AI Risk Management Framework

Introduction

This AI Risk Management Framework outlines the core principles and practices for the responsible development, deployment, and oversight of our AI-based healthcare technologies. This includes both the AI Summary Tool and the Ambient Listening Tool. The framework aligns with the NIST AI Risk Management Framework (AI RMF 1.0) and is designed to promote trustworthy AI by managing potential risks and adverse impacts across the AI lifecycle.

Scope and References

This framework applies to all AI-powered Decision Support Interventions (DSIs) used within our technology offerings. These tools include*:

- Summarize patient encounters (AI Summary Tool)
- Detect potential issues in a signed encounter (AI Summary Tool)
- Capture and process ambient clinical conversation (Ambient Listening Tool)

*Tools listed are RXNT current offerings, but are subject to updates, modifications, and additions.

Key References:

- 45 CFR 170.315(b)(11)
- NIST AI Risk Management Framework (AI RMF 1.0)
- HIPAA / HITECH and relevant amendments
- ONC Health IT Certification Program Resource Guide

---

1. Governance

Effective governance is essential for responsible AI development. We implement policies and controls to guide how data is **acquired, managed, and used** throughout the AI lifecycle. These practices are aligned with legal, regulatory, and ethical standards and are continuously adapted to reflect evolving expectations.

- ➢ AI governance roles and responsibilities are clearly defined and distributed across legal, engineering, operations, and clinical teams.

- ➢ Governance includes proactive engagement with stakeholders, including healthcare professionals and patients, to ensure transparency and accountability.

- ➢ The organization continuously monitors legal and regulatory updates relevant to AI in healthcare.

- ➢ Trustworthy AI characteristics are embedded in organizational decision-making, policy enforcement, and oversight.

- ➢ All data is acquired, managed, stored, and used pursuant to the Health Insurance Portability and Accountability Act (HIPAA). This includes implementing appropriate administrative, physical, and technical safeguards to protect the confidentiality, integrity, and availability of data throughout its lifecycle.

---

2. Risk Analysis

Risk analysis involves a structured assessment of the **potential** and **adverse impacts** that may arise from the use of AI tools in clinical environments associated with the following characteristics: **fairness, validity, reliability, intelligibility, safety, security, robustness,** and **privacy**.

- **Validity and Reliability:** Validity and reliability for deployed AI systems are assessed by ongoing testing and monitoring. We use our best efforts to ensure that our AI tools consistently support their intended clinical functions by conducting various tests to assess overall correctness and reliability.

- **Robustness and Safety:** Robustness is the ability of a system to maintain its level of performance under a variety of circumstances. Safety is, under defined conditions, not leading to a state in which life or health is endangered. To address robustness and safety, each system undergoes rigorous simulation and real-world testing to promote trusted performance under defined and variable conditions, including incomplete or noisy data inputs.

- **Fairness:** Fairness in AI includes concerns for equality and equity. We promote fairness by using randomized data sampling for model evaluations and conducting internal and manual reviews of training datasets and system outputs to detect and address potential bias.

- **Privacy and Security:** Privacy refers generally to the norms and practices that help to safeguard human autonomy, identity, and dignity. We implement strict data governance policies, addressing encryption, de-identification, access controls, and audit logs. Privacy risk assessments are conducted during the development and post-deployment phases.

- **Intelligibility:** Our AI tools are developed with a focus on transparency and interpretability. To support clinician understanding and trust, the system includes explanatory components that clarify how outputs are derived and how they should be interpreted within the clinical workflow.

- **Data Acquisition and Use:** The origin, integrity, and representativeness of data are carefully examined. Data is obtained through secure and compliant channels, and is governed by usage policies that ensure alignment with clinical and ethical standards.

- **Adverse Impact Analysis:** We proactively identify risks such as inaccurate transcripts, hallucinations, omissions, and automation complacency. Each identified risk is linked to specific clinical use cases to gauge severity and likelihood.

---

3. Risk Mitigation

Mitigation strategies aim to minimize identified risks by integrating protective mechanisms into every stage of the AI lifecycle.

- **Inappropriate Use Prevention:** Our systems include in-product messaging to reinforce the role of AI as a decision support tool. We provide guidance to discourage sole reliance on AI in clinical decisions.

- **Controls and Monitoring:** Continuous performance monitoring includes real-time checks, event logging, and scheduled audits. Anomalies trigger alerts and initiate root-cause investigations.

- **Performance Tracking:** We define quantitative benchmarks for accuracy, robustness, fairness, and intelligibility. Tools are re-evaluated upon any change in algorithm, data, or context of use.

- **Bias Mitigation:** Preprocessing techniques, sampling, and post hoc adjustments are applied aimed at identifying disparities, accuracy, and correctness in model behavior.

- **Feedback and Incident Management:** A structured feedback and escalation workflow is available to users. This includes triage, resolution, and retrospective analysis to ensure continuous improvement and risk reduction.

- **Governance Oversight:** Risk mitigation efforts are overseen by the AI Team, which meets as necessary to review outcomes, track residual risk, and recommend policy changes.

Conclusion

Through a structured approach to **governance**, **risk analysis**, and **mitigation**, we aim to promote safe, effective, and equitable use of AI in healthcare. Our commitment to managing risks associated with **validity, reliability, robustness, fairness, intelligibility, safety, security,** and **privacy** is ongoing and foundational to our product development and deployment. This includes continual attention to how data is **acquired, managed, and used**, as well as regular assessments of potential and adverse impacts.

This framework will be regularly reviewed and updated to reflect new evidence, stakeholder input, and regulatory guidance.

Last Updated: June 23, 2025