervention can be made	There will be time for reaction and response planning before serious consequences of the risk	There will be limited time for reaction and response planning before serious consequences of the risk	There will be very little or no time for reaction and response planning before serious consequences of the risk	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
09 Breadth Of Potential Harm The breadth of potential harm the AI solution could cause to patients if it performs incorrectly; Assess how broadly the AI solution is deployed across locations or institutions	Affects a single individual or a small number of patients in a limited number of settings (e.g., rare disease diagnostics, single-department pilot, one site, one clinic, or limited geographical area)	Affects a moderate number of patients (e.g., roughly half of patient population), possibly across multiple units or clinics (e.g., diabetes prediction across outpatient clinics)	Potential for widespread harm—across facilities, populations, or entire health systems (e.g., enterprise-wide triage algorithm, regional EMS AI for trauma prioritization)	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
10 Integrated Error Propagation Risk The degree to which the AI solution's integration within the broader health IT environment increases the potential for errors to cascade across systems, workflows, and clinical decisions. This includes both the breadth of technical integration and the depth of inter-dependence, reflecting how embedded the AI solution is and how errors in one part could propagate to others.	The AI solution is functionally isolated, with minimal integration into other digital systems or workflows. Errors are unlikely to spread beyond the immediate user or use case.	The AI solution is integrated into specific modules or workflows but has limited cross-functional connections. Errors could impact related components but are unlikely to cause widespread disruptions.	The AI solution is deeply embedded across multiple systems and workflows. Its outputs are widely relied upon and shared, increasing the chance that a single point of failure could cascade across care settings, decisions, or resource allocations.	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	
11 Population Sensitivity Or Disparity Risk The risk that the AI solution could exacerbate health disparities or biases affecting sensitive populations based on race, gender, SES, etc.	There is minimal to no risk of the AI solution's output contributing to health disparities.	There is risk of the AI solution contributing to health disparities, especially if mitigation strategies are not implemented effectively or continuously evaluated.	There could be significant risk of contributing to health disparities, such as high potential to cause harm through unequal diagnosis, treatment, or outcomes; the system could reinforce or worsen existing healthcare inequities, especially for vulnerable groups.	● Low ⬜ ● Medium ⬜ ● High ⬜		○ Low ○ Medium ○ High ○ N/A	○ Apply organizational low risk mitigation controls ○ Apply organizational medium risk mitigation controls ○ Apply organizational high risk mitigation controls	

Authors

Special thank you to the contributing members of the CHAI Risk Work Group, including:

Eric Henry, Brooke & Associates

Ashley Beecy, Sutter Health

Britt Anderson, UnitedHealth Group

Nicoleta Economou, Duke

Christina Silcox, Duke

Lindsay Mico, Providence

Vivek Tomer, Providence

Taylor Anderson, Stanford

Howard Strasberg, Wolters Kluwer

Noelle Vidal, University of California

Joshua Miller, University of Rochester Medical Center

Sam Pinson, Nixon Law

Taylor Rhoades, Mercy

Brenna Loufek, Mayo

Ben Kaplan, Mount Sinai

Christine Palermo, Encore Health

January Choy, Memorial Sloan Kettering

Brenton Hill, CHAI

Merage Ghane, CHAI

Anthony DiDonato, CHAI

Greg Shemancik, CHAI

Get in touch at
admin@chai.org

Visit our website

Connect on
LinkedIn

The previous content is work in-progress and in draft form available for public comment. The mention or sharing of any examples, products, organizations, or individuals does not indicate any endorsement of those examples, products, organizations, or individuals by the Coalition for Health AI (CHAI). Any examples provided here are still under review for alignment with existing standards and instructions. We welcome feedback and stress-testing of the tool in draft form.

The information provided in this document is for general informational purposes only and does not constitute legal advice. It is not intended to create, and receipt or review of it does not establish, an attorney-client relationship.

This document should not be relied upon as a substitute for consulting with qualified legal or compliance professionals. Organizations and individuals are encouraged to seek advice specific to their unique circumstances to ensure adherence to applicable laws, regulations, and standards.